Comparison of travel-time and geostatistical inversion approaches for hydraulic tomography: A synthetic modeling study

Huiyang Qiu[1], Rui Hu[2], NING LUO[3], and Walter Arthur Illman[3]

[1]Hohai University
[2]School of Earth Sciences and Engineering, Hohai University
[3]University of Waterloo

November 22, 2022

Abstract

Hydraulic tomography (HT) has been proven to be an effective approach in mapping the heterogeneity of hydraulic parameters. The travel-time based inversion (TTI) and geostatistical inversion (GI) approaches are two of several HT methods. In particular, the GI approach is used to compute heterogeneous hydraulic conductivity (K) and specific storage (S_s) tomograms, while the TTI approach yields diffusivity ($D = K/S_s$) tomograms. The main objective of this paper is to evaluate these two methods through a synthetic study. Two cases are designed based on different monitoring configurations. Two independent scenarios are designed by: providing the same data sets and providing all available data for calibration, while data selection follows recommended strategies utilized by the two approaches. Then, the estimated tomograms are evaluated by visual comparison of estimated parameter distributions and assessments of model calibration and validation results. Results show that the advantages of TTI are: (1) imaging of structural features representing high D zones; (2) requirement of less data for inverse modeling; and (3) rapid computational times. In contrast, the advantages of the GI approach are: (1) the direct characterization of both K and S_s distributions; (2) better drawdown predictions; and (3) a larger estimated area. Our study suggests that the TTI approach is suitable for rapid, coarse characterization of heterogeneity that could potentially be utilized for providing hydrogeological structures for an initial model for the GI approach. The GI approach, although significantly more computationally intensive, is more robust and preferable to applications that require higher accuracy in parameter estimation.

Hosted file

essoar.10509776.1.docx available at https://authorea.com/users/540246/articles/600169-comparison-of-travel-time-and-geostatistical-inversion-approaches-for-hydraulic-tomography-a-synthetic-modeling-study

Comparison of travel-time and geostatistical inversion approaches for hydraulic tomography: A synthetic modeling study

Huiyang Qiu[1,2], Rui Hu[1], Ning Luo[2], and Walter A. Illman[2]

1. School of Earth Science and Engineering, Hohai University, Nanjing, 210000, China

2. Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON, N2L 3G1, Canada

Corresponding author: Huiyang Qiu (qiuhuiyang@hhu.edu.cn)

For submission to *Water Resources Research*

Key words: Subsurface heterogeneity, Hydraulic tomography, Travel-time based inversion, Geostatistical inverse modeling

Key points:

- Comparison of tomograms from travel time and geostatistical inversion approaches;
- Both inversion approaches have varying abilities for characterizing heterogeneity;
- Geostatistical inversion yields tomograms with better performance, while travel time inversion has significantly lower computational cost.

ABSTRACT

Hydraulic tomography (HT) has been proven to be an effective approach in mapping the heterogeneity of hydraulic parameters. The travel-time based inversion (TTI) and geostatistical inversion (GI) approaches are two of several HT methods. In particular, the GI approach is used to compute heterogeneous hydraulic conductivity (K) and specific storage (S_s) tomograms, while the TTI approach yields diffusivity ($D = K/S_s$) tomograms. The main objective of this paper is to evaluate these two methods through a synthetic study. Two cases are designed based on different monitoring configurations. Two independent scenarios are designed by: providing the same data sets and providing all available data for calibration, while data selection follows recommended strategies utilized by the two approaches. Then, the estimated tomograms are evaluated by visual comparison of estimated parameter distributions and assessments of model calibration and validation results. Results show that the advantages of TTI are: (1) imaging of structural features representing high D zones; (2) requirement of less data for inverse modeling; and (3) rapid computational times. In contrast, the advantages of the GI approach are: (1) the direct characterization of both K and S_s distributions; (2) better drawdown predictions; and (3) a larger estimated area. Our study suggests that the TTI approach is suitable for rapid, coarse characterization of heterogeneity that could potentially be utilized for providing hydrogeological structures for an initial model for the GI

approach. The GI approach, although significantly more computationally intensive, is more robust and preferable to applications that require higher accuracy in parameter estimation.

1 Introduction

The accurate characterization of hydraulic properties is critical to groundwater resources management, contaminant transport and groundwater flow modeling. Traditionally, the aquifer is treated to be homogeneous and effective hydraulic parameters, such as hydraulic conductivity (K) and specific storage (S_s), are estimated using analytical solutions (e.g., Cooper & Jacob, 1946; Theis, 1935). However, the accurate prediction of groundwater flow and contaminant transport requires more accurate descriptions of subsurface heterogeneity (e.g., Maier et al., 2009; Zheng & Gorelick, 2003).

In comparison to homogeneous models, well-characterized heterogeneous models yield better results in predicting groundwater flow and solute transport (e.g., Luo et al., 2017, 2020; Zhao & Illman, 2021). In order to characterize subsurface heterogeneity, a novel approach termed hydraulic tomography (HT) has been developed. Gottlieb & Dietrich, (1995) considered HT to be a special application of impedance tomography. Nowadays, HT has been developed into a method for highly parameterized estimation of hydraulic properties regarding their heterogeneities. The robust performance of HT in revealing subsurface heterogeneity has been validated through numerous synthetic (Bohling et al., 2002; Qiu et al., 2019; Jim Yeh & Liu, 2000; Zhu & Yeh, 2006), laboratory (e.g., Brauchler et al., 2003; Illman et al., 2007, 2010, 2012a, 2015; Liu et al., 2002; Liu et al., 2007; Yin & Illman, 2009), and field studies (e.g., Berg & Illman, 2011; Bohling et al., 2007; Brauchler et al., 2013; Cardiff et al., 2020; Zha et al., 2015; Zhao & Illman, 2018). Over the past few decades, different methods have been proposed for subsurface heterogeneity characterization using data obtained from HT surveys. Geostatistical inversion (Snodgrass & Kitanidis, 1998), travel time inversion (Brauchler et al., 2003), steady shape inversion (Bohling et al., 2007), and pilot-point (Marsily et al., 1978) are common inverse methods based on different algorithms, but all aim to reconstruct the subsurface heterogeneity of hydraulic parameters.

Geostatistical inversion (GI) approach forms the backbone of a majority of HT algorithms (Illman et al., 2015). This method estimates hydraulic parameters as well as their uncertainties at all elements or cells in the model domain by adopting a highly parameterized heterogeneous conceptual model (Kitanidis & Vomvoris, 1983; Yeh & Liu, 2000). Successive linear estimator (SLE) is a geostatistical inverse algorithm (Yeh et al., 1995, 1996) that led to the development of a code to analyze HT data. With the development of the GI approach, Xiang et al. (2009) developed the simultaneous successive linear estimator (SimSLE), which can analyze transient hydraulic data from different pumping tests simultaneously. Compared to other inverse methods, GI approach has advantages on its ability: (1) to provide uncertainty estimates; (2) for data fusion (Zha et al., 2017); (3) to yield hydraulic parameter tomograms that are able to provide

adequate validation results (Luo et al., 2017). However, there are some drawbacks to the GI approach, such as: (1) the highly parameterized geostatistical model is restricted for large-scale site characterization due to the effect of computational intensity; and that (2) the results could be impacted by the accuracy of prior information when there is a lack of hydraulic data for inversion (Luo et al., 2017).

The steady shape analysis of HT data has been evaluated as an efficient and robust approach for subsurface heterogeneity characterization (Bohling et al., 2002, 2007). Specifically, this approach interprets transient hydraulic head data under steady shape condition, in which, hydraulic head data continuously change with time, while the hydraulic gradient remains constant (Bohling et al., 2002). The advantages associated with steady shape analysis of HT data can be summarized as high computational efficiency and low impact of uncertain boundary conditions on inversion results. However, steady shape analysis can only estimate the spatial distribution of K. Meanwhile, steady shape analysis cannot analyze the drawdown data prior to the stabilization of hydraulic gradient.

The pilot point approach was first proposed by Marsily et al. (1984), and then adopted for subsurface heterogeneity characterization through many studies (e.g., Lavenue & De Marsily, 2001; Vesselinov et al., 2001). In this approach, hydraulic property values are estimated at a set of points that are distributed throughout the model domain and subsequently interpolated (Doherty, 2003). The approach is amenable to regularization, which allows for the incorporation of structure data from available measurements to reduce the issue of nonuniqueness (Jiménez et al., 2013). However, the number of pilot points and their locations are considered to have effects on the characterization of subsurface heterogeneity. LaVenue & Pickens (1992) suggested that pilot points should be positioned at locations where model sensitivity to the observation data are highest. Some scholars have also attempted to guide the distribution of pilot points through geophysical or tracer tests showing promising results (Poduri & Kambhammettu, 2020; Sanchez-León et al., 2016).

Hydraulic travel-time based inversion (TTI) approach is based on solving an eikonal equation which is derived from the groundwater flow equation (Vasco et al., 2000; Brauchler et al., 2003; Vasco & Karasaki, 2006). The simultaneous iterative reconstruction technique (Trampert & Leveque, 1990) and ray-tracing technique (Moser, 1991) are commonly used to reconstruct the diffusivity (D) of an aquifer using the TTI approach. TTI approach is known to have some attractive advantages over other inversion methods which are: (1) computational efficiency; (2) less data required for inversion; and (3) less dependency on initial conditions. However, the eikonal equation which forms the basis of the TTI approach is built on many assumptions, which limit the range of its application. The limitations include: (1) variation of aquifer properties and natural hydraulic gradient should be smooth; (2) the inversion accuracy of grid decreases with lower ray density; and (3) effects of boundary conditions are ig-

nored. To overcome these limitations: (1) Vasco (2018) developed an extended trajectory mechanics approach, which is valid for arbitrary spatial variations of aquifer properties; (2) Brauchler et al., (2013) and Jiménez et al., (2013) utilized null-space energy map to extract the area with high reliability; and (3) Vasco et al., (2019) presented a travel-time based inversion to estimate D iteratively by coupling with a groundwater forward model. This coupling allowed for the estimation of D that considers initial and boundary conditions. There are some suggestions for improving TTI: Brauchler et al. (2007) suggested to use earlier time rather than peak arrival time to reconstruct D-tomograms and Hu et al. (2011) suggested to select data with constrained angle between pumping and monitoring intervals to improve the image of layer structures.

The major drawback of the TTI approach is that the estimated parameter is diffusivity (D), which is the ratio of hydraulic conductivity (K) to specific storage (S_s). To obtain the S_s distribution directly, Brauchler et al. (2013) presented a hydraulic attenuation tomography based on the eikonal equation. Several other studies have been published that have attempted to separate K and S_s from the estimated D values (Brauchler et al., 2011, 2013; Hu et al., 2011; Jiménez et al., 2013). However, it is still common to assume a homogeneous S_s when conducting site characterizations using the TTI approach (Vasco et al., 2019).

Other than the inverse methods mentioned above, HT data can also be interpreted using radial flow model coupled with Marquardt-Levenberg inversion (Paradis et al., 2016), hybrid inverse method (Wang et al., 2017), Nelder-Mead Simplex algorithm (Cardiff et al., 2020), and Gaussian mixtures (Minutti et al., 2020) for subsurface heterogeneity characterization. The performances of these approaches in revealing heterogeneity details of hydraulic parameters have been evaluated individually through many studies, but there is still a lack of study that shows a comprehensive comparison of different approaches.

TTI and GI approaches has some attractive and complementary characteristics on several aspects. The first is that geological information can be incorporated to GI approach, while TTI approach has been used to provide the structure information (Jiménez et al., 2013). Meanwhile, hydraulic response data from pumping/slug tests can be used for both TTI and GI approach. GI approach requires the hydraulic head data, while TTI approach requires the arrival time of the peak derivative of drawdowns. More importantly, TTI approach requires significantly less computational resources than GI approach. Meanwhile, the TTI approach only requires early time pumping or slug test data, which means that using the TTI approach has low cost regarding the collection of head response data in field. Last but not the least, the utilization of GI approach is able to estimate hydraulic parameters throughout the simulation domain along with their uncertainties, while the TTI approach can only be adopted to reconstruct D tomogram for the area between pumping and observation wells. To date, there are no study which objectively compares the performance of GI and TTI approaches with data collected from the same aquifer.

In this study, a comprehensive comparison between the GI and TTI approaches

is performed to evaluate their performances in revealing subsurface heterogeneities through analogue experiments. The utilization of an analogue model with known heterogeneity patterns, no measurement error, and controlled initial and boundary conditions is of critical importance in assessing the comparison results. Since the temporal sampling strategies for GI and TTI approaches are different, the comparison is performed under two scenarios based on the fairness and preferred sampling strategies for each approach. Under each scenario, two cases with different datasets are utilized to illustrate the effect of data selection on inversion results for each approach. The comparison results are then assessed through visual comparison of the estimated hydraulic parameter tomograms, and statistical analyses of validation results. This study makes a comprehensive comparison to identify advantages of TTI and GI approaches, which can be utilized in future studies on combining the strengths of each approach.

2 Inverse Modeling Approaches

Before introducing the datasets utilized and the procedures adopted for inversions, the fundamental knowledge of investigated approaches is briefly introduced in this section. Additional details about TTI and GI approaches can be found in Brauchler et al. (2003) and Xiang et al. (2009), respectively.

2.1 Travel-time based inversion

The start of the TTI approach is a line integral equation, which describes the relationship between the peak arrival time and diffusivity (Kulkarni et al., 2001; Vasco et al., 2000):

$$\sqrt{t_{\text{peak}}\left(x_2\right)} = \frac{1}{\sqrt{6}} \int_{x_1}^{x_2} \frac{\text{ds}}{\sqrt{D(s)}} \quad (1)$$

where $t_{\text{peak}}\left(x_2\right)$ is the peak arrival time of hydraulic signal form a source x_1 (pumping port) to a receiver x_2 (observation port), s is the trajectory path of signal, and $D(s)$ is the diffusivity along the path.

Technically, equation (1) is valid only for a Dirac pulse in a homogeneous medium. Vasco et al. (2000) demonstrated that Eqn. (1) is also valid with a Heaviside source. Stated another way, they concluded that Eqn. (1) was also valid for cases with constant pumping or injecting rates.

In this study, the eikonal solvers GEOTOM3D (Jackson & Tweeton, 1996) and TOMOGO (Qiu et al., 2019) are applied for the TTI approach, in which the SIRT and ray-tracing algorithms are utilized to reconstruct the D distribution of the investigated aquifer. The eikonal solver has been widely used in previous studies for subsurface heterogeneity characterization using the TTI approach (Brauchler et al., 2003, 2007, 2011, 2013; Hu et al., 2011; Jiménez et al., 2013). Ray-tracing is used to solve the trajectory or the propagation path of the hydraulic signal as it travels through the porous medium. Referring to the Fermat's principle (Brandstatter, 1974), Brauchler et al. (2003) assumed that hydraulic

signals prefer to travel along the fastest path between the source and receivers. The eikonal equation based trajectory x follows Eqn. (2) and (3) (Vasco, 2018):

$\frac{dx}{\mathrm{dt}} = \mathbf{p}$, (2)

$\frac{d\mathbf{p}}{\mathrm{dt}} = \nabla \left[\frac{S_s}{K}\right]$, (3)

where $\mathbf{p}$ is the slowness gradient and ∇ is the gradient operator.

It is important to note that eqns. (2) and (3) do not consider the influence from the hydraulic head field in which the signal propagates. Instead, Vasco (2018) notes that the gradient of hydraulic head will influence the propagation of the hydraulic signal.

When applying the TTI approach in this study, the following strategies and operations are considered to improve the inversion results: (1) polynomial interpolation of drawdown data; (2) utilization of earlier travel time than initially considered peak arrival time; and (3) omitting data collected at observation ports forming large angles between the horizontal line and trajectory to the pumping port.

Firstly, the polynomial interpolation of drawdown data can help to obtain the travel-time with high decimal precision. For instance, Fig. S1 in the supporting information section illustrates the difference between the derivative curves of drawdown with or without the polynomial interpretation of the original drawdown data. In this study, an 8^{th}-order polynomial model is used to fit each drawdown curve prior to the computation of derivatives following the work of Liu et al. (2007).

Secondly, Brauchler et al. (2007) found that the utilization of travel time prior to the arrival of peak derivatives (early time) yielded improved TTI results in terms of revealing preferential flow paths. In their study, Eqn. (1) was modified by including a transformation factor to approximate the early time as:

$$\sqrt{t_{\alpha,d}} = \frac{1}{\sqrt{6 f_{\alpha,d}}} \int_{x_1}^{x_2} \frac{\mathrm{ds}}{\sqrt{D(s)}}, \quad (4)$$

where $t_{\alpha,d}$ is the travel time of selected point of hydraulic signal and $f_{\alpha,d} = \frac{t_{\mathrm{peak}}}{t_{\alpha,d}}$, is the related transformation factor. Fig. S2a shows an example of early time, in which, the arrival time of the 20% amplitude of peak derivative is considered as $t_{20\%}$. However, utilization of early time data for TTI approach may not be appliable to cases with closely distributed sources and receivers in a high D zone, in which, the derivative of drawdown data may yield the "peak" arrival time as the first observation point (as shown in Fig. S2b).

Thirdly, Hu et al. (2011) showed that the selection of observation ports with low angles to the pumping port was advocated for imaging horizontal and layered structures using the TTI approach. A schematic diagram showing the consider-

ation of all observation ports or only observation ports with low angles to the pumping ports is illustrated in Fig. S3.

2.2 Geostatistical inversion

Different from TTI, the GI approach estimates hydraulic parameters through the interpretation of hydraulic head data. In this study, the 3D finite element model MMOC3 (Yeh et al., 1993) is applied to generate transient hydraulic data for a synthetic model, which are then subjected to HT analysis for heterogeneity characterization using different approaches.

Transient groundwater flow can be described by the following equation:

$$\nabla \bullet \left[K\left(\mathbf{x}\right)\nabla h\right] + Q\left(x_p\right) = S_s\left(\mathbf{x}\right)\frac{\partial h}{\partial t} \quad (5)$$

subject to the following initial and boundary conditions:

$$h|_{t=0} = h_0, \quad h|_{\Gamma_1} = h_1, \text{ and } \left[K\left(\mathbf{x}\right)\nabla h\right] \bullet \ \mathbf{n}|_{\Gamma_2} = q \quad (6)$$

where, in Eqn. (5), $K\left(\mathbf{x}\right)$ is hydraulic conductivity $\left(\frac{L}{T}\right)$, h is hydraulic head (L), $Q\left(x_p\right)$ is the pumping rate $\left(\frac{L^3}{T}\right)$ at location x_p, and $S_s(x)$ is specific storage $\left(\frac{1}{L}\right)$. In Eqn. (6), h_0 represents the initial hydraulic head (L), h_1 is a constant head (L) at boundary Γ_1, q is the specific discharge $\left(\frac{L}{T}\right)$ at Neumann boundary Γ_2, and $\mathbf{n}$ is a unit vector normal to Γ_2.

In this study, the estimation of hydraulic parameters using the GI approach is carried out by the Simultaneous Successive Linear Estimator (SimSLE) developed by Xiang et al. (2009). SimSLE characterizes the heterogeneous K and S_s fields of aquifers by analyzing transient hydraulic heads. Compared with sequential successive linear estimator (SSLE) (Yeh & Liu, 2000), SimSLE has several advantages (Xiang et al., 2009): (1) high efficiency on dealing with the adjoint state equations; (2) fast speed of convergence; and (3) as the data from different pumping tests are analyzed simultaneously, it is not affected by the sequence of datasets included in the inverse model. A brief description of SimSLE is provided below.

