

# Stochastic Subsurface Data Integration Assessing Aquifer Parameter and Boundary Condition Uncertainty

Dongdong Wang · Ye Zhang

Received: date / Accepted: date

**Abstract** This research successfully extends a deterministic, physically-based inverse theory that is capable of simultaneous parameter and boundary condition estimation to uncertainty quantification in inverting steady-state groundwater flow in a two-dimensional aquifer with facies heterogeneity. Using facies and dynamic flow measurements sampled at wells as observations, a stochastic subsurface data integration technique is proposed: (1) Sequential Indicator Simulation integrates facies data by characterizing its geostatistical parameters (experimental directional variograms and sample facies proportions) to generate correlated facies models; (2) for each facies model, hydraulic conductivities and flow field (including the unknown boundary conditions) are estimated via a direct inversion method; (3) uncertainty in inversion, including both uncertainties of the estimated hydraulic conductivities and the flow field, is evaluated by assessing the inversion outcome for all facies models. To test the proposed integration technique, a reference forward model provides both facies characterization and dynamic measurements at increasing sampling densities (i.e., data quantity) and measurement errors (i.e., data quality). Via smoothing and grid coarsening, alternative hydraulic conductivity parameterization is also tested in inversion. Uncertainty in the estimated conductivities and boundary conditions is then quantified against the reference model to evaluate the quality of inversion. Results suggest that for the ranges of tested variation in data quantity, quality, and inverse conductivity parameterization, (1) data quantity has the strongest impact on both inversion accuracy and precision; (2) data quality influences inversion accuracy; (3) inverse parameterization has the weakest influence on inversion as long as the overall facies pattern is captured (i.e., sufficient data quantity). A balance can thus be achieved between parameterization, computational efficiency, and inversion performance. For

---

Dongdong Wang  
Department of Geological Sciences, University of North Carolina, Chapel Hill, North Carolina,  
USA  
E-mail: wang1986@live.unc.edu

Ye Zhang  
1000 University Ave., University of Wyoming, Laramie, Wyoming, USA  
E-mail: yzhang9@uwyo.edu

the heterogeneity pattern investigated herein, by defining an acceptable margin of uncertainty for either conductivity or flow field estimation, optimal well spacing in relation to the characteristic length of heterogeneity can be determined under unknown boundary conditions. Finally, inversion domain should be closely defined by the measurement locations in order to minimize extrapolation errors.

**Keywords** Hydrogeological Modeling · Uncertainty quantification · Physically-based inversion · Boundary conditions

## 1 Introduction

Parameterizing hydraulic conductivities ( $Ks$ ) of an aquifer model is a challenging topic in many subsurface investigations. Groundwater modeling often suffers from significant uncertainty in depicting both the  $K$  values and its spatial distribution in the subsurface, as extensive drilling and sampling to collect hydrogeological data is impractical. However, both aquifer management and subsurface contaminant cleanup operations require the development of accurate and appropriately parameterized groundwater models. Decision-making further requires scientifically informed assessment of the uncertainty in the model predictions, which often arises from the uncertainty in inferring subsurface hydraulic conductivities and the associated flow field. In natural aquifers, conductivity distribution is heterogeneous at multiple scales and its direct measurement often suffers from the well-known “scale effect”, i.e.,  $Ks$  measured by different instruments change with the support volumes being tested [1][2][3]. On the other hand, by developing a numerical aquifer model,  $Ks$  of the model can be calibrated using inverse theory based on measurements of the hydrogeological state variables (e.g., hydraulic heads, flow rates). In particular, inverse methods have been developed to not only facilitate parameter estimation, but also to quantify the associated estimation uncertainty, which can lead to an assessment of the prediction uncertainty when models are used for decision-making.

Generally, inverse methods in hydrogeology can be categorized into indirect and direct methods. With the indirect inversion method, an objective function, typically defined as a mismatch between the measurement data and the corresponding model simulated values, is minimized [4]. During inversion, parameters are updated iteratively by running a forward model that provides a link between the parameters and the data. The updated parameters then lead to an updated forward model, with which the objective function is reevaluated. This process is repeated until a user-defined fitting criterion is reached. The indirect method is found flexible and efficient in calibrating many subsurface models, although a number of issues have been identified in its applications. Among them, ill-posedness is a well-known problem, which is manifested by instability (sensitivity of the estimated parameters to small changes in the observation data and their errors), nonuniqueness (more than one set of parameters can calibrate the model, with results depending on the starting points in the parameter space), and failure to converge to a reasonable solution. Because a forward model is needed to optimize the objective function, boundary conditions (BC) of the model are either assumed known, or are calibrated as part of the inversion process. However, BC of aquifers are often unknown or uncertain. As demonstrated by [5], different combinations

the

of parameters and BC can lead to the same objective function values, thus results of indirect inversion may become non-unique. (In transient problems, both aquifer initial and BC are unknown – this topic is addressed elsewhere [6] and is not evaluated in this work.) For aquifer parameter estimation, direct inversion methods have also been explored owing to their mathematical straightforwardness and computational efficiency. However, direct methods suffer from issues such as high sensitivity to observation errors at streamline locations [7][8][9], stringent error bounds on measurements [10], reduced accuracy with increasing problem dimensions [11], etc. As reviewed in [12], ~~these~~ methods also simulate the forward model requiring the specification of BC, they therefore can suffer from instability and non-uniqueness in inversion and are not commonly used in groundwater model calibration.

Recently, based on the Stress Trajectories Element Method developed for solving certain geophysical and fluid mechanics problems [13], a new direct method was proposed for steady-state aquifer inversion that allows the simultaneous estimation of aquifer conductivities, sources/sinks (e.g., areal recharge, pumping and injection), and state variables (e.g., hydraulic head and Darcy flux) [5][14][15][16]. Because the state variables are solutions of the inversion, the unknown aquifer boundary conditions can thus be estimated. The direct method honors the physics of flow using continuity principles, while it directly incorporates noisy observed data at the measurement locations in a single step, without using forward simulations and iterative parameter updates to optimize any objective function. Because it does not simulate the forward model, the direct method is computationally efficient, and it utilizes an “inversion grid” with highly flexible discretization and parameterization. However, previous works have several limitations that restrict the method to solving simple problems with small grid sizes. First of all, conductivity distribution and patterns were assumed to be known when formulating the inversion equations. In real aquifers,  $K$  distributions are typically unknown or highly uncertain. ~~On the other hand~~, aquifer characterization can provide information on  $K$  distribution at borehole locations. Secondly, inversion equations have so far been developed and tested for small problems: the maximum size of the inversion grid tested is  $16 \times 16$  (two-dimensional problems) or  $14 \times 14 \times 14$  (three-dimensional problems), which limits the inversion to resolving only coarse heterogeneity features in the  $K$  model. Conductivity in natural aquifers often exhibits irregular spatial variation, which will require the development of refined inversion grid with more grid cells. Thirdly, because the  $K$  patterns were deterministic and assumed fully known, inversion outcomes, including both the estimated  $K$  values and the flow field, were deterministic. Uncertainty in inversion thus could not be quantified.

To demonstrate a wider applicability of the new direct method, the above issues need to be addressed by: (1) incorporating direct or indirect information about aquifer heterogeneity, e.g., facies observations at borehole locations; (2) inverting larger grid sizes to better resolve detailed heterogeneities; (3) accounting for uncertainty in inversion, including both the uncertainty of the estimated parameters and the flow field. For steady-state aquifer flows, this research aims to address the above issues by developing and testing a stochastic subsurface integration technique that combines geostatistics (i.e., static data integration) with direct inversion (i.e., dynamic data integration) into a single estimation and uncertainty analysis framework. The previous inversion will thus be extended to include aquifer prob-

lems with realistic  $K$  heterogeneity, while the associated estimation uncertainty, including the uncertainty in the inferred flow field, can be quantified. To evaluate the applicability of the integration technique to field conditions with limited and noisy observations, this research first investigates the impact of data quantity (i.e., number of sampling wells) and quality (i.e., magnitude of measurement errors) on the accuracy and stability of inversion, with results leading to a set of sampling strategies that can render the proposed integration more cost-effective. In particular, by defining an acceptable margin of uncertainty in either  $K$  or flow field estimation, optimal well spacing in relation to the characteristic length of heterogeneity can be recommended. Here, the term “optimal” means sufficient static and dynamic measurements can be obtained from fewest wells (or largest well spacing) that can still provide constraints to inversion, resulting in an acceptable level of estimation uncertainty. Moreover, because inverse  $K$  parameterization is flexible and, as demonstrated by previous upscaling studies [17], high-resolution  $K$  heterogeneity is not indispensable for predicting bulk flow in aquifers, this research will parameterize the inversion grid with heterogeneity at varying degrees of smoothness. In particular, by coarsening the inversion grid, balance between computational efficiency and estimation accuracy/uncertainty is explored.

