Basin-scale transmissivity and storativity estimation using hydraulic tomography

Kristopher L. Kuhlman, Andrew C. Hinnell, Phoolendra K. Mishra, and Tian-Chyi Jim Yeh

University of Arizona, Department of Hydrology & Water Resources, 1133 East James E. Rodgers Way, Tucson, AZ 85721.

Abstract:

While tomographic inversion has been successfully applied to laboratory- and field-scale tests, here we address the new issue of scale that arises when extending the method to a basin. Specifically, we apply the hydraulic tomography concept to jointly interpret four multi-well aquifer tests in a synthetic basin to illustrate the superiority of this approach to a more traditional Theis analysis of the same tests. Transmissivity and storativity are estimated for each element of a regional numerical model using the geostatistically-based SSLE inverse solution method. We find that hydraulic tomography inversion is an effective strategy for incorporating data from potentially disparate aquifer tests into a basin-wide aquifer property estimate. The robustness of the SSLE algorithm is investigated by considering the effects of noisy observations, changing the variance of the true aquifer parameters, and supplying incorrect initial and boundary conditions to the inverse model. Ground water flow velocities and total confined storage are used as metrics to compare true and estimated parameter fields; they quantify the effectiveness of hydraulic tomography and SSLE compared to a Theis solution methodology. We discuss alternative software that can be used for implementing tomography inversion.

Introduction

Managing ground water resources requires knowledge of aquifer property distributions, since they affect water movement and solute transport. This understanding is often developed and tested with regional numerical ground water flow models, which are used for simulation, prediction, and scenario analysis. Regional models facilitate long-term management of water resources, where they can be used for both evaluation and mitigation of supply and quality issues.

In ground water model calibration we seek to best represent a complex natural system with an idealized numerical model at the appropriate scale of interest. The scale depends on the intended use of the calibrated model (e.g., flow vs. transport predictions) and the desired detail needed in the predictions. Many regional ground water studies do not attempt to build detailed heterogeneity into large scale (tens to hundreds of kilometers)flow models, due to the prohibitive costs of detailed sampling over large areas and the computational limits on calibrating multi-scale heterogeneity in the model. Regional geologic or hydrologic units are often treated as zones, assumed to be homogeneous with a single effective parameter value (e.g., Barlebo et al. [2004]). This zoned representation may offer computational advantages, but it can only yield large-scale effective properties, which are best for predicting "ensemble" behaviors of a ground water system [Yeh, 1992; Yeh et al., 2007].

In regional studies that include local-scale heterogeneity (i.e., heterogeneity smaller than the hydrologic unit, at the scale of several model cells), the parameter distribution is often estimated from a steady-state or pre-development head distribution (e.g., Yeh and Mock [1996]). Heterogeneous transmissivity fields are estimated by manually adjusting parameter values in model cells or zones to match simulated and observed hydraulic heads. More advanced approaches use automated calibration algorithms (e.g., PEST [Doherty, 2007] or UCODE [Poeter et al., 2005]) to minimize the residual between observed and simulated heads [Barlebo et al., 2004]. Steady-state calibrations are limited to estimating transmissivity (*T*), and few regional studies attempt to calibrate ground water flow models using transient head measurements due to the large increase in complexity and computational effort.

Basin-scale transient model calibrations are often ill-posed and non-unique due to difficulties collecting the necessary and sufficient information to make an inverse problem well-posed [Yeh et al., 2007]. As a result, there are many non-unique parameter distributions that equally fit sparse head observations. In other words, traditional inverse modeling efforts often yield ambiguous aquifer characterization. Because of the uncertainty inherent in aquifer parameter and boundary condition characterization, many modelers have developed misleading predictive models of ground water flow and contaminant migration. Because of this, some have seriously questioned the ability to validate ground water flow models at all [Konikow and Bredehoeft, 1992; Oreskes et al., 1994; Bredehoeft, 2003].

To improve our ability to adequately characterize and solve inverse ground water problems, we propose utilizing hydraulic tomography (HT). Many researchers have shown it can be used to characterize heterogeneous hydraulic properties, including Tosaka et al. [1993], Gottlieb and Dietrich [1995], Vasco et al. [2000], Yeh and Liu [2000], Bohling et al. [2002], Brauchler et al. [2003], and Zhu and Yeh [2005 and 2006]. HT involves collecting responses throughout an aquifer due to a sequence of overlapping aquifer tests, then calibrating a heterogeneous ground water flow model using the observed responses from all the tests. Multiple sets of aquifer tests and their observed responses improve the inverse problem, since tests cross-validate each other. As a result, the estimated hydraulic property fields become more detailed and less uncertain than those computed from a single set of data.

HT has been applied successively to small-scale synthetic aquifers [Yeh and Liu, 2000; Zhu and Yeh, 2005 and 2006; Hao et al., 2008], laboratory sandboxes [Liu et al., 2002; Liu et al., 2007; Illman et al., 2007], and plot-scale fields [Vesselinov et al., 2001; Bohling et al., 2007; Straface et al., 2007; Li et al, 2007]. In these small-scale studies it is possible to stress the entire domain with each pumping well, providing new information throughout the domain from each pumping event. We propose using regional-scale HT to estimate *T* and storativity (*S*) distributions for a regional flow model, where the main new challenge is determining how to adequately stress the entire aquifer. Unlike smaller-scale applications of HT, it is not possible to pump a single well causing a response throughout the aquifer; both the pumping rate and test length would be unreasonably large. Realistically, a single aquifer test can only stress a portion of a large aquifer and only cause measurable drawdown in a subset of a basin-wide observation network. At the regional scale, we reformulate HT as an interference problem; the head distribution due to multiple simultaneous pumping wells is observed using a monitoring well network as might be found in a municipal water supply or remedial well field (off-duty pumping wells can serve as observation wells). Rather than successively pumping from individual wells, we cycle through sets of pumping wells. In this way, the regional aquifer is repeatedly stressed to the fullest possible extent using existing wells.

We investigate the HT approach for estimating aquifer properties in a regional-scale ground water model; the method results in both more detailed (higher resolution) and more trustworthy (lower uncertainty) estimates. An improved estimate of aquifer properties is necessary to improve the reliability of predictions made with a calibrated model. Estimating aquifer parameters using the sequential successive linear estimator (SSLE, Zhu and Yeh, 2005) with tomographic test data leads to better predictions of flow velocities and estimates of total storage for the basin, compared to traditional methods. The numerical analysis in this study was completed on a personal computer, demonstrating that HT inversion can be implemented using existing computer resources.

In this work we use a synthetic regional confined aquifer to minimize unknown sources of error (e.g., measurement and model errors) that would complicate the analyses. Initially, we demonstrate that HT can be used on a regional scale, then we investigate the robustness of the method by changing the variance in the true field, adding random error to the head observations, and reducing the number of pumping events. Finally, HT was applied using observations of drawdown, rather than head, to investigate the effects of unknown initial and boundary conditions.

Methods

We solve the HT inverse problem using the SSLE algorithm which is similar to that developed by Yeh and Liu [2000] and Zhu and Yeh [2005]. The SSLE algorithm is an extension of the SLE (successive linear estimator), that was developed for solving spatially variable parameter inverse problems using a geostatistical framework [Yeh et al, 1996; Zhang and Yeh, 1997; Hughson and Yeh, 2000]. The implementation of the SSLE used here is coupled with the finite element flow model VSAFT2 [Yeh et al., 1993] (available for free download at http://www.hwr.arizona.edu/yeh). We qualitatively discuss the key features of this approach; details on the SSLE are found in Zhu and Yeh [2005].

Because high-resolution parameter estimates are the desired result of tomographic inversion, we independently estimate parameter values (*T* and *S*) in each model element; this leads to a large number of free parameters. The tomographic approach results in two hurdles to overcome: 1) the large computational effort required to estimate the sensitivity of model parameters model predictions at observation locations, and 2) the need for additional constraints to reduce the degrees of freedom in the solution, since there are more estimable parameters than calibration data (an ill-posed inverse problem).