The geostatistical inversion of transient hydraulic data using SimSLE treats the natural logarithm of K and S_s as multi-Gaussian, second-order stationary, and stochastic processes. With given unconditional means, variances, and correlation lengths of K and S_s, SimSLE starts with the cokriging of all observation data and the initial K and S_s values to create the first estimate of heterogeneous $\ln K$ and $\ln S_s$ maps. The hydraulic parameter fields are then updated using the successive linear estimator (SLE) (Yeh et al., 1996) built in SimSLE by comparing the differences between the simulated and observed hydraulic heads at

observation points, in which, the covariances of hydraulic parameters and the cross-covariances between the head measurements and estimated parameters are evaluated and updated as the inversion progresses.

The iteration stops if: (1) the spatial variance of the estimated parameters stabilizes; (2) the differences between simulated and observed heads are closer than the prescribed tolerance; and (3) iteration steps reach a user-defined maximum value.

For GI approach using transient head data, observation data at different time stages (early, intermediate, and late) are commonly suggested to fully capture the drawdown curve. However, such a temporal sampling strategy may require a great number of observation data, which in turn, would increase the computational effort for inversion. Sun et al. (2013) studied the sampling strategy for transient HT analysis using the GI approach and showed that the pairs of head data at the early time t_m and those at either the steady-state or late time could yield the best estimates of heterogeneous K and S_s. The t_m can be calculated by Eqn. (7):

$$t_m = \frac{r^2 S_s}{1.8K}, \quad (7)$$

where r is the distance between the pumping and observation ports. Note that, the value of t_m is larger than the peak arrival time ($t_{\text{peak}} = \frac{r^2 S_s}{6K}$) considered in the TTI approach.

3 Experimental setup

3.1 Model Setup

A synthetic heterogeneous aquifer was constructed in the laboratory through the cyclic-deposition of sediments through a sediment transport process in a real sandbox which was utilized previously for various HT studies (Berg & Illman, 2011; Illman et al., 2010, 2015; Luo et al., 2017; Zhao et al., 2016). The referenced sandbox (Fig. S4) is 193.0 cm in length, 82.6 cm in height, and has a thickness of 10.2 cm, which contains both high and low permeable layers. More detailed information on the synthetic aquifer and experiments conducted is available in Illman et al. (2010).

For the comparison of the TTI and GI approaches, we created a synthetic numerical model based on the above laboratory aquifer. The dimension of the analogue aquifer is 161.0 cm in length, 75.6 cm in height, and 10.2 cm in width. Based on the inversion results of Luo et al. (2017), hydraulic parameters estimated from the geological-based zonation model are utilized as the true fields for the analyses presented in this study. Figs. 1a, 1b, and 1c show the spatial distributions of hydraulic conductivity (K), specific storage (S_s), and diffusivity (D), respectively. Based on the setup of the sandbox, there are 48 ports that can be used to pump water and to record drawdowns within the aquifer (Fig.

S4). Fig. 1d shows ports of the synthetic experiments, which mimic the real experimental setup. The exact values of K, S_s, D of different layers for the analogue aquifer are provided on Table S1.

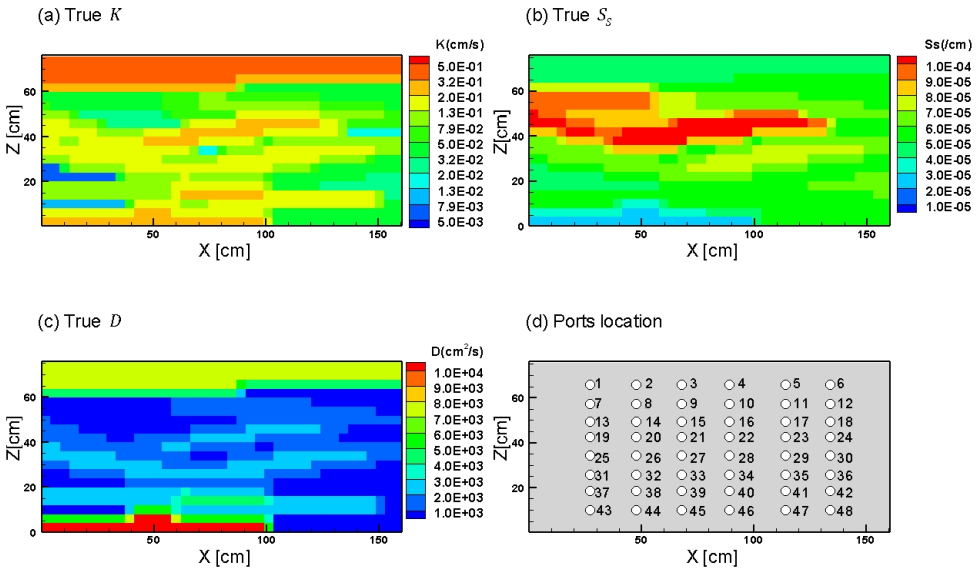

Figure 1. Referenced analogue model showing distributions of: (a) K; (b) S_s; (c) D; and (d) available ports.

The analogue model is constructed with 741 elements and 1,600 nodes. The single cuboid element of the model has dimensions of 4.1 cm $\times$ 4.1 cm $\times$ 10.2 cm. The size of the single element is maintained for both TTI and GI approaches to eliminate the influence of grid size on inversion results. As same as the sandbox, three constant head boundaries are set at left, right and top boundaries. The front, back, and bottom are treated as no-flow boundaries.

The pumping rate is set at 8 ml/s for all cases. In previous studies, the magnitude of pumping rates for both TTI and GI approaches were proven to be negligible in analogue experiments (Brauchler et al., 2003; Yeh & Liu, 2000). However, for laboratory and field experiments, signal strength should be strong enough so that it can be detected under the presence of various noises. The time interval of observation data is set as 0.1 s, which results in an observed data collection frequency of 10 Hz.

Two cases are designed to assess and compare the performances of the TTI and GI approaches, which are based on different number of observation ports. Note that, the design of the pumping and observation port arrangement avoids the

condition in which the observation port occurs on the trajectory of the other observation port (Fig. S5). This is because the TTI approach is found to yield poor results when the above condition is met in the sandbox case. The smaller distance between pumping and observation ports makes the accurate calculation of peak arrival time harder, while the smaller distance results in an earlier peak arrival time. At the sandbox scale, too early peak arrival time (Fig. S2b) brings issues which renders the deterioration of hydraulic parameter estimates. Further information of this issue is introduced subsequently.

For Case 1, eight pumping tests and data from 8 observation ports were utilized for inversion (Fig. 2a). Case 1 is a fundamental configuration which can be found in previous studies of TTI and GI approaches (e.g., Brauchler et al., 2013; Jiménez et al., 2013; Vasco, 2018; Yeh et al., 2000;). For Case 2, eight pumping ports were conducted at the middle of the aquifer, while two column ports on both sides are used as observation ports (Fig. 2b). Multiple observations ports that are installed along various boreholes within an aquifer are common for previous studies of the GI approach (Luo et al., 2017; Illman et al., 2015; Zhao et al., 2016). For the TTI approach, this type of arrangement is rare for a laboratory study because the use of multiple columns of boreholes results in the occurrence of the small distance condition which may cause some issues. For Case 2, the close distance between observation and pumping ports in this synthetic model causes excessively short peak arrival times that could result in numerical difficulties for the TTI approach as described earlier (Figs. S2b). Yet, similar configuration can also be found in the application of the TTI approach to large scale investigations (Brauchler et al., 2011, 2013; Hu et al., 2011) which provided sufficient distance between observation and pumping ports.

All inversion codes were run on the same PC with Intel core i7-7700 CPU and 16GB Random Access Memory.

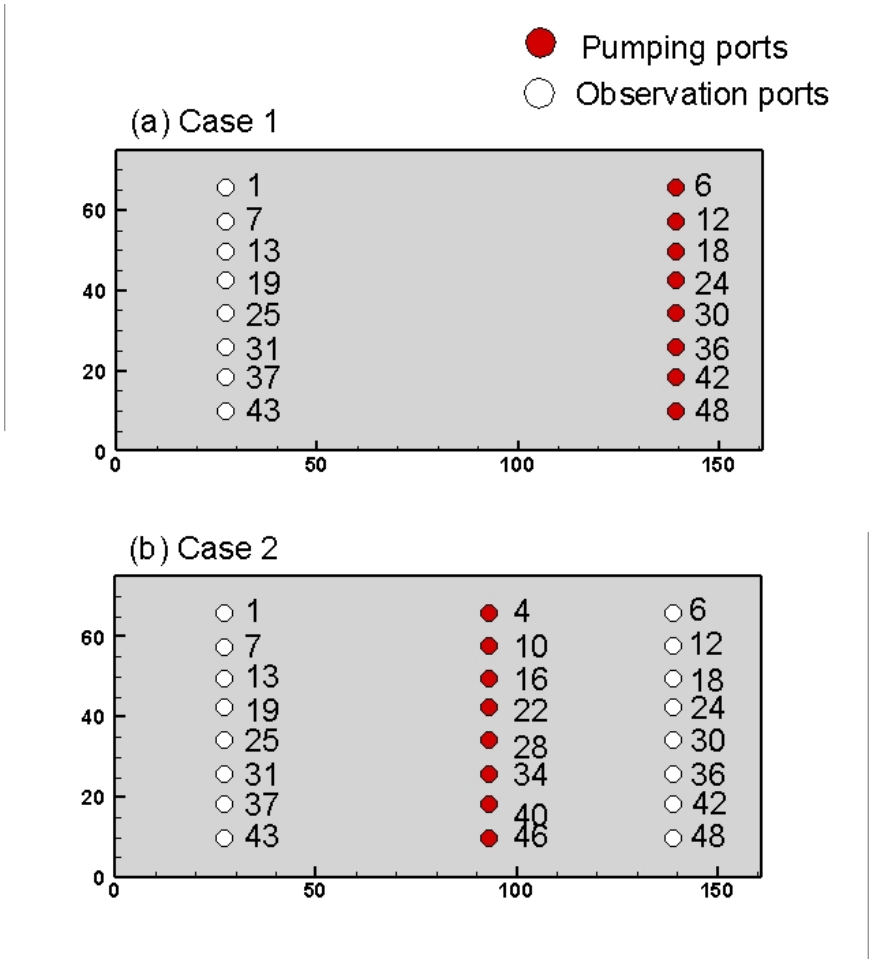

Figure 2. Sketch of sandbox and ports utilized for: (a) Case 1; and (b) Case 2. Red and white circles indicate pumping and observation ports, respectively.

3.2 Data Utilized for Inverse Modeling

Two scenarios were considered in order to achieve a fair comparison between the TTI and GI approaches, as well as to mimic a realistic HT survey. For scenario 1, data sets for different approaches were constructed with the same data points. Peak arrival times are selected as travel-time data for the TTI approach, while head data at peak arrival times are utilized for the GI approach. According to the definition of travel time, there is one travel time value in a drawdown curve generated from one constant rate pumping test. Hence, for scenario 1, on a pumping test, only one head data at the travel time at every observation port would be selected to the dataset. This is different from previous studies of the GI approach in which several data points were selected from a single pressure head curve to capture the salient features of the hydraulic behavior (Illman et

al., 2015; Luo et al., 2017; Zhao et al., 2015, 2016).

To mimic a realistic HT survey, for the second scenario, we assume that the entire drawdown data from which data can be selected for inversion. Yet, the construction of provided data sets for different approaches follows optimal operations and suggested strategies, which are previously introduced. For the TTI approach, the optimal practice involves: (1) selecting early travel-time data when the derivative rises to 20% of the maximum amplitude; and (2) maintaining three data points from each observation port with lower angle. Table S2 summarizes the data selection criteria used for the two scenarios considered by the TTI approach in this study. For the GI approach, the optimal practice is selecting sufficient data points from each observation head curve which captures the salient features of the transient drawdown behavior including inflections that may contain information on heterogeneity.

For scenario 2, the TTI approach utilizes 24 and 48 data points for Cases 1 and 2, respectively. In contrast, the GI approach utilizes 320 and 640 data points for Cases 1 and 2, respectively. The quantity of calibrated observation ports is the same for both approaches in order to maintain fairness.

In previous content, the small distance issue is introduced briefly which will be explained in further detail. The close distance between pumping and observation ports for Case 2 results in some peak arrival times to arrive as the first observation data points (Fig. S2b). When the peak arrival times occur at first observation data points, interpolation cannot be used to compute more accurate travel time at very early time, since the first data point of the derivative curve always has the maximum drawdown derivative. Generating drawdowns using a finer model discretization and selecting a smaller time interval can solve this issue, yet they will cause more problems: excessively small time steps and a very fine grid will result in significantly longer computational times. For example, in this case, a finer model grid with 6,669 (741×9) elements costs 80 times longer than the coarse grid model with 741 elements for forward simulation. More importantly, an excessively fine grid will also cause a higher degree of nonuniqueness during the inversion. Meanwhile, a time interval (0.1 s) selected for this study is small enough, while a smaller time interval (0.01 s) is impractical for existing observation equipment. For cases in which the peak times are not accurately computed, an accurate D-distribution cannot be reconstructed. However, these extreme early travel times carry information on high-D pathways that can provide structural information of the aquifer. Hence, in this study, the issue of very early travel time is considered further below.

Brauchler et al. (2007) found travel times from earlier period of drawdown derivative are better for charactering layers with higher D. However, when selecting earlier travel-time to estimate the D-tomograms, the inversion algorithm requires input travel times that come from same time period. For example, when t_{20} is selected, all data points are times when the derivative rises to 20% of the maximum amplitude. Hence, for Case 2, peak arrival times are still selected for both scenarios because the very early travel time makes the selection of earlier

period invalid. For Case 2, the numbers of data points with very early travel time were 19 and 12 for scenarios 1 and 2, respectively. Abandoning these very early travel times can make the selection of time period work. However, abandoning them would render the tomogram from the TTI approach to contain a larger area filled with mean D through the inversion algorithm. The mean D value is calculated by the average of input travel time data based on the eikonal equation. Maintaining these early travel time data in the TTI approach would provide more information for aquifer reconstruction resulting in a fair comparison between the TTI and GI approaches.

The overall workflow for both TTI and GI approaches in this study is briefly illustrated in Fig. 3. Initially, the forward simulation code MMOC3 is utilized to generate synthetic hydraulic head data. Then, several data sets are constructed based on the various cases and scenarios. TTI and GI approaches are then used to analyze the data sets to compute the estimated parameter fields or tomograms. Finally, the D tomogram from the TTI approach and K and S_s tomograms from the GI approach are evaluated by visual comparison as well as through quantitative assessment of model calibration and validation results.

In this study, a calibrated effective parameter model is used to provide homogeneous K and S_s estimates for different approaches. For TTI, during model validation, the homogeneous S_s is used to separate K ($D{=}\frac{K}{S_s}$) from estimated D for generating drawdown data by MMOC3. For the GI approach, the effective K and S_s estimates are used as initial values. The provided dataset for effective parameter estimation also follows the constraint of cases and scenarios.

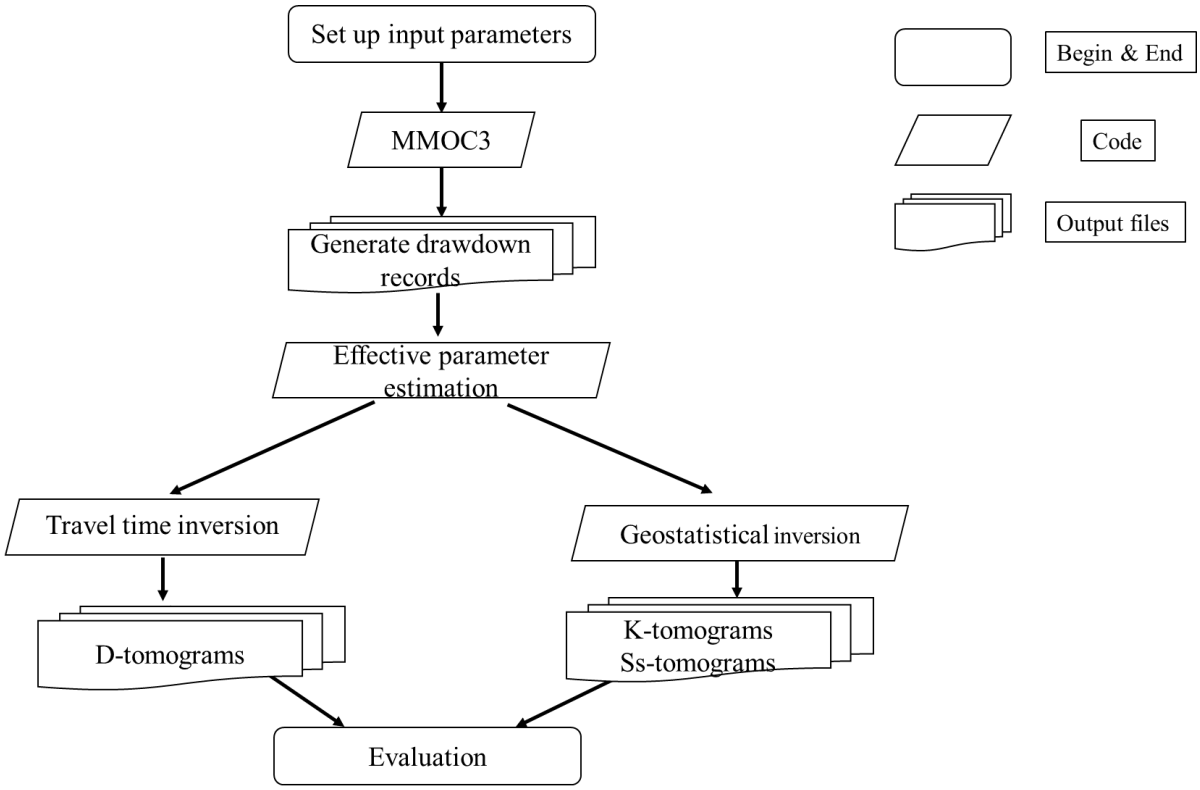

Figure 3. Work flow of TTI and GI approaches implemented in this study.

4 Results

4.1 Travel-time based inversion

Fig. 4 shows the reconstructed D tomograms from different comparison scenarios and cases along with the true D distribution of the synthetic model (Fig. 4e). To highlight the estimated structure mapped by the reconstructed D tomograms, different color scales were selected to present the results in comparison to the true D field. The estimated tomograms with the same color scale are provided in Fig. S6 of the supporting information section.

Comparing with the true D spatial distribution (as shown in Fig. 4e), most of estimated D tomograms (Figs. 4b, 4c and 4d) successfully reveal significant structural features of the aquifer through the TTI approach. The comparison of the D distribution from the true model with D tomograms from the TTI approach (Figs. 4b, 4c, 4d) reveals that the horizontal layer with high D can be well captured by the TTI approach in most situations. When data selection criteria of "constraining angle" and "earlier time selection" are considered, the estimated D tomogram from Case 1 (Fig. 4b) reveals distinct horizontal layers with high D on the top and middle of the aquifer. D tomograms estimated for

Case 2 (Figs. 4c and 4d) also show an obvious horizontal layer with high D on the top of the aquifer. This result is consistent with previous studies (Brauchler et al., 2003, 2007), which utilized the TTI approach to successfully locate high D areas.