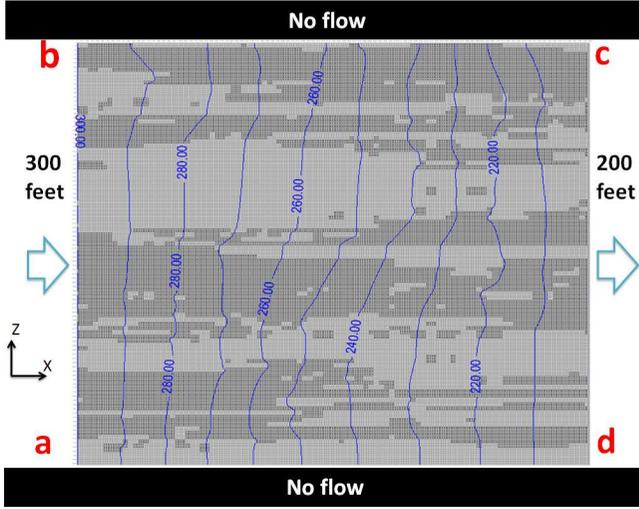
In the remainder of this article, the stochastic integration technique is introduced first, followed by description of a synthetic forward model that will be used as a reference model. The sampling design is discussed, which will provide inversion with facies and hydrogeological data borehole locations. Static data integration via geostatistical simulation is explained briefly, before dynamic data integration via the direct method is introduced. Because larger inverse problems are solved, matrix conditioning techniques are adopted to improve the computational efficiency of inversion. The results are presented in which the effects of data quality, quantity, and inverse parameterization on inversion accuracy and uncertainty are evaluated. Finally, implications of these results for field implementation of the proposed integration technique is discussed and future research indicated.

## 2 Method

In this work, a stochastic integration technique is proposed, consisting of 3 steps: (1) Sequential Indicator Simulation (SIS) integrates the facies data (i.e., experimental directional variograms and sample facies proportions computed from borehole measurements) to generate random models of correlated facies; (2) for each facies model, hydraulic conductivities and flow field (including the unknown BC) are estimated via direct inversion; (3) uncertainty in inversion including uncertainties of the estimated  $K$ s and the flow field is evaluated by assessing the outcomes for all facies models. Because a forward model is used as a reference model to test the quality of inversion (i.e., estimation accuracy and precision), below, the forward model is described first.

### 2.1 Forward (Reference) Model

To test and verify the proposed integration technique, a reference forward model, representing a two-dimensional synthetic aquifer transect, was created with known



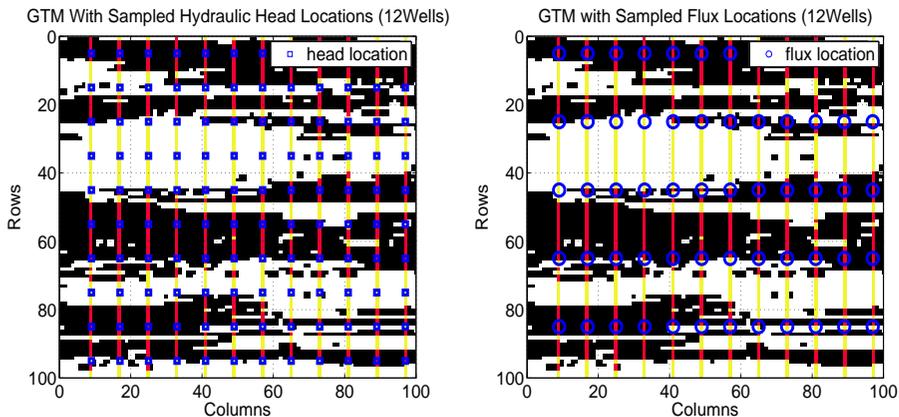
**Fig. 1** The reference model flow field containing two facies (dark gray: silty sand with  $K_1$ ; light gray: clean sand with  $K_2$ ). Given the specified BC and facies  $Ks$ , hydraulic head distribution is also shown. Four boundary segments are labeled: a-b denotes the inflow boundary, and c-d denotes the outflow boundary.

facies patterns and  $K$  values (Fig. 1). The groundwater flow equation of the forward model, representing a confined aquifer without sources/sinks, can be written as:

$$\begin{aligned} \nabla \cdot (\mathbf{q}) &= 0 \\ \mathbf{q} &= -K(x, z)\nabla h \quad (x, z) \in \Omega \end{aligned} \quad (1)$$

where  $\Omega$  is the spatial domain (assumed to be the same for the forward model and for inversion),  $x$  and  $z$  represent the horizontal and vertical axes in a Cartesian coordinate,  $\nabla$  is the gradient operator,  $h$  is hydraulic head,  $\mathbf{q}$  is Darcy flux, and  $K(x, z)$  denotes locally isotropic hydraulic conductivity which is parameterized as facies in this work. Note that though prior research has addressed inversion with source/sink effects [14][16], this topic is beyond the scope of the current study and will be left for another treatment.

To drive steady-state flow through the aquifer, no-flow boundaries are specified to the model top and bottom boundaries, while constant hydraulic heads are specified on the left and right sides of the model, i.e., 300 feet and 200 feet, respectively (Fig. 1). The map of facies  $K$  is adopted from [18], consisting of 100 pixels in both the horizontal and vertical directions. The facies (true) conductivities were set to  $K_1 = 1$  ft/yr (silty sand) and  $K_2 = 10$  ft/yr (clean sand) according to lithology type [19]. Note that the horizontal facies correlation range of the clean sand (light grey regions in Fig. 1) is approximately the lateral domain problem size, i.e., 100 pixels. The forward model is discretized with a uniform fine grid ( $N_x = 500$  and  $N_z = 500$ ; each facies pixel is thus represented by  $5 \times 5$  grid cells) and is solved with the finite difference method (FDM) with MODFLOW2000 [20]. (The forward model is also referred to herein as the FDM.) Mass balance reports from MODFLOW2000 suggest negligible errors on the order of  $10^{-7}$ , thus the simulated flow field is considered error-free.



**Fig. 2** Locations of 12 sampling wells placed uniformly across the aquifer transect. Along each well, head and flux sampling points are shown. The reference facies map (100 by 100 pixels) is shown in the background.

To provide the measurements for inversion, a set of synthetic observation wells, each with a width of one FDM grid cell ( $1/5$  pixel width in the facies map), was sampled for facies, hydraulic heads, and point-scale Darcy fluxes. Facies was sampled along the full length of each well, representing borehole measurements. Hydraulic head and flux samplings were made at discrete locations corresponding to a multilevel configuration, e.g., 10 hydraulic head and 5 flux measurements per well. A schematic diagram with 12 uniformly spaced sampling wells is shown in Fig. 2. Moreover, if the lateral domain size is used to approximate the horizontal facies correlation range ( $\lambda_H$ ), the lateral well spacing of Fig. 2 is thus  $\sim 1/12$  of  $\lambda_H$ . As part of the analysis on sampling density, the number of sampling wells will be reduced, i.e., 6 wells (well spacing is  $\sim 1/6$  of  $\lambda_H$ ) and 3 wells (well spacing is  $\sim 1/3$  of  $\lambda_H$ ). A horizontal sampling index (SI) can be defined as:  $\frac{\lambda_H}{well\ spacing}$ . For the different sampling schemes, SI = 12 (12 wells), 6 (6 wells), and 3 (3 wells). A larger SI value suggests higher sampling density, with correspondingly smaller well spacing in relation to the lateral characteristic length of the facies. Furthermore, twice and four times as many measurements were sampled when SI is 12, compared to when SI=6 and SI=3, respectively. For each sampling density, the reference model facies distribution, the facies  $Ks$ , and fluid flow BC will be later recovered using the proposed stochastic integration technique.

## 2.2 Static Data Integration

Indicator geostatistics is used to assimilate the observed facies along the wellboles [21]. Given the coded facies data (e.g., 0 for silty sand, 1 for clean sand), experimental variograms are computed along the horizontal and vertical directions. These variograms are fitted with an exponential model, which is found to best capture the facies correlation structure. When 12 wells were sampled, Fig. 3 (top row) presents the experimental facies variograms for “clean sand” and the fitted