The SSLE approach addresses both of these problems. First, the parameter-observation sensitivities required for the inverse problem are computed using the adjoint approach [Sykes et al., 1985; Sun and Yeh, 1992], rather than using the perturbation approach (as in PEST or UCODE). The perturbation approach changes each parameter independently, running the model forward to compute the corresponding model prediction change. With the perturbation inverse approach, a problem with 500 estimable parameters would require 501 (forward difference) or 1001 (central difference) independent forward model runs per iteration. For the adjoint approach, the effort to compute the model sensitivities is proportional to the number of observation data. This benefits problems with a large number of parameters and sparse observations, allowing sensitivities to be computed more efficiently. Secondly, due to the geostatistical foundation upon which SSLE is built, the parameters being estimated (*T* and *S*) are not allowed to vary arbitrarily in space, but rather their distribution follows a geostatistical framework. Regularization (the observation that parameters vary “smoothly” in space [Tikonov and Arsenin, 1977]) is also a means of constraining the spatial distribution of parameters (implemented in PEST). The difference between the geostatistical and regularization approaches is analogous to the distinction between kriging and inverse distance as interpolation schemes. Both kriging and SSLE incorporate additional geostatistical knowledge into their estimates, while Tikhonov-style regularization and inverse distance squared are purely empirical approaches. The geostatistical framework does have additional requirements (estimates of the mean, variance and directionality of *T* and *S*), but the accuracy of these a priori estimates is not essential to the success of the algorithm in HT analyses [Yeh and Liu, 2000].

Description of synthetic problem

The synthetic confined aquifer used here was designed to be realistically complex, while simple enough to allow straightforward interpretation of the results and the timely execution of many runs required for the robustness analysis. The 2D model represents a depth-averaged heterogeneous 54 km × 27 km aquifer, bounded by a river flowing west to east on the western, northern and eastern boundaries and a mountain block on the southern boundary (see Figure 1). The aquifer has two large bedrock outcrops which are represented in the model by "islands" of no-flow cells (in Figure 1 inactive cells are gray). The finite element mesh consists of 519 active square elements, each 1200 m on a side . The river is a specified head boundary condition, ranging linearly from 1015 m to 1000 m from west to east (dashed boundaries in Figure 1). Specified flux boundary conditions (inflow) were used in four separate sections along the southern boundary to simulate fluxes into the model domain from neighboring basins (dotted boundaries in Figure 1).

A random true T field (Figure 2) was generated with an arithmetic mean of 300 m^2/day and variance of $\ln(T)$ of 2.0, while the S field had an arithmetic mean of 0.001 and variance of $\ln(S)$ of 2.0. Transmissivity was assumed to be isotropic at the scale of the model elements ($T_x = T_y$), but both the T and S fields were assumed statistically anisotropic at the scale of the domain. The correlation scale was 20 km in the east-west direction and 8 km in the north-south direction. The random T and S fields are uncorrelated; they utilized different random seeds during their generation. Initial conditions were the results of a steady-state simulation with no pumping.

HT was used to estimate the T and S fields by stressing the aquifer simultaneously with multiple pumping wells in a manner analogous to municipal pumping or a "pump and treat" remediation system. The synthetic well field was comprised of 70 wells: 20 pumping and 50 observation wells. All wells were located randomly within the domain using a Latin hypercube approach to limit spatial clustering. While no specific effort was made to optimize the wellfield for aquifer parameter estimation, the subset of 20 pumping wells was visually selected to provide good spatial distribution of pumping wells throughout the synthetic aquifer. Pumping wells were assigned to one of four events such that each pumping event stressed most of the aquifer, resulting in overlap between the stressed areas of different pumping events. The pumping well locations are shown as open symbols in Figure 1: triangles are wells pumped in pumping event one, squares in event two, stars in event three, and circles in event four. Each pumping event is an aquifer test that lasted 14 days, during which each of the five wells was pumped at 2000 m^3/day (367 gpm). The initial hydraulic head distribution for each pumping event was the steady-state head distribution. The aquifer response to each pumping event was observed at 50 observation wells (filled dots in Figure 1). For this example we did not include the pumping wells from the other pumping events in the set of observation wells, although in reality one would include as many observation wells as possible.

Hydraulic head was sampled continuously at each observation well, but only four observation times from each pumping event were used in the inversion: three at early time and one at late time. These observation times were chosen to minimize the computational effort in the SSLE inversion, while providing sufficient information to constrain the aquifer parameter estimates. For noise-free data, observations through time at one location are highly correlated and each new temporal observation contributes little new information [Zhu and Yeh, 2005].

The true and estimated parameter fields were compared using spatial distribution maps, scatter plots, and summary statistics. Good parameter estimates produce distributions that are visually "similar" and scatter plots with data clustered along the 1:1 diagonal. High correlation coefficient (ρ) indicates a significant linear relationship between the values of the two datasets, while high rank correlation coefficient (ρ_{rank}) indicates patterns of highs and lows are well correlated, regardless of numerical values [Isaaks and Srivastava, 1989]. The L1 and L2 norms indicate the differences in the log mean (bias), and log standard deviation of the two datasets respectively [Yeh and Liu, 2000] (low norm values indicate better fit).

Quantitative comparisons were also made between the true and predicted overall storage for the entire basin. When managing ground water basins resources, accurate information regarding the amount of water available from storage is essential. Lastly, we compared observed and simulated velocity fields (v_x and v_y), which are required in transport simulations. While head is diffuse by nature and therefore easy to match, leading to non-unique solutions, solute transport is governed by advection (flow velocity, the gradient of head), which is much more sensitive to aquifer property distributions. Two head distributions can match observed point head measurements equally well, but their corresponding flux distributions (and solute transport behaviors) may be very different. Velocity field comparisons provide a measure of how useful the simulation would be for making transport predictions.

Results

Estimation of T and S using Theis solution

Aquifer parameters (*T* and *S*) are often estimated for real-world applications using the Theis solution for drawdown from a pumping well, even if some of its fundamental assumptions are known to be violated. The Theis solution is 2D (depth-averaged) and assumes an infinite homogeneous aquifer. We modeled the drawdown observed during each pumping event using the Theis solution to both illustrate the inappropriateness of a homogeneous solution for interpreting heterogeneous regional scale pumping tests and to provide a comparison to the HT results.

We estimated *T* and *S* values from the “observed” model drawdown at observation wells. For simplicity we assigned the estimated values to the location of the observation well, resulting in 50 estimates of *T* and *S* for each of the four pumping events. Drawdown at each observation well is due to pumping at five pumping wells. *T* and *S* are estimated by matching the observed drawdown to the drawdown predicted by summing the Theis solutions for the five pumping wells in a homogeneous infinite aquifer. PEST was used to minimize the sum of squared residuals between the observed and the Theis-simulated drawdown. The estimated parameter values at all 50 observation locations, for each pumping event, were then kriged to the flow simulation grid to generate the eight estimated parameter fields shown in Figures 3a and 3b. The eight model variograms used for kriging were derived by least-squares fitting an anisotropic exponential model to the experimental variograms created from the Theis results.

Due to the large domain, the simultaneous pumping of the five wells during each pumping period does not cause significant interference between the wells. The radius of influence of the pumping wells after 14 days of pumping (distance from the well to 1 cm of drawdown) varies between 4 km and 15 km. However, more importantly for HT, at least 1 cm of drawdown was observed in 46/50 of the observation wells during at least two of the pumping events and drawdown was observed in 26/50 of the observation wells for all of the pumping events.

We assigned parameters to observation locations, rather than pumping locations because the latter would have resulted in 50 parameter estimates associated with 20 locations, requiring cokriging or additional averaging to be utilized in the flow model. T and S estimates could also have been attributed to a "representative" volume or location in the aquifer, but for heterogeneous aquifers Theis-predicted values may change with time, orientation, and location [Wu et al., 2005]. This makes interpretation of a representative location or volume difficult, especially with the presence of boundaries. Through the kriging of the intermediate point results onto the final flow simulation grid, we effectively volume-averaged the Theis results in an objective and straightforward manner.