However, there are three obvious drawbacks to the TTI approach based on the results obtained in this study. Firstly, the estimated D values at some areas (e.g. the top high zone) of the simulation domain are far from the true values, which is consistent with the finding by Jiménez et al. (2013). This drawback was fully exposed by Jiménez et al. (2013) that found obvious difference between real and estimated D values in some areas. However, the above study suggested to extract mainly the structural information rather than studying the issue further.

During the early stages of demonstrating the TTI approach (Brauchler et al., 2003, 2007), research focused on the ability of this approach in mapping the structural pattern, while revealing that the estimated D values are different from the true values. In a following synthetic study, the smoother variation of aquifer heterogeneity resolved this problem (Hu et al., 2011) showing that the TTI approach can yield improved D estimates. Then Vasco et al. (2019) compared two algorithms for trajectory calculation and found that the magnitude issue regarding estimated D values would be worsened when there are sharp changes to heterogeneity patterns, causing the assumption behind the implementation of the eikonal equation for the TTI approach to break down.

In this study, we speculate that this issue is directly caused by the fact that travel times are calibrated by the eikonal equation (Eqn. (1)) iteratively for this algorithm. The eikonal equation ignores the effects of boundary and initial conditions in the groundwater flow problem, while they obviously affect the travel times in this sandbox case which has a much smaller scale than field. In addition, the extreme small distance between pumping and observation ports and the close presence of boundaries severely undermines the infinite domain assumption embedded in the eikonal equation. There are many assumptions embedded in the eikonal equation based TTI approach, and some of them are hard to maintain. Hence, a previous study (Jiménez et al. 2013) preferred to use the TTI approach to extract geological structure information rather than absolute values of D.

The second drawback is an anomalous 'X' pattern that is visible in the D tomograms for Scenario 1 (Figs. 4a and 4c). This anomalous pattern is relieved for Scenario 2, after constraining the angle between the pumping and observation intervals (Figs. 4b, 4d). This phenomenon was fully discussed by Hu et al. (2011). In particular, Hu et al. (2011) found that the image quality of horizontally arranged layers can be significantly improved when using data subsets with small source-receiver angles. Therefore, subsequent research on the TTI approach adopted this practice of selecting data subsets with small source-receiver angles to improve the estimation of D tomograms (Brauchler et al., 2013; Jiménez et al., 2013; Jimenez et al., 2015).

The third drawback of tomograms from the TTI approach is that the estimated D tomograms between closely distributed pumping and observation ports tend to appear blurry as shown in the right half side of Figs. 4c and 4d. In particular, examination of the estimated D tomograms of Case 2 (Figs. 4c, 4d) reveals that the right half side is comprised of a blurry high D zone on the top and a low D zone on the bottom of aquifer. And a boundary which is inaccurate compared with true D distribution splits the high D zone and low D zone, on the right part of Figs. 4c, 4d. As explained earlier, the TTI approach estimates the D distribution by analyzing travel-times between pumping and observation ports. If pumping and observation ports are very close to each other and, then the travel times are almost the same so that it is hard to reveal the heterogeneity. Meanwhile, the close distance with high D results in peak arrival times that are so small that some observations receive the peak derivative of drawdowns at the first observation time. These reasons explain the poor estimated D of some areas on the right half of Figs.4c and 4d that are blurry and significantly different from the true D. On Figs.4c and 4d, the 'X' patterns are shown more obviously in the right half of the tomograms. Note, the constraining angle requirement limits the resolution since smaller angle means using less data. Hence, for TTI, Scenario 2 provides three travel time data which are close to the horizontal observation ports for every pumping ports. Resulting from a closer distance, the angle between the connection line of pumping-observation ports and horizontal line is higher on the right half of Figs.4c and 4d than the left. This phenomenon informs us that too close of a distance between pumping and observation ports would result in the poor characterization of heterogeneity by the TTI approach. In addition, the high D zone on the right upper part of the tomogram (Fig.4c and 4d) would have a higher value, which is closer to the true D, if the peaks at earlier time are available.

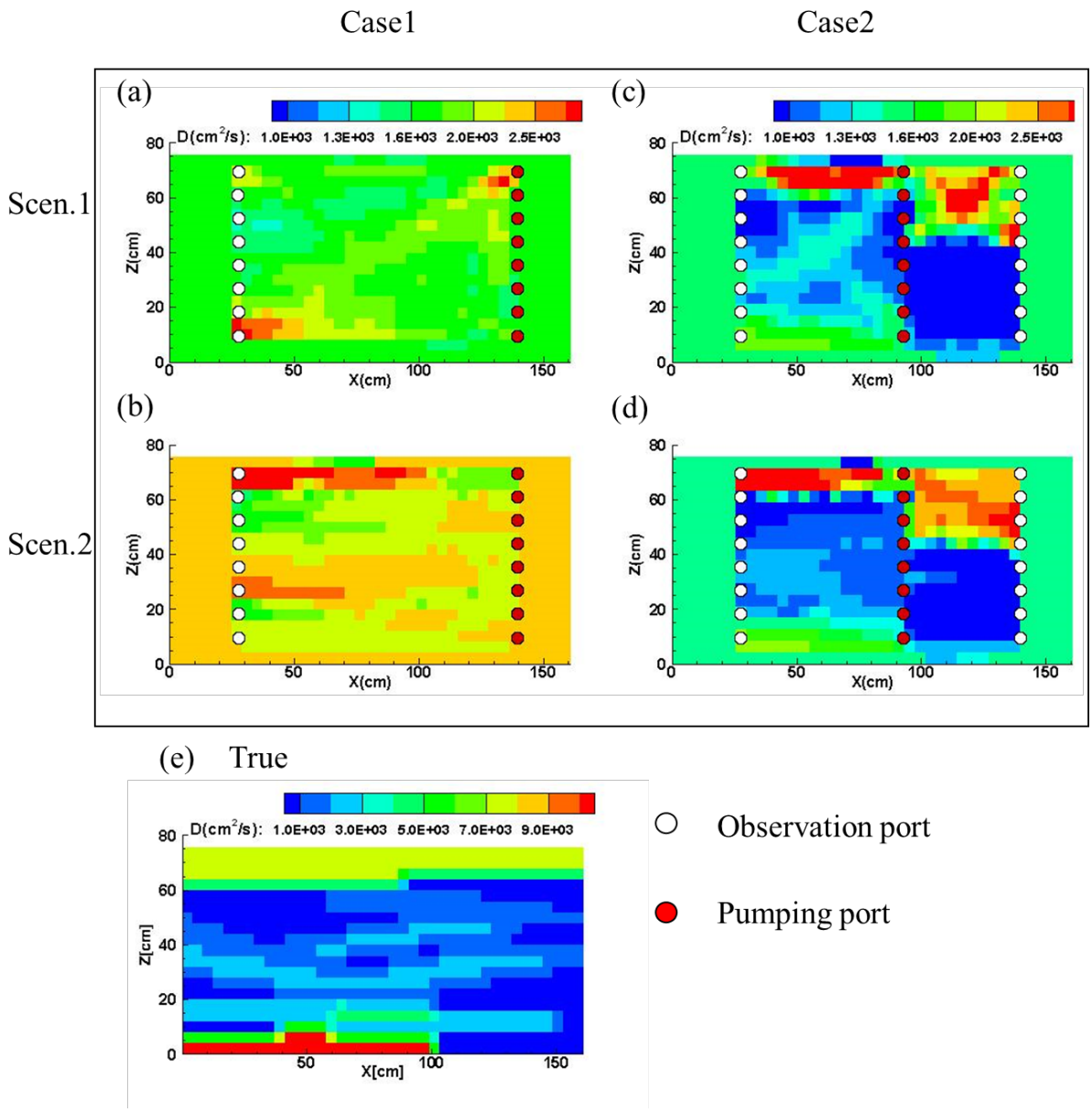

Figure 4. Comparison of reconstructed D tomograms from the TTI approach to the true D distribution: (a) Case 1 Scenario 1; (b) Case 1 Scenario 2; (c) Case 2 Scenario 1; (d) Case 2 Scenario 2; (e) True D. Reconstructed D tomograms use different color scales with True D to highlight the estimated structure.

4.2 Geostatistical inversion

In this study, the calibrated effective parameter model is utilized as initial K

and S_s fields (homogeneous) for the GI approach. Following the suggestions made by Xiang et al. (2009), inversion results are selected when the L_2 norm stabilizes. Here, stabilization in the L_2 norm is meant that the variation of adjacent iterations is smaller than 1×10^{-4}.

Fig. 5 shows the estimated K and S_s tomograms for Case 1 under two comparison scenarios along with the true K and S_s fields (Figs. 5a and 5b). Fig. 6 shows the same, but for Case 2. For Scenario 1, the K tomogram (Figs. 5c, 6c) reveals obvious heterogeneity structure with high K. For Scenario 2, K tomograms (Figs. 5e, 6e) show more details to the low K layer than the high K layer, but there is some loss of heterogeneity details in terms of the high K layer on the top and middle of the aquifer. Comparison of estimated tomograms shows that the GI approach for Case 2 Scenario 2 (Fig. 6e) has the best heterogeneity characterization of K which is close to the true field.

The ln K and ln S_s variance maps for Case 1 are provided in Fig. S7, while Fig. S8 shows the same for Case 2. Smaller variances indicate higher confidence of estimated parameters. For each case and scenario, relatively small variances are obtained near pumping and observation ports. In general, for each case and scenario, ln S_s variances are larger than those of ln K, which reveals that the reliability of S_s estimates are lower in comparison to the K estimates from the GI approach. Comparing with the true model, the S_s tomograms lose greater details in terms of stratigraphic features than K tomograms for each case and scenario. Combining these phenomena, we conclude that estimating S_s is harder than K through the GI approach, which is consistent with findings from previous studies (e.g., Luo et al., 2017; Zhao & Illman, 2021).

In order to compare the results obtained from the TTI approach, the D tomograms are calculated based on the estimated K and S_s values ($D = \frac{K}{S_s}$) from the GI approach. Fig. S9 shows the estimated distributions of D for two cases (Cases 1 and 2) under two scenarios (Scenarios 1 and 2) based on the GI results.

Examination of Fig.S9 reveals that the D tomograms obtained from the GI approach capture the heterogeneity patterns with satisfactory accuracy, although the GI approach is not developed for estimating D. Similar to the estimated K tomograms, D tomograms in Scenario 1 show a more obvious structure of high D layer than Scenario 2 on the top and middle of the aquifer (Figs.S9b, S9d). The estimated S_s tomogram does not show distinct heterogeneity features, thus the estimated K tomogram dominates the D distribution. This is the reason why the calculated D tomograms from the GI approach also show the structure of the aquifer. The comparison of visual quality and accuracy of estimated hydraulic parameters from the TTI and GI approaches are presented in the next section.

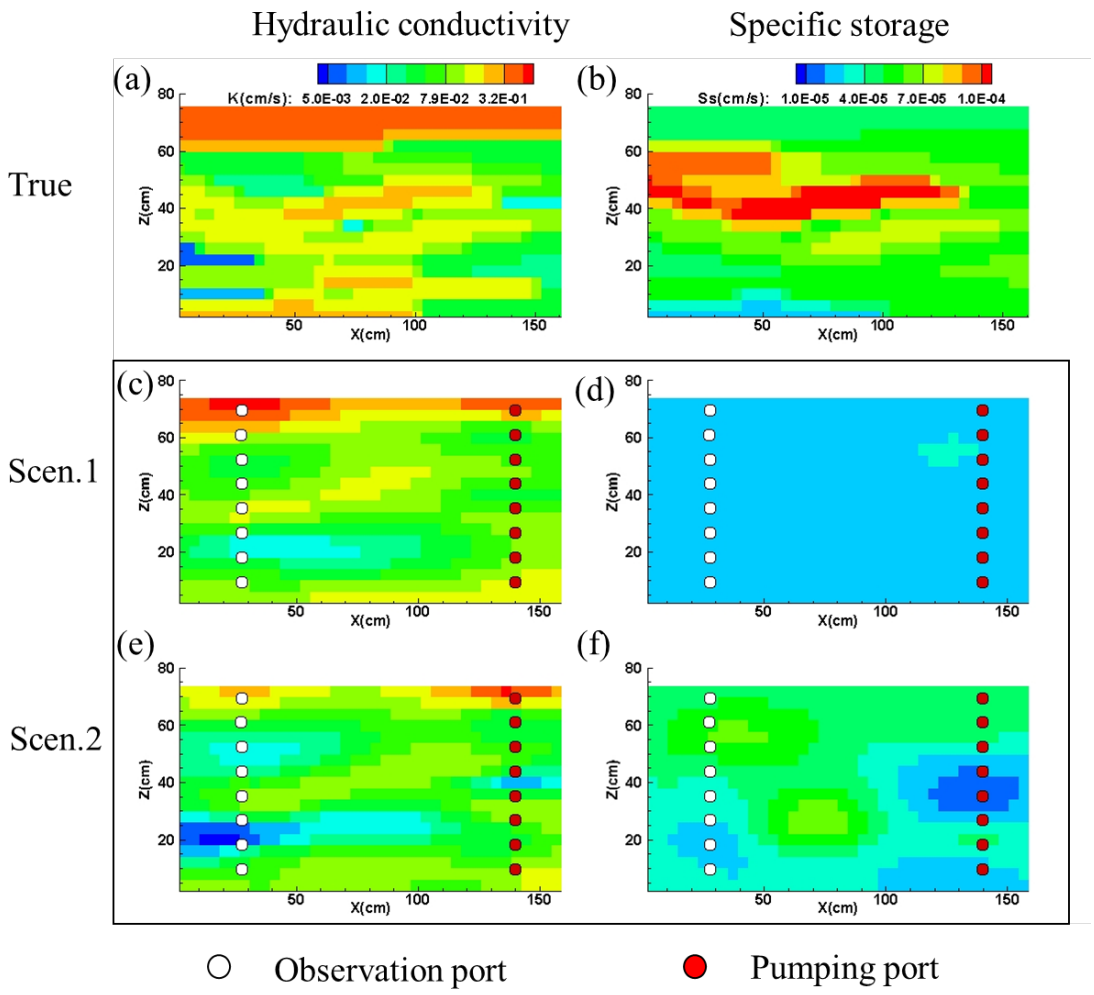

Figure 5. K and S_s tomograms from the 'true' model and geostatistical inversion for Case 1. True model: (a) K tomogram; (b) S_s tomogram. Estimated tomograms: (c) K under Scenario 1; (d) S_s under Scenario 1; (e) K under Scenario 2; and (f) S_s under Scenario 2.

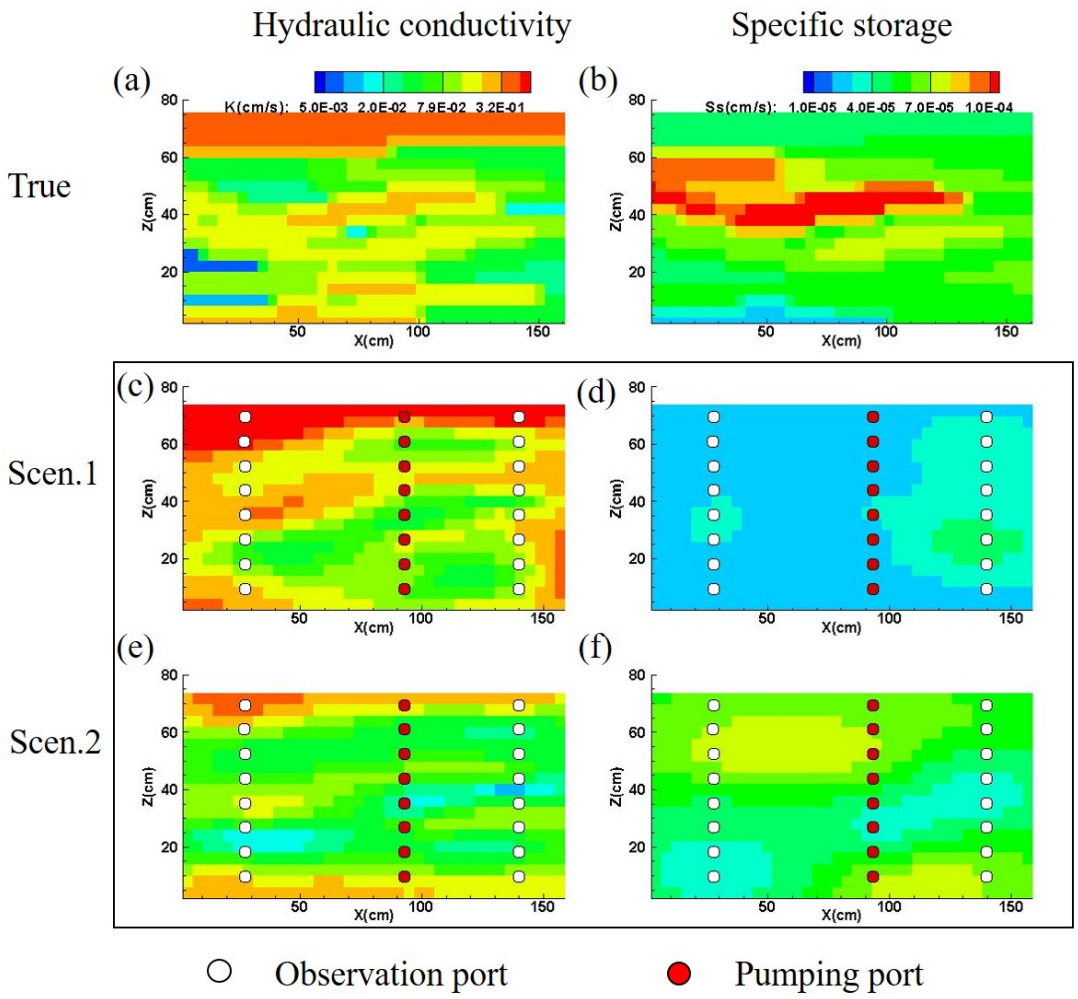

Figure 6. K and S_s tomograms from the 'true' model and geostatistical inversion for Case 2. True model: (a) K tomogram; (b) S_s tomogram. Estimated tomograms: (c) K under Scenario 1; (d) S_s under Scenario 1; (e) K under Scenario 2; and (f) S_s under Scenario 2.

Computational requirement is vital for HT as an inversion tool for mapping aquifer heterogeneity. Table 1 summarizes the time cost for all model runs and shows that: (1) the TTI approach requires considerably less computational time; (2) the TTI approach is not sensitive to the addition of more data; and (3) the GI approach costs significantly more computational time with more input data. For TTI, the selection of iteration step is based on the residual error between the calculated travel time and observed travel time. In this study, we selected the step when the residual error is lower than 1×10^{-2}. It is worth

mentioning that the time cost of the GI approach is also affected by the number of maximum iteration steps. In this study, the default maximum step is set as 100. As mentioned previously, the GI approach would converge when one of the three criteria is met. In a practical sense, based on suggestions from Xiang et al. (2009), users often select the iteration step when the L_2 norm stabilizes to a small value. Hence, there is a waste of computational power for continuing to run the GI approach after the L_2 norm stabilizes, and the solution can also diverge. For instance, in Case1 scenario 1, the GI approach with SimSLE runs over 100 iteration steps in total, yet the 33rd step is selected as the L_2 norm has stabilized for this case. Hence, the computational cost after 33 steps is wasted. This waste should be avoided, because the computational cost of the GI approach is expensive for a 3D field scale project.