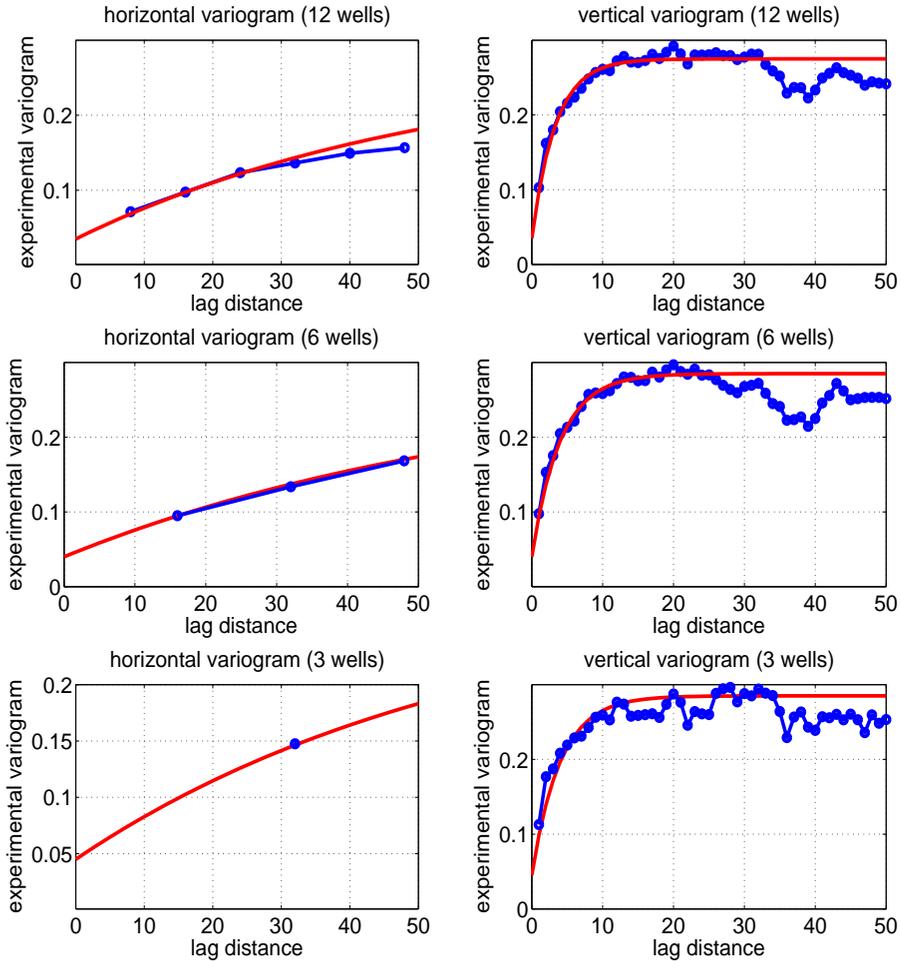
models, which are also presented below:

$$\gamma(|l|) = \begin{cases} 0.035 + 0.24[1 - e^{(-\frac{l_x}{60})}] \\ 0.035 + 0.24[1 - e^{(-\frac{l_z}{4})}] \end{cases} \quad (2)$$

where  $\gamma$  is variogram and  $l$  is lag distance in pixels ( $l_x$  is horizontal lag;  $l_z$  is vertical lag). In fitting the model, geometric anisotropy is assumed, thus the modeled horizontal sill (not reached in Fig. 3; top row) is the same as the modeled vertical sill. The fitted horizontal facies correlation range is  $\sim 105$  pixels, which is greater than the fitted vertical facies correlation range ( $\sim 15$  pixels). Both parameters are consistent with the average lateral extent and thickness of stratification as exhibited by “clean sand” (Fig. 1).

Given the fitted variogram models and the sampled indicator histograms, Sequential Indicator Simulation (SIS) was used to generate 100 facies realizations, which were conditioned to the observed wellbore facies (Fig. 4). Each facies realization has the same resolution as the original facies map, i.e.,  $100 \times 100$  pixels. When well density is changed, the above analysis, including both variogram modeling and SIS, is repeated. Fig. 3 also presents the variogram models fitted when fewer wells were sampled: not only will the fitted models be less accurate, the number of conditioning facies data for SIS is significantly reduced. As a result, as the number of wells is reduced, the SIS realization of the facies has become less accurate (Fig. 4).

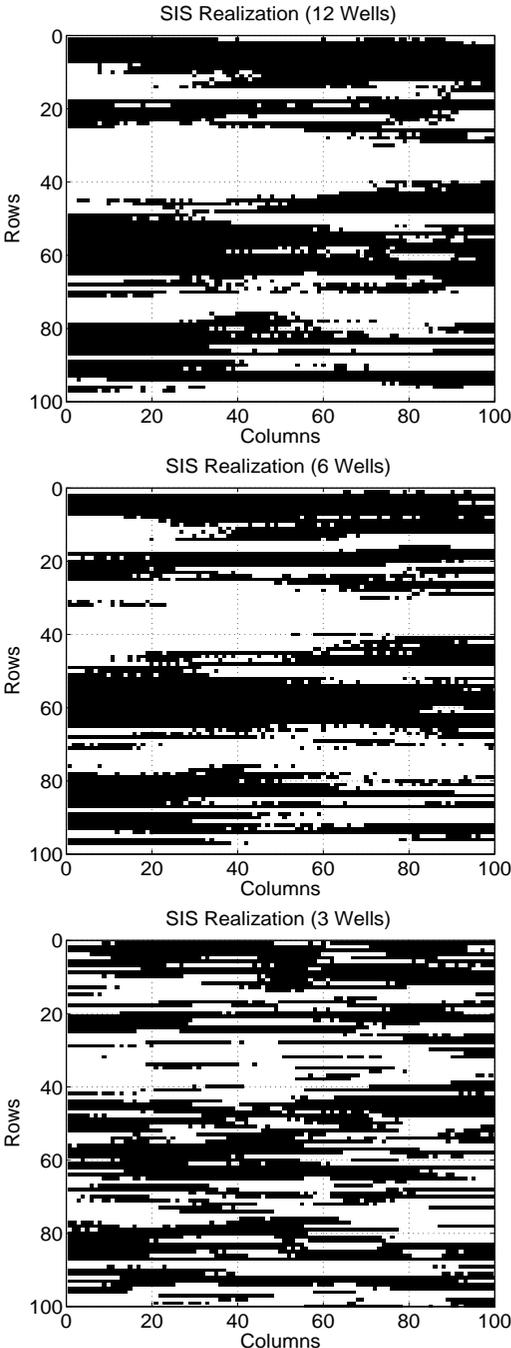
For a given sampling density, 100 facies realizations will be created. Using direct inversion, 100 flow fields and 100 sets of facies  $K$  values will be determined, one set for each realization. To drive inversion, dynamic observations (i.e., hydrogeological data) will be sampled from the same wells in the forward model.



**Fig. 3** Variogram analysis of the “clean sand” facies when a different number of wells was sampled. The dotted lines denote the experimental variograms and the smooth curves are the fitted exponential models. To eliminate edge effect, the maximum lag in both horizontal and vertical directions is half of the corresponding domain length scales, i.e., 50 pixels.

### 2.3 Dynamic Data Integration

For a given facies realization, which delineates the  $K$  distribution, this research adopted the direct method of [5] to estimate facies  $K$  values and head distribution throughout the domain. To apply the direct method, a set of fundamental solutions of inversion is fitted locally to the observed dynamic data, while flow continuity is honored over all inversion grid cells. The fundamental solutions of inversion are derived by solving the steady-state groundwater flow equation to obtain a set of local analytical solutions assuming that local  $K$  of a subdomain or a single inversion grid cell is homogeneous. However, unlike [5], whose fundamental solutions yielded only a single  $K$  value (i.e., ratios between facies  $K$ s were assumed known as a set of prior information constraints for inversion), the fundamental so-



**Fig. 4** SIS realizations given 12, 6, and 3 sampling wells. For each sampling density, one facies realization (out of 100 realizations) is shown.

lutions have been modified to allow the simultaneous estimation of both  $K$ s of the reference model (this approach is extendable to any number of facies):

$$\begin{aligned}\tilde{h}(x, z) &= a_0 + a_1x + a_2z + a_3xz + a_4(x^2 - z^2) \\ \tilde{q}_x(x, z) &= -K(a_1 + a_3z + 2a_4x) \\ \tilde{q}_z(x, z) &= -K(a_1 + a_3x - 2a_4z) \quad (x, z) \in \Omega_i\end{aligned}\quad (3)$$

where  $\tilde{h}$  denotes the approximate hydraulic head,  $(\tilde{q}_x, \tilde{q}_z)$  denote the approximate groundwater fluxes,  $a_i$  ( $i = 0, \dots, 4$ ) denote a set of coefficients that locally define these approximate solutions,  $K$  is local hydraulic conductivity:  $K \in (K_1, K_2, \dots)$  of the facies, and  $\Omega_i$  is a subdomain of the problem, here corresponding to an inversion grid cell.

The continuity equations, which penalize the mismatch between the fundamental solutions at the interface between adjacent inversion grid cells, can be written as:

$$\begin{aligned}\delta(p_j(x_j, z_j) - \epsilon)(K_1\tilde{h}^{(k)}(x, z) - K_1\tilde{h}^{(l)}(x, z)) &= 0, \quad \forall^{(k,l)} \in K_1 \\ \delta(p_j(x_j, z_j) - \epsilon)(K_2\tilde{h}^{(k)}(x, z) - K_2\tilde{h}^{(l)}(x, z)) &= 0, \quad \forall^{(k,l)} \in K_2 \\ \delta(p_j(x_j, z_j) - \epsilon)(K_m\tilde{h}^{(k)}(x, z) - K_m\tilde{h}^{(l)}(x, z)) &= 0, \quad \forall^{(k)} \in K_1, {}^{(l)} \in K_2, m \in (1, 2) \\ \delta(p_j(x_j, z_j) - \epsilon)(\tilde{q}_n^{(k)}(x, z) - \tilde{q}_n^{(l)}(x, z)) &= 0, \quad \forall K^{(k)} \neq K^{(l)} \\ \delta(p_j(x_j, z_j) - \epsilon)(\tilde{q}_t^{(k)}(x, z) - \tilde{q}_t^{(l)}(x, z)) &= 0, \quad \forall K^{(k)} = K^{(l)} \\ \delta(p_j(x_j, z_j) - \epsilon)(\tilde{q}_n^{(k)}(x, z) - \tilde{q}_n^{(l)}(x, z)) &= 0, \quad \forall K^{(k)} = K^{(l)}\end{aligned}\quad (4)$$

where  $p_j(x_j, z_j)$  denotes the  $j$ th collocation point, which lies on the interface between grid cells ( $k$ ) and ( $l$ ),  $\tilde{q}_n$  is normal flux at  $p_j$ ,  $\tilde{q}_t$  is tangential flux at  $p_j$ ,  $\delta(p_j(x_j, z_j) - \epsilon)$  is a Dirac delta weighting function [5] that samples the mismatch between the fundamental solutions at  $p_j(x_j, z_j)$ . The relation between  $(\tilde{q}_n, \tilde{q}_t)$  and  $(\tilde{q}_x, \tilde{q}_z)$  can be determined using the angles between the interface and the global coordinate axes.