Visually comparing the Theis-estimated and true parameter fields (Figures 3 and 2) one can see that each pair of estimated T and S fields is different and a poor estimate of the true fields. As expected, the results of the Theis analysis are sensitive to boundaries. While Li et al [2007] indicate Theis-based analyses can lead to estimates that agree, on average, with tomographic results, they did not have significant boundary conditions in their problem.

The curve matching produced high values of T near the boundaries of the domain (see Figure 3a) while high values of S were consistently predicted in the south-west corner of the domain (see Figure 3b). The Theis analysis has produced four different distributions of T and S which represent the head observations associated with the four pumping events. It is clear that the Theis solution doesn't give any useful information regarding the distribution of parameters [Li et al. 2007], because it is a homogeneous model. The hydrogeologist is left to average or decide which estimated parameter field they feel best represents the true field. More realistically, Theis analyses would be performed individually on each pumping test, potentially using distance drawdown to incorporate multiple observation wells at one time, but the hydrogeologist may be unaware that different overlapping tests can lead to markedly different results using a homogeneous model such as the Theis solution.

Although there are obvious limitations to using the Theis solution to analyze drawdown in a finite, heterogeneous domain, the exercise was done to illustrate two points. First, the results from the four pumping events, which used different pumping wells (but had many observation wells in common), do not produce identical or even similar results. This illustrates the fact that the Theis solution doesn't simply "average out" the heterogeneity around the pumping well [Wu et al, 2005]. Aside from averaging or possibly cokriging, there is no straightforward way to combine the data collected in the four pumping events into a single estimate (kriging does not allow for multiple values at the same location). Secondly, although the shortcomings of the Theis solution are "obvious" in this synthetic example, it is common practice to use Theis type curve analysis, with far less data, to analyze aquifer test results. In a real-world case, unconfined, leakage, skin, wellbore storage or partial penetration effects would also be compounded upon the boundary and heterogeneity artifacts seen here. Here, these effects can truly be ignored, because the data are synthetic. The effects of ignoring wellbore storage or unconfined behavior may have a larger impact on predictions than the effects of ignoring distant boundary conditions, depending on field conditions.

Full tomography results

T and *S* fields were estimated using SSLE with error-free observations of head from 50 observation wells divided into the same four pumping events used in the Theis analysis. The estimated fields (Figure 4) compare favorably with the true fields (Figure 2). Scatter plots of the true and estimated parameters (Figure 5) show a low degree of bias (small L1 norm). Outlier parameter estimates are primarily located where the model is insensitive to parameter values: in cells adjacent to specified head boundaries and far from observation wells. The true and estimated *T* and *S* are well correlated (large ρ and ρ_{rank}). Summary statistics are listed for all the SSLE scenarios in Tables 1a (for *T*) and 1b (for *S*). Columns 1-4 give statistics for the base case, with no data noise and correctly specified boundary and initial conditions. Columns 5 and 6 show the effects of adding noise to observations and using the wrong boundary conditions (using all four pumping events). The last two columns give the statistics corresponding to the scenarios where the true random *T* and *S* fields were generated using the same random seed, but different variances.

Since ground water velocity controls the advective transport of solutes, velocity fields were compared as a means to quantify the quality of the SSLE calibration. The x and y components of the velocity are well correlated with small L1 and L2 norms (Figure 6). The SSLE-estimated T and S fields would produce a reasonable estimate of advective solute transport, since accurate flow velocities are the most important part of a solute transport model.

Comparing the estimated and true total confined storage for the entire basin is another form of model validation. This quantity is found by summing the product of S and the area for each element, for the all elements in the domain. The area of all 519 model elements is 7.4736×10^{8} m², while the sum of the S (ΣS) in all elements is 0.4858 for the true field, giving a true total storage of 3.63×10^{8} m³ (8225 acre-ft). The results of SSLE inversion gave ΣS = 0.6425 (overestimation by 32%), while the Theis approach gave ΣS = 3.029, 2.302, 2.436, and 3.646 for events 1 through 4, respectively; the average Theis result is 2.853 (overestimation by 587%). While SSLE does overestimate S, it is an order of magnitude better than the Theis solution. We interpret that because of the lack of boundary conditions in the Theis solution, it must overcompensate by overestimating the amount of water coming from aquifer storage. Overestimating available storage could easily lead to fallacious management decisions, by basing long-term strategies on misinformation regarding available ground water.

Robustness analysis

We tested the robustness of the HT inversion by changing several aspects of the synthetic example. First, we repeated the analysis with fewer pumping events, illustrating how HT leads to an improved estimate with additional information. Second, we added zero-mean Gaussian noise with a standard deviation of 0.1 m to the head data, to better replicate field-measured observations. Third, the HT analysis was repeated with true T and S fields with variances of half ($\sigma^2_{\ln(T)} = \sigma^2_{\ln(S)} = 1$) and 1.5 times ($\sigma^2_{\ln(T)} = \sigma^2_{\ln(S)} = 3$) the levels of the original analysis. Finally, we reformulated the HT problem in terms of drawdown to minimize the effects of potentially unknown initial and boundary conditions.

Decreasing number of pumping events

One of the main strengths of HT is the ability to use multiple datasets to estimate a single coherent parameter set. To illustrate the improvements from inverting multiple tests together, the analysis was repeated, each time removing more pumping events from the analysis. Inversion was performed using pumping events one through three, one and two, and pumping event one on its own. The scatter plots of true versus estimated T and S for each analysis are presented in Figure 7, while the results from using all four pumping events are shown in Figure 4.

The T estimate improved as more pumping events (each with different pumping wells but the same observation locations) are inverted together. The cloud of points, representing T in each element of the flow model, moves closer to the 1:1 line, as two and three pumping events are jointly inverted. This type of improvement is typical when inverting tomographic aquifer tests. Each pumping event adds new information to the overall estimate of the aquifer parameters, but no single pumping event by itself results in better parameter estimates than analyzing two datasets simultaneously. The addition of each pumping event to the inversion process produces a smaller incremental improvement to the estimated parameters than the last addition, illustrating the diminishing returns off including similar data. Addition of a fourth pumping event noticeably decreases the quality of the estimated *T* field, while the quality of the estimated S field remains approximately the same, as can be seen in the summary statistics in Table 1. Using all four pumping events together may not produce optimal results for both parameters (in a non-synthetic case this would be difficult to quantify), but the SSLE results remain a very good estimate of the parameter distributions. In all scenarios, we used an SSLE convergence criterion of a 5% relative change in the estimated parameter variance.

Random error added to observations

For the baseline analysis, the observations were noise-free. In this case we corrupted the data with unbiased Gaussian noise with a standard deviation of 0.1 m, to simulate more realistic observations. Corrupting the observations smooths the parameter estimates, however, the estimated parameter fields still generally agree with the true fields (Figure 8a), as can be seen by the high ρ_{rank} values. Corrupting the data effectively decreases the pumping well radius of influence (decreasing the signal to noise ratio), resulting in fewer observation wells with significant drawdown signal, and increasing the scatter of the predicted flow velocities (Figure 8b). For the noisy data analysis the same four observation times were used from each observation well and the same convergence criterion was used. This criterion aims at avoiding perfect fits between the observed and simulated heads at the observation wells; this is useful when the observations are noisy.

The data can be smoothed before using them in the inversion process (e.g., with a moving average or wavelet smoothing) or the forward and inverse models will effectively do the smoothing, because the models cannot perfectly match noisy data. To improve the convergence of the inverse method, unexplained (especially biased) noise should be investigated and dealt with if possible, to reduce its impact on the inverse solution [Xiang et al, 2008].