Table 1: Comparison of computational times for the TTI and GI approaches.

	scenario	method	data points	time cost at selected step	Selected Iteration step
Case1	1	TTI	64	2s	7
		GI	64	53 mins	33
	2	TTI	24	2s	7
		GI	320	122 mins	70
Case2	1	TTI	128	2s	7
		GI	128	70 mins	82
	2	TTI	48	2s	7
		GI	640	170 mins	52

4.3 Visual comparison

In this section, the obtained spatial distributions of hydraulic parameters from the TTI and GI approaches are compared visually and quantitatively. Since the estimated S_s tomograms from GI approach are overall smooth and do not show any distinct layer structures, the following analyses are based on the comparison of D tomograms from the TTI approach and K as well as D tomograms from the GI approach.

On the one hand, there are some similarities between the TTI and GI approaches. In general, both methods have the ability to characterize aquifer heterogeneity (Figs. 4 - 6), although the visual quality is different for the two approaches. Meanwhile, the utilization of more observation data from additional locations can improve the visual quality of tomograms for both approaches (Figs. 4 and 6).

On the other hand, there are some differences between the TTI and GI approaches. Firstly, the GI provides estimates of heterogeneity for a larger area (as shown in Figs. 5 and 6) in comparison to the TTI approach (as shown in Fig, 4). In particular, the TTI approach only reconstructs the area between pumping and observation ports, while the GI approach extends the estimation

area to the lateral boundaries of the simulation domain. This is because the GI approach is based on the groundwater flow model and allows for the adjustment of K and S_s estimated at each element of the simulation domain. In contrast, the TTI approach only adjusts the elements between the pumping (source) and observation (receiver) ports since other elements outside the cross-well area have a little effect on the travel-time of hydraulic signal, which travels between the source (the pumping port) and receiver (the observation port). The calculation logic of the TTI approach is also supported by previous studies (Sun et al., 2013; Vasco, 2018). In particular, Sun et al. (2013) found high cross-correlation values between head and transmissivity to be confined to the area between the pumping and observation locations, while Vasco (2018) found that the sensitivities of travel-times is concentrated in a narrow zone between the pumping and observation ports. Secondly, K tomograms from the GI approach (Figs. 5e, 6e) have clearer patterns in comparison to the D tomograms obtained from the TTI approach (Fig. 4), especially in the area surrounding pumping and observation ports. We also note that the magnitude of estimated K and S_s values from the GI approach is closer to the true model (Figs. 5 and 6), while the estimated D values from TTI approach are far from the D of true model in some areas, e.g., the area of high D on top zone (Fig. 4).

In order to make a further quantitative assessment of the estimated parameters, scatterplots of the estimated versus true ln D values are plotted for both the TTI and GI approaches. Some quantitative indices, such as mean absolute error (L_1), mean square error (L_2), and the coefficient of determination (R^2) are used to quantitatively assess the comparison results. In each scatterplot, a linear model that fits all data and a 45^o line are included, along with the computed L_1, L_2, and R^2 values.

L_1, L_2 and R^2 are defined as:

$$L_1 = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}_i| \quad (8)$$

$$L_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 \quad (9)$$

$$R^2 = \left[\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_x)(\widehat{x_i}-\mu_{\widehat{x}})}{\sqrt{\sum_{i=1}^{n}(x_i-\mu_x)^2\frac{1}{n}\sum_{i=1}^{n}(\widehat{x}_i-\mu_{\widehat{x}})^2}}\right]^2 \quad (10)$$

where n is total number of data, i indicates the data number, x_i and $\hat{x}_i$ represent the value from estimated and true models, respectively. μ_x and $\mu_{\widehat{x}}$ represent averaged values from the inversion and true models, respectively.

As mentioned above, the TTI approach only reconstructs the D tomogram between pumping and observation ports, while the GI approach is able to reveal the heterogeneity throughout the simulation domain. Thus, a selected same area is adopted for the comparison between the TTI and GI approaches (estimated area in Fig. S10). Fig. 7 shows that, in each case and scenario, the linear model of fit to scatterplots of results from the GI approach has smaller intercepts and slopes that are closer to 1.0. Meanwhile, the R^2 values for the cases inverted

with the GI approach are larger than those associated with those from the TTI approach, indicating that GI models have better matches between the estimated and true D values. These quantitative indices reveal that, for the same case, GI models are more accurate than TTI models in estimating hydraulic parameters not only for Scenario 1, but also for Scenario 2.

Previously, we showed that the TTI approach has good performance on showing the structure of the top zone of the aquifer with high D (Fig.4), which is referred to as the top high D zone in this study. Meanwhile, if the pumping and observation ports are vertically distributed, the structure of the horizontal high D zone was found to be better captured than other structures in previous studies of the TTI approach (Brauchler et al., 2007; Hu et al., 2011). So, we highlight the scatters of estimated $\ln(D)$ on the top high D zone (Fig. 7). Fig. S10 reveals the compared area and the top high D zone. On the top high D zone which is highlighted as a red square, results from the GI approach also are closer to the 45^{o} line than the TT. This is also supported by the $L_1$and L_2 norms for the TTI and GI approaches for the top high D zone (Table S3). Table S3 shows that the GI approach has smaller L_1 and L_2 norms for the top high D zone which is considered as the more easily captured structure for the TTI approach. Combining above indices, on top high D zone, estimated parameters from GI approach is obviously closer to the true model than the TTI approach in three situations (Case 1 Scenario 1, Case 1 Scenario 2, Case 2 Scenario 2).

However, we notice that the configuration of ports affects the performance of TTI and GI approaches. As mentioned previously, the configuration of ports on Case 2 is not suggested for the TTI approach in a laboratory scale study, since the ray tracing technique and the simultaneous iterative reconstruction technique are commonly used for larger scale inversion with smooth variation of heterogeneity. In this study, Fig. 4 reveals that tomograms from the TTI approach have only an incremental improvement in Case 2 than Case 1, and even worse for the right blurry part of the tomograms in Case 2, which is as introduced earlier, caused by the utilization of inaccurate peak arrival times (first derivative data already as peak). This comparison is confirmed by Fig. 7c and 7d that indices (intercept, slope of linear model, L_1, L_2 and R^2) have small improvements than Fig. 7a and 7b. For the top high D zone, the performance of TTI tomograms is lower for Case 2 than Case 1. Table S3 indicates that the TTI approach for Case 1 has lower L_1, L_2 than Case 2 for the top high D zone. For the GI approach, Case 2 with optimal operations in Scenario 2 render the GI approach to generate the best tomograms, which has the best quantitative indices (Fig. 7). Fig 7f shows that only providing sufficient data points on drawdown curves for the GI approach is not sufficient for generating accurate tomograms of high resolution. For obtaining high quality tomograms from the GI approach, incorporating a sufficient number of ports and data points from drawdown curves are both vital (Fig.7h). Furthermore, for Case 1 Scenario 2, tomograms from the GI approach yields worse results for the top high D zone than Case 1 Scenario 1. This is against common sense because Case 1 Scenario 2 provides more data which includes the data on Case 1 Scenario 1, yet it yields

worse results for the top high D zone than Case 1 Scenario 1.

To study the reason of this phenomenon, an additional numerical experiment is conducted. In this experiment, the provided dataset for GI approach on Case 1 Scenario 1 is replaced with data at different times (0.5, 1, 1.5, 2, 3, 4, 5 s), while other conditions are as same as Case 1 Scenario 1. Then, the estimated K (Fig. 8) and estimated S_s tomograms (Fig. S11) are obtained which utilize different data from various times. The comparison of tomograms at different times reveals that providing earlier time data to the GI approach can improve the characterization of layer structure with high K when one data point is provided at each observation port. This phenomenon is obvious for the top zone of the aquifer with high K and S_s. And, it is consistent with the finding by Brauchler et al. (2007) for the TTI approach, which used earlier travel times to improve the characterization of the target zone (between pumping and observation ports). We speculate that the improvement of the result from incorporating earlier data for this study is caused by higher pressure head sensitivity. Vasco (2018) demonstrated the high sensitivity of earlier time pressure head is more focused on the narrow area between pumping and observation ports. Likewise, the higher sensitivity of pressure head to K for the area between pumping and observation ports may result in the improved estimation of layer structure for the GI approach.

In subsequent sections, the performances of the TTI and GI approaches are evaluated through qualitative and quantitative assessments of model calibration and validation results.

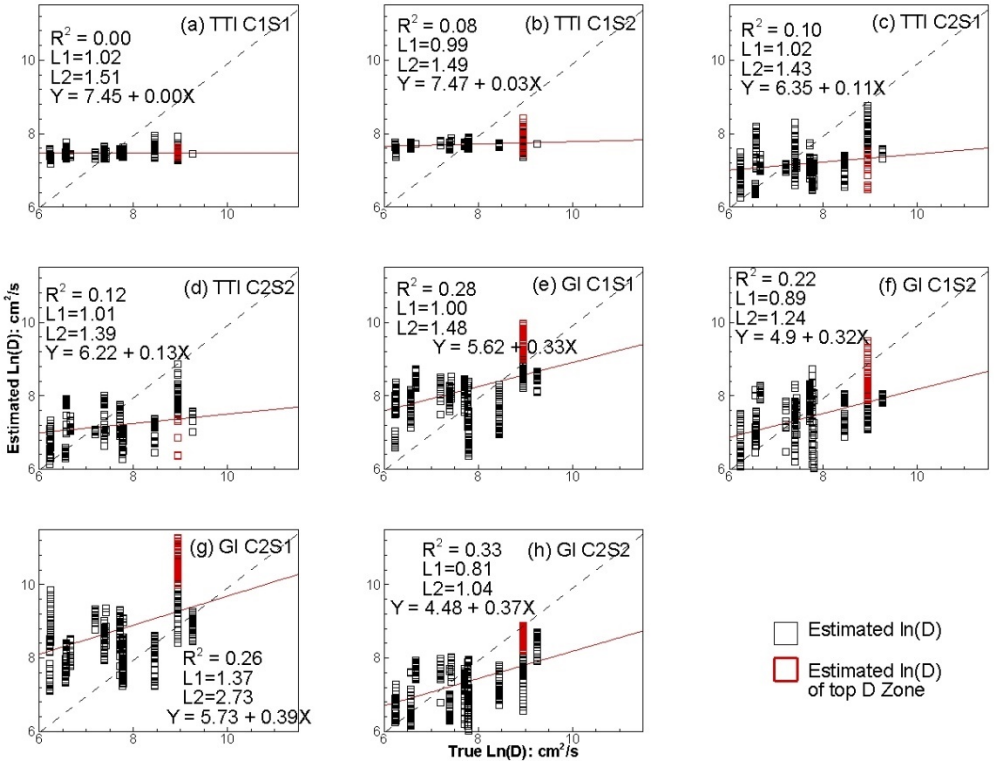

Figure 7: Comparison scatterplots of true $\ln(D)$ versus estimated $\ln(D)$ for all TTI and GI cases and scenarios. Red square is the estimated $\ln(D)$ of top high D zone. Only the area between pumping and observation ports is compared.

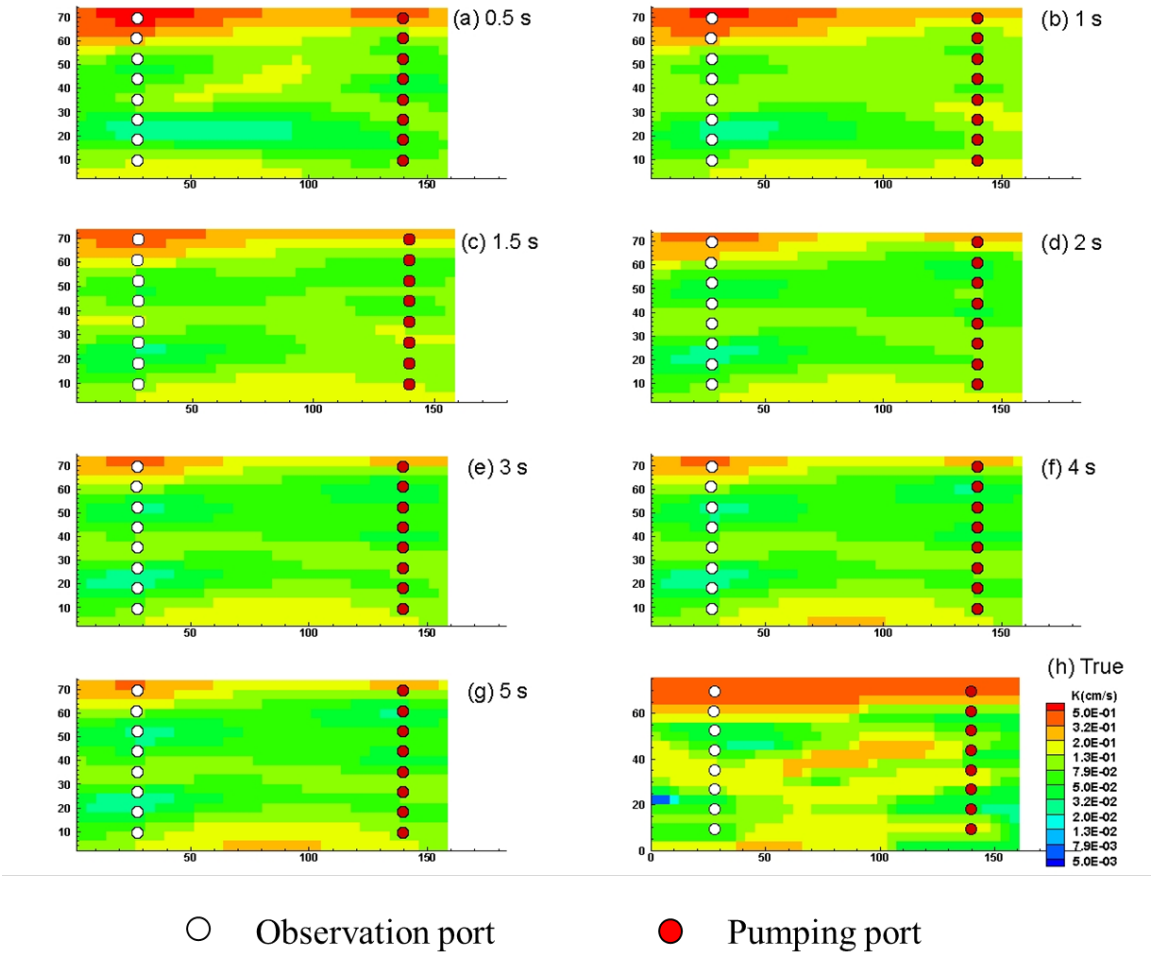

Figure 8: K tomograms from the 'true' model and geostatistical inversion for Case1 Scenario 1 with data at various times. Estimated K tomograms from GI at (a): 0.5 s; (b) 1 s; (c) 1.5 s;

(d) 2 s; (e) 3 s; (f) 4 s; (g) 5 s, while (h) is the true K.

4.4 Model calibration and validation

The calibration and validation results associated with the GI approach are assessed by comparing the scatterplots of simulated values from the estimated hydraulic parameter tomograms versus simulated values from the true model. The mean absolute error (L_1), mean square error (L_2), and the coefficient of determination (R^2) are used to quantitatively assess the calibration and validation results. In each scatterplot, a linear model that fits all data points is provided along with the corresponding R^2 value.

For the TTI approach, the simulated travel-times based from the reconstructed D tomograms are compared to the ones obtained from the true model. The residual error is used to characterize the performance of model calibration. For each case and scenario, the residual error of TTI model is lower than 1×10^{-2} s.

For the GI approach, the simulated drawdown data based on the estimated tomograms are compared to the ones obtained from the true model. Fig. S12 reveals that all geostatistical models are well calibrated. In the following section, the performance of estimated tomograms is tested through model validation. For classifying results briefly, the abbreviation of C.X-S.X is used to represent Case X Scenario X in the following content. For example, C.1-S.1 means Case 1 Scenario 1.

For model validation, the observation time interval is 0.5 s, and the simulation period is set to be 5 s, as the synthetic aquifer system reaches steady-state after 5 s after pumping begins. Two validation approaches are designed to test the performance of estimated tomograms from the TTI and GI approaches. The L_1 and L_2 norms of validation results for two plans are summarized in Tables S4 and S5 of the Supporting Information section.

The first approach is to utilize the data from observation ports which have not been used in calibration, while retaining the pumping ports the same as those designed for model calibration. The configuration of used ports for validation approach one is shown in Fig. S13.

Fig. S14 displays scatterplots of the TTI and GI approaches for Case 1 under validation approach one. Figs. S14a and S14b show that, for C.1-S.1 and C.1-S.2, the TTI approach generates reasonable drawdowns based on a near unit slope of the linear model and R^2 values. However, the simulated drawdowns of the TTI approach are biased for Case 1 with the intercepts of the linear models being 0.17 and 0.16 for scenarios 1 and 2, respectively (Fig. S14a and S14b). Fig. S14c shows that the GI approach for C.1-S.1 yields a less than satisfactory validation result based on the small slope (0.65).

Fig. S15 shows the scatterplots of the TTI and GI approaches for Case 2 under validation Plan one. Fig. S15a shows, for C.2-S.1, that the TTI model yields very biased prediction of drawdowns, which has a very large slope (2.35) and intercept (0.43) for the linear model fit. In contrast, the GI approach on C.2-S.1 yields very biased drawdowns based on a small slope (0.53) of the linear model fit in Fig. S15c. For validation approach one, the GI approach on scenario 2 yields acceptable drawdowns for both Cases 1 and 2. This is supported by the near unit slope and small intercept on Figs. S14d and S15d.

Performance metrics of the above scatterplots are compared in histograms (Fig. 9), which includes the slope, intercept, L_1 and L_2 norms, as well as R^2values. In Figs. 9d 9e, 9i and 9j, it is obvious that the optimal strategy of data incorporation for both approaches improve the tomograms and ultimately drawdown predictions. For example, the TTI approach for C.1-S.2 yields a slightly better prediction of drawdowns than C.1-S.1, while the TTI approach for C.2-S.2

yields better prediction of drawdowns than C.2-S.1, which are supported by the lowering of L_1, and L_2 norms. Similarly, the GI approach for C.1-S.2 yields a better validation result than C.1-S.1, which is supported by an improved slope (Fig. 9b) and smaller L_1 and L_2 norms (Figs. 9d and 9e). The GI approach for C.2-S.2 also yields a better validation result than C.2-S.1, which is supported by an improved slope (Fig. 9g) and smaller L_1 and L_2 norms (Figs. 9i and 9j). It is worth mentioning that the TTI approach for scenario 2 yields smaller simulated drawdowns than scenario 1 when examining each case individually, which is supported by the slope of scenario 2 being lower than scenario 1 for each case (Figs. 9b, 9g). This phenomenon is consistent with the earlier finding that the TTI approach for Scenario 2 yields better tomograms than Scenario 1, which eliminate the 'X' pattern with high diffusivity.

According to this validation for Scenario 2, the GI approach yields the best result not only for Case 1, but also for Case 2. For both cases, scatterplots of the GI approach reveal significant advantages in terms of L_1 and L_2 norms (Figs. 9d, 9e, 9i, 9j) as well as the intercept (Figs. 9c, 9h) than the TTI approach. However, in terms of the slope and R^2 values, the GI approach exhibits similar performance compared to the TTI approach.