Inversion further satisfies a set of data constraints which can be written as:

$$\begin{aligned}\delta(p_t - \epsilon)(K_m\tilde{h}^{(k)}(x_t, z_t) - K_m h^o(x_t, z_t)) &= 0 \quad m \in (1, 2) \\ \delta(p_t - \epsilon)(\tilde{q}_n^{(k)}(x_t, z_t) - q_n^o(x_t, z_t)) &= 0 \\ \delta(p_t - \epsilon)(\tilde{q}_t^{(k)}(x_t, z_t) - q_t^o(x_t, z_t)) &= 0\end{aligned}\quad (5)$$

where  $\delta(p_t - \epsilon)$  is the Dirac delta weighting function, which reflects confidence in the observed data (e.g., it can be inversely proportional to the measurement error variance),  $(x_t, z_t)$  denotes the location where an observed datum was sampled, and  $h^o$ ,  $q_n^o$ ,  $q_t^o$  are the observations,  $K_m$  denotes the conductivity of the facies which contains the observations. Flux measurements are used here to provide flow rate related information for inversion, because conductivity cannot be uniquely identified from hydraulic head observations alone. If subsurface flow rate measurements are available, however, the flux conditioning equations can be integrated to enforce conditioning by flow rates [5]. As was discussed in [15], measurements of in-situ fluxes and flow rates can be made with a variety of borehole and water balance techniques.

Based on the proposed fundamental solutions [Eqn. (3)], by assembling the continuity [Eqn. (4)] and the data equations [Eqn. (5)], a system of linear equations was developed:

$$\mathbf{A} \cdot \mathbf{x} \approx \mathbf{b} \quad (6)$$

where  $\mathbf{A}$  is the coefficient matrix,  $\mathbf{b}$  is the right hand side vector, and  $\mathbf{x}$  is the solution vector, where  $\mathbf{x} = [K_1, K_2, a_0^{(k)}, a_1^{(k)}, a_2^{(k)}, a_3^{(k)}, a_4^{(k)}, \dots]$ ,  $k = 1, 2, \dots, M$  (number of cells in the inversion grid).

Previous research [14] suggests that if dynamic observations are too few resulting in under-determined inversion systems of equations, stability of inversion can suffer. In this work, a sufficient number of head and flux observations was sampled (Fig. 2), thus Eqn. (6) is over-determined and is solved with a least-squares minimization technique as implemented by the LSQR algorithm [22]. For smaller problems reported in previous research, LSQR was found to be an efficient solver. However, problems inverted here are larger with many more unknowns, for which the speed of convergence was found to be sensitive to the condition number of  $\mathbf{A}$ . To improve matrix conditioning, preconditioning techniques including matrix scaling and Gaussian Noise Perturbation were used before the matrix solve [23]. Depending on the problem, pre-conditioning has improved the solver speed by up to 100 times (generally, greater speedup is achieved with larger matrices), without suffering ill-effects in inversion accuracy. After pre-conditioning, inversion of a single realization with the SIS grid will take  $\sim 10$  min. The hydraulic head field can then be recovered piecewise from the coefficients defining the fundamental solutions at each grid cell; from the estimated  $K$ s, Darcy flux field can be similarly recovered. The hydraulic head BC can be obtained by sampling the appropriate  $\tilde{h}(x, z)$  at the boundary locations. Similarly, flux BC can be obtained by sampling the appropriate  $\tilde{q}_x(x, z)$  and  $\tilde{q}_z(x, z)$  at the boundaries. In this work, the inverted BC are presented as hydraulic head values.

Because forward model is not solved, the direct method can employ flexible grids with various parameterizations for representing conductivity. In general, the SIS-generated facies realizations exhibit small-scale artifacts which lead to “patchiness” in the simulated facies distributions (Fig. 4). This may in turn impact the quality of inversion, as a greater degree of patchiness will lead to fewer continuity equations (i.e., the number of flux continuity constraints will be reduced). Three  $K$  parameterizations are tested in this work: (1) inversion directly adopts the noisy SIS parameterization with an identical grid ( $N_x = 100$ ,  $N_z = 100$ ); (2) inversion adopts the same SIS grid ( $N_x = 100$ ,  $N_z = 100$ ), but the SIS-simulated facies are smoothed by simulated annealing (SA)[24]; (3) inversion adopts a coarsened grid ( $N_x = 50$ ,  $N_z = 50$ ) whose facies parameterization is created by upscaling the SA field. For ease of reference, these parameterizations are labeled as “SIS”, “SA”, and “coarsened” inversion grids (example of each grid can be seen in [25]). For the latter two parameterizations, SIS realizations are first smoothed, and then coarsened, before dynamic integration is carried out to estimate the  $K$  and flow field ensembles. For a given sampling density, by comparing the inversion outcomes using all grids, the impact of inverse  $K$  parameterization on the quality of inversion is assessed. Moreover, if reduced resolution in the parameterization does not significantly degrade inversion performance (i.e., accuracy and uncertainty), lower resolutions will be preferred due to their greater computational efficiency.

## 2.4 Uncertainty Quantification

Uncertainty in the proposed stochastic integration derives from both static and dynamic data uncertainty, which comes from limited borehole sampling and measurement errors (if imposed). For facies sampling along boreholes, no measurement errors are imposed, i.e., observed facies type is assumed error-free. For sampling the dynamic data (heads, fluxes) from the same wells, random measurement errors are imposed. Because the FDM is the reference model, measurements sampled from this model are error-free, that is, “true” heads or “true” fluxes. To impose errors, the true heads are corrupted by random noises:  $h^m = h^{FDM} \pm \delta h$ , where  $h^m$  is measured head provided to inversion,  $h^{FDM}$  is true head, and  $\delta h$  is hydraulic head measurement error. The highest  $\delta h$  is  $\pm 10\%$  of the total hydraulic head variation in the forward model, with an absolute error up to  $\pm 10$  feet. This is reasonable considering that modern measurement techniques can determine water levels with a precision as fine as 1 cm [26]. Flux measurements, though also amenable to errors, are assumed error-free. The effect of imposing errors on fluxes is similar to imposing errors on heads, while imposing both errors at the same time can lead to inversion outcomes that are difficult to interpret, e.g., positive error in observed head gradient can be canceled by negative error in a nearby flux measurement.

Given a given facies parameterization (SIS, SA, or coarsened), 100 inversion systems of equations were assembled and solved, resulting in a set of ensemble solutions including the inverted  $K$ s, flow fields (i.e., heads and fluxes), and  $BC$ s. The accuracy of each solution can be assessed by comparing the estimated parameters and flow fields to those of the reference model. Because a given flow field can be uniquely determined from  $K$  distribution and BC, only the estimated  $K$ s and the recovered BC are compared to those of the forward model. The inverted ensemble flow fields will not be presented.

Uncertainty in the inverted conductivity of each facies is evaluated by a set of ensemble error statistics or performance metrics:

$$\begin{aligned}\varepsilon_K &= \left| \frac{E[K^{inv}] - K^{true}}{K^{true}} \right| \times 100\% \\ \sigma_K &= \left| \frac{K_{max}^{inv} - K_{min}^{inv}}{4K^{true}} \right| \times 100\%\end{aligned}\tag{7}$$

where  $K^{inv}$  is the inverted facies conductivity,  $K^{true}$  is the corresponding true conductivity, and  $E[\cdot]$  is expectation, taken here as arithmetic mean.  $E[K^{inv}]$  is thus the ensemble mean of the inverted facies conductivity.  $K_{max}^{inv}$  and  $K_{min}^{inv}$  denote the maximum and minimum inverted facies conductivity from the ensemble.