Different variances in true T and S fields

In the previous cases, the true T and S fields were generated for a variance in $\ln(T)$ and $\ln(S)$ of 2. Here we examine the effect of using a smaller and larger log variance ($\sigma^2_{\ln(T)}$ and $\sigma^2_{\ln(S)}$ of 1 and 3). Increasing the variance in the true field resulted in much poorer parameter estimates. Both ρ and ρ_{rank} are smaller and the norms and larger (see columns 7 and 8 in Tables 1a and 1b). The larger T and S parameter ranges associated with the larger variances are more difficult to estimate. As expected, the parameter estimates from the case with a lower variance are more accurately estimated (the true parameter fields are smoother) due to less nonlinear relationship between the head and the parameters [Yeh et al, 1996].

Drawdown-based estimation

For all previous analyses, the true initial conditions were used and the boundary conditions used to generate the initial condition were also used in the inverse model. In a real world case, aquifer tests are rarely begun from equilibrium and the aquifer's boundary conditions are often poorly known, therefore a scenario was performed where these were specified incorrectly.

A zero-drawdown specified head boundary condition was specified at all elements around the outside edge of the domain – even for the specified flux and no-flow boundary conditions in the true model (the two bedrock outcrops were still specified as no-flow). At all the observation locations, the drawdown from the pre-test condition was used in place of the simulated head. The results from this exercise, shown in Figure 9a and summarized in column 6 of Tables 1a and 1b, indicate that very good results are still obtainable, even when the initial or boundary conditions are poorly known. The predicted velocity components (Figure 9b) are not as good as in the case where the initial condition and boundary conditions are perfectly known, but the prediction is still reasonable, indicated by the high ρ_{rank} values.

Discussion

While a synthetic study can never take into account all the uncertainty potentially present in real-world field problems, such as the potential mis-characterization of a hydrologic system, it can isolate the issues related to data availability and aquifer test design. In this case we have used the same model type and grid to compute the “true” and inverse solutions, therefore there is no estimation error due to epistemic uncertainty.

Viability of other methods

This work stresses the benefits of using tomographic aquifer tests, and their inversion can be carried out with a variety of different tools. All the results computed here were done using SSLE and the finite element 2D flow model VSAFT2. Less than 10 iterations in SSLE were needed to meet the specified convergence criterion.

Qualitative comparisons of the possible combinations of different "machinery" that could be used to implement the HT inversion outlined here is beyond the scope of this paper, but a similar implementation could be done using public domain software such as MODFLOW [Harbaugh, 2005], PEST, or UCODE which utilize the perturbation approximation to the sensitivity.

If other methods are used and aquifer properties in each element of the forward model are estimated, then a regularization technique must be employed to reduce the effects of over-parameterization. One could effectively increase the number of observations by adding regularization "observations" that the parameter distribution is smooth. Alternatively, one could decrease the number of parameters being estimated. This can be accomplished using a pilot point method [RamaRao et al., 1995], where kriging fills in the model grid with aquifer parameters from a smaller set of estimated values. Another means of accomplishing this is through the singular value decomposition threshold method [Doherty, 2007], where only those parameters with large singular values in the estimation process are included. This reduces the dimensionality of the inverse problem without choosing a priori which parameters are more important, or where pilot points should be located.

Kalman filters are another class of candidate inversion algorithm; they are popular in control and systems engineering, and have been applied hydrologic problems in different ways [Chen and Zhang, 2006; Goegenbeur and Pauwels, 2007]. They are more general than non-linear least squares, since model and measurement noise can be incorporated directly into the inversion process, obviating the need for smoothing noisy data, but they do not have any means of incorporating the spatial correlation between the parameters into the estimation process, as SSLE does.

Conclusions

Based on the numerical experiments performed on the given synthetic regional domain, transient HT inversion using the SSLE is shown to work well for estimating the aquifer parameters T and S on a regional scale. While all the simulations performed in this work have been done using the SSLE adjoint-based inverse method, this is not the only option.

We address the test scale issue that arises from applying HT to a basin-scale problem by using multiple wells distributed across the basin in each pumping event. We feel this is a realistic way to address the scale problem in a manner that can potentially be applied to monitored municipal or treatment wellfields.

The tomographic approach to analyzing aquifer test data could potentially be used on existing monitoring data. In many basins there are collections of operational data and numerous aquifer tests which have been conducted through time, which may not provide a great deal of useful basin-wide information individually, but when analyzed together, they can create a whole which is greater than the sum of the parts. Results of this study appear to echo the call by Yeh and Lee [2007]: It is time to change the way we collect and analyze data for aquifer characterization.

Acknowledgments

This is an extension of a term project for the advanced subsurface hydrology class (HWR535) offered at Department of Hydrology and Water Resources at the University of Arizona by the last author; it was partially funded by a SERDP grant (ER-1365) subcontracted through University of Iowa, and NSF IIS-0431079. We would like to thank Michael Fienen and one anonymous reviewer for their detailed and insightful comments.

References

Barlebo, H. C., M. C. Hill, and D. Rosbjerg (2004), Investigating the Macrodispersion Experiment (MADE) site in Columbus, Mississippi, using a three-dimensional inverse flow and transport model, *Water Resources Research*, 40(4), W04211.

Bohling, G. C., X. Zhan, J. J. Butler Jr., and L. Zheng (2002), Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities, *Water Resources Research*, 38(12), 1324.

Bohling, G. C., J. J. Butler Jr., X. Zhan, and M. D. Knoll (2007), A field assessment of the value of steady shape hydraulic tomography for characterization of aquifer heterogeneities, *Water Resources Research*, 43(5), W05430.

Brauchler, R., R. Liedl, and P. Dietrich (2003), A travel time based hydraulic tomographic approach, *Water Resources Research*, 39(12), 1370.

Bredehoeft, J. D. (2003), From models to performance assessment: the conceptualization problem, *Ground Water*, 41(5), 571–577.

Chen, Y., and D. Zhang (2006), Data assimilation for transient flow in geologic formations via ensemble Kalman filter, *Advances in Water Resources.*, 29, 1107–1122.

Doherty, J. (2007), *PEST: Model-independent Parameter Estimation User Manual,* edition 5, Watermark Numerical Computing.

Goegbeur, M., and V. R. N. Pauwels (2007), Improvement of the PEST parameter estimation algorithm through extended Kalman filtering, *Journal of Hydrology*, 337, 436–451.

Gottlieb, J., and P. Dietrich (1995), Identification of the permeability distribution in soil by hydraulic tomography, *Inverse Problems*, 11, 353–360.

Hao, Y., T.-C. J. Yeh, J. Xiang, W. A. Illman, K. Ando, K.-C. Hsu (2008), Hydraulic tomography for detecting fracture connectivity, *Ground Water*, in press.

Harbaugh, A. W. (2005), The US Geological Survey modular ground-water model, the ground-water flow process, in *USGS Techniques and Methods* 6-A16.

Hughson, D.L., and T.-C.J. Yeh (2000), An inverse model for three-dimensional flow in variably saturated porous media, *Water Resources Research*, 36(4), 829–839.

Illman, W. A., X. Liu, and A. Craig (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), 222–244.

Isaaks, E. H., and R. M. Srivastava (1989), *An Introduction to Applied Geostsatistics*, Oxford.

Konikow, L. F., and J. D. Bredehoeft (1992), Groundwater models cannot be validated, *Advances in Water Resources*, 15(1), 75–83.

Li, W., A. Englert, O. A. Cirpka, J. Vanderborght, and H. Vereecken (2007), Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests, *Water Resources Research, 43(4), W04433.*

Liu, S., T.-C. J. Yeh, and R. Gardiner (2002), Effectiveness of hydraulic tomography: sandbox experiments, *Water Resources Research*, 38(4)WR000338.