It is unfair to compare the estimated tomograms for different cases by validation approach one, because the utilized pumping and observation ports of two cases are different for approach one. To further evaluate the estimated tomograms for the various cases, validation approach two is conducted.

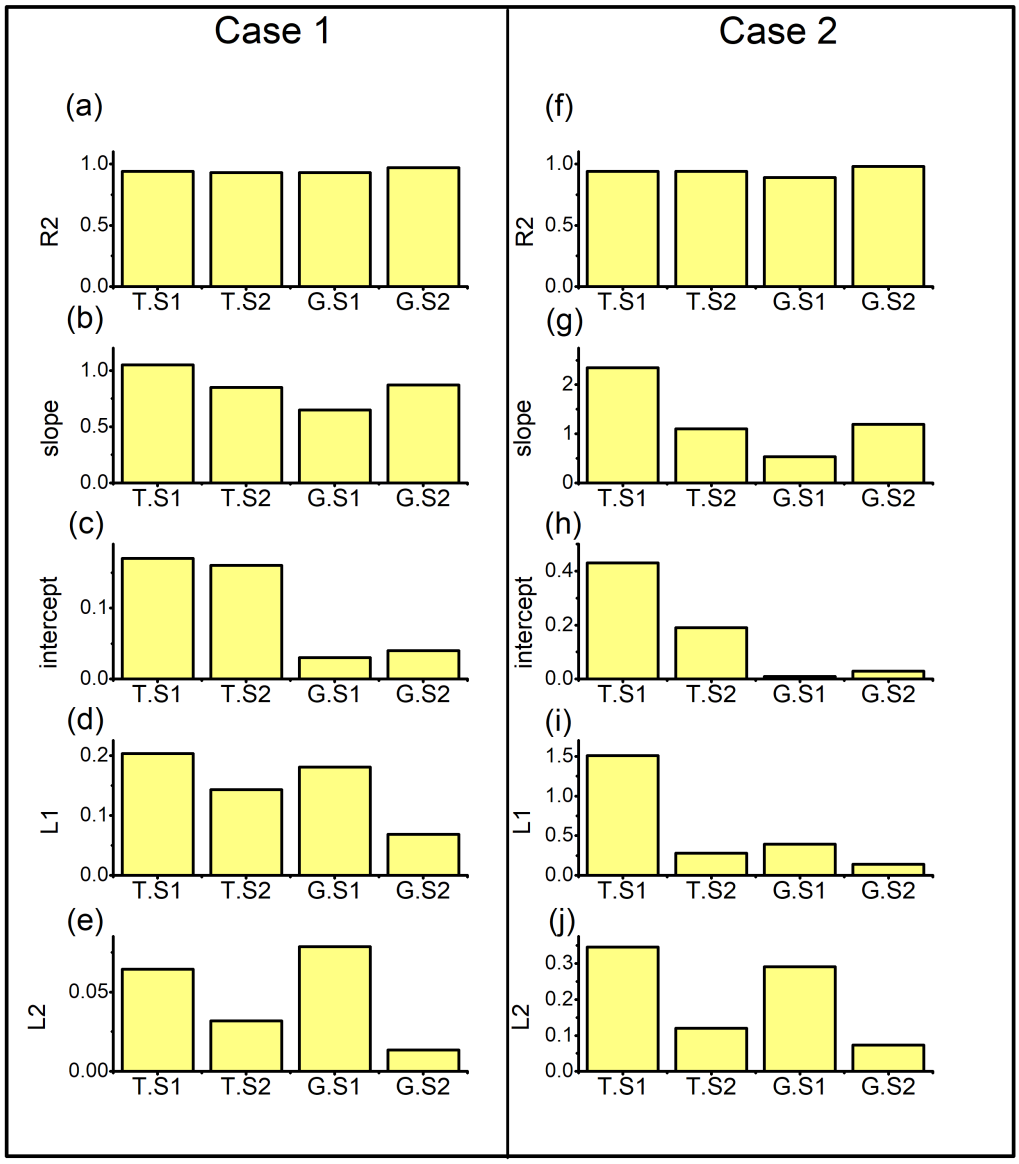

Figure 9: Comparison of performance metrics of linear model fits to scatterplots from validation approach one. Case 1: (a) R^2values; (b) slope; (c) intercept; (d) L_1 norm and (e) L_2 norm. Case 2: (f) R^2values; (g) slope; (h) intercept; (i) L_1 norm and (j) L_2 norm. The abbreviation of T/G. SX means TTI/GI in Scenario X.

For the second validation approach, pumping tests are conducted at 16 ports

(ports 3, 5, 9, 11, 15, 17, 21, 23, 27, 29, 33, 35, 39, 41, 45, 47) and data that are not used for model calibration are used for model validation. All ports (1-48) are utilized as observation ports for each pumping test. The configuration of validation approach two is shown in Fig. S16. For this approach, all estimated tomograms from the TTI ad GI approaches are validated under a fair setting with the same data sets.

Fig. S17 shows the validation scatterplots of TTI and GI approaches with models from various cases and scenarios on validation approach two. Similar with validation approach one, the performance metrics from scatterplots (Fig. S17) for validation approach two are compared through histograms (Fig. 10). Undoubtedly, the estimated parameters of GI approach from C.2-S.2 have the best validation result among all investigated cases and scenarios, yielding the lowest intercept (0.01), acceptable slope (1.21) of linear model fit, highest R^2 (0.94), and small $L_1(1.90\times10^{-1})$ and L_2 (1.00×10^{-1}) norms. However, the slope of the linear model (1.21) is larger than 1 revealing that the predicted drawdowns is slightly larger than true drawdowns. The worst model is C.2-S.1 of the TTI approach, which has the largest slope (2.48), high intercept (0.21) of linear model fit and largest L_1 (1.26×10^{0}) and L_2 norms (3.35×10^{0}). For C.1-S.2, the performance of the TTI approach is more acceptable than C.2-S.2. Except the large intercept (0.28) of the linear model fit, other performance metrics are satisfying, i.e., the slope (1.01), $R^2(0.89)$, L_1 (3.17×10^{-1}) and L_2 (1.58×10^{-1}) norms. It reveals the TTI approach can yield an acceptable validation result for specific cases, but the performances of other cases and scenarios are not satisfactory. The comparison of performance metrics shows that the GI approach has better performance on the intercept of linear model fit (Fig. 10c). This phenomenon reveals that compared to the TTI approach, the GI approach yields more accurate hydraulic heads data when drawdown value is small (0~0.5s cm). The port configuration and scenarios result in the provided data for the GI approach always includes small drawdown, which is the primary reason of this phenomenon.

As previously mentioned, the effect from the port configuration on 'visual comparison', Fig. 10 reveals that the port configuration ultimately affects validation results. As shown previously in Fig.2, the multiple port configuration (Case 2) is rare for a TTI study conducted at the laboratory scale (or small scale). On Fig. S17c, C.2-S.1 of the TTI approach has the worst validation result with the largest slope (Fig.10b) as well as L_1 and L_2 norms (Figs. 10d and 10e). Even after using optimal operation for data selection, C.2-S.2 of the TTI approach also has room to improve, which is supported by the fact that it has higher intercept than the GI approach in Fig.10c. Compared with Case 2, Case 1 is more suitable for the TTI approach, though Case 1 has less observation ports than Case 2. The sufficient distance between pumping and observation ports result in a better result for the TTI approach. For the GI approach, the configuration of Case 2 is better. With a reasonable strategy, the GI approach for C.2-S.2 yields the best estimated tomogram among all models. The comparison between C.1-S.1 and C.1-S.2 shows that providing more data points from drawdown curves do not always improve the GI validation result. The GI approach

for C.1-S.2 has higher L_1 and L_2 norms, worse slope, and intercept of linear model fit (Fig. 10) than C.1-S.1. For validation approach two, with Case 1 configuration of monitoring ports, providing more data from drawdown curves results in a worse GI validation result (Fig. S17f) than only selected one data point (Fig. S17e).

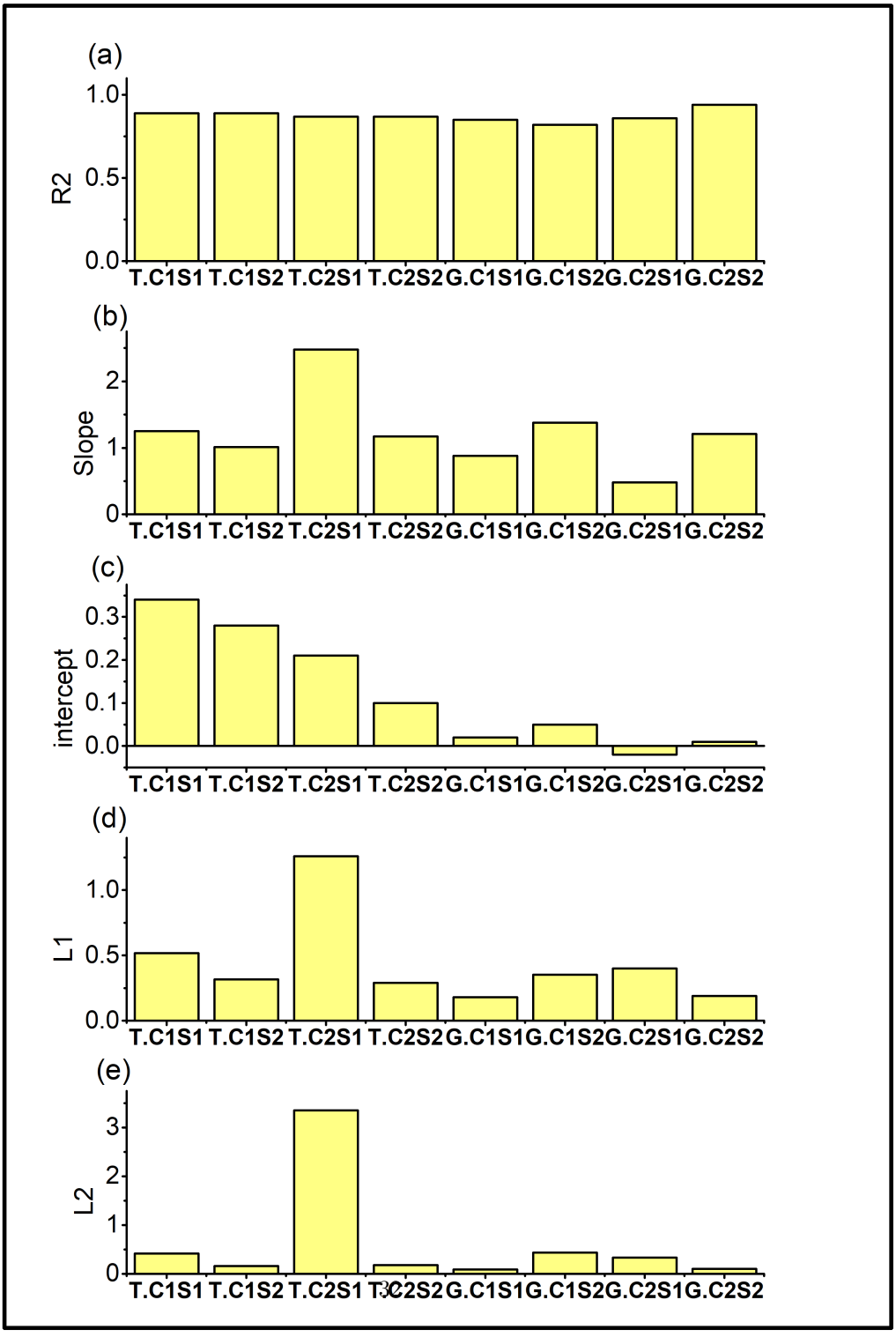
(a)
R2
1.0
0.5
0.0
T.C1S1 T.C1S2 T.C2S1 T.C2S2 G.C1S1 G.C1S2 G.C2S1 G.C2S2
(b)
Slope
2
1
0
T.C1S1 T.C1S2 T.C2S1 T.C2S2 G.C1S1 G.C1S2 G.C2S1 G.C2S2
(c)
intercept
0.3
0.2
0.1
0.0
T.C1S1 T.C1S2 T.C2S1 T.C2S2 G.C1S1 G.C1S2 G.C2S1 G.C2S2
(d)
L1
1.0
0.5
0.0
T.C1S1 T.C1S2 T.C2S1 T.C2S2 G.C1S1 G.C1S2 G.C2S1 G.C2S2
(e)
L2
3
2
1
0
T.C1S1 T.C1S2 T.C2S1 T.C2S2 G.C1S1 G.C1S2 G.C2S1 G.C2S2

Figure 10 Comparison of performance metrics of linear model fits to scatterplots from validation approach one: (a) R^2values; (b) slope; (c) intercept; (d) L_1 norm and (e) L_2 norm. The abbreviation of T/G. CXSX means TTI/GI in Case X Scenario X.

5 Discussion

In this study, as the two common inversion methods utilized for HT survey, the TTI and GI approaches are compared with each other based on the visual comparison of computed tomograms, as well as through evaluation of model calibration and validation based on model performance metrics.

The results have shown that both TTI and GI approaches have the ability to characterize the heterogeneity with satisfactory accuracy (Figs. 4 - 6), yet the TTI approach costs significantly less computational resources. The TTI approach only requires head data during a much earlier stage of pumping/slug test, thus the time cost of field tests for the TTI approach is significantly lower than the GI approach. Hence, we suggest to utilize the TTI approach when the aim is to characterize heterogeneity rapidly, focusing on the main hydrogeological structures, especially for larger, 3D field cases. However, early field data may be impacted by wellbore storage (Wu et al., 2017). When conducting a HT survey at a field site, this issue may reduce the TTI approach performance. Thus, efforts should be made to reduce wellbore storage effects in pumping and observation wells. For the GI approach, the computational requirements should be considered when using SimSLE for a large 3D field scale. For example, two days were required by Zhao & Illman (2018), who used a PC-cluster with 16 processors and 192 GB of RAM to run SimSLE for estimating K and S_s tomograms for a highly heterogeneous aquifer-aquitard system at a field site.

Numerical models with accurate input parameters achieved through HT approaches are proven to provide better prediction of groundwater flow and transport. In this study, the GI approach can yield the best validation result for Case 2 Scenario 2 among all cases and scenario. Hence, we suggest the application of GI approach for the objective of achieving accurate values of hydraulic parameters. The TTI approach is not recommended for obtaining the accurate hydraulic parameters due to the biased D-reconstruction, if the travel time data are limited or inaccurate due to cross-well distance.

The validations have shown that the reconstructed D values from the TTI approach can vary much from the true values for some areas, resulting in inaccurate estimated K values that are not appropriate for groundwater flow and solute transport predictions. Another reason of inaccurate head prediction is that, the TTI approach is based on the transformation of groundwater flow equation into an eikonal equation and directly related to the D values. Therefore, the TTI approach is not directly sensitive to K and $S_s$values, which are the two main parameters that control the groundwater flow behavior, including drawdown induced through pumping tests. Even if an accurate D distribution is obtained, the issue of separating K and S_s still exists. However, we suggest the

TTI approach to be utilized as an initial estimation with key structural information and zonation for other inversion methods (e.g., Jiménez et al., (2013)), because of its ability in mapping structural features of aquifers at a significantly less computational cost and much less field testing time compared to the GI approach.

With regard to the built-in result analysis, providing the uncertainty analysis of estimated tomograms is a featured function of the GI approach. In this study, Figs. S7, S8 provide variances of hydraulic parameters for the GI approach. As for the TTI approach, it can also provide information such as ray density and ray trajectory to reveal the reliability of pixel values on tomograms. Some previous studies also calculate the null space energy map to characterize the reliability of obtained tomograms (Brauchler et al., 2013; Jiménez et al., 2013).

In the 'visual comparison' section, tomograms from the GI approach have shown that on the heterogeneity details with high K can be better characterized for Case 1 Scenario 1 than Case 1 Scenario 2. In the 'model calibration and validation' section, the GI approach for Case 1 Scenario 1 has a better result than Case 1 Scenario 2 according to the second validation approach, while the GI approach for Case 1 Scenario 1 has the worse result than Case 1 Scenario 2 according to the first validation approach. Obviously, the second validation approach is consistent with 'visual comparison'. Different conclusions from the two validation approaches reveal the fact that the data from new/unused pumping ports is necessary for model validation. In particular, without data from new/unused pumping ports, the performance of the inverse model may be overestimated for model validation.

In terms of future improvement, the TTI approach still has room for providing reliable estimates of D. Replacing the eikonal equation (Eqn. (1)) with a new trajectory algorithm is promising in resolving this problem (Vasco et al., 2019). And apart from the D values, less study on the estimation of S_s is performed based on the TTI approach. In order to take advantage of travel-time based tomography, Brauchler et al. (2011) presented an attenuation integral which relates the S_s of aquifer with the amplitude attenuation of the hydraulic signal. Based on this method, Brauchler et al. (2013) estimated S_s successfully at a field site. However, the issue of ignoring initial and boundary conditions still exists. With further development of this method, an accurate S_s distribution achieved through travel time derived method could help to obtain the K value based on the D values from travel time inversion ($K = D \times S_s$). It is worth mentioning that the effect of prior information is not considered in this study, because the utilized code for the TTI approach does not have the capability in its current form to add geological information or reliable prior information from other sources (e.g., geophysical tests). This additional functionality to the TTI code could also be a significant improvement that we consider to be highly beneficial for future development.

For the GI approach, the future improvement should consider: (1) finding an optimal strategy of configuration of pumping and observation ports; (2) incor-

porating data from other tests (e.g., geophysical survey, tracer test) (Illman, 2014); and (3) adding constraints to mitigate ill-posed and nonunique solutions. One promising idea is the combination of these two methods, in particular, by rapidly obtaining the hydro-/geological zonation information through the TTI approach, which in turn is used as an initial estimation for the GI approach. Zhao and Illman (2018) demonstrated that the use of a reliable geological model as initial guess can yield both inter- and intralayer heterogeneity of both aquifer and aquitard units at a highly heterogeneous field site. We are trying to present a method to build an initial model for the GI approach based on the D or/and S_s tomogram from the TTI approach, which could assist the GI approach in yielding the aquifer reconstruction that preserves geological features more accurately and more efficiently.

6 Summary, conclusions and outlook

Through various studies, tomograms obtained from HT surveys have demonstrated significant advantages over traditional characterization approaches, which include: (1) the ability in charactering aquifer heterogeneity including the connectivity of flow/transport pathways between boreholes; (2) the performance of predicting groundwater flow and solute transport (Illman et al., 2012; Jimenez et al., 2015; Luo et al., 2020); (3) providing information on uncertainty or reliability of estimated tomograms (Brauchler et al., 2013; Illman et al., 2010); and (4) demonstrating acceptable cost performance compared to site heterogeneity characterization based on a large number of small-scale samples (Illman et al., 2012b; Yeh & Liu, 2000).

In this study, numerical experiments were conducted with a synthetic heterogeneous aquifer, derived from a laboratory sandbox aquifer to compare two kinds of transient hydraulic tomography (THT) algorithms from TTI and GI approaches.

