AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2017WR020459

Key Points:
- A stochastic method is developed to fuse gravity measurements and hydraulic tomography to estimate heterogeneous aquifer properties
- This data fusion method extends the ability of hydraulic tomography to characterize kilometer-scale aquifers with high accuracy
- Numerical experiments demonstrate that the estimated parameters fields can predict head and gravity measurements in independent events

Supporting Information:
- Supporting Information S1

Correspondence to:
T.-C. J. Yeh,
ybiem@mac.hwr.arizona.edu

Citation:
Tsai, J.-P., Yeh, T.-C. J., Cheng, C.-C., Zha, Y., Chang, L.-C., Hwang, C., . . . Hao, Y. (2017). Fusion of time-lapse gravity survey and hydraulic tomography for estimating spatially varying hydraulic conductivity and specific yield fields. *Water Resources Research*, *53*, 8554–8571. https://doi.org/10.1002/2017WR020459

Received 23 JAN 2017
Accepted 1 SEP 2017
Accepted article online 12 SEP 2017
Published online 30 OCT 2017

Fusion of Time-Lapse Gravity Survey and Hydraulic Tomography for Estimating Spatially Varying Hydraulic Conductivity and Specific Yield Fields

Jui-Pin Tsai[1], Tian-Chyi Jim Yeh[2,3], Ching-Chung Cheng[1], Yuanyuan Zha[4], Liang-Cheng Chang[1], Cheinway Hwang[1], Yu-Li Wang[3], and Yonghong Hao[2]

[1]Department of Civil Engineering, National Chiao-Tung University, Hsinchu, Taiwan, [2]Key Laboratory for Water Environment and Resources, Tianjin Normal University, Tianjin, China, [3]Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, AZ, USA, [4]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, China

Abstract Hydraulic conductivity (K) and specific yield (S_y) are important aquifer parameters, pertinent to groundwater resources management and protection. These parameters are commonly estimated through a traditional cross-well pumping test. Employing the traditional approach to obtain detailed spatial distributions of the parameters over a large area is generally formidable. For this reason, this study proposes a stochastic method that integrates hydraulic head and time-lapse gravity based on hydraulic tomography (HT) to efficiently derive the spatial distribution of K and S_y over a large area. This method is demonstrated using several synthetic experiments. Results of these experiments show that the K and S_y fields estimated by joint inversion of the gravity and head data set from sequential injection tests in unconfined aquifers are superior to those from the HT based on head data alone. We attribute this advantage to the mass constraint imposed on HT by gravity measurements. Besides, we find that gravity measurement can detect the change of aquifer's groundwater storage at kilometer scale, as such they can extend HT's effectiveness over greater volumes of the aquifer. Furthermore, we find that the accuracy of the estimated fields is improved as the number of the gravity stations is increased. The gravity station's location, however, has minor effects on the estimates if its effective gravity integration radius covers the well field.

1. Introduction

Hydraulic conductivity (K) and specific yield (S_y) are important hydrogeological parameters that are pertinent to groundwater resources assessment and protection (e.g., Schwartz & Zhang, 2003). Conventionally, K and S_y are derived from a cross-well pumping test in an unconfined aquifer, which is time consuming and expensive due to the costs of well installations. As a result, estimates of detailed K and S_y spatial distribution are often limited. In addition, traditional analyses of cross-well pumping tests rely on homogeneous aquifer assumption, even though aquifers are inherently heterogeneous at many scales. The K and S_y parameters derived from the traditional pumping tests thus represent effective properties of an equivalent homogeneous aquifer. These effective parameters have been shown to vary with time although approaching the geometric means of the spatially varying K and S_y in the field, if the field is stationary (i.e., no spatial trends) (Wu et al., 2005). On the other hand, reliable assessment and management of groundwater resources demand the knowledge of the detailed distribution of these parameters of aquifers. Because of this reason, development of cost-effective methods for mapping their spatial distributions is essential.

Hydraulic tomography (HT) is a recently developed cost-effective approach (Yeh & Liu, 2000; Zhu & Yeh, 2005, and many others) for estimating the spatial distribution of aquifer parameters. Specifically, HT is a sequential pumping test that collects nonredundant information of aquifer heterogeneity by sequentially stressing the aquifer at different wells in a well field and collecting aquifer responses at other wells during each stress. This information is then analyzed using an inverse model to delineate the spatial distribution of aquifer properties (e.g., Bohling et al., 2007; Brauchler et al., 2003; Cardiff & Barrash, 2011; Yeh & Liu, 2000; Zhu & Yeh, 2005). The widely used inverse approach for HT analysis is the successive linear estimator (SLE) developed by Yeh et al. (1996), Yeh and Liu (2000), and Zhu and Yeh (2005), and it has been proven to be

effective (e.g., Hao et al., 2008; Illman et al., 2010; Illlman et al., 2012; Li & Yeh, 1999; Liu et al., 2002; Mao et al., 2013; Ni et al., 2009; Tso et al., 2016; Yeh & Lee, 2007; Yeh & Liu, 2000; Zha et al., 2015, 2016). SLE is a geostatistical-based interpretation technique that not only assimilates the head data sets from HT survey but also considers spatial statistics that describes the overall characteristics of heterogeneity of aquifers. It provides the unbiased estimates and addresses uncertainty associated with the estimates. Nevertheless, most previous HT studies have focused on confined aquifers and estimation of specific storage and saturated hydraulic conductivity. Only a few HT studies were applied in unconfined aquifers and to the estimation of K and S_y (e.g., Cardiff & Barrash, 2011). Recently, Mao et al. (2013) developed an HT analysis based on a three-dimensional variably saturated flow model to estimate variably saturated hydraulic properties (i.e., unsaturated hydraulic conductivity-pressure function and moisture release function) of the unconfined aquifer, instead of K and S_y.

Previous HT studies rely on the head/drawdown observations from the well field and prior knowledge of geologic structures via spatial statistics. Using only these data, HT often leads to a relatively high uncertainty of the estimates if the number of observation wells is insufficient or if the estimates are far away from the well field. In order to improve the K and S_y estimate in large areas with few observations, this study proposed a method that integrates time-lapse gravity survey with HT. The rationale for using gravity data is that the gravity is a direct measure of the mass change due to pumping or injection of water into or from the aquifer, and able to represent the hydrological state variables such as groundwater storage (Christiansen et al., 2011) as well as provides the groundwater mass information of kilometer-scale aquifers.

In the past, many studies (e.g., Christiansen et al., 2011; Gehman et al., 2009; Howle et al., 2003; Leiriao et al., 2009) have used gravimeters to measure the groundwater mass change to improve the accuracy of the groundwater storage monitoring and exploration. In addition, Blainey et al. (2007), Christiansen et al. (2011), Damiata and Lee (2006), Gehman et al. (2009), and Howle et al. (2003) have applied the gravity measurements to the aquifer parameter estimation. For example, Howle et al. (2003) measured the head and gravity changes for a 5 month injection to estimate the specific yield. Damiata and Lee (2006) developed a mathematical model to simulate the gravitational response to the hydraulic tests of unconfined aquifers. Blainey et al. (2007) applied Damiata and Lee (2006)'s model to the analysis of the uncertainty of the hydraulic conductivity and S_y estimates. These studies indicate that joint interpretation of drawdown and gravity data results in more unbiased and precise estimates than using gravity data alone. More recently, Wilson et al. (2012) applied a superconducting gravimeter for the time-lapse gravity measurement in the recharge zone of the karstic Edwards Aquifer. They identified the specific yield based on the time series of head and gravity. Christiansen et al. (2011) reported that gravity is sensitive to the hydrogeological parameters, especially S_y, and S_y estimates can be better constrained by gravity and head data than gravity alone.

Overall, previous studies have proven that integration of gravity and head data can improve the accuracy of the aquifer parameter estimation. However, these studies did not consider heterogeneity. For this reason, in this study, we developed a stochastic approach to estimate both heterogeneous K and S_y fields by jointly interpreting the head and gravity data sets derived from sequential injection tests in an unconfined aquifer. The proposed method is developed based on simultaneously successive linear estimator (SimSLE) with a 2-D groundwater flow model (VSAFT2; Yeh et al., 1993) and a gravity integration model.

Two numerical experiments are used to test the proposed method. Experiment #1 is designed to investigate the improvement on the S_y estimate due to the addition of gravity measurements, and the Experiment #2 is to examine the improvement on the S_y and K estimates due to the fusion of HT and gravity data. Lastly, the evaluation of these cases is undertaken by examining the ability of the estimated S_y and K fields to predict heads and gravity measurements during an independent pumping test.

2. Methodology

2.1. Equation of Flow in Unconfined Aquifers

For the simulation of the two-dimensional, depth-averaged, flow fields of unconfined aquifers with heterogeneous hydraulic properties, the Boussinesq equation is used, which described below:

$$\nabla_{xy}\left[KH\nabla_{xy}H\right]+Q(x_w,y_w)=(n_e+S_sH)\frac{\partial H}{\partial t}=S_y\frac{\partial H}{\partial t} \quad (1)$$

subject to the boundary and initial conditions

$$H|_{\Gamma_1}=H_1,\ -KH\nabla_{xy}H|_{\Gamma_2}=Q_2,\ H|_{t=0}=H_o,$$

where ∇_{xy} is the differential operator in the x and y directions of a horizontal plane, H denotes the saturated thickness of the aquifer, which is a function of x and y as well as time, t. K is the depth-averaged hydraulic conductivity, and S_y is the specific yield of the aquifer, which includes the effective porosity, n_e, and the product of the specific storage, S_s, and H. The effective porosity is used to denote the effect of filling or draining the pores due to the rise or fall of the water table. The S_sH term represents the elastic characteristic of porous media and water on the storage due to changes in the water table. This term is generally small in comparison with the porosity. Both K and S_y are spatial variables depending on location x and y. $Q(x_w,y_w)$ represents the pumping/injection rate per unit area at location (x_w,y_w)—the subscript w denotes the pumping/injection well. As shown in Yeh et al. (2015a), S_y is a byproduct of the depth-averaged procedure and the assumption of instantaneous drainage or filling of pore water as the water table falls or rises as well as the omission of the effects of the unsaturated zone above the water table. This assumption is likely not realistic in most geologic media (see Mao et al., 2011).

In this application, we modify the code of VSAFT2 (Variably Saturated Flow and Transport 2-D) developed by Yeh et al. (1993) to solve equation (1). VSAFT2 was developed for two-dimensional cross-sectional or cross-sectional asymmetrical variably saturated geologic media or two-dimensional depth-averaged fully confined aquifers. In order to adapt VSAFT2 for our study, specific storage is replaced by specific yield (S_y), and the K is multiplied by the saturated thickness of the aquifer, H. Because of the nonlinearity of equation (1), it is solved iteratively.