Liu, X., W. A. Illman, A. J. Craig, J. Zhu, and T.-C. J. Yeh (2007), Laboratory sandbox validation of transient hydraulic tomography, *Water Resources Research*, 43(5), W05404.Oreskes, N., K. Shrader-Frechette, and K. Belitz (1994), Verification, validation and confirmation of numerical models in the earth sciences, *Science*, 243(5147), 641--646.

Poeter, E. P., M. C. Hill, E. R. Banta, S. Mehl, and S. Christensen (2005), UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation, in *USGS Techniques and Methods* 6-A11.

RamaRao, B. S., A. M. LaVenue, G. de Marsily, and M. G. Marietta (1995), Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1, Theory and computational experiments, *Water Resources Research*, 31(3), 475–493.

Straface, S., T.-C. J. Yeh, J. Zhu, S. Troisi, and C. H. Lee (2007), Sequential aquifer tests at a well field, Montalto Uffugo Scalo, Italy, *Water Resources Research*, 43, W07432.

Sun, N.-Z., and W. W.-G. Yeh (1992), A stochastic inverse solution for transient groundwater flow: Parameter identification and reliability analysis, *Water Resources Research*, 28(12), 3269–3280.

Sykes, J. F., J. L. Wilson, and R. W. Andrews (1985), Sensitivity analysis for steady state groundwater flow using adjoint operators, *Water Resources Research*, 21(3), 359–371.

Tikhonov, A. N., and V. Y. Arsenin (1977), *Solutions of Ill-Posed Problems*, Wiley.

Tosaka, H., K. Masumoto, and K. Kojima, (1993), Hydropulse tomography for identifying 3-D permeability distribution, in *Proceedings of the 4th Annual International Conference on High Level Radioactive Waste Management*.

Vasco, DW, H, Keers, and K. Karasaki (2000), Estimation of reservoir properties using transient pressure data: An asymptotic approach, *Water Resources Research*, 36

(12): 3447–3465.

Vesselinov, V. V., S. P. Neuman, and W. A. Illman (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), 3019–3042.

Wu, C. -M., T.-C. J. Yeh, J. Zhu, T. H. Lee, N.-S. Hsu, C.-H. Chen, and A. Folch-Sancho (2005), Traditional analysis of aquifer tests: Comparing apples to oranges?, *Water Resources Research*, 41(9), W09402.

Xiang, J., T.-C. J. Yeh, C.-H. Lee, K.-C. Hsu, and J.-C. Wen (2008), A new estimator and a guide for hydraulic tomography analysis, *Ground Water* (in press).Yeh, T.-C.J. (1992), Stochastic modeling of groundwater flow and solute transport in aquifers, *Journal of Hydrologic Processes*, 6, 369–395.

Yeh, T.-C. J., R. Srivastava, A. Guzman, T. Harter (1993), A numerical-model for water-flow and chemical-transport in variably saturated porous-media. *Ground Water*. 31 (4): 634–644.

Yeh, T.-C. J., M. Jin, and S. Hanna (1996), An iterative stochastic inverse method: conditional effective transmissivity and hydraulic head fields, *Water Resources Research*, 32(1), 85–92.

Yeh, T.-C. J., and C. H. Lee (2007), Time to change the way we collect and analyze data for aquifer characterization, *Ground Water* 45 (2): 116–118.

Yeh, T.-C. J., C. H. Lee, K.-C. Hsu, and Y.-C. Tan (2007), “Fusion of Active and Passive Hydrologic and Geophysical Tomographic Surveys: The Future of Subsurface

Characterization" in *Data Integration in Subsurface Hydrology*, Edited by D. W. Hyndman, F. D. Day-Lewis, and K. Singha, AGU monograph.

Yeh, T.-C. J., and S. Liu (2000), Hydraulic tomography: Development of a new aquifer test method, *Water Resources Research*, 36(8), 2095–2105.

Yeh, T.-C. J., and P. A. Mock (1996), Structured approach for calibrating steady-state ground-water flow models, *Ground Water*, 34(3), 444–450.

Zhang, J., and T.-C. J. Yeh (1997), An iterative geostatistical inverse method for steady flow in the vadose zone, *Water Resources Research*, 33(1), 63–71.

Zhu, J. and T.-C. J. Yeh (2005), Characterization of aquifer heterogeneity using transient hydraulic tomography, *Water Resources Research*, 41(7), W07028.

Zhu, J. and T.-C. J. Yeh (2006), Analysis of hydraulic tomography using temporal moments of drawdown recovery data, *Water Resources Research*, 42(2), W02403.

TABLE

1a	$\sigma^2_{\ln(T)} = 2.0$							
	Event 1	Events 1–2	Events 1–3	Events 1–4	Noisy observations	Drawdown + Incorrect BC	$\sigma^2_{\ln(T)} = 1.0$	$\sigma^2_{\ln(T)} = 3.0$
ρ	0.53	0.77	0.87	0.73	0.69	0.82	0.86	0.011
ρ_{rank}	0.77	0.86	0.89	0.85	0.82	0.84	0.93	0.033
L1	0.85	0.56	0.46	0.58	0.72	0.60	0.27	1.39
L2	1.25	0.58	0.42	0.64	0.89	0.87	0.13	2.90

1b	$\sigma^2_{\ln(S)} = 2.0$							
	Event 1	Events 1–2	Events 1–3	Events 1–4	Noisy observations	Drawdown + Incorrect BC	$\sigma^2_{\ln(S)} = 1.0$	$\sigma^2_{\ln(S)} = 3.0$
ρ	0.37	0.77	0.81	0.80	0.65	0.33	0.87	0.083
ρ_{rank}	0.67	0.73	0.74	0.77	0.76	0.71	0.87	0.031
L1	1.14	0.75	0.67	0.70	0.75	0.95	0.33	1.38
L2	2.20	1.13	1.01	0.96	1.05	1.46	0.22	3.07

Table 1Comparison of summary T (1a) and S (1b) statistics for different SSLE inverse solutions, ρ and ρ_{rank} are the correlation and rank correlation coefficients, L1 and L2 are norms indicating bias and error in standard deviation respectively.

FIGURES

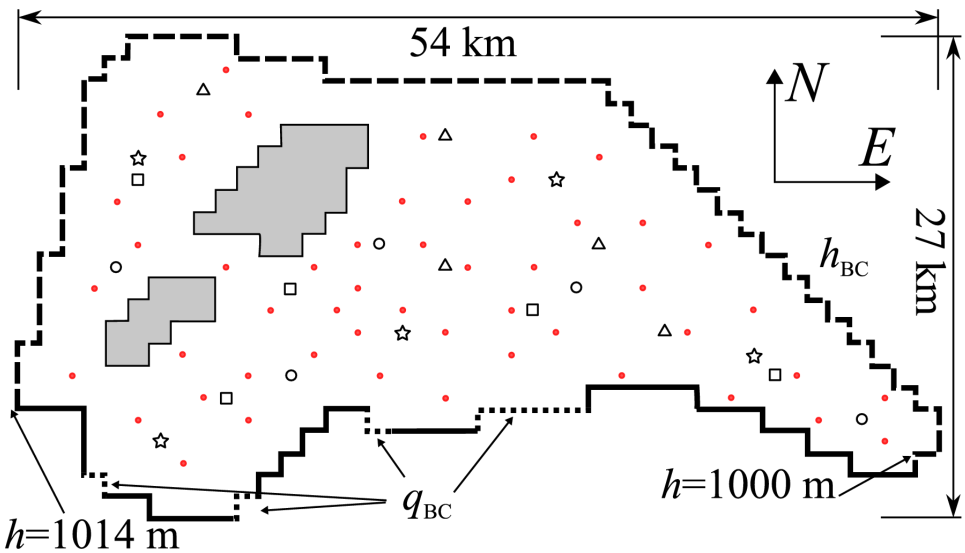

Fig 1 Map of domain showing pumping (open) and observation (solid) locations; Δ=event 1, □=event 2, star=event 3, ○=event 4. Dashed boundary is specified head, dotted boundary is specified non-zero flux, and solid boundary is no-flow.