Two cases with different configurations of pumping and observation ports were designed with two scenarios for each case. For Case 1, eight pumping and eight observation ports were set at opposing sides of the sandbox. For Case 2, eight pumping ports were set in the middle part of the aquifer, while 8 observation ports were set at each side with a total of 16 observation ports.

The two scenarios in each case were designed to compare the estimation performance and to study the effects of using different selected data. For Scenario 1, the same whole data set was provided to both approaches. For Scenario 2, only selected data from the whole drawdown curve was utilized for both inversion approaches.

For the TTI approach, the optimal data selection strategy involves the incorporation of data with small angles between pumping and observation ports, as well as the selection of earlier travel-times from the drawdown curves to estimate D. In contrast, for the GI approach, the second scenario provides sufficient data points to capture the salient features of drawdown curves that contain information on both K and S_s heterogeneity.

The estimated tomograms from both approaches were compared through visual comparisons and qualitative performance statistics computed for model validation. For every inverse model, two validation approaches were used to compare the performance of inversion. One validation approach was to utilize data from observation ports not used for calibration, while retaining the pumping ports the same as those designed for model calibration. Another validation approach was to obtain data from 16 additional pumping tests at 48 observation ports. For both approaches, the estimated tomograms were utilized to predict drawdowns and compared with true drawdowns.

Our study led to the following findings and conclusions:

1. In terms of visual comparison, both approaches demonstrated the ability to characterize subsurface heterogeneity with regard to key hydraulic parameters. The TTI approach showed good performance in imaging layered structures with high D, especially when the operation of constraining angle is implemented. But in some areas, the reconstructed D from the TTI approach is far from the true D values. In general, the K and S_s tomograms from the GI approach, were considered to be closer to true parameters.

2. In terms of model validation and head predictions, the GI approach yields the best validation result when head data and observation ports are sufficient. The TTI approach with optimal strategy also displayed satisfactory validation results, when the distance between pumping and observation ports is sufficient. However, the TTI approach without the application of the optimal strategy yielded the worst validation result, for the case when the pumping ports are very close to observation ports.

3. The TTI approach can provide tomograms with satisfactory characterization of heterogeneity with the optimal strategy. Meanwhile, this approach requires significantly less computational times than GI approach. Hence, it is advantageous to apply the TTI approach for the characterization of heterogenous aquifer, especially during the early stage of site characterization studies or when field characterization budget is limited.

4. The GI approach can provide satisfactory validation results when sufficient data is available. Yet, good results from the GI approach require sufficient data and computational resources to calibrate models well. Sufficient data includes ample long term data collected at high frequency to capture the entire drawdown curves with a sufficient number of observation ports. Inverse modeling of large data sets from large, 3D field scale problems with the GI approach often requires significant computational resources such as a computer-cluster with large amounts of memory.

5. The optimal strategies of using selected data, i.e. "constraining angle" and "earlier travel-time" are found to be helpful in obtaining good validation results for the TTI approach. Yet, these optimal strategies cannot solve the problem that the reconstructed D is far from real value in some certain areas as shown through this study. This phenomenon is due to the utilized TTI algorithm that

ignores the impacts of initial and boundary conditions, which are crucial factors for HT data collected in a small domain such as the small-scaled sandbox in this study.

6. For a fair comparison to TTI approach, a dataset is constructed with providing just one data point of observation drawdown curve for the GI approach, which is different from common interpretation scenarios. Based on various time selection for the inversion, different tomograms have shown that the selected time point of provided data clearly affects the performance of the GI approach. Providing early time data can also improve its performance on characterizing high K areas, which is similar with a previous study based on the TTI approach (please provide reference).

7. The configuration of pumping and observation ports are important for both approaches. The TTI approach prefers a configuration with sufficient distance between pumping and observation ports. The very close distance for TTI approach may cause the failure of recording very early travel-time data, which worsen the performance severely. The GI approach prefers a configuration with more ports that are spread around the domain.

Based on this study, the GI approach yields different results when compared to the TTI approach as both approaches take advantage of different data sets from a given pumping test. This suggests that the two approaches can be used in a complementary fashion. More work is required in taking advantage in strengths of both approaches for improved high-resolution characterization of hydraulic parameters such as the rapid imaging of preferential flow/transport pathways. An attractive topic is the combination of these two methods, which is currently studied to build an initial model for the GI approach with the zonal information that is based on the D tomogram from the TTI approach. We believe that with a successful combination of these two methods, the HT survey can have a promising efficiency and accuracy for aquifer characterization.

Acknowledgements

This work was supported by the Ministry of Science and Technology of China through the Program "Driving process and mechanism of three dimensional spatial distribution of high risk organic pollutants in multi field coupled sites" (Project Code: 2019YFC1804303). Walter A. Illman acknowledges the partial support from the Discovery grant funded by the Natural Sciences and Engineering Council of Canada which facilitated this collaboration. The first author acknowledges the financial support of "China Scholarship Council" 201906710028.

Open Research

There is an open-source version of TOMOGO via Github: https://github.com/wichniarek/SIRT-Variants.git. The data and configuration for inverse code are available at Github through: https://github.com/qiuhuiyang516/Comparison-of-TT-and-GI.git. All input files and data can be found through the above link.

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*, W10507. doi:10.1029/2011WR010616Bohling, 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), 1324. doi:10.1029/2001wr001176Bohling, G. C., Butler, J. J., Zhan, X., & Knoll, M. D. (2007). A field assessment of the value of steady shape hydraulic tomography for characterization of aquifer heterogeneities. *Water Resources Research*, *43*(5), 1–23. doi:10.1029/2006WR004932Brandstatter, J. J. (1974). Dynamic programming, Fermat's principle, and the eikonal equation for anisotropic media. *Journal of the Optical Society of America*, *64*(3), 317. doi:10.1364/JOSA.64.000317Brauchler, R., Liedl, R., & Dietrich, P. (2003). A travel time based hydraulic tomographic approach. *Water Resources Research*, *39*(12), 1–12. doi:10.1029/2003WR002262Brauchler, R., Cheng, J.-T., Dietrich, P., Everett, M., Johnson, B., Liedl, R., & Sauter, M. (2007). An inversion strategy for hydraulic tomography: Coupling travel time and amplitude inversion. *Journal of Hydrology*, *345*(3–4), 184–198. doi:10.1016/j.jhydrol.2007.08.011Brauchler, R., Hu, R., Dietrich, P., & Sauter, M. (2011). A field assessment of high-resolution aquifer characterization based on hydraulic travel time and hydraulic attenuation tomography. *Water Resources Research*, *47*, W03503. doi:10.1029/2010WR009635Brauchler, R., Hu, R., Hu, L., Jiménez, S., Bayer, P., Dietrich, P., & Ptak, T. (2013). Rapid field application of hydraulic tomography for resolving aquifer heterogeneity in unconsolidated sediments. *Water Resources Research*, *49*(4), 2013–2024. doi:10.1002/wrcr.20181Cardiff, M., Zhou, Y., Barrash, W., & Kitanidis, P. K. (2020). Aquifer Imaging with Oscillatory Hydraulic Tomography: Application at the Field Scale. *Groundwater*, *58*(5), 710–722. doi:10.1111/gwat.12960Cooper, H. H., & Jacob, C. E. (1946). A generalized graphical method for evaluating formation constants and summarizing well-field history. *Transactions, American Geophysical Union*, *27*(4), 526. doi:10.1029/TR027i004p00526Doherty, J. (2003). Ground Water Model Calibration Using Pilot Points and Regularization. *Ground Water*, *41*(2), 170–177. doi:10.1111/j.1745-6584.2003.tb02580.xGottlieb, J., & Dietrich, P. (1995). Identification of the permeability distribution in soil by hydraulic tomography. *Inverse Problems*, *11*(2), 353–360. doi:10.1088/0266-5611/11/2/005Hu, R., Brauchler, R., Herold, M., & Bayer, P. (2011). Hydraulic tomography analog outcrop study: Combining travel time and steady shape inversion. *Journal of Hydrology*, *409*(1–2), 350–362. doi:10.1016/j.jhydrol.2011.08.031Illman, 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. doi:10.1016/j.jhydrol.2007.05.011Illman, W. A., Zhu, J., Craig, A. J., & Yin, D. (2010). Comparison of aquifer characterization approaches through steady state groundwater model validation: A controlled laboratory sandbox study. *Water Resources Research*, *46*(4), 1–18. doi:10.1029/2009WR007745Illman, W. A., Berg, S. J., & Yeh, T.-C.

J. (2012a). Comparison of Approaches for Predicting Solute Transport: Sandbox Experiments. *Ground Water*, *50*(3), 421–431. doi:10.1111/j.1745-6584.2011.00859.xIllman, W. A., Berg, S. J., & Alexander, M. (2012b). Cost Comparisons of Aquifer Heterogeneity Characterization Methods. *Ground Water Monitoring & Remediation*, *32*(2), 57–65. doi:10.1111/j.1745-6592.2011.01376.xIllman, W. A., Berg, S. J., & Zhao, Z. (2015). Should hydraulic tomography data be interpreted using geostatistical inverse modeling? A laboratory sandbox investigation. *Water Resources Research*, *51*(5), 3219–3237. doi:10.1002/2014WR016552Jackson, M. J., & Tweeton, D. R. (1996). *3DTOM, three-dimensional geophysical tomography.*Jiménez, S., Brauchler, R., & Bayer, P. (2013). A new sequential procedure for hydraulic tomographic inversion. *Advances in Water Resources*, *62*(Part A), 59–70. doi:10.1016/j.advwatres.2013.10.002Jiménez, S., Brauchler, R., Hu, R., Hu, L., Schmidt, S., Ptak, T., & Bayer, P. (2015). Prediction of solute transport in a heterogeneous aquifer utilizing hydraulic conductivity and specific storage tomograms. *Water Resources Research*, *51*(7), 5504–5520. doi:10.1002/2014WR016402Kitanidis, P. K., & Vomvoris, E. G. (1983). A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resources Research*, *19*(3), 677–690. doi:10.1029/WR019i003p00677Kulkarni, K. N., Datta-Gupta, A., & Vasco, D. W. (2001). A Streamline Approach for Integrating Transient Pressure Data Into High-Resolution Reservoir Models. *SPE Journal*, *6*(03), 273–282. doi:10.2118/74135-PALaVenue, A. M., & Pickens, J. F. (1992). Application of a coupled adjoint sensitivity and kriging approach to calibrate a groundwater flow model. *Water Resources Research*, *28*(6), 1543–1569. doi:10.1029/92WR00208Lavenue, M., & de Marsily, G. (2001). Three-dimensional interference test interpretation in a fractured aquifer using the Pilot Point Inverse Method. *Water Resources Research*, *37*(11), 2659–2675. doi:10.1029/2000WR000289Liu, S., Yeh, T.-C. J., & Gardiner, R. (2002). Effectiveness of hydraulic tomography: Sandbox experiments. *Water Resources Research*, *38*(4), 5-1-5–9. doi:10.1029/2001WR000338Liu, 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), 1–13. doi:10.1029/2006WR005144Luo, 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. doi:10.1016/j.jhydrol.2017.09.045Luo, 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. doi:10.1016/j.jhydrol.2020.125438Maier, U., DeBiase, C., Baeder-Bederski, O., & Bayer, P. (2009). Calibration of hydraulic parameters for large-scale vertical flow constructed wetlands. *Journal of Hydrology*, *369*(3–4), 260–273. doi:10.1016/j.jhydrol.2009.02.032de Marsily, G, Ledoux, E., Levassor, A., Poitrinal, D., & Salem, A. (1978). Modelling of large multilayered aquifer systems: Theory and applications. *Journal of Hydrology*, *36*(1–2), 1–33.

doi:10.1016/0022-1694(78)90034-3de Marsily, Ghislain, Lavedan, G., Boucher, M., & FASAMINO, G. (1984). Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In *Geostatistics for natural resources characterization. NATO advanced Study Institute* (pp. 831–849).Minutti, C., Illman, W. A., & Gomez, S. (2020). A New Inverse Modeling Approach for Hydraulic Conductivity Estimation Based on Gaussian Mixtures. *Water Resources Research*, *56*(9), 1–21. doi:10.1029/2019WR026531Moser, T. J. (1991). Shortest path calculation of seismic rays. *GEOPHYSICS*, *56*(1), 59–67. doi:10.1190/1.1442958Paradis, D., Gloaguen, E., Lefebvre, R., & Giroux, B. (2016). A field proof-of-concept of tomographic slug tests in an anisotropic littoral aquifer. *Journal of Hydrology*, *536*, 61–73. doi:10.1016/j.jhydrol.2016.02.041Poduri, S., & Kambhammettu, B. V. N. P. (2020). On the Performance of Pilot-Point Based Hydraulic Tomography with a Geophysical a Priori Model. *Groundwater*. *59*(2), 214–225. doi:10.1111/gwat.13053Qiu, P., Hu, R., Hu, L., Liu, Q., Xing, Y., Yang, H., et al. (2019). A Numerical Study on Travel Time Based Hydraulic Tomography Using the SIRT Algorithm with Cimmino Iteration. *Water*, *11*(5), 909. doi:10.3390/w11050909Sanchez-León, E., Leven, C., Haslauer, C. P., & Cirpka, O. A. (2016). Combining 3D Hydraulic Tomography with Tracer Tests for Improved Transport Characterization. *Groundwater*, *54*(4), 498–507. doi:10.1111/gwat.12381Snodgrass, M. F., & Kitanidis, P. K. (1998). Transmissivity identification through multi-directional aquifer stimulation. *Stochastic Hydrology and Hydraulics*, *12*(5), 299. doi:10.1007/s004770050023Sun, 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. doi:10.1002/wrcr.20337Theis, C. V. (1935). The relation between the lowering of the Piezometric surface and the rate and duration of discharge of a well using ground-water storage. *Transactions, American Geophysical Union*, *16*(2), 519–524. doi:10.1029/TR016i002p00519Trampert, J., & Leveque, J.-J. (1990). Simultaneous iterative reconstruction technique: Physical interpretation based on the generalized least squares solution. *Journal of Geophysical Research*, *95*(B8), 12553–12559. doi:10.1029/JB095iB08p12553Vasco, D. W. (2018). An Extended Trajectory Mechanics Approach for Calculating the Path of a Pressure Transient: Derivation and Illustration. *Water Resources Research*, *54*(4), 2642–2660. doi:10.1002/2017WR021360Vasco, D W, & Karasaki, K. (2006). Interpretation and inversion of low-frequency head observations. *Water Resources Research*, *42*(5). doi:10.1029/2005WR004445Vasco, D W, Keers, H., & Karasaki, K. (2000). Estimation of reservoir properties using transient pressure data: An asymptotic approach. *Water Resources Research*, *36*(12), 3447–3465. doi:10.1029/2000WR900179Vasco, Donald W, Doetsch, J., & Brauchler, R. (2019). An extended trajectory-mechanics approach for calculating the path of a pressure transient: travel-time tomography. *Hydrology and Earth System Sciences*, *23*(11), 4541–4560. doi:10.5194/hess-23-4541-2019Vesselinov, V. V, Neuman, S. P., & Illman, W. A. (2001). Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 2. Equivalent parameters, high-resolution stochastic

imaging and scale effects. *Water Resources Research*, *37*(12), 3019–3041. doi:10.1029/2000WR000135Wang, X., Jardani, A., & Jourde, H. (2017). A hybrid inverse method for hydraulic tomography in fractured and karstic media. *Journal of Hydrology*, *551*, 29–46. doi:10.1016/j.jhydrol.2017.05.051Wu, Y.-X., Shen, J. S., Cheng, W.-C., & Hino, T. (2017). Semi-analytical solution to pumping test data with barrier, wellbore storage, and partial penetration effects. *Engineering Geology*, *226*, 44–51. doi:10.1016/j.enggeo.2017.05.011Xiang, 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), 1–14. doi:10.1029/2008WR007180Yeh, T.-C. Jim, Srivastava, R., Guzman, A., & Harter, T. (1993). A Numerical Model for Water Flow and Chemical Transport in Variably Saturated Porous Media.pdf. *Ground Water*, *31*, 634–644. doi:10.1111/j.1745-6584.1993.tb00597.xYeh, T-C Jim, Gutjahr, A. L., & Jin, M. (1995). An Iterative Cokriging-Like Technique for Ground-Water Flow Modeling. *Ground Water*, *33*(1), 33–41. doi:10.1111/j.1745-6584.1995.tb00260.xYeh, T-C Jim, Jin, M., & Hanna, S. (1996). An Iterative Stochastic Inverse Method: Conditional Effective Transmissivity and Hydraulic Head Fields. *Water Resources Research*, *32*(1), 85–92. doi:10.1029/95WR02869Yeh, T.-C. Jim, & Liu, S. (2000). Hydraulic tomography: Development of a new aquifer test method. *Water Resources Research*, *36*(8), 2095–2105. doi:10.1029/2000WR900114Yin, D., & Illman, W. A. (2009). Hydraulic tomography using temporal moments of drawdown recovery data: A laboratory sandbox study. *Water Resources Research*, *45*(1). doi:10.1029/2007WR006623Zha, Y., Yeh, T.-C. J., Illman, W. A., Tanaka, T., Bruines, P., Onoe, H., & Saegusa, H. (2015). What does hydraulic tomography tell us about fractured geological media? A field study and synthetic experiments. *Journal of Hydrology*, *531*, 17–30. doi:10.1016/j.jhydrol.2015.06.013Zha, Y., Yeh, T.-C. J., Illman, W. A., Onoe, H., Mok, C. M. W., Wen, J.-C., et al. (2017). Incorporating geologic information into hydraulic tomography: A general framework based on geostatistical approach. *Water Resources Research*, *53*(4), 2850–2876. doi:10.1002/2016WR019185Zhao, 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. doi:10.1016/j.jhydrol.2018.02.024Zhao, Z., & Illman, W. A. (2021). On the importance of considering specific storage heterogeneity in hydraulic tomography: Laboratory sandbox and synthetic studies. *Journal of Hydrology*, *593*, 125874. doi:10.1016/j.jhydrol.2020.125874Zhao, Z., WalterA., I., Yeh, T. J., Berg, S. J., & Mao, D. (2015). Validation of hydraulic tomography in an unconfined aquifer:A controlled sandbox study. *Water Resources Research*, *51 (6)*, 4137–4155. doi:10.1002/2015WR017200.AZhao, 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*. doi:10.1016/j.jhydrol.2016.08.061Zheng, C., & Gorelick, S. M. (2003). Analysis of Solute Transport in Flow Fields Influenced by Preferential Flowpaths at the Decimeter Scale. *Ground Water*, *41*(2), 142–155. doi:10.1111/j.1745-6584.2003.tb02578.xZhu, J., & Yeh, T.-C. J. (2006). Anal-

ysis of hydraulic tomography using temporal moments of drawdown recovery data. *Water Resources Research*, *42*(2), 1–11. doi:10.1029/2005WR004309

Water Resources Research

Supporting Information for

Comparison of travel-time and geostatistical inversion approaches for hydraulic tomography: A synthetic modeling study

Huiyang Qiu[1,2], Ning Luo[2], R. Hu[1], and Walter A. Illman[2]
1. School of Earth Science and Engineering, Hohai University, Nanjing, 210000, China

2. Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON, N2L 3G1, Canada

Contents of this file

Introduction

The supporting information contains five additional tables (Tables S1-S5), and seventeen additional figures (Figs. S1 – S17).