To evaluate the uncertainty in the inverted BC, two performance metrics were adopted:

$$\begin{aligned}\varepsilon_{BC} &= (1/n) \sum_{i=1}^n \left| \frac{(E[h_i^{inv}] - h_i^{true})}{h_i^{true}} \right| \times 100\% \\ \sigma_{BC} &= \sqrt{(1/n) \sum_{i=1}^n \left( \frac{(h_i^{max} - h_i^{min})}{4h_i^{true}} \right)^2} \times 100\%\end{aligned}\tag{8}$$

where  $h_i^{inv}$  is the inverted hydraulic head at boundary cell location  $i$  ( $n$  is the number of cells along the domain boundary),  $h_i^{true}$  represents the true head at the

same location,  $h_i^{max}$  and  $h_i^{min}$  denote the maximum and minimum inverted heads at location  $i$  from the ensemble solution, respectively, and  $E[h_i^{inv}]$  is the ensemble mean of the recovered head at location  $i$ . Note that by summing and averaging over all boundary cells,  $\varepsilon_{BC}$  and  $\sigma_{BC}$  represent the average BC estimation error and spread surrounding the ensemble mean.

In the following, inversion accuracy is defined by  $\varepsilon_K$  and  $\varepsilon_{BC}$ : higher values indicate a bias in the ensemble mean, thus poorer accuracy [4]. On the other hand,  $\sigma_K$  and  $\sigma_{BC}$  reflect the spread of the ensemble from the true parameters and flow field, respectively, and are thus linked to the precision of inversion [4]. Higher precision is reflected by lower  $\sigma_K$  and  $\sigma_{BC}$ , and thus lower estimation uncertainty. The quality of inversion is considered high if  $\varepsilon$  and  $\sigma$  are both small for conductivity estimation and BC recovery.

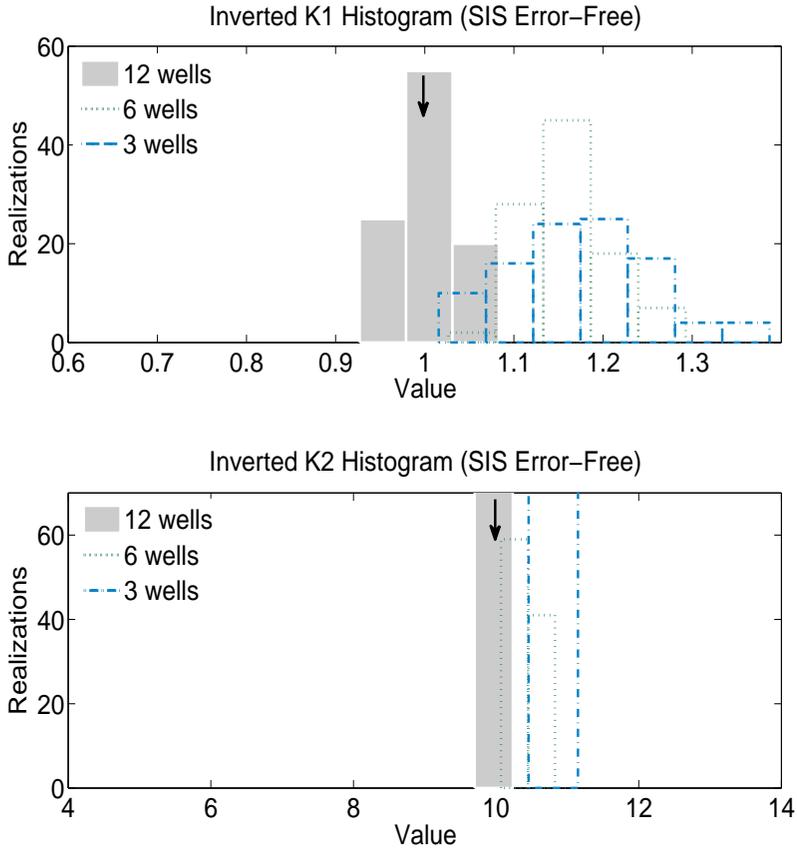
### 3 Results

In the uncertainty study, the accuracy and precision of the ensemble inverse solutions are examined when measurement quantity, quality, and inverse parameterization are varied. (1) Data quantity was varied by reducing the number of sampling wells from 12 to 6, and then to 3 wells, from which static and dynamic data were sampled from the FDM. The observed hydraulic heads were then subject to increasing measurement errors ( $\pm 1, \pm 2, \pm 5, \pm 10\%$ ), while inverse parameterization adopted the SIS (noisy) facies realizations with  $100 \times 100$  grid cells. (2) From the above, data quality was examined closely to evaluate the effect of measurement errors on inversion. (3) Using error-free measurements from the 12 wells, inverse parameterization was then varied, i.e., dynamic integration was performed using the *same* observed heads and fluxes, but with facies patterns embodied by the SIS, SA, and coarsened grids. (4) A co-effect study was carried out varying two factors at the same time.

In the following, units of the relevant quantities are:  $K$  in ft/yr (1 ft/yr = 0.305 m/yr),  $q$  in ft<sup>3</sup>/yr (1 ft<sup>3</sup>/yr = 0.305 m<sup>3</sup>/yr),  $h$  in ft (1 ft = 0.305 m), flow rate in ft<sup>3</sup>/yr (1 ft<sup>3</sup>/yr = 0.028 m<sup>3</sup>/yr). Alternatively, a consistent set of units can be assumed and all unit information can be removed [27].

#### 3.1 Data Quantity

When 12 wells were sampled from the forward model, static and dynamic data from these wells were provided to the proposed data integration: (1) given the observed facies at the wells, variogram model was fitted, with which 100 facies realizations were generated with SIS; (2) for each realization, dynamic data, sampled at the same wells, were inverted to obtain facies  $K$ s and the associated flow field; (3) for all realizations, an ensemble of inversion outcomes was generated, and Eqn.(7) and Eqn.(8) were used to assess the quality of inversion. The dynamic data were initially error-free, before increasing measurement errors —  $\pm 1, \pm 2, \pm 5, \pm 10\%$  — were imposed (for the 5 levels of measurement errors tested, 500 inversions were performed). When the sampling density was reduced to 6, and then to 3, the above steps were repeated (an additional 1,000 inversions were performed).



**Fig. 5** Inverted distributions of  $K1$  and  $K2$  obtained with different sampling densities when measurements were error-free. For each facies, arrow points to the true conductivity value. Inverse parameterization is SIS.

The results for conductivity estimation are summarized in Table 1. Several observations can be made: (1) for a given magnitude of the measurement error, the quality of inversion is highest when 12 wells were sampled and lowest when 3 wells were sampled (Fig. 5). Clearly, when sampling density is reduced, fewer static and dynamic data are available for the integration and, accordingly, inversion outcome worsens. (2) for a given sampling density, the quality of inversion worsens with increasing measurement errors (Table 1). When the observed heads were imposed with  $\pm 10\%$  errors, conductivity estimation error becomes very high suggesting strong biases: when 12 wells were sampled,  $\varepsilon_K$  is 76% for  $K1$  and 27% for  $K2$ ; when 6 wells were sampled,  $\varepsilon_K$  is 102% for  $K1$  and 40% for  $K2$ ; when 3 wells were sampled,  $\varepsilon_K$  is 156% for  $K1$  and 34% for  $K2$ . When 6 and 3 wells were sampled,  $E[K1]$  even becomes negative. This occurs when head gradients yield the wrong sign due to noisy measurements, but the correct flux measurements were still used in inversion. (3) an optimal sampling density for *accurate*  $K$  estimation depends on the magnitude of the measurement error (that is expected in a field problem) and an acceptable level of estimation error (that is user defined). For example, if

$\pm 2\%$  is considered a realistic measurement error (absolute measurement error is  $\pm 2$  ft), 3 wells (SI=3) will provide sufficient accuracy for inversion if  $K$  estimation error of 12% is considered acceptable. However, for the same level of measurement error, 6 wells (SI=6) will be needed if  $K$  estimation error of 5% is considered acceptable. (4) an optimal sampling density for *precise*  $K$  estimation also depends on the magnitude of the measurement error and an acceptable level of estimation error. For example, if  $\pm 2\%$  is considered a realistic measurement error, 6 wells will provide sufficient precision if  $\sigma_K$  of 7% is considered acceptable. However, 12 wells will be needed if  $\sigma_K$  of 5% is considered acceptable.

The results for BC estimation are summarized in Table 2, where ensemble statistics are compiled for 4 boundary segments (Fig. 1):  $\mathbf{a} - \mathbf{b}$ ,  $\mathbf{b} - \mathbf{c}$ ,  $\mathbf{c} - \mathbf{d}$ , and  $\mathbf{d} - \mathbf{a}$ . Along the same segment, the inverted boundary heads are compared to the true heads (Fig. 6 presents results given error-free measurements). Similar to  $K$  estimation, for a given measurement error, the quality of BC estimation (i.e., accuracy and precision) is highest when 12 wells were sampled and lowest when 3 wells were sampled. For a given sampling density, the quality of inversion worsens with increasing measurement errors (more on this later). For BC estimation, the optimal number of sampling wells also depends on the magnitude of the measurement error and an acceptable level of BC estimation error. For each boundary segment, distinct behavior in the ensemble statistics is also observed. For example, head recovery along different segments is sensitive to well density to different degrees. Both accuracy and precision of head recovery along the inflow boundary ( $\mathbf{a} - \mathbf{b}$ ) are not significantly affected by the decreasing well density (Fig. 6). However, inversion along the outflow boundary ( $\mathbf{c} - \mathbf{d}$ ) degrades appreciably when the number of sampling well is reduced. Inversion quality is spatially non-uniform, despite the fact that the sampling wells are uniform. This suggests the importance of identifying optimal sampling locations. To understand this issue, future work will combine inversion with sensitivity analysis [4][28].