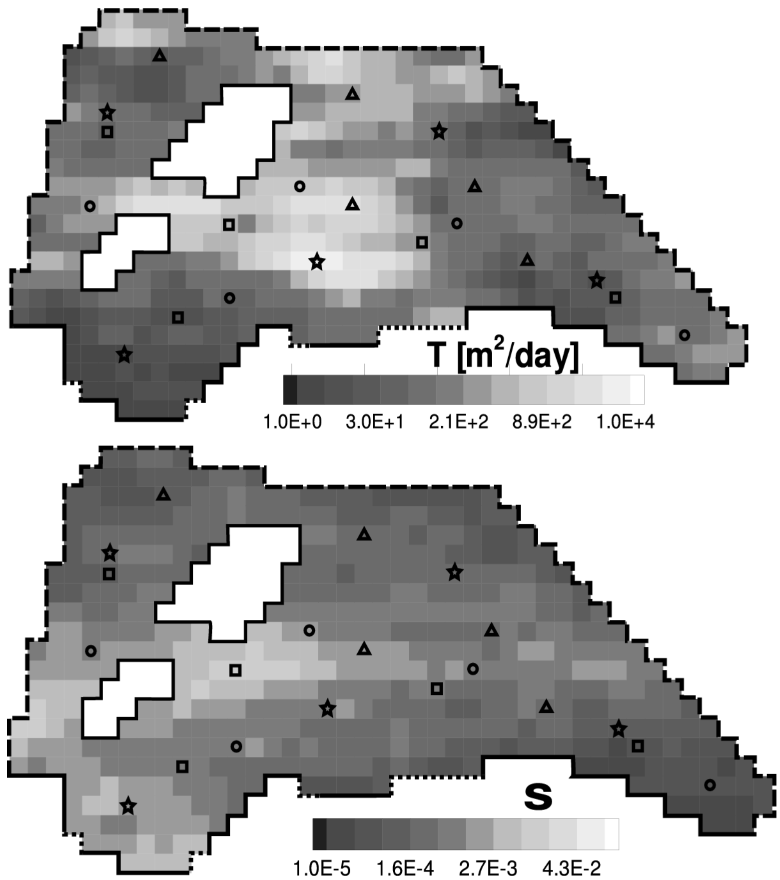

Fig 2: Map of true randomly-generated T field with $\sigma^2_{\ln(T)}$=2.0 and mean=300m²/day, and S field with $\sigma^2_{\ln(S)}$=2 and mean=0.0001.