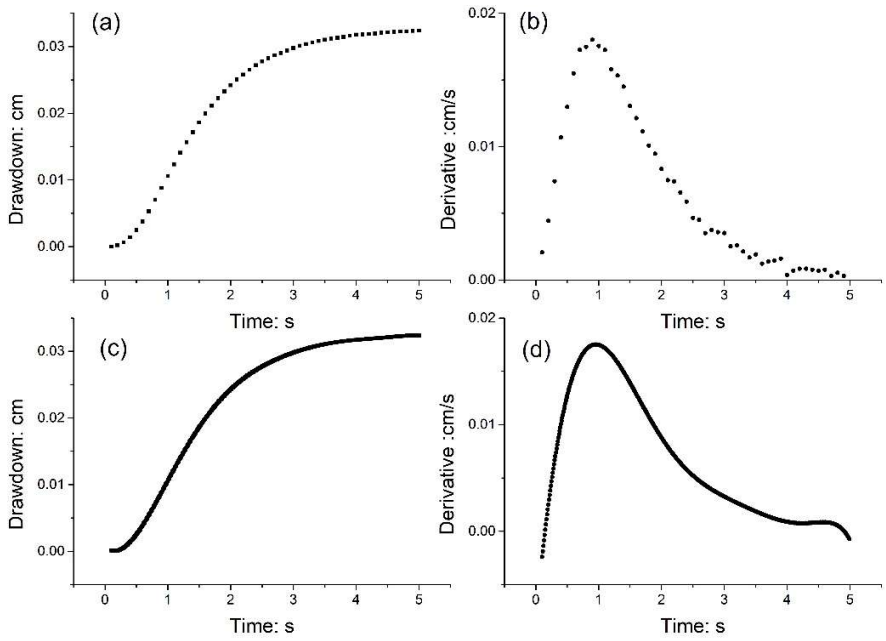

Figure S1. Illustration of hydraulic travel-time: (a) drawdowns at an observation port; (b) derivative of drawdown with respect to time calculated from the original data; (c) drawdowns after polynomial interpolation; (d) derivative of drawdowns with respect to time calculated from the data after polynomial interpolation.

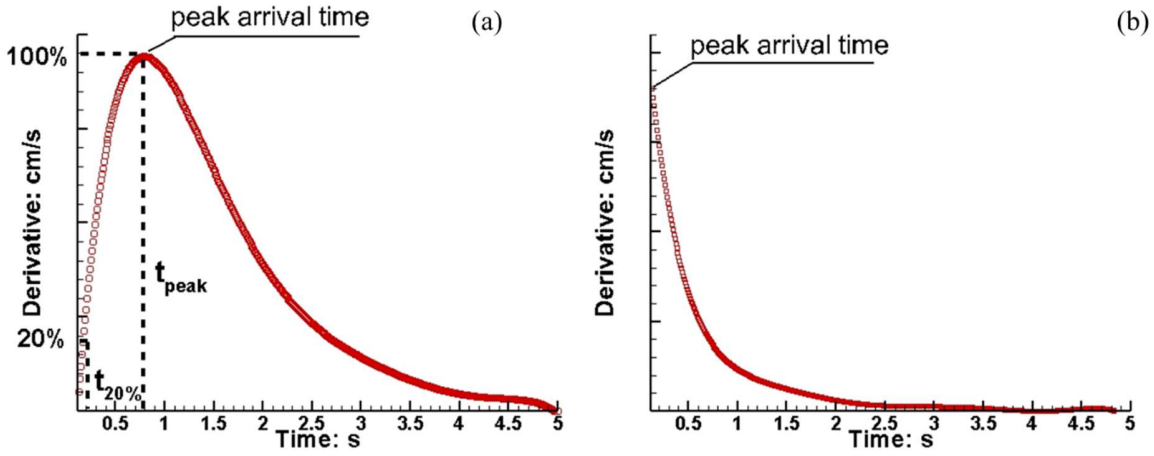

Figure S2. Derivative peak of drawdowns: (a) the common situation; (b) the situation in which the peak arrives at the first observation point.

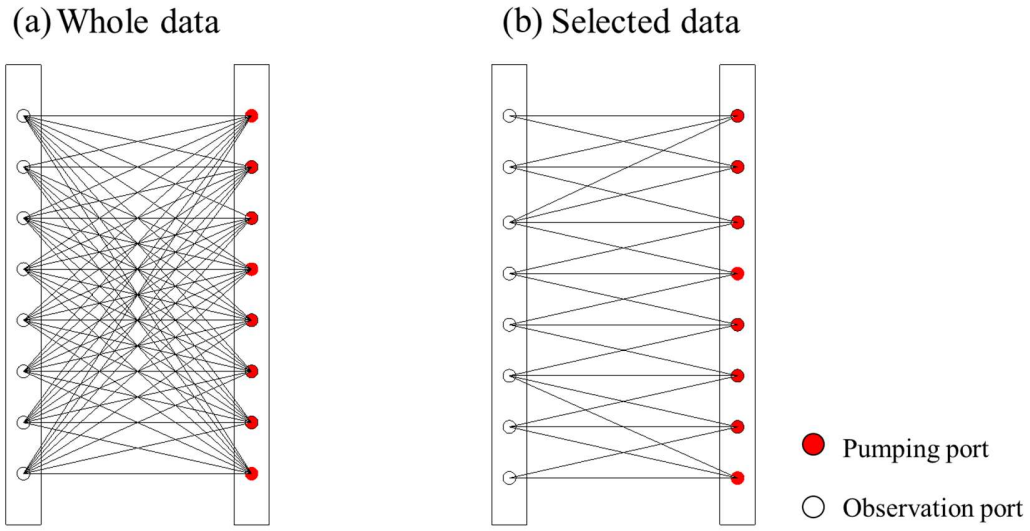

Figure S3. Ray path configurations for two different data sets: (a) whole original data set; and (b) selected data set eliminating source and receiver pairs with large angles.

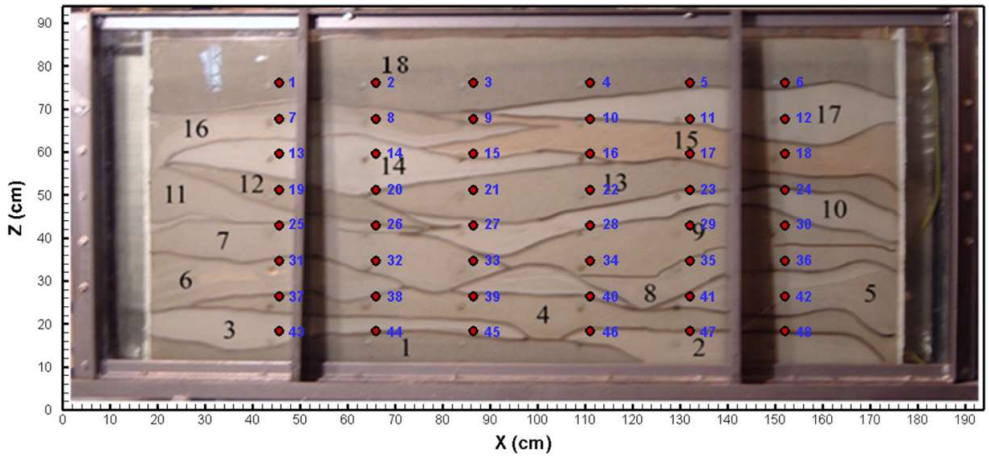

Figure S4. Photograph of the analogue heterogeneous aquifer with layers (black) and port numbers (blue) (Luo et al., 2017). Red circles indicate the 48 ports installed on the aquifer.

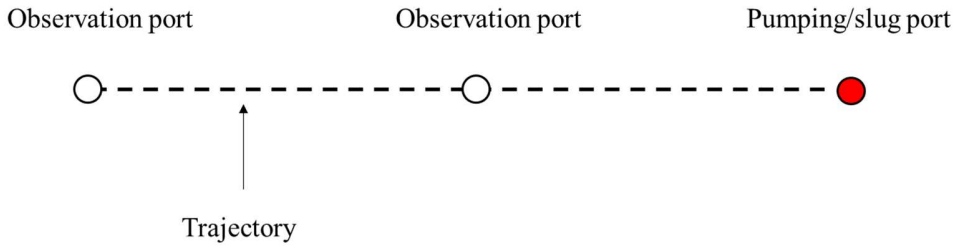

Figure S5. The condition to avoid for TTI.

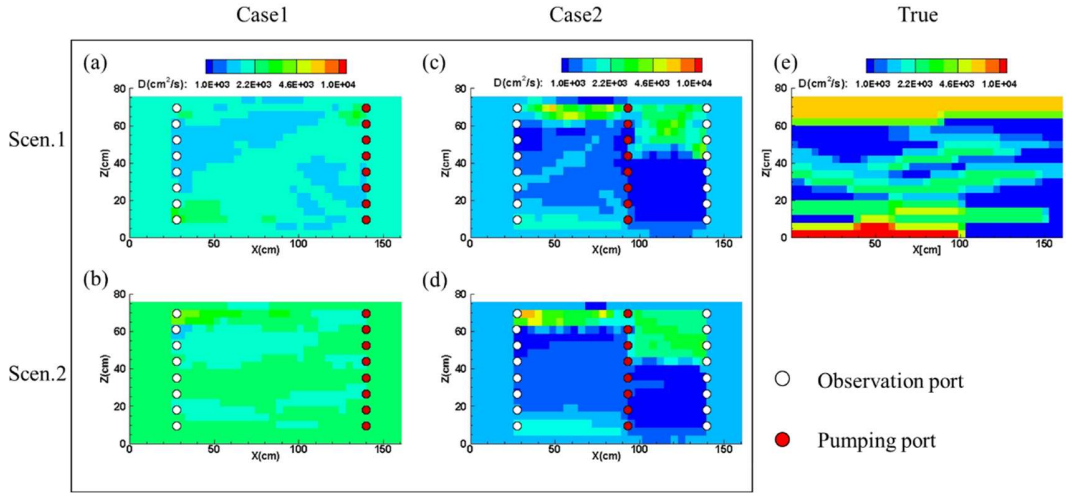

Figure S6. Comparison of reconstructed *D* tomograms from the TTI approach to the true *D* distribution: (a) Case 1 Scenario 1; (b) Case 1 Scenario 2; (c) Case 2 Scenario 1; (d) Case 2 Scenario 2; (e) True *D*. Reconstructed D tomograms use same color scales with True D.

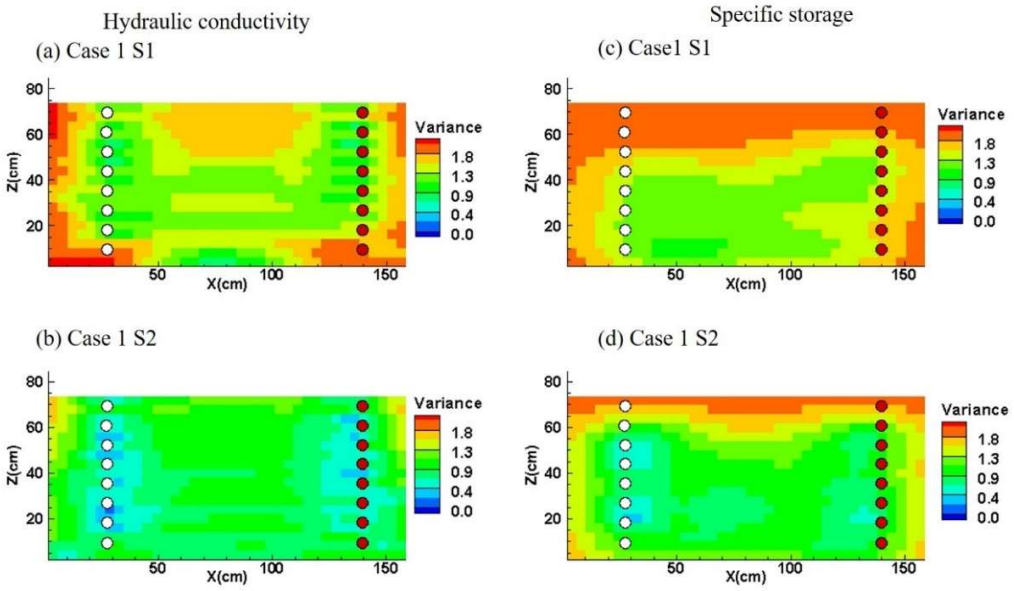

Figure S7. On Case 1, *K* tomogram variances under: (a) Scenario 1; (b) Scenario 2. On Case 1, $\boldsymbol{S_s}$ tomogram variances under: (c) Scenario 1; (d) Scenario 2.

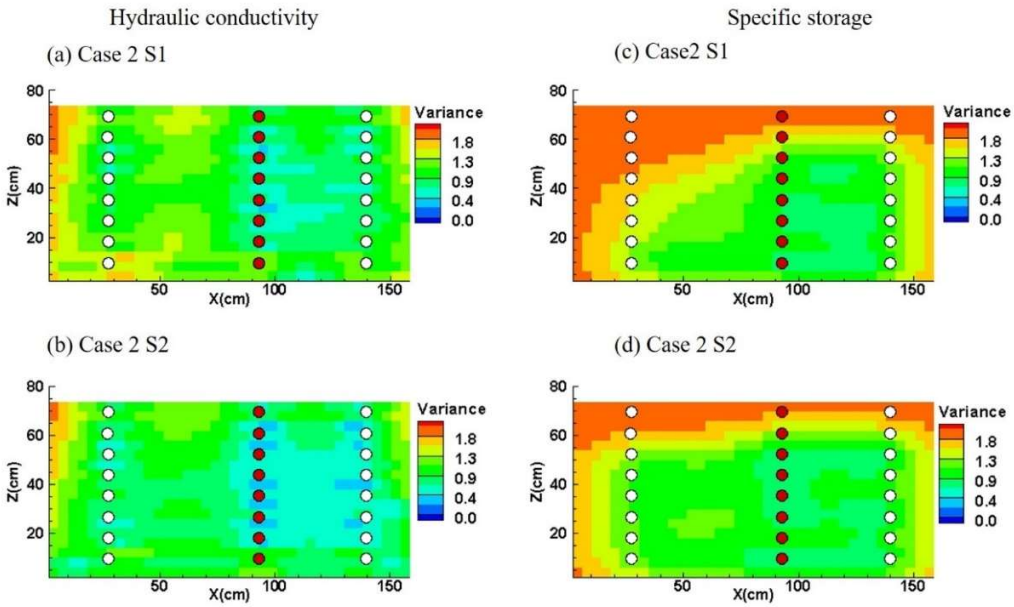

Figure S8. On Case 2, *K* tomogram variances under: (a) Scenario 1; (b) Scenario 2. On Case 2, $\boldsymbol{S_s}$ tomogram variances under: (c) Scenario 1; (d) Scenario 2.

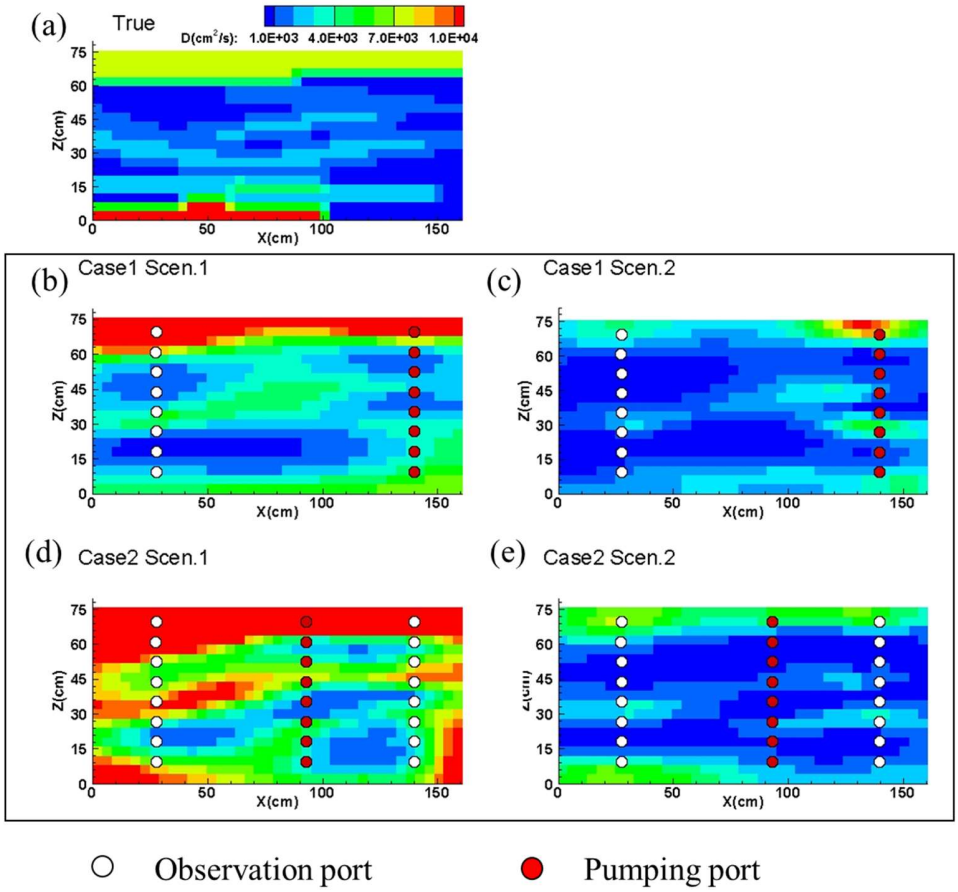

Figure S9. *D* tomograms from the 'true' model and the geostatistical inversion approach. True model: (a) *D* tomogram. Calculated *D* tomograms under: (b) Case 1 Scenario 1; (c) Case 1 Scenario 2; (d) Case 2 Scenario 1; and (e) Case 2 Scenario 2.

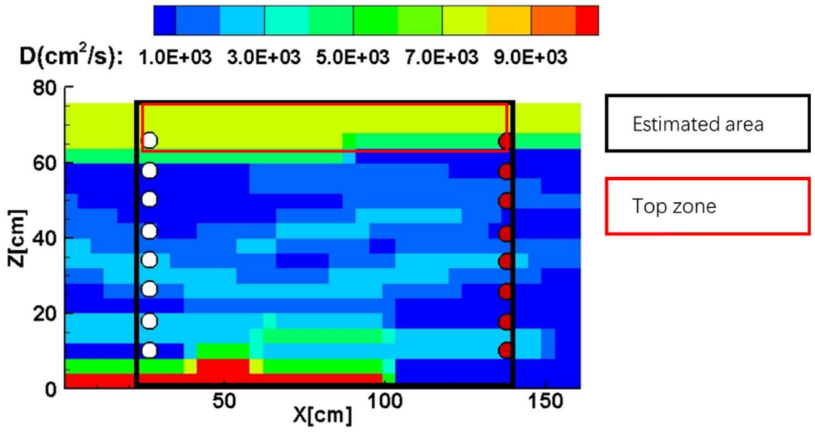

Figure S10. Real D of the true model. The black square is the area of aquifer which is used to compare the TTI and GI approaches. The red square area is the top high D zone of the aquifer.