For a given sampling density, hydraulic head recovery along the inflow ( $\mathbf{a} - \mathbf{b}$ ) and outflow ( $\mathbf{c} - \mathbf{d}$ ) boundaries are generally less accurate and precise, when compared to head recovery along the model top and bottom boundaries ( $\mathbf{b} - \mathbf{c}$ ,  $\mathbf{d} - \mathbf{a}$ ). This can be explained by extrapolation of the fundamental solutions in the model domains between the inflow boundary and the first sampling well, and between the outflow boundary and the last sampling well (Fig. 2). Because no measurements were made there, the inverse solution, which is conditioned at the well locations, is extrapolated towards the boundaries. Inversion quality along model top and bottom is higher, because both boundaries are closely conditioned by well data (Fig. 2). This suggests that (1) inversion domain should be defined by the well locations to reduce extrapolation errors, and (2) sampling along  $\mathbf{a} - \mathbf{b}$  and  $\mathbf{c} - \mathbf{d}$ , in effect providing (partial) BC to inversion, will likely improve inversion at these boundary locations.

### 3.2 Data Quality

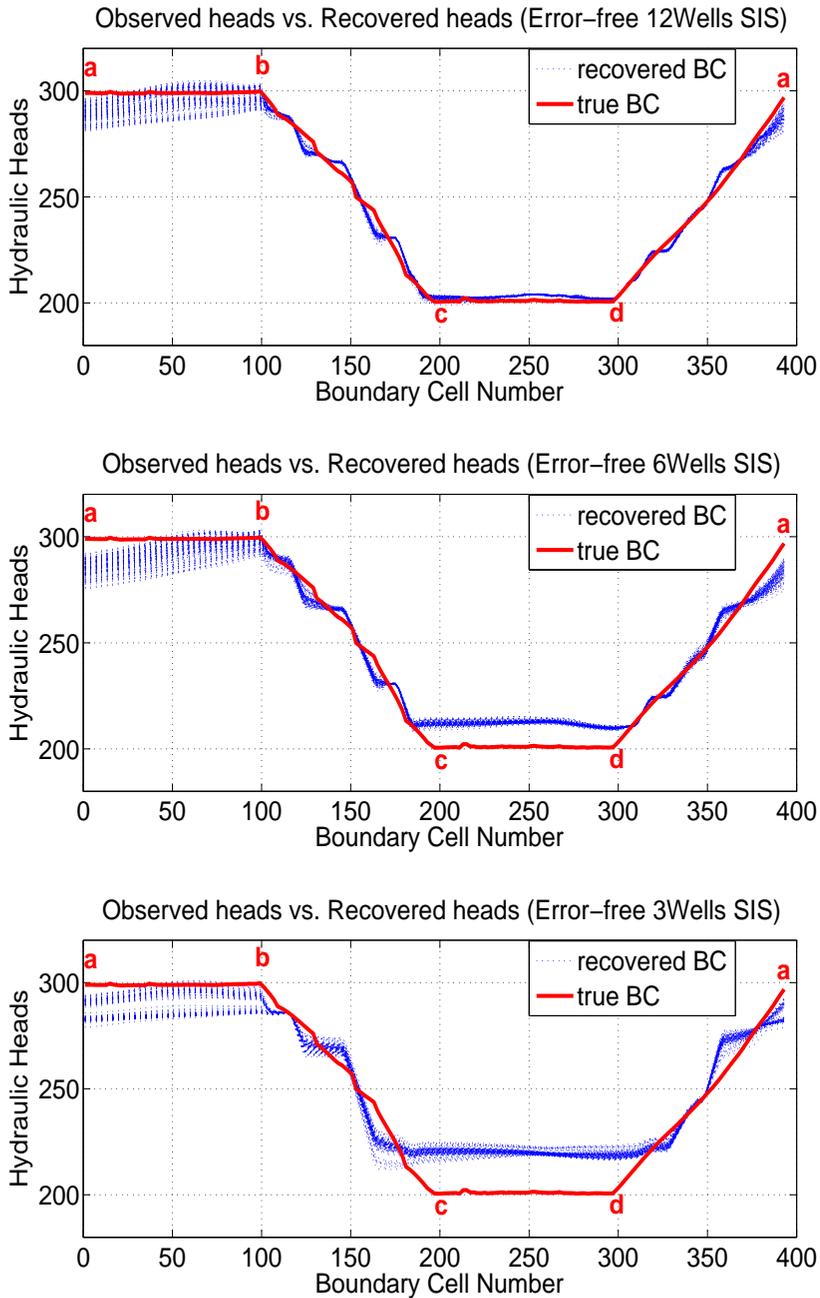
Results from the previous section are reexamined to assess the effect of measurement errors on inversion. Again, as reported in Table 1 and Table 2, quality of the inverted  $K$ s and  $BC$ s degrades with increasing head measurement errors and decreasing sampling density. Fig. 7 presents the distributions of the estimated  $K$ 1

**Table 1** Inversion accuracy and precision for the estimated facies conductivities.

		SIS grid						3 wells					
		12 wells			6 wells			Accuracy			Precision		
		Accuracy	Precision		Accuracy	Precision		Accuracy	Precision		Accuracy	Precision	
		$K_1$	$K_2$										
Error	$\epsilon(\%)$	0.0	0.3	3.8	1.2	15.0	4.2	7.0	1.9	18.0	8.3	10.3	1.7
	$\sigma(\%)$	3.0	1.3	4.3	1.5	13.0	3.6	5.8	2.0	17.0	8.4	11.5	1.9
$\pm 1\%$		7.0	2.9	4.8	1.5	4.0	1.6	6.3	1.8	12.0	7.1	13.8	2.2
$\pm 5\%$		29.0	10.3	6.5	2.2	42.0	12.1	8.3	3.4	42.0	4.7	24.5	3.6
$\pm 10\%$		76.0	27.4	7.3	2.5	102.0	40.0	15.5	4.4	156.0	33.7	31.5	4.1
Error-free Observed Data													
SIS													
SA													
Coarsened													
		Accuracy			Precision			Accuracy			Precision		
		$K_1$	$K_2$										
Wells	$\epsilon(\%)$	0.0	0.3	3.8	1.2	3.0	0.2	5.75	1.65	6.0	3.9	7.5	2.08
	$\sigma(\%)$	15.0	4.2	7.0	1.875	27.0	6.9	8.75	2.725	3.0	0.6	8.0	1.725
3		18.0	8.3	10.25	1.725	6.0	4.2	21.0	3.45	34.0	21.6	20.75	5.75
12 wells (Error-free)													
Accuracy													
Precision													
		section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>	section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>	section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>
		$\epsilon(\%)$	$\epsilon(\%)$	$\epsilon(\%)$	$\epsilon(\%)$	$\sigma(\%)$							
Grids		2.17	0.924	0.335	0.044	1.502	1.821	1.11	0.772	0.874	1.145	0.874	
SIS		1.63	0.084	0.165	0.18	1.60	1.948	1.145	0.874	1.145	0.874	0.874	
SA		3.39	0.352	0.6	0.772	0.809	0.95	0.709	0.669	0.95	0.709	0.669	
Coarsened													

**Table 2** Inversion accuracy and precision for the estimated hydraulic head BC. The bold alphabets denote different boundary segments (Fig. 1).