2.2. Gravity Simulation

To evaluate the gravitational attraction due to the groundwater storage of an aquifer, we develop a gravity simulation model based on the formulations proposed by Sato (1984). That is, we discretize the aquifer into three-dimensional finite rectangular blocks. Each block, b, (Figure 1) is defined by eight vertices denoted by $\left(x_{i,b},y_{j,b},z_{k,b}\right)$, where the subscripts, i, j, and k stand 1 or 2. Specifically, x_1 denotes the x coordinate of the vertices on the left-hand side of the block, and x_2 denotes the x coordinate of the vertices on the right-hand side of the block. The same notation applies to y and z coordinates. The term, $P\left(x_p,y_p,z_p\right)$, denotes the coordinates of the gravity station at a site. The distances from the gravity station to a vertex of each block are described by the vector $\left(r_{i,b},r_{j,b},r_{k,b}\right)$. θ is the polar angle between the vector and the fixed zenith direction. As a result, the vertical component of gravitational attraction can be written as

$$g_{z,b}=G\rho_wS_{y,b}\int_1^2\int_1^2\int_1^2\frac{r_{k,b}}{\sqrt{\left(r_{i,b}^2+r_{j,b}^2+r_{k,b}^2\right)^3}}dr_{i,b}dr_{j,b}dr_{k,b}, \quad (2)$$

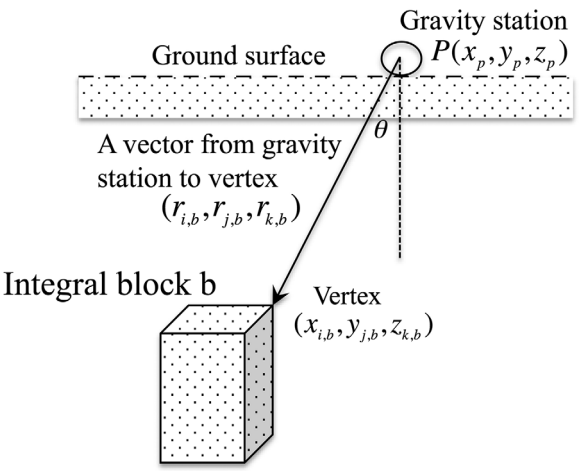

Figure 1. Cartesian coordinates of an observation point P and a rectangular prismatic block b.

where $r_{i,b}=x_{i,b}-x_p$, $r_{j,b}=y_{j,b}-y_p$, and $r_{k,b}=z_{k,b}-z_p$; ρ_w stands the groundwater density; and $S_{y,b}$ is the specific yield of the block b. G is the universal gravitational constant.

According to MacMillan (1958), Nagy (1966), and Banerjee and Das Gupta (1977), the integration of equation (2) can be evaluated using

$$g_{z,b}=\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}(-1)^{i+j+k}\times$$
$$G\rho_wS_{y,b}\left[r_{i,b}\log\left(\sqrt{D_b}+r_{j,b}\right)+r_j^b\ \log\left(\sqrt{D_b}+r_{i,b}\right)+2|r_{k,b}|\tan^{-1}\frac{\sqrt{D_b}+r_{i,b}+r_{j,b}}{|r_{k,b}|}\right], \quad (3)$$

where $D_b=(r_{i,b})^2+(r_{j,b})^2+(r_{k,b})^2$. Equation (3) determines the gravitational attraction along the z axis contributed by the groundwater mass of the block to the gravity station. In order to accurately calculate the gravity measurement due to

the mass of the groundwater from each location of the groundwater flow model, we divide the volume of water at each element into small blocks. That is, we first average the head values at the four nodes of each element to derive the average thickness of saturation. The volume defined by the thickness and the area of the element is then divided into a large number of small blocks. The net effect exerted on site P can be further determined by the summation of contributions from all blocks defined in the study area (N). That is,

$$g_z = \sum_{b=1}^{N} g_{z,b}. \tag{4}$$

To evaluate the sensitivity of gravity to specific yield, we replace the terms in the square bracket of equation (3) with a function, $C(r_{i,b}, r_{j,b}, r_{k,b})$. Thus, equation (3) can be organized as follows:

$$g_{z,b} = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} (-1)^{i+j+k} \times G\rho_w S_{y,b} C(r_{i,b}, r_{j,b}, r_{k,b}). \tag{5}$$

By taking the derivative of equation (5), we have the sensitivity of gravity to S_y:

$$\partial g_{z,b} / \partial S_{y,b} = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} (-1)^{i+j+k} \times G\rho_w C(r_{i,b}, r_{j,b}, r_{k,b}). \tag{6}$$

Since the size of a computational grid of groundwater model is much greater than that of gravity simulation, the sensitivity of gravity to the specific yield over a finite element groundwater model grid can be derived based on Taylor series expansion theory (Appendix A):

$$\partial g_z^f / \partial S_y^f = \sum_{p=1}^{p=w} \left(\partial g_{z,b} / \partial S_{y,b} \right)_p, \tag{7}$$

where f denotes a groundwater flow grid and w denotes the number of gravity simulation grid inside a flow computational grid.

2.3. Relationship Between Gravity Measurements and Mass in Space

Before discussing the fusion algorithm that integrates the gravity measurements and head data to derive the specific yield parameter distribution, we will illustrate the influence of mass at a location in the aquifer on the gravity measurement at a given station first. Since the groundwater mass of a volume of the aquifer (m) at a given location is unknown and it can vary from location to location, we will treat it as a spatial random field. As a result, the gravity measurement at a given location is also a random variable, and it changes with the location. Therefore, the measurement is also a random field. Suppose the cross correlations between the groundwater mass and the measured gravity (g) at the gravity station (Cor_{gm}) is

$$Cor_{gm} = \frac{\boldsymbol{R}_{gm}}{\sigma_{gg}\sigma_{mm}}, \tag{8}$$

where σ_{gg} and σ_{mm} are the standard deviations of gravity measurements and groundwater mass, respectively. $\boldsymbol{R}_{gm}(n_g \times n)$ denotes the cross-covariance matrix between the measured gravities and groundwater mass of the computational elements. n_g and n are the numbers of the observed gravities and discretized material blocks of the simulation area, respectively. Note that the matrices and vectors are shown in bold characters, and the dimension of each matrix is shown in parentheses.

A first-order analysis was used to estimate these covariance and cross-covariance matrixes $\boldsymbol{R}_{gg}$, $\boldsymbol{R}_{mm}$, and $\boldsymbol{R}_{gm}$, respectively. That is, $\boldsymbol{R}_{mm}$ is

$$\boldsymbol{R}_{m_i m_j} = h_i h_j \boldsymbol{R}_{P_i P_j} + P_i P_j \boldsymbol{R}_{h_i h_j} + 2h_i P_j \boldsymbol{R}_{p_i h_j}, \tag{9}$$

where h stands head and P denotes $\ln(S_y)$; the subscript i and j are the computational element number; $\boldsymbol{R}_{P_i P_j}$ denotes the covariance of $\ln(S_y)$ between element i and j; $\boldsymbol{R}_{h_i h_j}$ is the covariance of head observations between the element i and j; and $\boldsymbol{R}_{p_i h_j}$ stands the cross-covariance of $\ln(S_y)$ and head between the element i and j.

$\boldsymbol{R}_{gg}$ was derived by equation (10):

$$\boldsymbol{R}_{gg}=\boldsymbol{J}_{gm}\boldsymbol{R}_{mm}\boldsymbol{J}_{gm}^{T}, \tag{10}$$

where $\boldsymbol{J}_{gm}$ is the sensitivity of the gravity at the measured point to the groundwater mass of an element and is derived from a perturbation method. $\boldsymbol{R}_{gm}$ can be obtained by using equation (11):

$$\boldsymbol{R}_{gm}=\boldsymbol{J}_{gm}\boldsymbol{R}_{mm}. \tag{11}$$

The cross correlation between the gravity measurements and the mass of each computational element (equation (8)) is the spatial relationship between the mass and the gravity measurement. If the gravity measurement is available, this relationship allows one to estimate the likely mass distribution.

2.4. Fusion of Head and Microgravity Measurements on Specific Yield Estimation

The flow chart of the algorithm for assimilating HT with gravity measurements to estimate the spatial distribution of S_y is illustrated in supporting information Figure S1. The fundamental mathematical algorithm of this data fusion is based on the successive linear estimate (SLE), developed by Dr. Yeh and his colleagues (Yeh et al., 1995; Yeh & Hanna, 1996). SLE has evolved into sequential successive linear estimator (SSLE) and simultaneous successive linear estimator (SimSLE). Both have been the algorithm of HT for estimating aquifer parameters with the head data sets from sequential pumping tests (e.g., Mao et al., 2013; Xiang et al., 2009; Yeh & Liu, 2000; Zha et al., 2015, 2016; Zhu & Yeh, 2008; Zhu et al., 2011, and many others). Overall, SLE is a versatile data fusion method, based on the stochastic concept.

To implement data fusion algorithm in this paper, we first adopt a highly parameterized heterogeneous conceptual flow model (Yeh et al., 2015a, 2015b). This model discretizes a flow domain into n finite elements, and each element has a hydraulic parameter. The parameter (e.g., hydraulic conductivity or specific yield) over the entire domain are treated as spatial stochastic processes or random fields. This stochastic conceptualization allows us to characterize the spatial variability and uncertainty of the parameters.

In this study, we consider spatially varying natural log of hydraulic conductivity ($Y(\mathbf{x})=\ln K(\mathbf{x})$) and the natural log specific yield ($P(\mathbf{x})=\ln S_y(\mathbf{x})$) as two mutually independent stochastic processes in space (or spatial random fields), which can be expressed as

$$Y(\mathbf{x})=\overline{Y}(\mathbf{x})+y(\mathbf{x}) \text{ and } P(\mathbf{x})=\overline{P}(\mathbf{x})+p(\mathbf{x}), \tag{12}$$

where $\overline{Y}(\mathbf{x})$ and $\overline{P}(\mathbf{x})$ are the unconditional mean of the Y and P at the location $\mathbf{x}$, respectively; $y(\mathbf{x})$ and $p(\mathbf{x})$ are their perturbations. The vector $\mathbf{x}$ denotes the location of the center of a finite element. Note that the choice of the natural log of hydraulic conductivity and specific yield is to avoid negative K and S_y estimates. The perturbation has a zero mean and is characterized by their unconditional variances and a same exponential correlation function with the same isotropic correlation scales. Similarly, the gravity measured at any location is a stochastic process.

2.4.1. Cokriging

To jointly use gravity and head data set to estimate S_y field, SLE starts with cokriging. That is, we use a guess or known mean value of S_y to simulate the flow in an equivalent homogeneous aquifer under a given stress. With the knowledge of the variation of the saturated thickness of the aquifer, we then use equations (3) and (4) to calculate the gravity measurements $g(\mathbf{x}_k,t)$ at the measurement stations $\mathbf{x}_k$ and time t. Next, if measurements of S_y at $\mathbf{x}_j$ ($j=1,...n_s$) are available, we calculate their perturbations, $p^*(\mathbf{x}_j)$. Suppose we have an actual measurement of gravity at n_g stations and times from 1 to τ. Subsequently, using a stochastic linear estimator (cokriging) and measurements, $p^*(\mathbf{x}_j)$, we derive our first estimate of perturbation of S_y at every element of the groundwater simulation domain: $\hat{\boldsymbol{p}}(\mathbf{x}_\ell)$, $\ell=1,\ldots n$. That is,

$$\hat{\boldsymbol{p}}(\mathbf{x}_\ell)=\sum_{j=1}^{n_s}\boldsymbol{\lambda}^T p^*(\mathbf{x}_j)+\sum_{k=n_s+1}^{n_s+n_g}\sum_{t=1}^{t=\tau}\boldsymbol{\beta}^T(g^*(\mathbf{x}_k,t)-g(\mathbf{x}_k,t)),\ \ell=1,\ldots n, \tag{13}$$

where superscript T denotes transpose; $\hat{\boldsymbol{p}}$ is a $n\times1$ vector; $\boldsymbol{p}^*$ is a $n_s\times1$ vector, $g^*(\mathbf{x}_k,t)$ is the gravity measurements at the location $\mathbf{x}_k$ at time t; and $g(\mathbf{x}_k,t)$ is the simulated gravity at location $\mathbf{x}_k$ and time t. $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$ are a $n_s\times n$ and a $(n_g\cdot\tau)\times n$ weight matrix, respectively. They can be derived by solving equation (14):