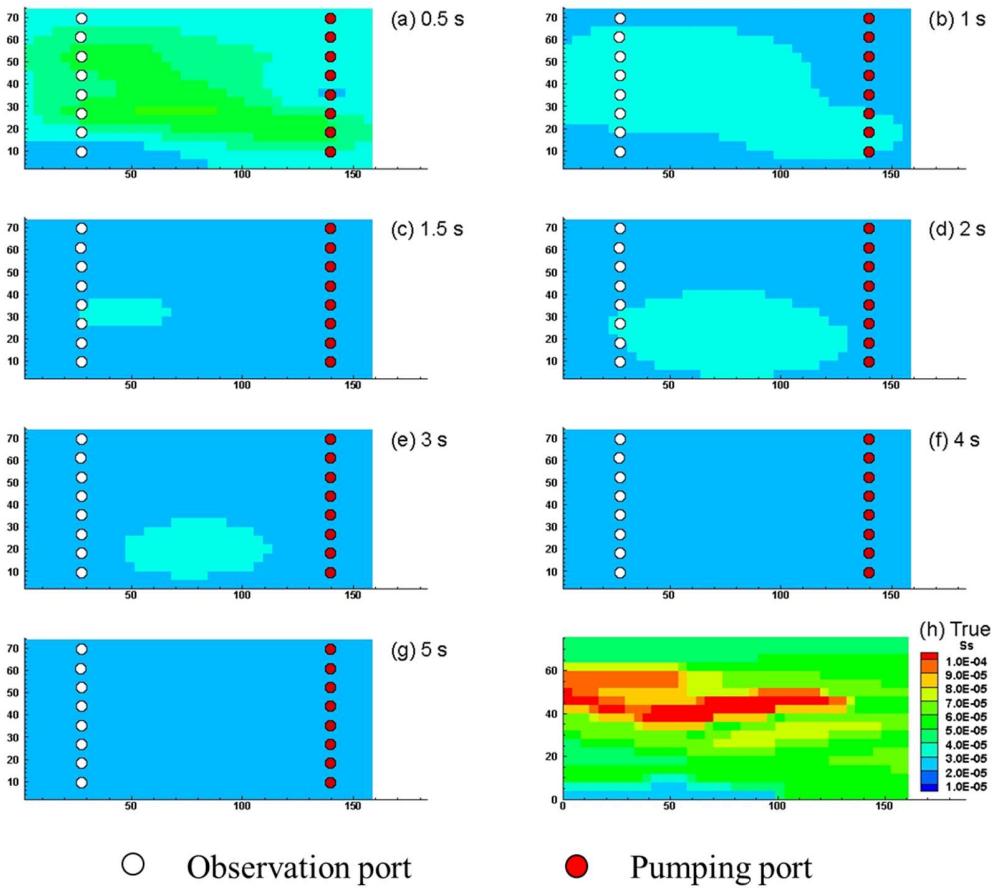

Figure S11 $\boldsymbol{S_s}$ tomograms from 'true' model and the geostatistical inversion approach for Case1 Scenario 1 with data at various times. Estimated $\boldsymbol{S_s}$ tomograms from GI at (a): 0.5 s; (b) 1 s; (c) 1.5 s; (d) 2 s; (e) 3 s; (f) 4 s; (g) 5 s; while (h) is the true $\boldsymbol{S_s}$.

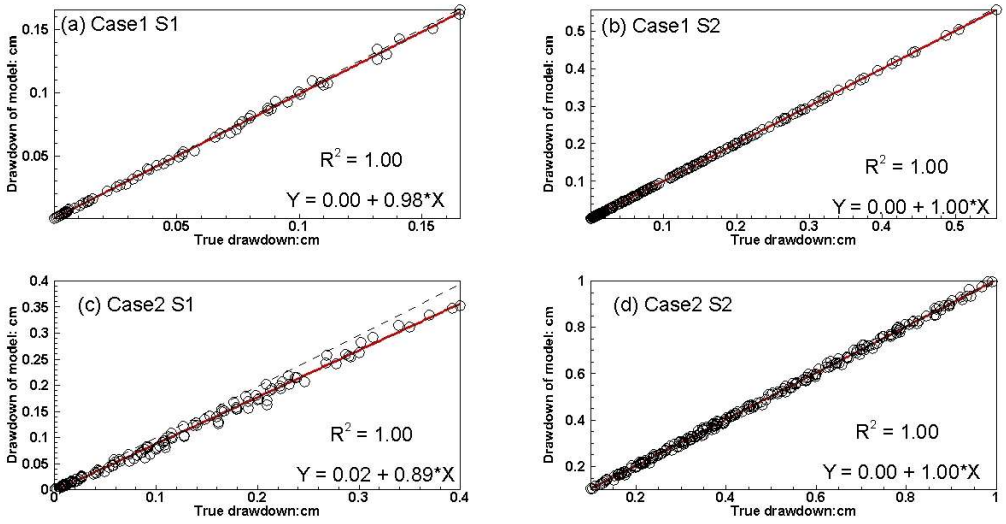

Figure S12. Calibration scatterplots of true drawdowns versus drawdowns simulated by estimated parameters from GI approach under: (a) Case 1 Scenario 1; (b) Case 1 Scenario 2; (c) Case 2 Scenario 1; (d) Case 2 Scenario 2.

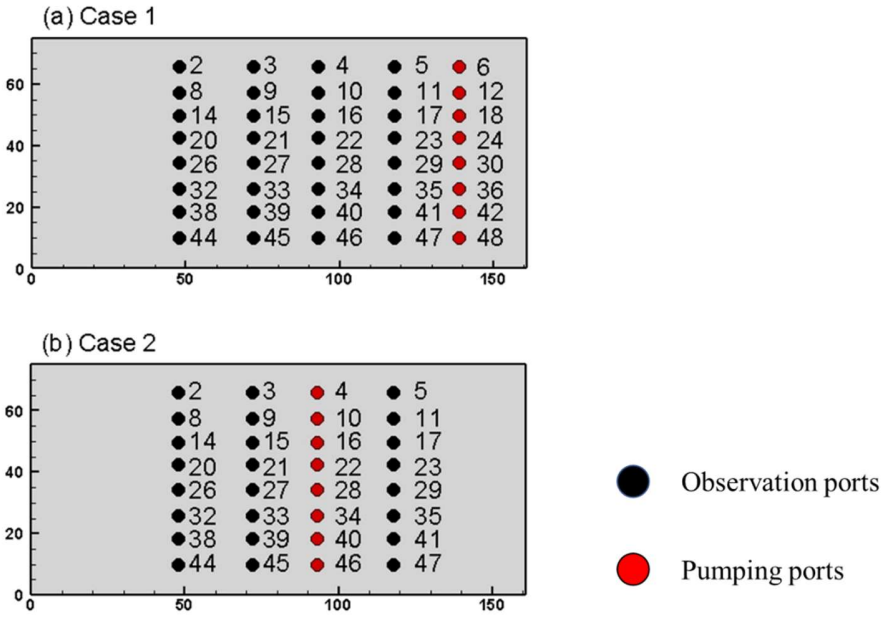

Figure S13. Schematic diagram showing the ports used for model validation on approach one: (1) Case 1; (b) Case 2.

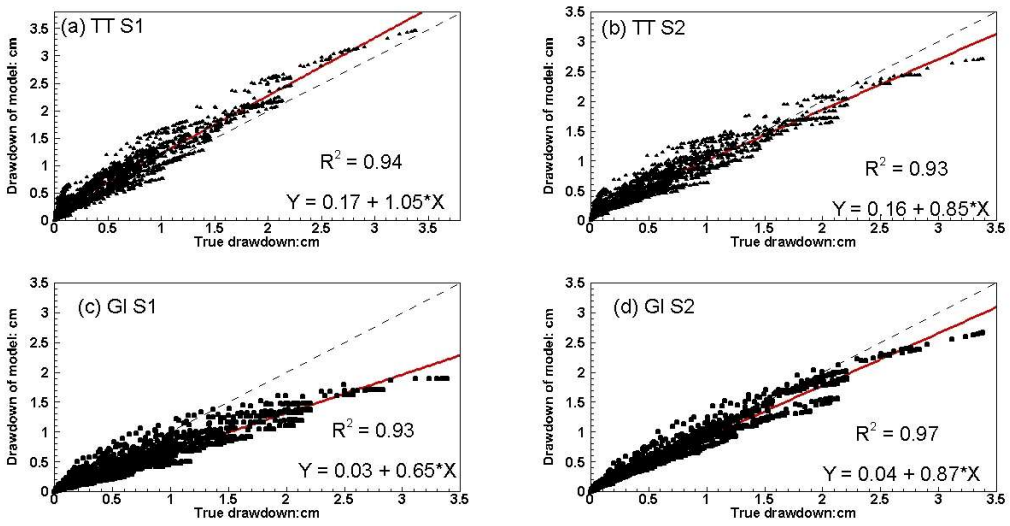

Figure S14. Comparison scatterplots of true drawdowns versus drawdowns simulated by estimated parameters from different models for Case 1 on validation approach one: (a) TTI with scenario 1; (b) TTI with scenario 2; (c) GI with scenario 1; and (d) GI with scenario 2.

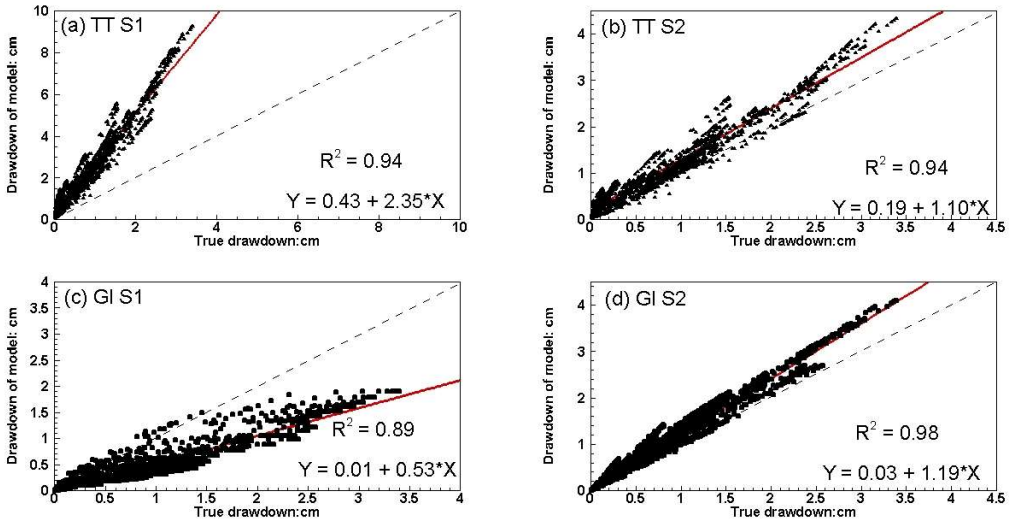

Figure S15. Comparison scatterplots of true drawdowns versus drawdowns simulated by estimated parameters from different models for Case 2 on validation approach one: (a) TTI with scenario 1; (b) TTI with Scenario 2; (c) GI with Scenario 1; and (d) GI with Scenario 2.

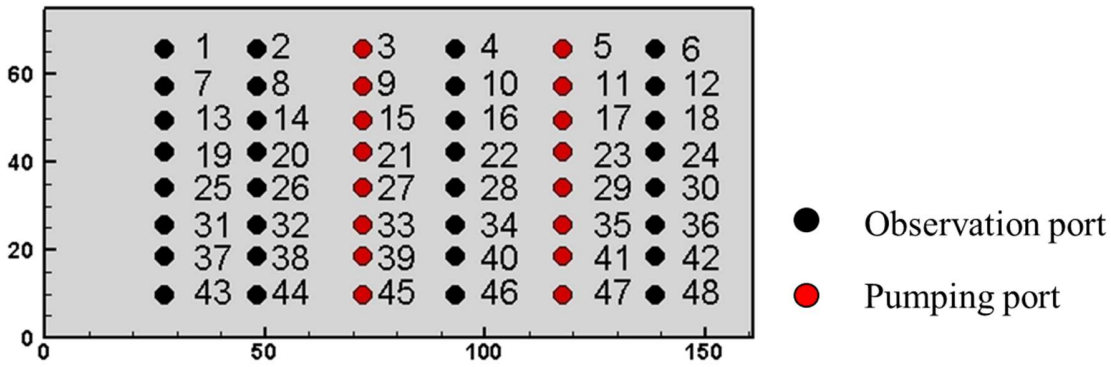

Figure S16. Schematic diagram showing the ports used for model validation for approach two.

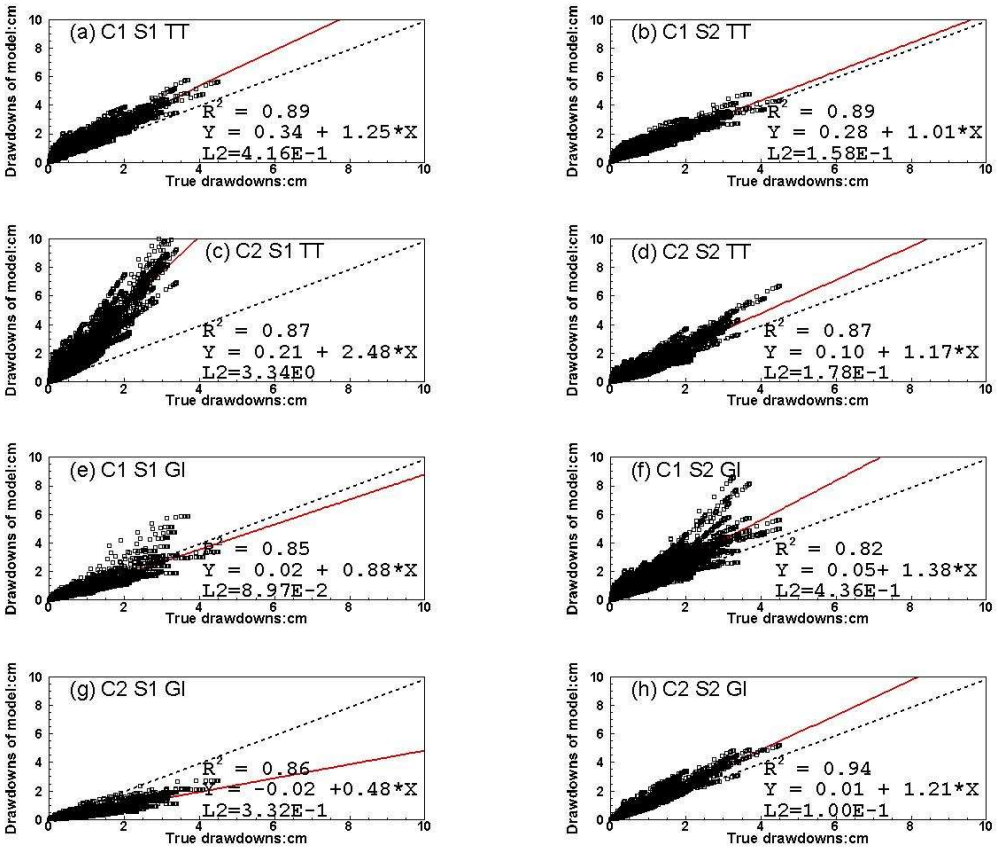

Figure S17. Comparison scatterplots of true drawdowns versus drawdowns simulated by estimated parameters from different models on validation approach two: (a) TTI with Case 1 Scenario 1; (b) TTI with Case 1 Scenario 2; (c) TTI with Case 2 Scenario 1; (d) TTI with Case 2 Scenario 2; (e) GI with Case 1 Scenario 1; (f) GI with Case 1 Scenario 2; (g) GI with Case 2 Scenario 1; (h) GI with Case 2 Scenario 2.

Table S1. Input values of K, S_s, and D for the analogue model.

Layer	K (cm/s)	S_s (/cm)	D (cm^2/s)
1	2.49×10^{-1}	2.39×10^{-5}	1.04×10^{4}
2	4.50×10^{-2}	5.75×10^{-5}	7.83×10^{2}
3	1.02×10^{-2}	5.37×10^{-5}	1.90×10^{2}
4	2.35×10^{-1}	5.03×10^{-5}	4.67×10^{3}
5	3.08×10^{-2}	6.06×10^{-5}	5.08×10^{2}
6	5.68×10^{-3}	4.68×10^{-5}	1.21×10^{2}
7	1.62×10^{-1}	6.68×10^{-5}	2.43×10^{3}
8	3.87×10^{-2}	5.48×10^{-5}	7.06×10^{2}
9	1.82×10^{-1}	7.71×10^{-5}	2.36×10^{3}
10	1.12×10^{-1}	6.69×10^{-5}	1.67×10^{3}
11	1.66×10^{-2}	5.21×10^{-5}	3.19×10^{2}
12	1.48×10^{-1}	1.11×10^{-4}	1.33×10^{3}
13	2.74×10^{-1}	1.23×10^{-4}	2.23×10^{3}
14	3.10×10^{-2}	8.73×10^{-5}	3.55×10^{2}
15	1.03×10^{-1}	6.37×10^{-5}	1.62×10^{3}
16	4.79×10^{-2}	9.30×10^{-5}	5.15×10^{2}
17	4.11×10^{-2}	5.78×10^{-5}	7.11×10^{2}
18	3.60×10^{-1}	4.71×10^{-5}	7.64×10^{3}
Geometric mean	6.78×10^{-2}	6.22×10^{-5}	1.08×10^{3}

Table S2. summarizes the data criteria used for two scenarios considered by TTI

	Scenario 1		Scenario 2	
	Selecting early travel-time	Constraining angle	Selecting early travel-time	Constraining angle
Case 1	×	×	√	√
Case 2	×	×	×	√

*√used; × abandoned.

Table S3. $L_1$and L_2 norms of top high D zone for various cases and scenarios.

	Travel time inversion		Geostatistical inversion	
	L_1	L_2	L_1	L_2
Case 1 Scenario 1	1.48	2.19	0.51	0.36
Case 1 Scenario 2	1.14	1.37	0.61	0.45

Case 2 Scenario 1	1.72	3.01	1.59	2.65
Case 2 Scenario 2	1.68	2.86	0.43	0.22

Table S4. Summary of L_1 and L_2 norms of validation for Cases 1 and 2 on validation approach one.

	Scenario	Method	Validation (drawdown)	
			L1	L2
Case 1	1	TTI	2.03×10^{-1}	6.45×10^{-2}
		GI	1.81×10^{-1}	7.86×10^{-2}
	2	TTI	1.43×10^{-1}	3.19×10^{-2}
		GI	6.88×10^{-2}	1.36×10^{-2}
Case 2	1	TTI	1.51×10^{0}	3.46×10^{-1}
		GI	3.95×10^{-1}	2.91×10^{-1}
	2	TTI	2.80×10^{-1}	1.20×10^{-1}
		GI	1.42×10^{-1}	7.29×10^{-2}

Table S5. Summary of $\boldsymbol{L_1}$ and $\boldsymbol{L_2}$ norms of validation for Cases 1 and 2 on validation approach two.

	Scenario	Method	Validation (drawdown)	
			L1	L2
Case 1	1	TTI	5.18×10^{-1}	4.15×10^{-1}
		GI	1.80×10^{-1}	8.98×10^{-2}
	2	TTI	3.17×10^{-1}	1.58×10^{-1}
		GI	3.52×10^{-1}	4.36×10^{-1}
Case 2	1	TTI	1.26×10^{0}	3.35×10^{0}
		GI	4.00×10^{-1}	3.31×10^{-1}
	2	TTI	2.89×10^{-1}	1.77×10^{-1}
		GI	1.90×10^{-1}	1.00×10^{-2}