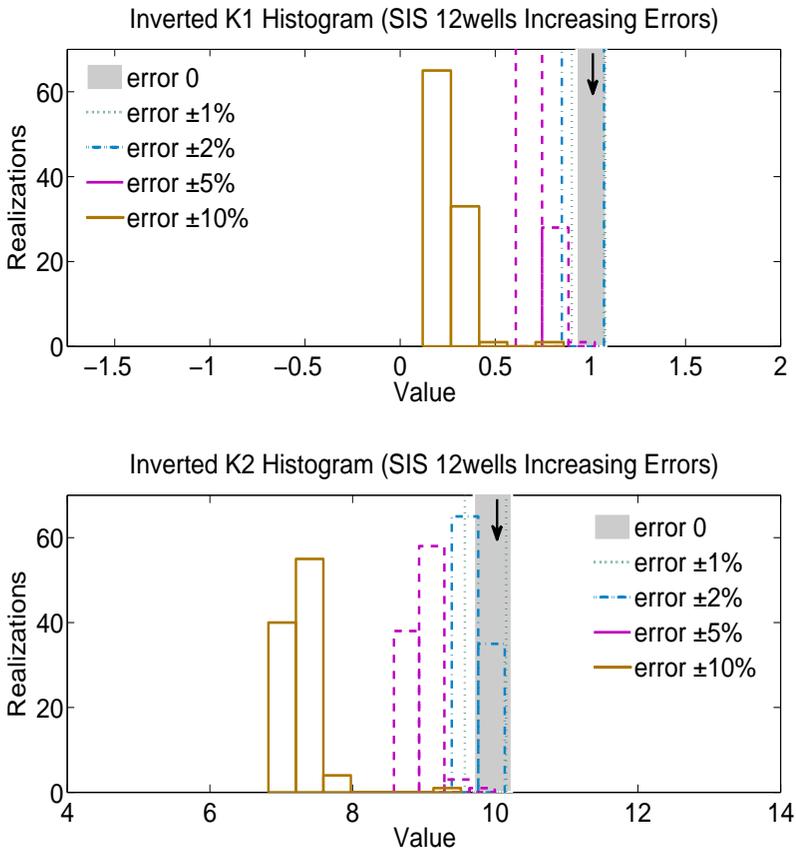
12 wells (SIS grid)											
Accuracy						Precision					
section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>	section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>	section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>
Error	$\epsilon$ (%)	$\epsilon$ (%)	$\epsilon$ (%)	$\epsilon$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)
0	2.170	0.924	0.335	0.044	1.502	1.821	1.114	0.772	1.114	1.300	0.733
$\pm 1\%$	2.187	0.792	0.260	0.008	1.568	1.901	1.300	0.733	1.901	1.300	0.733
$\pm 2\%$	2.183	0.660	0.195	0.060	1.632	1.980	1.488	0.721	1.980	1.488	0.721
$\pm 5\%$	2.020	0.260	0.055	0.208	1.899	2.295	2.064	0.973	2.295	2.064	0.973
$\pm 10\%$	1.273	0.412	0.020	0.452	2.958	3.627	3.111	1.648	3.627	3.111	1.648
6 wells (SIS grid)											
Accuracy						Precision					
section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>	section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>	section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>
Error	$\epsilon$ (%)	$\epsilon$ (%)	$\epsilon$ (%)	$\epsilon$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)
0	3.433	1.396	0.565	0.736	1.591	1.906	1.263	0.794	1.906	1.263	0.794
$\pm 1\%$	2.79	0.800	0.120	0.792	1.570	1.891	1.441	0.834	1.891	1.441	0.834
$\pm 2\%$	2.040	0.200	0.285	0.844	1.793	2.164	2.019	0.868	2.164	2.019	0.868
$\pm 5\%$	0.707	1.620	1.270	0.976	2.443	3.470	3.884	0.948	3.470	3.884	0.948
$\pm 10\%$	6.790	4.740	2.300	1.184	4.526	6.689	7.616	1.588	6.689	7.616	1.588
3 wells (SIS grid)											
Accuracy						Precision					
section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>	section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>	section <b>a-b</b>	section <b>b-c</b>	section <b>c-d</b>	section <b>d-a</b>
Error	$\epsilon$ (%)	$\epsilon$ (%)	$\epsilon$ (%)	$\epsilon$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)	$\sigma$ (%)
0	3.647	3.692	0.855	0.724	1.462	1.759	1.304	1.796	1.759	1.304	1.796
$\pm 1\%$	3.893	3.828	0.850	0.400	1.431	1.716	1.300	1.642	1.716	1.300	1.642
$\pm 2\%$	4.057	3.956	0.925	0.004	1.460	1.675	1.365	1.589	1.675	1.365	1.589
$\pm 5\%$	4.040	4.272	1.635	1.600	1.922	1.666	1.704	1.797	1.666	1.704	1.797
$\pm 10\%$	2.383	4.600	4.295	5.516	3.598	2.782	2.840	2.863	2.782	2.840	2.863



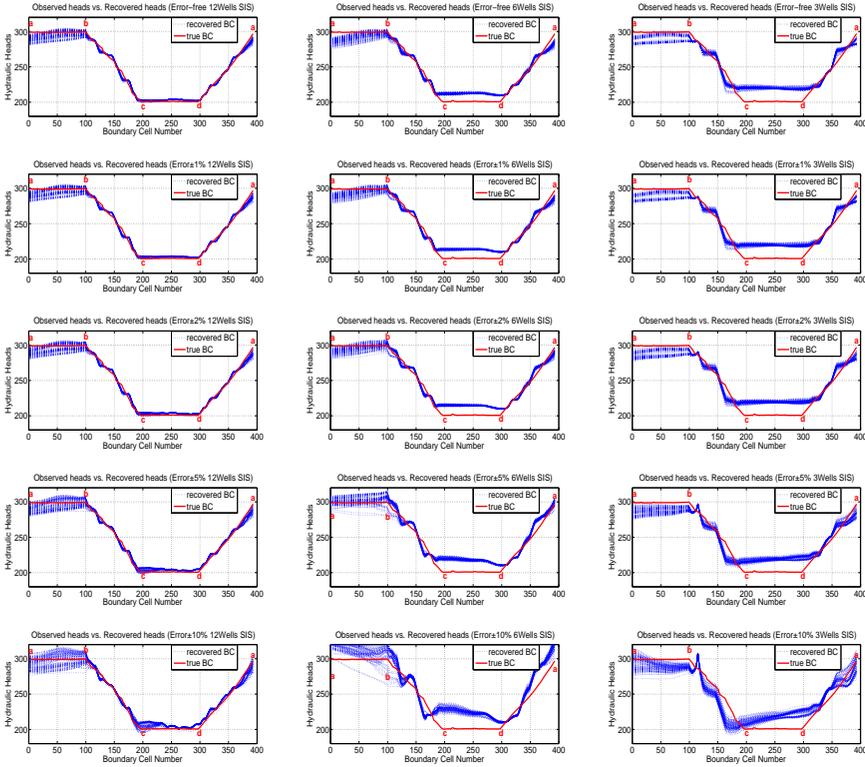
**Fig. 6** Inverted distributions of boundary heads obtained with different sampling densities when measurements were error-free. The true BC are also shown. The bold alphabets denote the boundary locations (Fig. 1).

and  $K2$  under increasing measurement errors. Compared to  $K2$  (clean sand), estimation of  $K1$  (silty sand) degrades more quickly with increasing errors. For the measurement errors of  $0, \pm 1, \pm 2, \pm 5,$  and  $\pm 10\%$  of the total head variation,  $\varepsilon_K$  for  $K1$  is  $0.0, 3.0, 7.0, 29.0, 76.0\%$ , and for  $K2$ , it is  $0.3, 1.3, 2.9, 10.3, 27.4\%$ ;  $\sigma$  for  $K1$  is  $3.8, 4.3, 4.8, 6.5, 7.3$ , and for  $K2$ , it is  $1.2, 1.5, 1.5, 2.2, 2.5\%$ . When measurement errors are low, inversion performance in estimating both conductivities is similar, but when errors grow higher,  $K2$  estimation is more accurate and precise. Thus, conductivity of low- $K$  facies may become more difficult to identify when significant errors exist in measurements. This effect may be attributed to the fact that flow in low- $K$  facies is more sensitive to variation in the hydraulic head and its spatial gradient.

The inverted boundary heads under increasing measurement errors are presented in Fig. 8. When 12 wells were sampled, increasing measurement errors only leads to a slight degradation of the inverted boundary heads. For example, when measurement errors are progressively increased, slightly increased  $\sigma_{BC}$  is observed along the inflow boundary ( $\mathbf{a} - \mathbf{b}$ ). When 6 wells were sampled, this degradation,



**Fig. 7** Distribution of the estimated  $K$  for each facies under increasing measurement errors. (top) inverted  $K1$  histograms; (bottom) inverted  $K2$  histograms. Arrows point to the true values.



**Fig. 8** Recovered boundary head ensembles under decreasing sampling density: (first column) 12 wells, (second column) 6 wells, (third column) 3 wells. In each column, boundary heads are inverted under increasing measurement errors: (first row) 0%, (second row)  $\pm 1\%$ , (third row)  $\pm 2\%$ , (fourth row)  $\pm 5\%$ , (fifth row)  $\pm 10\%$ .

in terms of inversion accuracy and precision, is more pronounced. When 3 wells were sampled, a similar degree of degradation as that of the 6 wells is observed. Overall, for the problem investigated here, BC recovery is not very sensitive to the imposed random measurement errors when a sufficient number of wells (12 wells) were sampled. On the other hand, for a practical situation where measurement errors can be quantified (e.g.,  $\pm 2\%$ ), optimal sample spacing can again be identified by setting acceptable tolerances for  $\varepsilon_{BC}$  and  $\sigma_{BC}$ .