$$\begin{bmatrix} \boldsymbol{R}_{p^*p^*} & \boldsymbol{R}^{1}_{p^*g^*} & \boldsymbol{R}^{\bullet}_{p^*g^*} & \boldsymbol{R}^{\tau}_{p^*g^*} \\ \boldsymbol{R}^{1}_{g^*p^*} & \boldsymbol{R}^{11}_{g^*g^*} & \boldsymbol{R}^{1\bullet}_{g^*g^*} & \boldsymbol{R}^{1\tau}_{g^*g^*} \\ \boldsymbol{R}^{\bullet}_{g^*p^*} & \boldsymbol{R}^{\bullet 1}_{g^*g^*} & \boldsymbol{R}^{\bullet\bullet}_{g^*g^*} & \boldsymbol{R}^{\bullet\tau}_{g^*g^*} \\ \boldsymbol{R}^{\tau}_{g^*p^*} & \boldsymbol{R}^{\tau 1}_{g^*g^*} & \boldsymbol{R}^{\tau\bullet}_{g^*g^*} & \boldsymbol{R}^{\tau\tau}_{g^*g^*} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\beta}^{1} \\ \boldsymbol{\beta}^{\bullet} \\ \boldsymbol{\beta}^{\tau} \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_{p^*\hat{p}} \\ \boldsymbol{R}_{g^*\hat{p}}{}^{1} \\ \boldsymbol{R}_{g^*\hat{p}}{}^{\bullet} \\ \boldsymbol{R}_{g^*\hat{p}}{}^{\tau} \end{bmatrix}, \tag{14}$$

where $\boldsymbol{R}_{p^*p^*}$ is an $n_s \times n_s$ unconditional covariance matrix between measurement locations of p^*; $\boldsymbol{R}^{1}_{p^*g^*}$, $\boldsymbol{R}^{\bullet}_{p^*g^*}$, and $\boldsymbol{R}^{\tau}_{p^*g^*}$ are $n_p \times n_g$ cross-covariances between p^* and g^* measurements at $t = 1$, $\bullet$, and τ, respectively. Note that $\bullet$ presents any times between $t = 1$ and τ. $\boldsymbol{R}^{11}_{g^*g^*}$ is a $n_g \times n_g$ covariance matrix of measurements g^* at different locations at $t = 1$; the covariance between measurement locations of g^* at $t = 1$ and $t = \bullet$ is denoted by $\boldsymbol{R}^{1\bullet}_{g^*g^*}$, while the covariance between g^* at $t = 1$ and $t = \tau$ is represented by $\boldsymbol{R}^{1\tau}_{g^*g^*}$. The same superscript notation applies to other $\boldsymbol{R}_{g^*g^*}$. Now, $\boldsymbol{R}_{p^*\hat{p}}$ is a $n_s \times n$ matrix, representing the covariance between $\hat{\boldsymbol{p}}$ (the parameter values at the locations where they are to be estimated) and $\boldsymbol{p}^*$. The covariance between the gravity measurements $g^*(\mathbf{x}_k, k{=}1, \ldots n_g)$ at $t = 1$ and the estimates, $\hat{\boldsymbol{p}}$, is represented by $\boldsymbol{R}_{g^*\hat{p}}{}^{1}$, which is a $n_s \times n$ matrix. The superscript notation used previously applies here as well.

A first-order analysis derives $\boldsymbol{R}_{g^*g^*}$ and $\boldsymbol{R}_{pg^*}$ in equation (14). That is,

$$\begin{aligned} \boldsymbol{R}_{g^*g^*}(\mathbf{x}_i, \mathbf{x}_j, t_k, t_l) &= \boldsymbol{J}_{g^*p}(\mathbf{x}_i, \mathbf{x}_j, t_k)\boldsymbol{R}_{pp}(\mathbf{x}_i, \mathbf{x}_j)\boldsymbol{J}^{T}_{g^*p}(\mathbf{x}_i, \mathbf{x}_j, t_l) \\ \boldsymbol{R}_{pg^*}(\mathbf{x}_i, \mathbf{x}_j, t_k, t_l) &= \boldsymbol{J}_{g^*p}(\mathbf{x}_i, \mathbf{x}_j, t_k)\boldsymbol{R}_{pp}(\mathbf{x}_i, \mathbf{x}_j) \end{aligned}. \tag{15}$$

In equation (15), $\boldsymbol{R}_{pp}$ is the unconditional exponential covariance of p at all elements, while $\boldsymbol{J}_{g^*p}$ is the sensitivity of g at measurement location with respect to p at each element in the domain and is derived from equations (6) and (7). The residual covariance of the estimate is derived from equation (16).

$$\varepsilon^{(1)}_{pp} = \boldsymbol{R}_{pp} - \boldsymbol{\lambda}\boldsymbol{R}_{pp^*} - \boldsymbol{\beta}\boldsymbol{R}_{pg^*}, \tag{16}$$

in which ε_{pp} is a $n \times n$ matrix and it reflects the reduction in the uncertainty of the estimated parameter field due to the inclusion of $p^*(\mathbf{x}_j)$ and gravity measurements, $g^*(\mathbf{x}_k, t)$.

2.4.2. HT Analysis

Afterward, the estimated $\hat{\boldsymbol{p}}$ is added to its unconditional mean and converted to S_y. Along with its residual covariance (i.e., equation (16)), they are input to HT analysis to include the head measurements for improving the estimate of $\hat{\boldsymbol{p}}$ and estimating $\hat{\mathbf{y}}$. SimSLE used in HT analysis has been widely accepted and reported by Yeh and his colleagues. Here we only briefly discuss some key procedures for deriving $\hat{\boldsymbol{p}}$ below.

Because the head is nonlinearly correlated with S_y, successive iterations of the linear estimator are required to extract all possible information from head measurements about S_y. That is,

$$\hat{\boldsymbol{P}}^{(r+1)} = \hat{\boldsymbol{P}}^{(r)} + \omega^{(r)}_{h}(\boldsymbol{h}^*(\mathbf{x}_m, t_n, q_k) - \hat{\boldsymbol{h}}^{(r)}(\mathbf{x}_m, t_n, q_k)), \tag{17}$$

where r are the iteration indexes of SimSLE. $\hat{\boldsymbol{P}}^{(r)}$ is the S_y estimates from the previous iteration. $\hat{\boldsymbol{P}}^{(r+1)}$ is the improved estimates of S_y after inclusion the head measurements, $\boldsymbol{h}^*$, at an observation well, $\mathbf{x}_m$, at the time, t_n, during injection q_k. $\hat{\boldsymbol{h}}^{(r)}$ is the simulated head based on $\hat{\boldsymbol{P}}^{(r)}$, and $\omega^{(r)}_{h}$ is the weighting matrix. The residual covariance of $\boldsymbol{p}$, $\varepsilon^{(r)}_{pp}$, is updated according to equation (18), and the diagonal term of the updated residual covariance matrix ($\varepsilon^{(r+1)}_{pp}$) is the uncertainty of the estimate of each element.

$$\varepsilon^{(r+1)}_{pp} = \varepsilon^{(r)}_{pp} \text{-} \boldsymbol{\omega}^{(r)}\varepsilon^{(r)}_{h^*p}, \tag{18}$$

where $\varepsilon^{(r)}_{h^*p}$ is the cross-covariance between $\hat{\boldsymbol{h}}^{(r)}$ and $\hat{\boldsymbol{P}}^{(r)}$.

2.4.3. Update S_y Estimate Using Gravity Difference

After SimSLE's solution converges, and the end of the oth outer iteration is reached, the convergence criteria of gravity are checked. If the convergence criteria of gravity are met, the estimation procedure stops. Otherwise, the gravity residuals and residual covariance ε^{o}_{pp} are updated and input to equation (21) to produce a new estimate of $\boldsymbol{p}$ as the initial estimate of SimSLE again:

$$\hat{\boldsymbol{P}}^{o+1} = \hat{\boldsymbol{P}}^{o} + \omega^{o}_{g}(\boldsymbol{g}^* - \hat{\boldsymbol{g}}^{o}). \tag{19}$$

In equation (19), $\hat{\boldsymbol{P}}^{o}$ is the conditional estimate of the $\ln(S_y)$ field at the end of the oth outer iteration. $\hat{\boldsymbol{P}}^{o+1}$ is the new estimate for the SimSLE at $(o + 1)$th outer iteration and $\boldsymbol{g}^{(r,o)}$ is the simulated gravity data set

obtained from the conditional mean equation (i.e., equation (5)) using the estimate of the conditional mean $\boldsymbol{P}_c^o$ at the end of the oth outer iteration. $\boldsymbol{\omega}_g^o$ is derived from equation (20):

$$\boldsymbol{\omega}_g^o \varepsilon_{gg}^o + \theta_g \boldsymbol{I} = \varepsilon_{gp}^o, \tag{20}$$

where $\boldsymbol{I}$ is the identity matrix and θ_g is the stabilizer. They are used to improve the conditional number of equation (20). θ_g is determined dynamically according to a specified multiplier and the maximum value of the diagonal terms of ε_{gg}^o at the end of the oth outer iteration. ε_{gg}^o ($n_g \times n_g$) is the conditional covariance between observed gravity data set; ε_{gp}^o ($n_g \times n$) is the conditional cross-covariance between the observed gravity data set and the parameter $\boldsymbol{p}$ representing the conditional perturbation of $\ln(S_y)$ field (see equation (12)) at the end of the oth outer iteration. ε_{gg}^o and ε_{gp}^o can be obtained from equation (21):

$$\begin{aligned} \varepsilon_{gg}^o &= \boldsymbol{J}_g^o \varepsilon_{pp}^o \boldsymbol{J}_g^{oT} \\ \varepsilon_{pg}^o &= \boldsymbol{J}_g^o \varepsilon_{pp}^o \end{aligned}, \tag{21}$$

where $\boldsymbol{J}_g^o$ stands the Jacobin (or sensitivity) matrix for the gravity data with respect to $\boldsymbol{p}$ using the parameters estimated at the end of the oth outer iteration and is derived from perturbation method. The residual covariance of $\boldsymbol{p}$ is updated according to equation (22),

$$\varepsilon_{pp}^{o+1} = \varepsilon_{pp}^o - \boldsymbol{\omega}_g^o \varepsilon_{gp}^o, \tag{22}$$

where the residual covariance ε_{pp}^o was derived from the final estimate of SimSLE (equation (18)) at the end of the oth outer iteration. The updated residual covariance is input to the $(o+1)$th SimSLE to calculate the covariance of the observed head and cross-covariance between $\boldsymbol{p}$ and observed head (equation (18)). The outer iteration and SimSLE continue until the convergence criteria are met, i.e., the maximum gravity residual divided by its corresponding observed gravity at $(o+1)$th outer iteration is less than 0.001 or the number of the outer iteration is greater than 100. A flowchart of the proposed algorithm is given in Figure S1 in the supporting information.

Note that if the necessary conditions for the estimation problem to be well defined are not given, the number of possible values for each parameter will be infinite and the solution is uncertain (Yeh et al., 2015b). For such ill-defined estimation problems, the SimSLE seeks the most likely, conditional, effective S_y fields, which will honor the measurements of the parameters (conditioning data) and the measurements of aquifer responses (such as heads and gravity) at sampling locations. In addition, these conditional, effective parameter fields will produce a statistically unbiased prediction of flow and gravity fields under different pumping events.

3. Numerical Experiments

3.1. Aquifer Characteristics and HT and Gravity Survey Setups

Two numerical experiments are conducted to test the proposed method. Experiment #1 addresses the issue if the gravity measurements can improve the HT's S_y estimates when the K field is known and not estimated. Experiment #2 investigates the improvements on S_y and K estimates by the method.