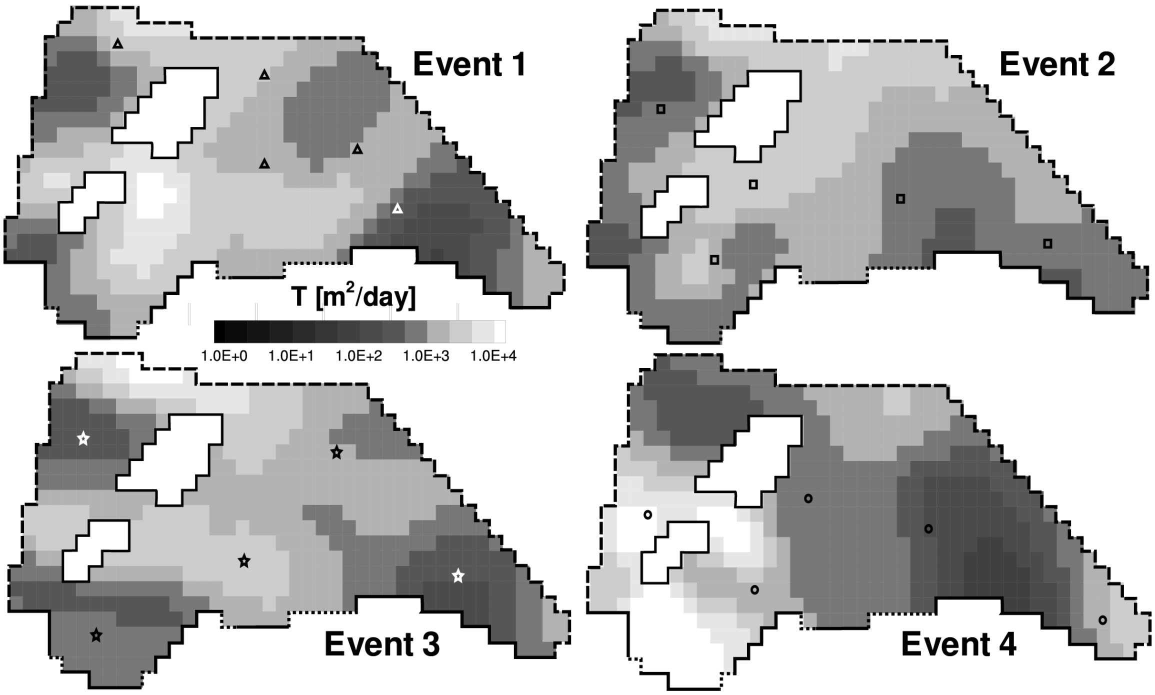

Fig 3a Maps of estimated *T* using Theis analysis

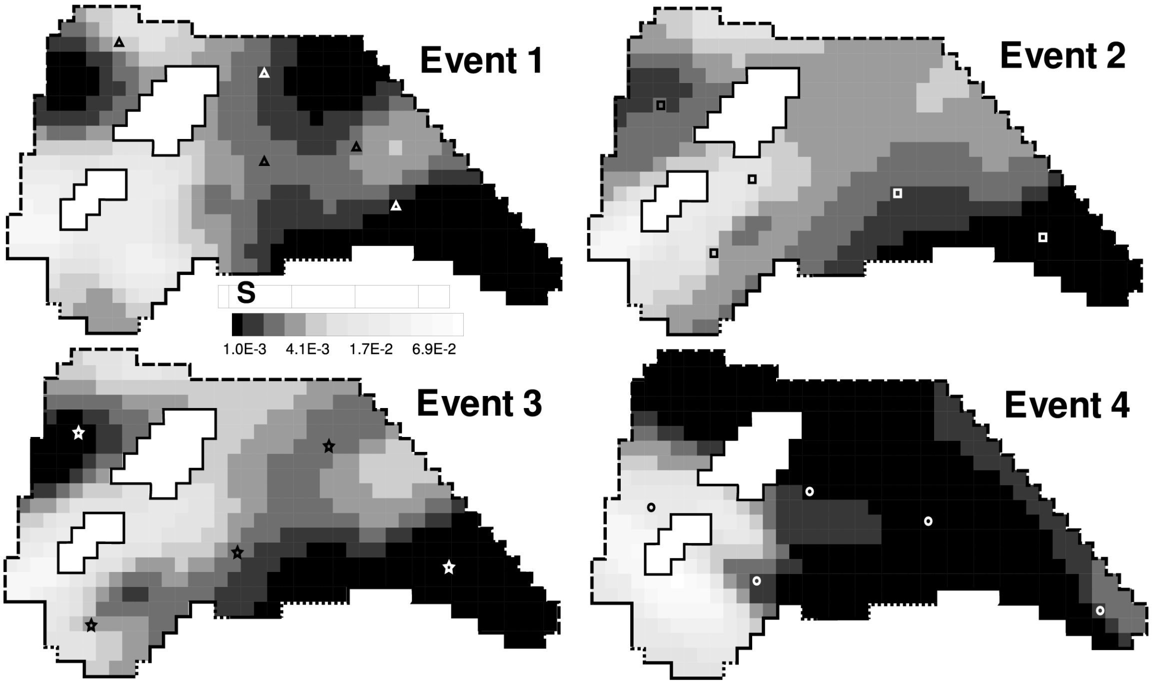

Fig 3b Maps of estimated *S* using Theis analysis

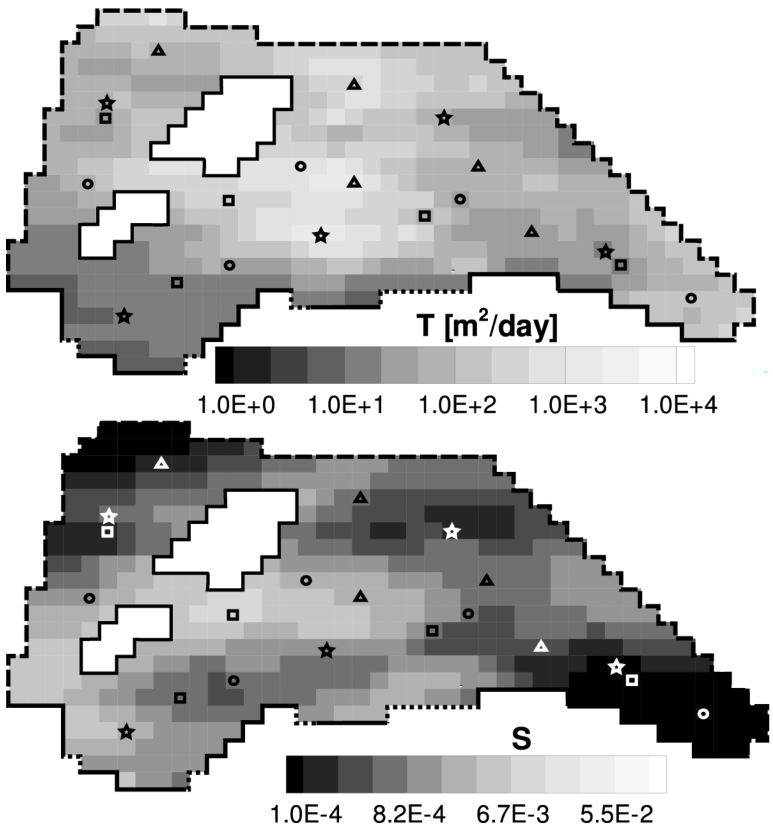

Fig 4 Maps of estimated *T* and *S* using SSLE method (all 4 pumping events)

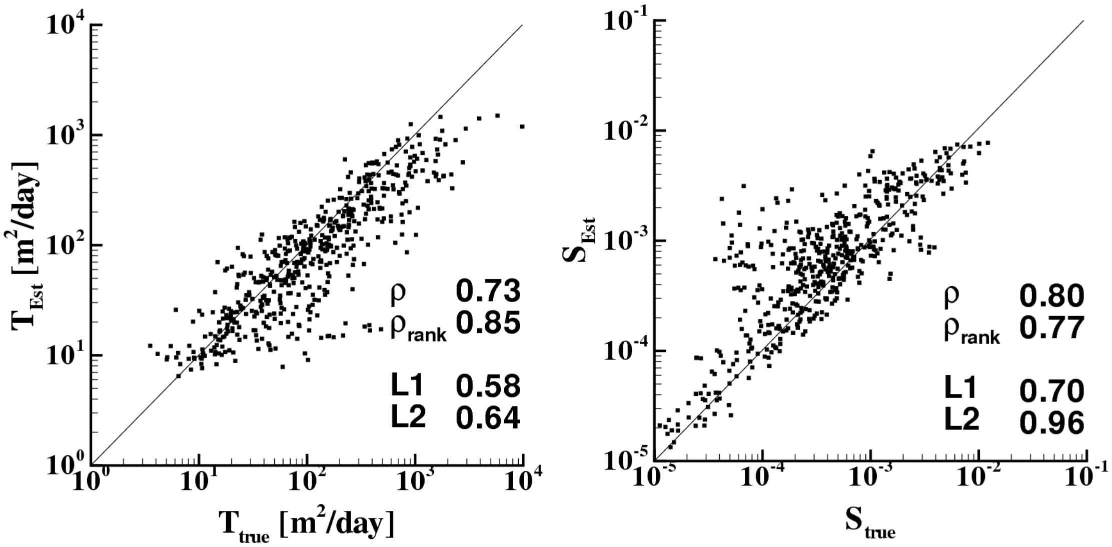

Fig 5 Scatterplots of estimated *T* and *S* using SSLE

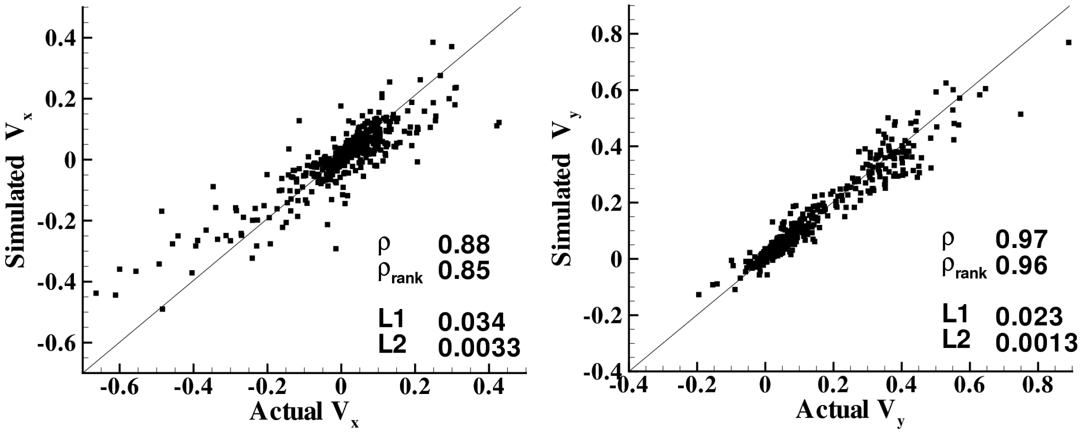

Fig 6 Scatterplots of *x* and *y* velocity components for SSLE-estimated *T* and *S*

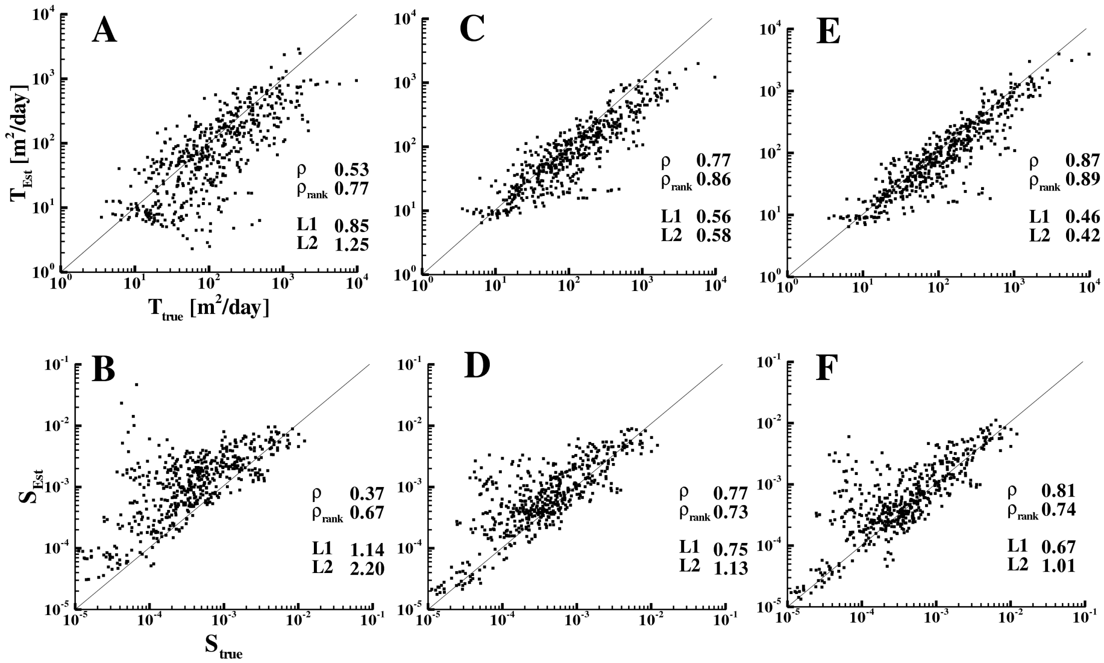

Fig 7 Scatterplots of SSLE inversion using different pumping events (top row is *T*, bottom row is *S*); A,B = event 1 only; C,D = events 1–2; E,F = events 1–3; see Figure 5 for events 1–4.