In both experiments, we use a two-dimensional horizontal synthetic aquifer of 3,000 × 3,000 m, which is discretized into grids of 100 × 100 m in dimension. A constant total head boundary is assigned to the left and the right boundaries with a value of 10 m, and the no-flow boundary is assigned to the upper and bottom boundaries. The initial head everywhere is 10 m. The simulation periods of the injection events of Experiment #1 are 60 days, and that of Experiment #2 is 1,000 days in order to reach a steady state condition. The reference K field of the synthetic aquifer in Experiment #1 is homogenous and has a value of 50 m/d, and that of Experiment #2 is a random field with a mean of 50 m/d, characterized by an exponential correlation structure with a variance of 10 m^2/d^2, and an isotropic correlation scale of 500 m (Figure 2a). For the reference S_y field of Experiments #1 and #2, the random field is generated with a mean of 0.13, characterized by an exponential correlation structure with a variance of 0.001, and a correlation scale of 500 m (Figure 2b).

The hydraulic tomography (HT) survey uses nine wells, installed in the aquifer (solid black circle in Figure 2), and the distance between each of two neighbor wells is 400–500 m. Nine injection tests are sequentially conducted at the nine wells with a constant injection rate of 4,000 m^3/d. During each injection test, head

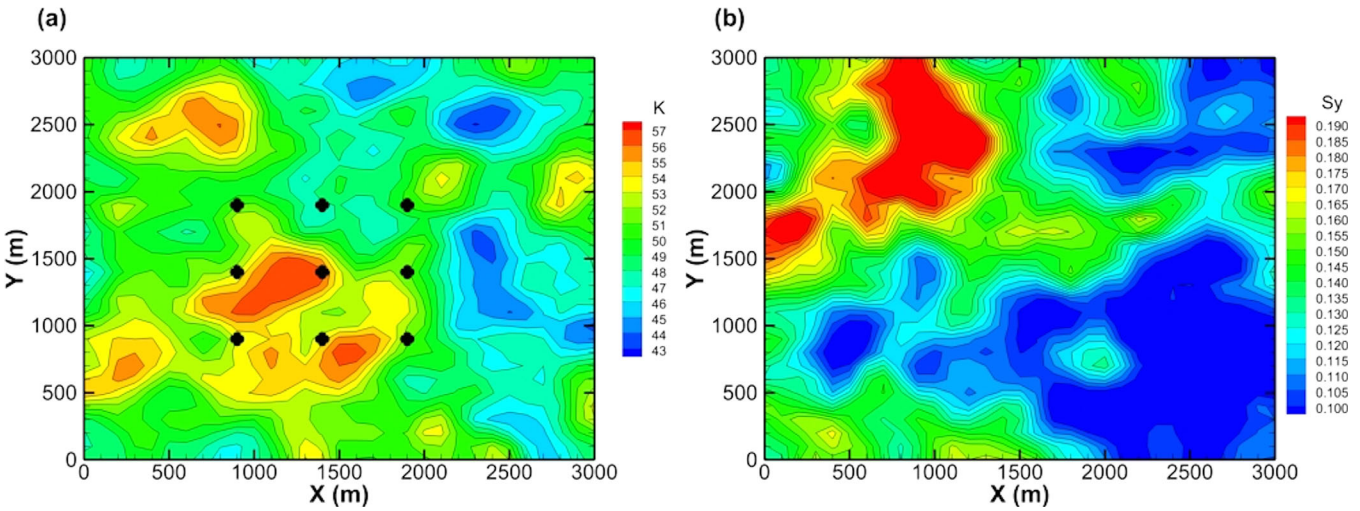

Figure 2. (a) Reference K random field. (b) Reference S_y random field for Experiment #2.

data sets are collected from the nine wells, which are subsequently used in HT analysis to estimate S_y in Experiment #1 and both K and S_y fields in Experiment #2.

Because we use 2-D depth-averaged groundwater model, the actual ground surface is not specified in the synthetic aquifer. However, the gravity simulation requires the ground surface elevation. Further, in order to make sure the elevation of the ground surface is higher than simulated heads, we add 1 m to the highest groundwater level of the last time step among all the injection events as the ground surface, which is set as the elevation of gravity stations. This height is the thickness of the synthetic water table aquifer.

Before conducting the two experiments, we investigate the effective gravity integration radius. That is, we aim to determine the radius of the gravity integration, which will capture the most mass of water in the aquifer. We use a synthetic, homogenous aquifer of a size of 3,000 m $\times$ 3,000 m without any injection (a uniform saturated thickness of 15 m) to calculate the gravity measured at the center of these aquifers with different gravity integration radius. For this gravity simulation, we discretize the aquifer with a grid size of 50 cm $\times$ 50 cm $\times$ 50 cm to assure the accuracy of the gravity integration (i.e., equation (3)). Figure 3 shows that the simulated gravity measurement at the center of the domain increases with the integration radius from 100 to 1700 m. It reaches asymptotically a stable value as the integration radius greater than 1000 m. The simulated gravity by using the effective integration radius of 1000 m is 115.07 *ugal* (a gravimetric unit), which contributes 99.76% of the total gravity (i.e., 115.35 *ugal*) to the gravity station. Therefore, we set the effective integration radius as 1,000 m for gravity integration calculation.

3.2. Data Selection

The cross-correlation analysis by Sun et al. (2013) reported that S_y highly correlates to the early time periods of the well hydrograph. As a result, we sample the head at early time steps 0.1, 0.2, 0.5, 1, 20, 40, and 60 days at the nine wells in each of the nine injection tests in Experiment #1. For Experiment #2, where we simultaneously estimate S_y and K fields, and we sample the head according to the characteristics of each observed well hydrography. That is, the sampling times are different for each well at early times, and at 900 and 1,000 day when the groundwater flow reaches steady state condition.

The gravity data are sampled at five gravity stations (white circles in Figure 5b) at the same time steps as head observations (i.e., only the early time steps) in Experiment #1. A total of 567 head observations and 35 gravity data are used for the joint inversion of the S_y field. For Experiment #2, the gravity data are sampled at the same locations and time steps as in Experiment #1, and a total of 480 head observations and 35 gravity data sets are used to simultaneously estimate S_y and K fields.

The SimSLE algorithm requires the input of the initial mean value, correlation scales, and spatial variance of K and S_y. Specifically, the mean of K is given as 50 m/d, correlation scale is 500 m, and the variance is 10 m^2/d^2. For S_y, the mean is 0.13, correlation scale is 500 m, and the variance is 0.001, respectively. Previous

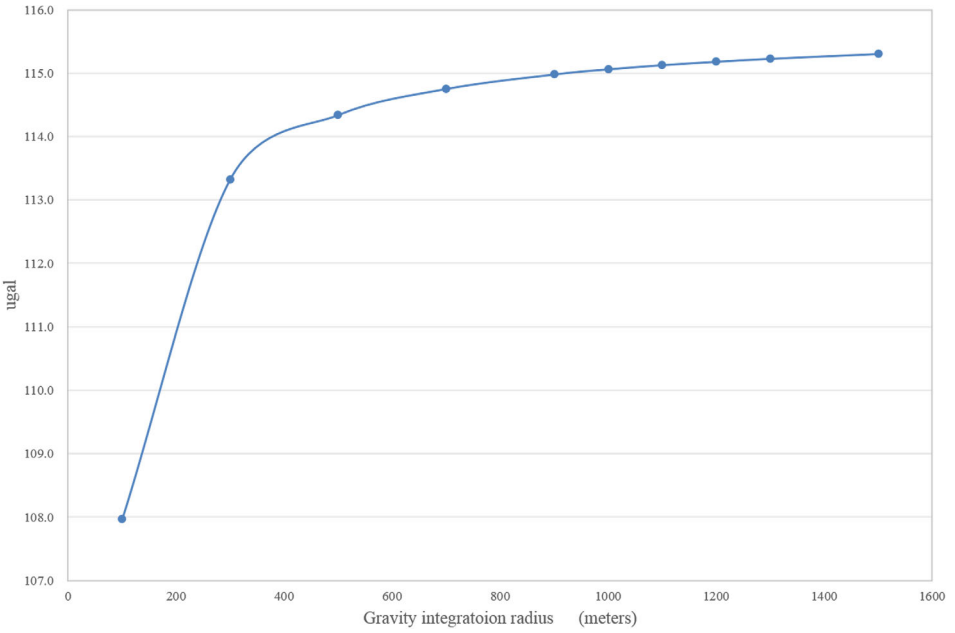

Figure 3. Effects of gravity integration radius from 100 to 1,700 m by using use a synthetic uniform saturated aquifer of a thickness of 15 m and a size of 3,000 m × 3,000 m without any injection.

HT studies report that the accuracy of these spatial statistics does not influence the results of stochastic estimator since abundant head data set are used in the HT analysis (Mao et al., 2013; Yeh & Liu, 2000).

All data in this study are assumed to be error-free to examine the validity of the proposed algorithm. This is because the robustness of the SimSLE with data infested with noise for flow in saturated aquifers has been demonstrated and discussed by Xiang et al. (2009). Further, Mao et al. (2011) demonstrated that as long as sufficient time data at a given observation location are used in inverse modeling of unsaturated flow problems, the effects of noise can be reduced due to least-square nature of the SimSLE algorithm.

3.3. Evaluation Criteria

The estimated parameter or head field is evaluated by comparing with its reference field, using average absolute error norm L_1, mean-square error norm L_2, which are defined as follows:

$$L_1=\frac{1}{n}\sum_{i=1}^{n}|\hat{Z}_i-Z_i|,\quad L_2=\frac{1}{n}\sum_{i=1}^{n}\left(\hat{Z}_i-Z_i\right)^2, \tag{23}$$

where $\hat{Z}_i$ and Z_i denote the estimated and the known value of parameters or heads or gravity measurements (depending on the quantity to be evaluated) in the reference field, respectively. The R^2 are derived by the following equation:

$$R^2=1-\left(\sum_{i=1}^{n}(Z_i-\bar{Z})^2/\sum_{i=1}^{n}(\hat{Z}_i-\bar{\hat{Z}})^2\right), \tag{24}$$

where $\bar{\hat{Z}}$ and $\bar{Z}$ stand for the mean of the estimated and the true fields, respectively. In addition, a linear regression line is used in the scatterplot; the slope and intercept are used to evaluate the bias of the estimates or predictions. These criteria are also used to evaluate the predicted head and gravity based on the estimated parameter fields at the validation step (section 4.4).

4. Results and Discussion

4.1. Cross-Correlation Analysis

The SimSLE algorithm is built upon the spatial cross correlation between the parameters to be estimated and the measurements of the physical process. Note that the spatial cross correlation comprises the

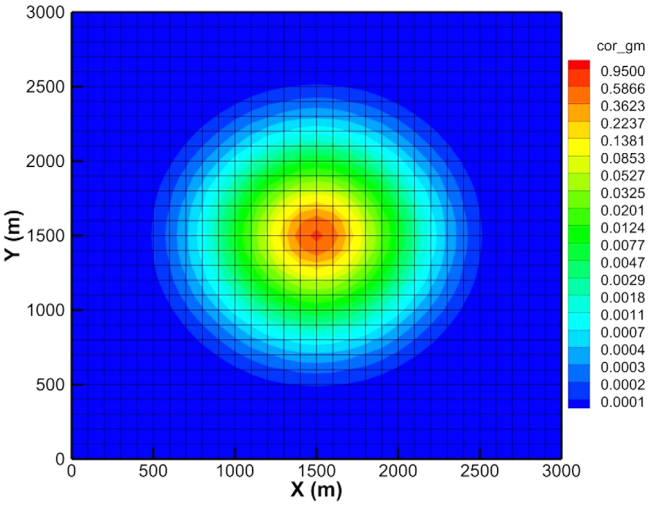

Figure 4. Contour map of the cross correlation between the groundwater mass everywhere and the gravity station at the center of the simulation area.

sensitivity of observed state variables at a location to the change of the parameters at any location in the model domain, and spatial correlation structure of the parameters (spatial statistics). In order to illustrate the spatial relationship between a gravity measurement at a station and S_y of each computational element in the study domain, we conduct a cross-correlation analysis based on the Experiment #1. The calculated cross-correlation map at the injection time of 0.1 days is shown in Figure 4. The value of the contour lines decays radially away from the gravity measurement point. The concentric circles of the contour map suggest that an anomaly anywhere at the same radial distance can contribute equally to the gravity measurement. That is to say, a gravity measurement represents the total groundwater mass/storage of the entire synthetic aquifer, and it cannot differentiate the detailed location where the anomaly is.

According to the above analysis, we conclude that the gravity measurement is a regional mass measurement. It can only be used as the total mass constraint for the hydrogeological inverse modeling. For delineating the detailed parameter spatial pattern, collecting the spatial head data set from sequential injection test becomes essential since it can reveal spatial heterogeneity of the parameter. For this reason, we believe that a joint inversion of the gravity and head data sets could yield more accurate estimates of K and S_y fields than HT alone.

4.2. Experiment #1: Improvement on S_y Estimates Due To Addition of Gravity Measurements

In this experiment, two test cases are examined. In Case (1-1), only head data sets from HT are used to identify S_y field. In Case (1–2), we identify the S_y by fusion of head and gravity data set. Specifically, this case uses the head measurements from the nine injection tests and the gravity measurements (see section 3.2) during the first injection test (i.e., the injection well located at $x = 1{,}400$ and $y = 1{,}400$) to identify the S_y field. The K fields are assumed known and are excluded from estimation in these two cases, and improvements of the S_y estimate at near and far fields are discussed.

By comparing the S_y estimates at the near field (i.e., the elements inside the red hollow rectangle in Figures 5a and 5b) of Case 1-1 and Case 1–2 with the reference field in Figure 2b, we observe that both HT and HT with gravity measurements capture the spatial trend of the true S_y field. Nonetheless, the S_y estimate from the fusion of HT and gravity is better than that of HT. For example, the pattern of relatively high and low S_y values (e.g., 0.12– and 0.18+) of Figure 5b are closer to those of the true S_y field than those of Figure 5a. At the far field (i.e., the elements outside the rectangle), the S_y estimates close to the five gravity stations in

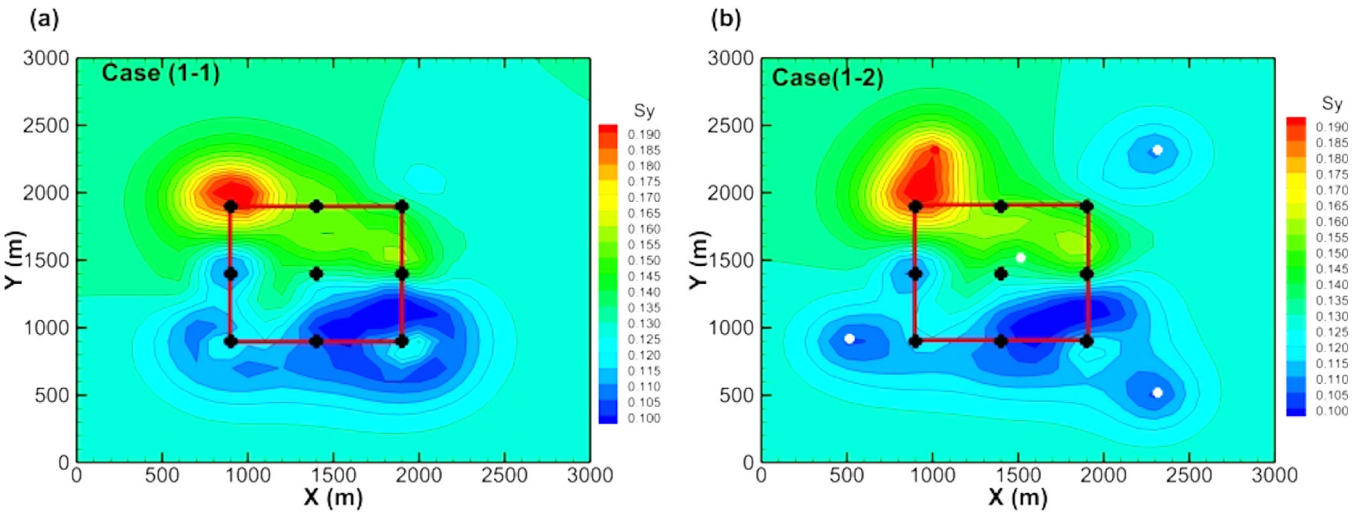

Figure 5. The spatial patterns of S_y estimates obtained from (a) HT without gravity measurements, Case (1-1) and (b) with gravity measurements, Case (1–2) of Experiment #1.

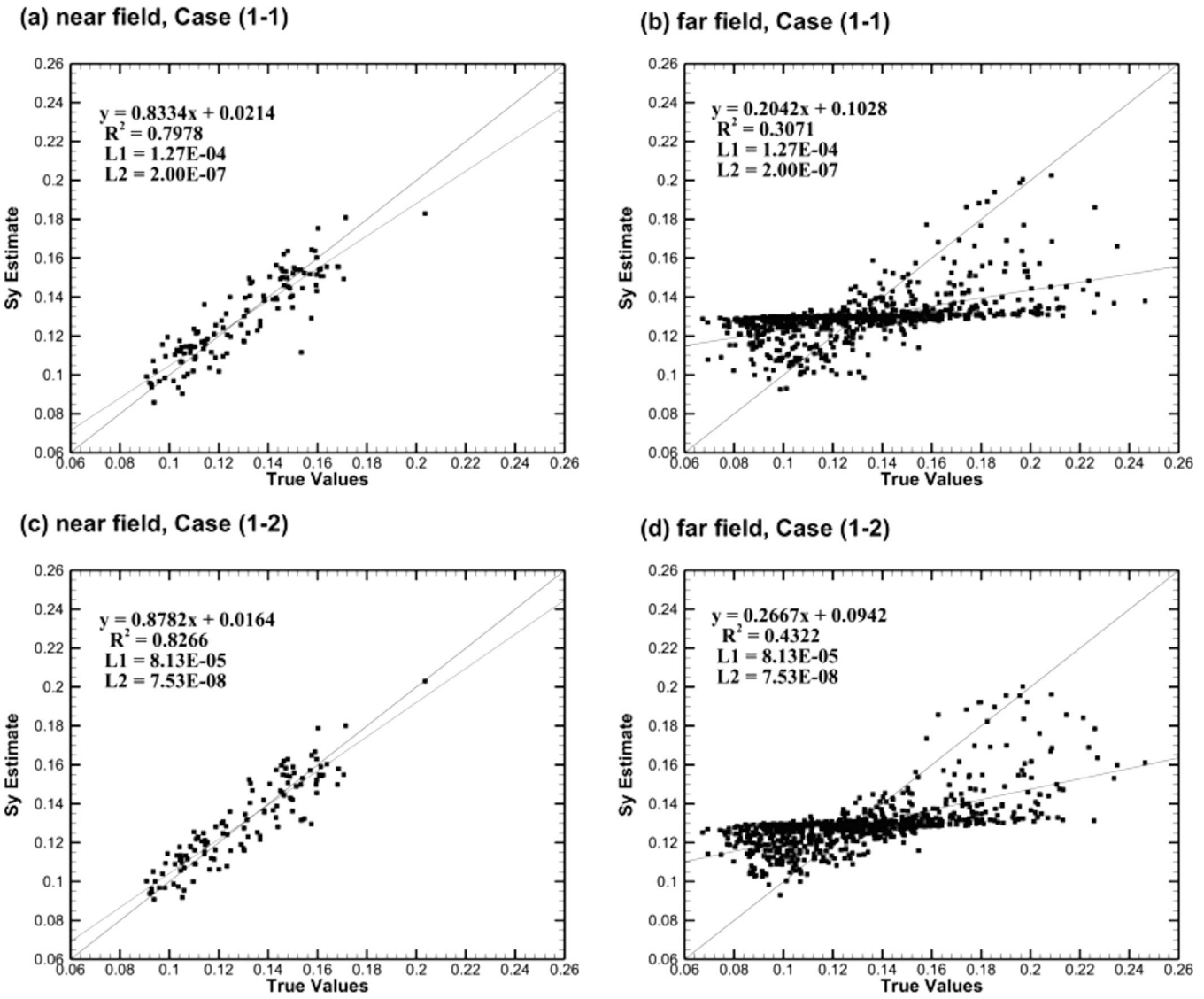

Figure 6. Scatterplots of the estimated versus true S_y of Cases (1-1) (a) near field and (b) far field; Those for Case (1–2) (c) near field and (d) far field.

Case (1–2) (see Figure 5b) are in agreement with the true values (Figure 2b]) and are better than the estimates by HT alone.

The scatterplots of the estimated versus true S_y estimates at the near fields of these two cases, and the evaluation criteria are illustrated in Figures 6a and 6c. These figures show that the regression line of the estimates of Case (1–2) is slightly closer to the 45° line than that of Case (1-1). The scattering of points around the regression lines of Case (1–2) is less than that of Case (1-1). As a result, the HT's S_y estimates at the near field are improved by the proposed method.

For the S_y estimate at the far field, the improvement of the estimate is relatively significant. Examining the spatial distribution of the S_y estimate of these two cases (Figures 5a and 5b) and the reference field (Figure 2b), we observe that the estimates close to the gravity stations are almost identical to the true values, especially for the estimate at the grids where the gravity stations are. This result indicates that parts of the S_y estimates at the far field are indeed improved.

The scatterplots and evaluation statistics of the estimated versus true S_y estimates at the far fields of these two cases are shown in Figures 6b and 6d. While the estimates at the far fields are not as good as those in the near fields in both cases, evaluation statistics indicate that the estimates in Case (1–2) are slightly better than those in Case (1-1).

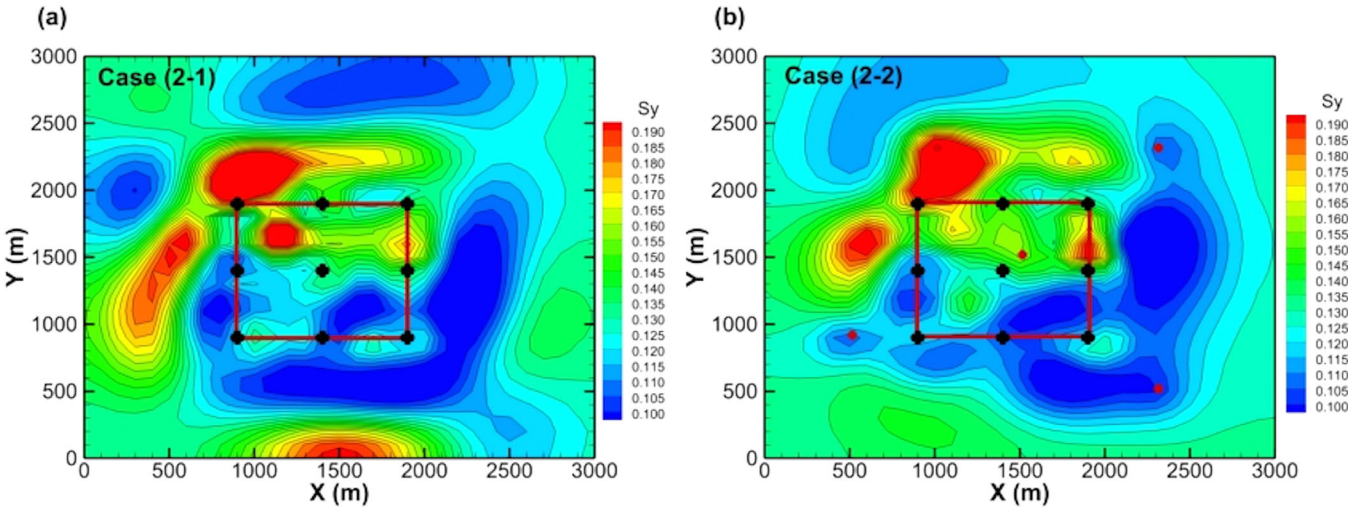

Figure 7. (a) The spatial patterns of S_y estimates obtained from HT alone, i.e., Case (2-1) and (b) S_y estimate derived from the fusion of HT and gravity data, i.e., Case (2-2).