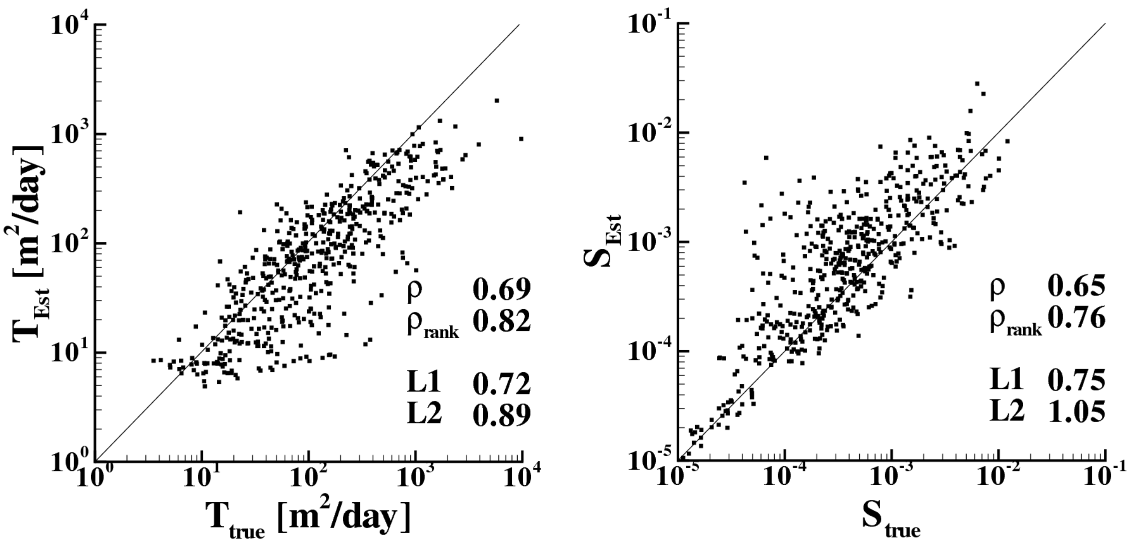

Fig 8a Scatterplots of SSLE-estimated *T* and *S* with noisy (σ=0.1m) head observations

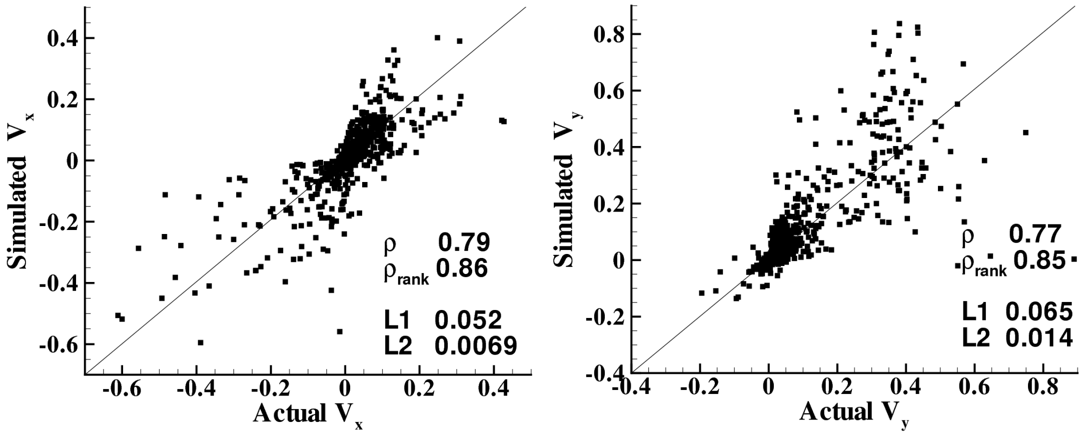

Fig 8b Scatterplots of *x* and *y* velocity components for SSLE-estimated *T* and *S* using noisy observations

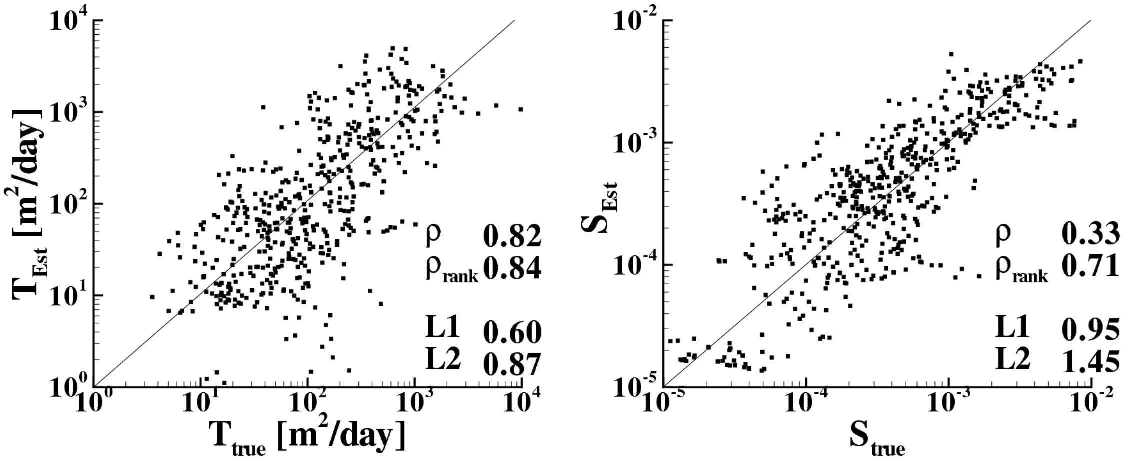

Fig 9a Scatterplots of SSLE-estimated *T* and *S* using drawdown and incorrect boundary conditions

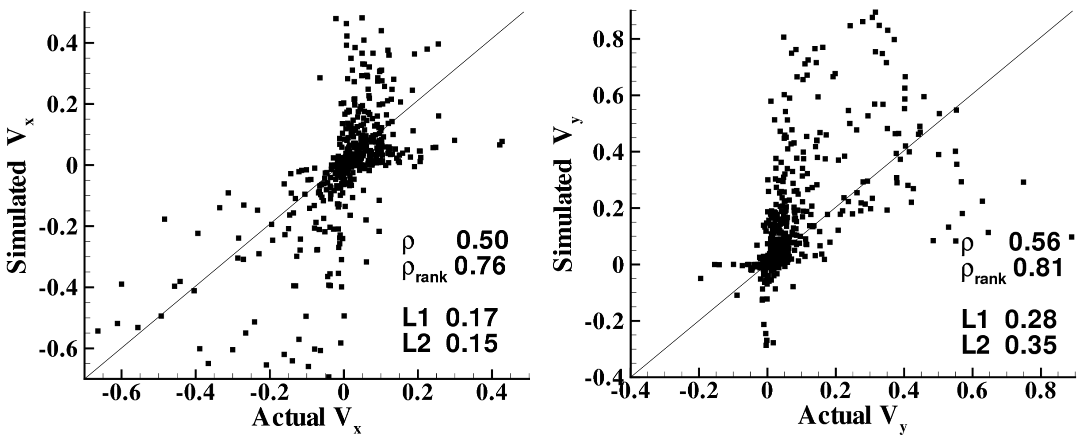

Fig 9b Scatterplots of *x* and *y* components of velocity for SSLE-estimated *T* and *S* using drawdown and incorrect boundary conditions