Note that this estimation problem is poorly defined according to Yeh et al. (2015b). In other words, the number of parameters to be estimated in heterogeneous aquifers is larger than the number of constraints given. The gravity information, however, provides an additional constraint such that it results in better estimates of S_y. This result agrees with the findings of Christiansen et al. (2011). That is, the S_y estimate derived from joint interpretation of gravity and the head data set is more accurate than that derived from gravity or head data set alone.

In addition to these results, we investigate the effects of the number and location of the gravity stations on the estimated S_y field. We find that the S_y estimates are improved as the number of the gravity stations is increased, and the gravity station's location has a minor effect on the S_y estimate if its effective gravity integration radius covers the well field. For more details, please see the supporting information sections S2 and S3.

4.3. Experiment #2: Improvement on the *K* and S_y Estimate Due To Addition of Gravity Measurements

Similar to Experiment #1, this experiment also considers two cases. In Case (2-1) only head data sets from HT are used to identify S_y and K field, while in Case (2-2), both head and gravity data sets are employed. The sampling layout and procedures are the same as those of Experiment #1.

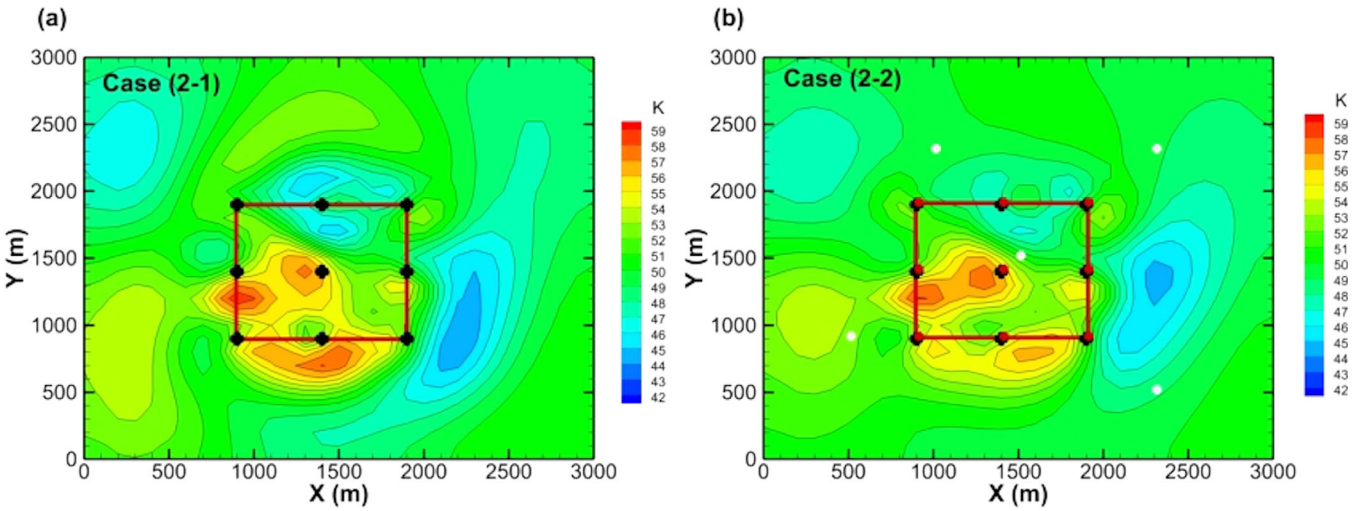

Figure 8. The spatial pattern of K estimates, (a) Cases (2-1), using HT only and that and (b) Case (2-2), using the joint inversion in Experiment #2.

A comparison of the S_y estimates of these two cases (i.e., Figures 7a and 7b) with the reference field (Figure 2b) reveals that both HT and HT with gravity measurements capture the spatial trend of the true S_y field, but the fusion of HT and gravity yields a better result at the near field than HT only. For example, the relatively high and low S_y values (e.g., 0.12– and 0.18+) in Figure 7b are much closer to the true S_y field (Figure 2b) than those in Figure 7a. For the far field, the S_y estimates close to the five gravity stations in Case (2-2) are in agreement with the true values (Figure 7b) and overall better than the estimates from HT alone.

The spatial distributions of the K estimates of Cases (2-1) and Case (2-2) are illustrated in Figures 8a and 8b, respectively. The estimated K patterns of the two cases are similar, but the K estimates of Case (2-2) is notably better.

The scatterplots of the estimated versus true K and S_y values at the near fields of these two cases as well as evaluation statistics are shown in Figures 9a–9d. The evaluations of the estimates at far fields for the two cases are displayed in Figures 10a–10d. By comparing these scatterplots and the statistics, one is certain that the K and S_y estimates at the near field of Case (2-2) are closer to the true values than those of Case (2-1). In particular, the R^2 values confirm that the S_y estimate from the fusion of HT and gravity is significantly improved.

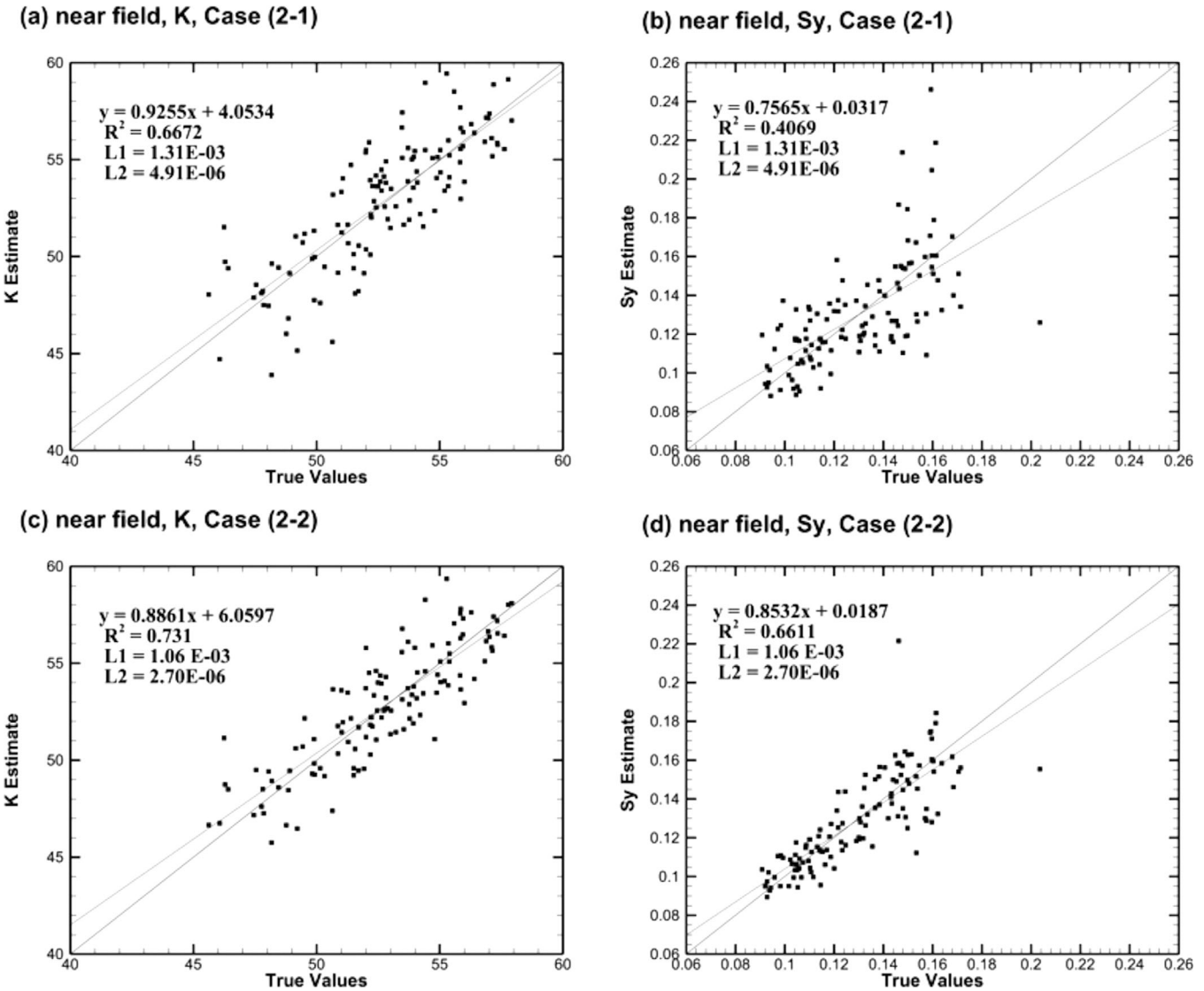

Figure 9. Scatterplots of the estimated versus true K and S_y fields at near field of Experiment #2, (a, b) for Case (2-1) and (c, d) for Case (2-2).

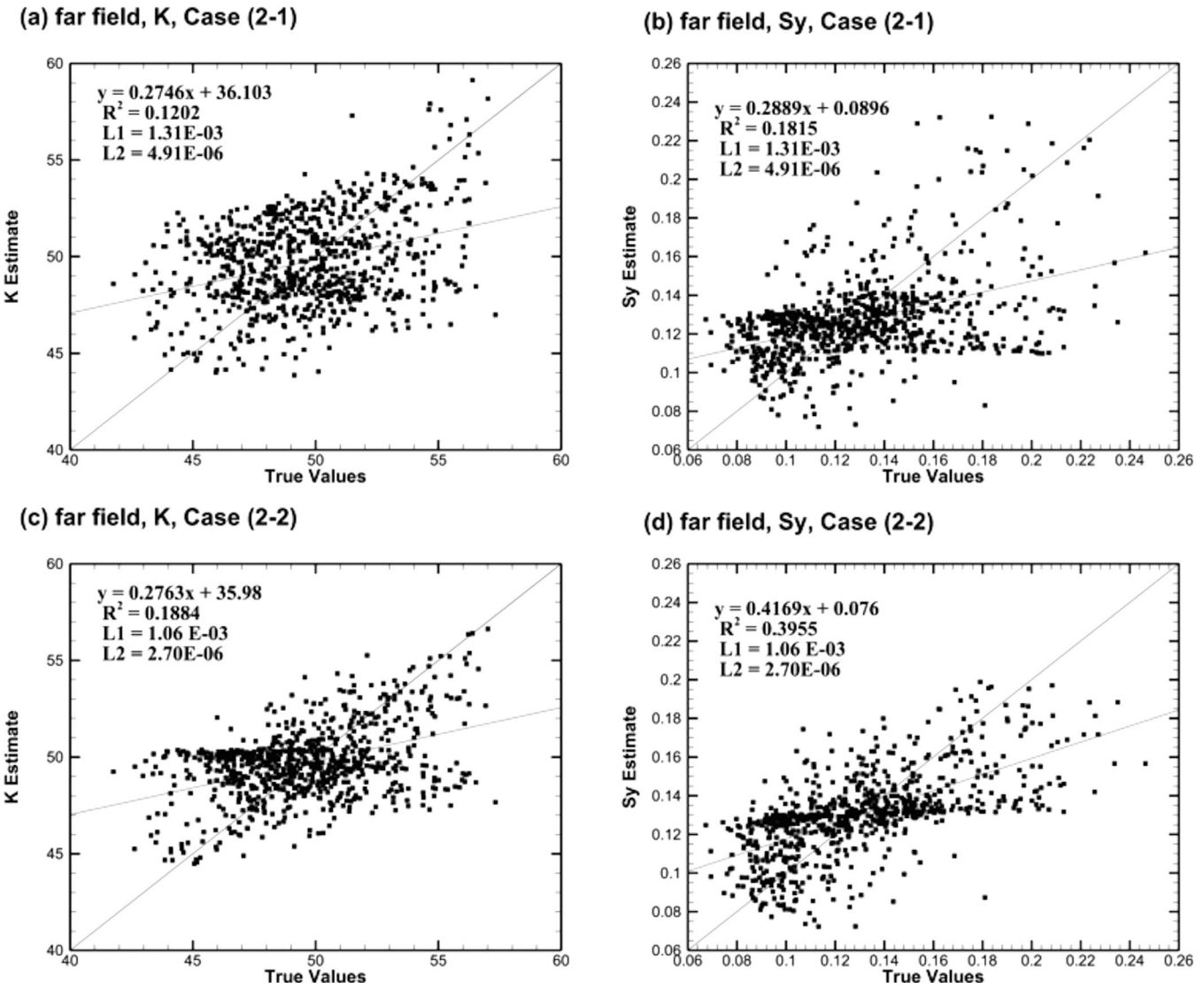

Figure 10. Scatterplots of the estimated versus true K and S_y fields at far field of Experiment #2, (a, b) for Case (2-1) and (c, d) for Case (2-2).

In addition to the improvement of the S_y estimate, the results also indicate that the estimate of K is slightly improved by considering the gravity into HT, even though the gravity measurements are not directly used to estimate the K field. This is because the estimates of K and S_y are mutually dependent during the inverse modeling. The S_y estimate of Case (2-2) is better than that of Case (2-1), resulting in a better K estimate of Case (2-2). Therefore, the gravity measurements can improve not only the S_y estimate but also the estimate of K.

4.4. Validation of Joint Inversion Results (K and S_y)

The ultimate goal of characterization of hydraulic properties of an aquifer is to provide an accurate prediction of the flow field in the aquifer under any stress. As a matter of fact, the above HT and joint inversion have predicted flow fields induced by several injection tests. The HT and joint inversion, however, have adjusted K and S_y parameters to ensure satisfactory predictions (i.e., model calibration). These predictions, therefore, cannot be viewed as a validation of the estimated parameter fields. In order to properly validate the estimated parameter fields, predictions of a new flow field using these calibrated parameter fields are necessary. As explained in Huang et al. (2011), Mao et al. (2013), and Sun et al. (2013), the heads at the observation wells may change if the location of an injection well is altered. For this reason, a new injection test at a location different from the injection locations used in HT analysis is created. This new well is

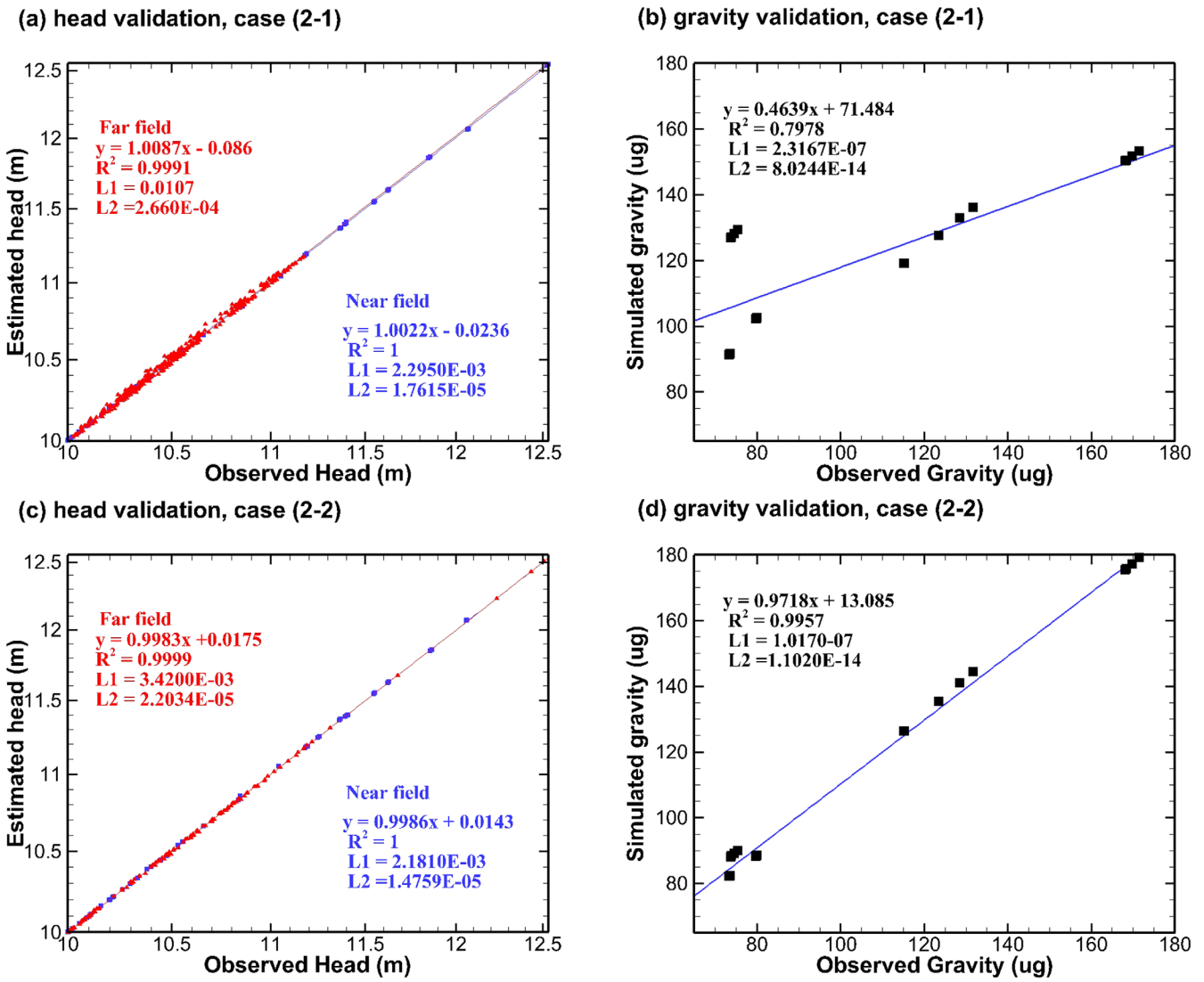

Figure 11. Scatterplots of the estimated versus (a, c) observed head and (b, d) gravity data of validation results for Case (2-1) and Case (2-2), respectively.

installed at the location $X = 1{,}200$ and $Y = 1{,}500$ with a constant injection rate of 4,000 m^3/d. During this new test, the simulated head data sets, using the reference parameters, are collected at the nine wells and other selected locations at a regular grid of 500 m over the entire domain, at times equal to 1, 10, 50, 100, 300, 500, 700, and 1,000 days. Likewise, simulated gravity measurements of the modeled flow field using the reference fields are collected at the five observation locations at 0.1, 0.2, 0.5, 1, 20, 40, and 60 days. These data sets are compared with those simulated using both estimated K and S_y fields from HT and those from the proposed joint inversion.

The validation results (scatterplots and evaluation statistics) for heads are illustrated in Figures 11a and 11c for Cases (2-1) and (2-2), respectively. The results for gravity are displayed in Figures 11b and 11d for Cases (2-1) and (2-2), respectively. Overall, the validation of the parameter fields from all cases are satisfactory in terms of the head prediction, but the evaluation statistics of Case (2-2), in which parameter estimation includes gravity measurements, are notably superior to that of Case (2-1). In other words, the K and S_y estimates of Case (2-2) are closer to the reference fields than those of Case (2-1). Further, we notice that the predicted heads at the far field of Case 2-2 (red dots in Figure 11c) are more accurate (i.e., less scattering) than those in Case 2-1 (Figure 11a) as indicated by the evaluation statistics, while those at the near fields are similar when the gravity measurements are included. This seems consistent with the fact that head data

set used in HT mainly improve the accuracy of the estimate in the near field, but that of gravity data set is able to enhance the accuracy of the estimates of HT over the greater extents. This is due in part to the five gravity stations are distributed over the entire study area, and each of their measurements detects the mass response of the groundwater storage for the area with a radius of 1,000 m. The improvement due to the gravity measurements is clearly notable in Figure 11d for Case (2-2), where the scatterplot of the predicted vs. true gravity values at the stations are significantly less biased and scattered than that in Figure 11b for Case (2-1).

Lastly, an interesting revelation of these results is that gravity measurements reflect the integrated change of groundwater mass due to pumping or injection over a large area without identifying the spatial distribution of the change. On the other hand, HT is known for its ability to map detailed aquifer heterogeneity using a limited number of observation wells and a sequential pumping or injection test. HT, however, yields estimates away from the observation network with large uncertainties since it deals with ill-defined inverse problems (Yeh et al., 2015b). For this reason, the gravity measurements, as demonstrated in this study, restrain the range of the hydraulic parameter values at the near and the far field identified by HT alone and reduce the uncertainty of HT estimates.

5. Conclusion and Future Works

This study has developed a new method to estimate heterogeneous K and S_y field by fusion of gravity measurements and head data sets from hydraulic tomography. This new approach expands the effectiveness of HT to an area from hundred-meters to kilometers with better accuracy and performance. These advantages stem from the fact that not only can the gravity measurements provide a good initial estimate of S_y field for SimSLE of HT analysis, but also provide the total mass measurements to constrain the estimation. For these reasons, both the estimates of K and S_y, in particular, the accuracy of the S_y estimates, are improved in comparison with those based on HT only. Besides, the accuracy of the S_y estimates can be further improved as the number of the gravity stations increases. The study also reveals that the location of a gravity station has minor effects on the S_y estimates if the effective radius of the gravity measurement covers the well field.

While we believe that these results are exciting, we must acknowledge that they are based on numerical experiments. Moreover, the numerical flow model solves the Boussinesq equation, which ignores processes in the unsaturated zone and assumes instantaneous fill or drainage of pore water as the water table rises or falls, and considers only the depth-averaged flow field (Yeh et al., 2015a, 2015b). To be realistic, a three-dimension variably saturated flow model, which considers the unsaturated hydraulic properties in heterogeneous unconfined aquifers, should be used as advocated by Yeh et al. (2015a, 2015b). Most importantly, the proposed methodology must be tested and validated in the fields as have done by Zha et al. (2015, 2016), Illman et al. (2012), Berg and Illman (2011), and Zha et al. (2017). In spite of these limitations, the study demonstrates that the proposed method is a promising step toward the development of a useful data fusion concept and method to probe subsurface and to better characterize the spatial variability of hydrogeological parameters in unconfined aquifers. We believe this study will lead to improvement of the reliability of the groundwater-related predictions in the future.

Appendix A: Mapping the Sensitivity of Gravity to S_y to a Groundwater Model Grid

Assuming a computational grid of groundwater flow includes n gravity simulation grids. For a gravity simulation element, the Taylor expression of the gravity variation resulted from the S_y perturbation can be expressed as follows:

$$g(S_y+\Delta s)_{e_g}=g(S_y)_{e_g}+\frac{\partial g(S_y)}{(\partial S_y)_{e_g}}\Delta s_{e_g}, \tag{A1}$$

where Δs denotes the perturbation of S_y, and subscript e_g stands a gravity simulation element. Since a groundwater simulation element (or material block) contains n groundwater simulation elements, the Taylor expression of the gravity variation resulted from the S_y perturbation can be expressed from equation (A2) to equation (A3):

$$g(S_y+\Delta s)_{e_f}=g(S_y)_{e_f}+\frac{\partial g(S_y)}{(\partial S_y)_{e_f}}\Delta s_{e_f}, \tag{A2}$$

$$g(S_y+\sum_{e_g=1}^{e_g=n}\Delta s_{e_g})_{e_f}=g(S_y)_{e_f}+\sum_{e_g=1}^{e_g=n}\frac{\partial g(S_y)}{(\partial S_y)_{e_g}}\Delta s_{e_g}, \tag{A3}$$

where subscript e_f denotes the flow simulation element.

Assuming each S_y perturbation (Δs_{e_g}) of the gravity simulation element being identical, equation (A2) is simplified as

$$g(S_y+\sum_{e_g=1}^{e_g=n}\Delta s_{e_g})_{e_f}=g(S_y)_{e_f}+\left[\sum_{e_g=1}^{e_g=n}\frac{\partial g(S_y)}{(\partial S_y)_{e_g}}\right](n\Delta s). \tag{A4}$$

Comparing equations (A2) and (A4), one can know that the sensitivity of gravity to the specific yield of a flow simulation element $\frac{\partial g(S_y)}{(\partial S_y)_{e_g}}$ is the accumulation of the sensitivity of gravity to the specific yield of gravity simulation elements, shown as equation (A5).

$$\begin{cases}\dfrac{\partial g(S_y)}{(\partial S_y)_{e_f}}=\displaystyle\sum_{e_g=1}^{e_g=n}\frac{\partial g(S_y)}{(\partial S_y)_{e_g}}.\\ \Delta s_{e_f}=n\Delta s\end{cases} \tag{A5}$$

Acknowledgments

The authors thank the Ministry of Science and Technology, Taiwan, for financially supporting this research under grants MOST 104–2917-I-564-085, 105–2221-E-009–054-MY3, and 105–2811-E-009–018. The second author acknowledges the support from the Strategic Environmental Research and Development Program (SERDP) (grant ER-1365); the Environmental Security Technology Certification Program (ESTCP) (grant ER201212); and the US National Science Foundation-Division of Earth Sciences (grant 1014594). The second author also acknowledges the Outstanding Oversea Professorship award through Jilin University from Department of Education, China as well as the Global Expert award through Tianjin Normal University from the Thousand Talents Plan of Tianjin City. All data from this work are available on request from the first author J.-P. Tsai (skysky2cie@gmail.com). We acknowledge the constructive comments by the Associate Editor and three reviewers, which helped in improving the final manuscript

References

Banerjee, B., & Das Gupta, S. (1977). Gravitational attraction of a rectangular parallelepiped. *Geophysics*, *42*(5), 1053–1055.

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. https://doi.org/10.1029/2011WR010616

Blainey, J. B., Ferré, T., & Cordova, J. T. (2007). Assessing the likely value of gravity and drawdown measurements to constrain estimates of hydraulic conductivity and specific yield during unconfined aquifer testing. *Water Resources Research*, *43*, W12408. https://doi.org/10.1029/2006WR005678

Bohling, 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*, W05430. https://doi.org/10.1029/2006WR004932

Brauchler, R., Liedl, R., & Dietrich, P. (2003). A travel time based hydraulic tomographic approach. *Water Resources Research*, *39*(12), 1370. https://doi.org/10.1029/2003WR002262

Cardiff, M., & Barrash, W. (2011). 3-D transient hydraulic tomography in unconfined aquifers with fast drainage response. *Water Resources Research*, *47*, W12518. https://doi.org/10.1029/2010WR010367

Christiansen, L., Binning, P. J., Rosbjerg, D., Andersen, O., & Bauer-Gottwein, P. (2011). Using time-lapse gravity for groundwater model calibration: An application to alluvial aquifer storage. *Water Resources Research*, *47*, W06503. https://doi.org/10.1029/2010WR009859

Damiata, B. N., & Lee, T.-C. (2006). Simulated gravitational response to hydraulic testing of unconfined aquifers. *Journal of Hydrology*, *318*(1), 348–359.

Gehman, C. L., Harry, D. L., Sanford, W. E., Stednick, J. D., & Beckman, N. A. (2009). Estimating specific yield and storage change in an unconfined aquifer using temporal gravity surveys. *Water Resources Research*, *45*, W00D21. https://doi.org/10.1029/2007WR006096

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

Howle, J. F., Phillips, S. P., Denlinger, R. P., & Metzger, L. F. (2003). *Determination of specific yield and water-table changes using temporal microgravity surveys collected during the second injection, storage, and recovery test at Lancaster, Antelope Valley, California, November 1996 through April 1997* (Water Resources Investigations Rep. 03–4019, pp. 6–26).

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

Illman, W. A., Berg, S. J., & Yeh, T.-C. J. (2012). Comparison of approaches for predicting solute transport: Sandbox experiments. *Ground Water*, *50*(3), 421–431.

Illman, 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*, W04502. https://doi.org/10.1029/2009WR007745

Leiriao, S., He, X., Christiansen, L., Andersen, O. B., & Bauer-Gottwein, P. (2009). Calculation of the temporal gravity variation from spatially variable water storage change in soils and aquifers. *Journal of Hydrology*, *365*(3), 302–309.

Li, B., & Yeh, T.-C. J. (1999). Cokriging estimation of the conductivity field under variably saturated flow conditions. *Water Resources Research*, *35*(12), 3663–3674.

Liu, S., Yeh, T.-C. J., & Gardiner, R. (2002). Effectiveness of hydraulic tomography: Sandbox experiments. *Water Resources Research*, *38*(4), 1034. https://doi.org/10.1029/2001WR000338

MacMillan, W. D. (1958). *The theory of the potential*. New York: Dover Publications.

Mao, D., Wan, L., Yeh, T. C. J., Lee, C. H., Hsu, K. C., Wen, J. C., & Lu, W. (2011). A revisit of drawdown behavior during pumping in unconfined aquifers. *Water Resources Research*, *47*, W05502. https://doi.org/10.1029/2010WR009326

Mao, D., Yeh, T.-C. J., Wan, L., Wen, J. C., Lu, W., Lee, C. H., & Hsu, K. C. (2013). Joint interpretation of sequential pumping tests in unconfined aquifers. *Water Resources Research*, *49*, 1782–1796. https://doi.org/10.1002/wrcr.20129

Nagy, D. (1966). The gravitational attraction of a right rectangular prism. *Geophysics*, *31*(2), 362–371.
Ni, C.-F., Yeh, T.-C. J., & Chen, J.-S. (2009). Cost-effective hydraulic tomography surveys for predicting flow and transport in heterogeneous aquifers. *Environmental Science & Technology*, *43*(10), 3720–3727.
Sato, T. (1984). A program for the Computation of Oceanic Loading Effects 'GOTIC.' *Pub. Int. Lat. Obs., Mizusawa*, *18*, 29–47.
Schwartz, F. W., & Zhang, H. (2003). *Fundamentals of ground water*. New York: John Wiely & Sons, Inc.
Sun, R., Yeh, T.-C. J., Mao, D., Jin, M., Lu, W., & Hao, Y. (2013). A temporal sampling strategy for hydraulic tomography analysis. *Water Resources Research*, *49*, 3881–3896. https://doi.org/10.1002/wrcr.20337
Tso, M. C.-H., Zha, Y., Yeh, T.-C. J., & Wen, J. C. (2016). The relative importance of head, flux, and prior information in hydraulic tomography analysis. *Water Resources Research*, *52*, 3–20. https://doi.org/10.1002/2015WR017191
Wilson, C. R., Scanlon, B., Sharp, J., Longuevergne, L., & Wu, H. (2012). Field test of the superconducting gravimeter as a hydrologic sensor. *Ground Water*, *50*(3), 442–449.
Wu, C. M., Yeh, T.-C. J., Zhu, J., Lee, T. H., Hsu, N. S., Chen, C. H., & Sancho, A. F. (2005). Traditional analysis of aquifer tests: Comparing apples to oranges? *Water Resources Research*, *41*, W09402. https://doi.org/10.1029/2004WR003717
Xiang, J., Yeh, T.-C. J., Lee, C. H., Hsu, K. C., & Wen, J. C. (2009). A simultaneous successive linear estimator and a guide for hydraulic tomography analysis. *Water Resources Research*, *45*, W02432. https://doi.org/10.1029/2008WR007180
Yeh, T. C. J., Gutjahr, A. L., & Jin, M. (1995). An iterative cokriging-like technique for ground-water flow modeling. *Ground Water*, *33*(1), 33–41.
Yeh, T.-C. J., Jin, M., & Hanna, S. (1996). An iterative stochastic inverse method: Conditional effective transmissivity and hydraulic head fields. *Water Resources Research*, *32*(1), 85–92.
Yeh, T.-C. J., Khaleel, R., & Carroll, K. C. (2015a). *Flow through heterogeneous geologic media*. New York: Cambridge University Press.
Yeh, T.-C. J., & Lee, C. H. (2007). Time to change the way we collect and analyze data for aquifer characterization. *Ground Water*, *45*(2), 116–118.
Yeh, T.-C. J., & Liu, S. (2000). Hydraulic tomography: Development of a new aquifer test method. *Water Resources Research*, *36*(8), 2095–2105.
Yeh, T.-C. J., Mao, D., Zha, Y., Wen, J.-C., Wan, L., Hsu, K.-C., & Lee, C.-H. (2015b). Uniqueness, scale, and resolution issues in groundwater model parameter identification. *Water Science and Engineering*, *8*, 175–194. https://doi.org/10.1016/j.wse.2015.08.002
Yeh, T.-C. J., Srivastava, R., Guzman, A., & Harter, T. (1993). A numerical model for water flow and chemical transport in variably saturated porous media. *Ground Water*, *31*(4), 634–644.
Zha, Y., Yeh, T.-C. J., Illman, W. A., Tanaka, T., Bruines, P., Onoe, H., & Saegusa, H. (2015). What does hydraulic tomography tell us about fractured geological media? A field study and synthetic experiments. *Journal of Hydrology*, *531*, 17–30.
Zha, Y., Yeh, T.-C. J., Illman, W. A., Tanaka, T., Bruines, P., Onoe, H., . . . Wen, J. C. (2016). An Application of hydraulic tomography to a large-scale fractured granite site, Mizunami, Japan. *Ground Water*, *54*, 793–804.
Zha, Y., Yeh, T. C. J., Illman, W. A., Onoe, H., Mok, C. M. W., Wen, J. C., . . . Wang, W. (2017). Incorporating geologic information into hydraulic tomography: a general framework based on geostatistical approach. *Water Resources Research*, *53*(4), 2850–2876.
Zhu, J., & Yeh, T.-C. J. (2005). Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resources Research*, *41*, W07028. https://doi.org/10.1029/2004WR003790
Zhu, J., & Yeh, T.-C. J. (2008). *Hydraulic tomography to characterization of heterogeneity of unconfined aquifers*. Paper presented at the AGU Fall Meeting Abstracts, San Francisco, CA.
Zhu, J., Yeh, T.-C. J., & Mao, D. (2011). Hydraulic tomography to characterize heterogeneity of unconfined aquifers. *Journal of Nanjing University*, *47*(3), 252–264.