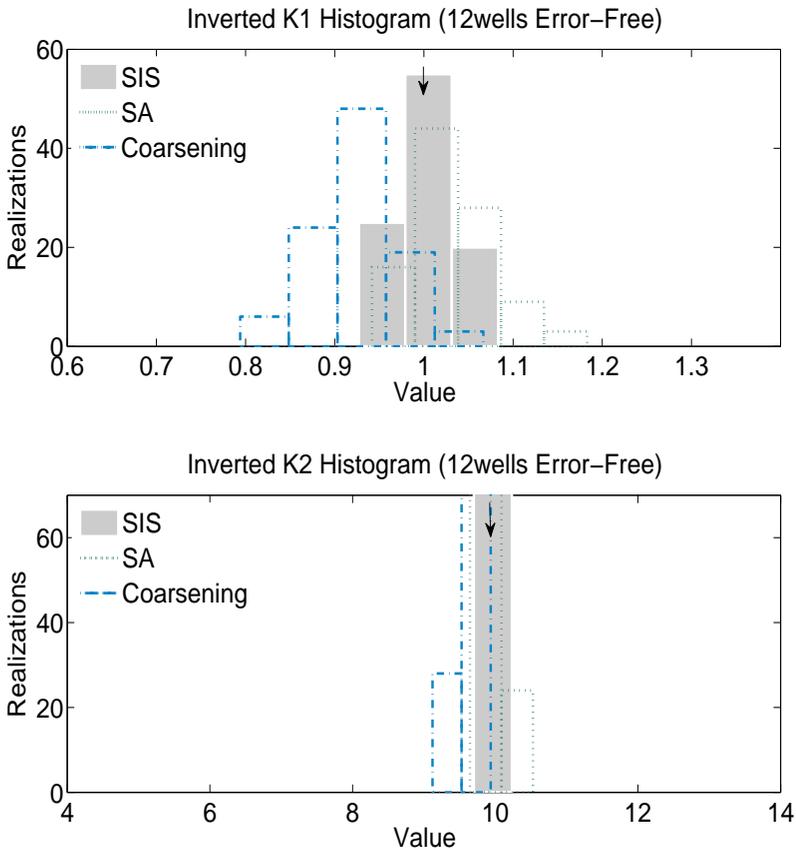
### 3.3 Inverse Parameterization

Given error-free measurements from 12 wells, inversion outcome is presented when facies is parameterized alternatively with SIS, SA, and the coarsened grids (Fig. 9, Fig. 10). When the SIS grid is used, mean conductivity prediction is extremely close to the true value, while  $\sigma_K$  remains very small. When the SA grid is used, both accuracy and precision of the inversion suffer slightly. When the coarsened grid is used, inversion quality is comparably the worst. The average CPU time for solving an inverse problem was 550 s (SIS), 450 s (SA), and 330 s (coarsened), thus speedup due to heterogeneity smoothing and further coarsening is around 18% and

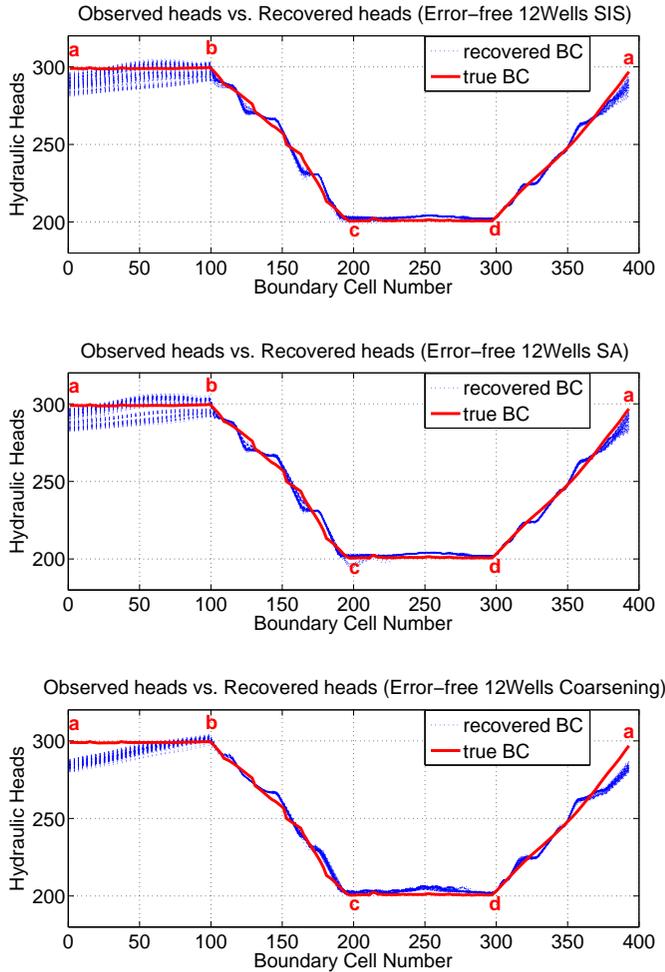
40%, respectively. When BC recovery is compared, however, inversion performance does not vary significantly with the variation in parameterizations (Fig. 10). The above findings suggest that  $K$  estimation is sensitive to its parameterization (i.e., smoothness in heterogeneity), while hydraulic head recovery is not. It is well-known that hydraulic head prediction is not sensitive to small-scale  $K$  variation because the groundwater flow equation acts as a filter: the high wave number components corresponding to small-scale heterogeneity are generally filtered out. Moreover, if a 10%  $K$  estimation error is considered acceptable, inversion using the coarsened grid is optimal, as the speed of solver with this smaller grid is faster by approximately 40% (compared to the SIS grid) and 27% (compared to the SA grid).

### 3.4 A Co-Effect Study

In the previous sections, inversion accuracy and precision were analyzed separately against changing observation quantity, quality, and inverse parameterization. To



**Fig. 9** Distribution of the estimated  $K$  for each facies under different inverse parameterizations. (top) inverted  $K1$  histograms; (bottom) inverted  $K2$  histograms. Arrows point to the true values.



**Fig. 10** Recovered ensembles of boundary heads along “a-b-c-d-a” under different inverse parameterizations: (top) SIS, (middle) SA, (bottom) coarsened grid.

understand the relative importance of each factor to the performance of inversion, a co-effect analysis was conducted. To determine the relative ranking, three separate two-dimensional experiments varying two factors at a time were designed, i.e., data quantity versus parameterization, data quantity versus data quality, data quality versus parameterization. For example, when the co-effect of data quantity and parameterization is of interest, three different parameterizations were employed for a given sampling density, resulting in 3 sets of stochastic inversions (i.e., 300 ensemble solutions). For the 3 sampling densities (12 versus 6 versus 3 wells), 900 inversion equations were thus solved. The inversion outcomes — ensemble means and standard deviation of  $K$  or BC estimation, were then analyzed in light of the joint variation of the two factors. For a given sampling density, this analysis reveals that parameterization with the coarsened grid always leads to less accuracy and lower precision in inversion, as expected. The same can be said when fewer wells are

provided to inversion under a fixed parameterization. To understand the relative importance of each factor on inversion, matrices of the ensemble outcomes can be created and analyzed [25]. For the ranges of tested variation in data quantity and inverse parameterization, data quantity is found to be the dominant factor that influences both the accuracy and precision of  $K$  and BC estimations.

Similar two-dimensional experiments were carried out to evaluate the relative importance of data quantity versus data quality, and data quality versus parameterization (an additional 1,800 ensemble solutions were obtained and analyzed). Detail about these experiments and the co-effect study is provided in [25], and will not be presented. Here, we summarize the insights gained from this analysis. For the ranges of tested variation in data quantity, quality, and inverse parameterization, data quantity plays the dominant role in determining the accuracy and precision of the inverted  $K$ s. Data quality is important to the accuracy of  $K$  inversion, but has limited influences on its precision. Compared to data quantity and quality, the impact of inverse parameterization on the quality of inversion is generally smaller, suggesting the viability of using upscaled parameter fields with coarse grids to achieve a greater computation efficiency. That is, a resolution-accuracy trade-off exists for parameter estimation.

For BC and thus flow field estimation, the most important factor that influences the quality of inversion is data quantity, while data quality and heterogeneity resolution have lesser influences. And, regardless of the combinations examined, regions of the inversion domain near the inflow and outflow boundaries, when compared to the interior regions, always correspond to lower inversion accuracy and precision. This is a result of extrapolation errors as no measurements exist along these boundary segments. This is consistent with the earlier results with deterministic inversion [16].

## 4 Conclusions

In this research, a stochastic subsurface data integration technique is proposed by combining geostatistical simulation with a direct inverse method to facilitate parameter and flow field estimation in steady-state aquifer inversion while quantifying the associated estimation uncertainty. The observation data are obtained from a set of sampling wells, and include static data (i.e., facies types) and dynamic data (i.e., hydraulic heads and flux measurements). The static data are assumed error-free, while increasing measurement errors are imposed onto the observed hydraulic heads. First, sequential indicator simulation is used to integrate the static data, which leads to a set of correlated facies realizations with different hydraulic conductivity distributions. This facies ensemble is then provided to direct inversion which utilizes the dynamic data sampled from the same wells to estimate a set of ensemble parameters and flow fields. With these ensemble inversion outcomes, the mean (expected) parameters and flow fields and their estimation uncertainty can then be determined. Because the above stochastic integration requires ensemble inversions on many facies realizations, computational efficiency is of interest. Via smoothing and grid coarsening, facies parameterization by SIS can be modified by reducing the resolution of the facies model prior to inversion. The smoothing of the facies can improve matrix conditioning in inversion, while smoothing combined

with grid coarsening can lead to a reduced equation size. Both approaches can reduce the computational time needed to solve the inversion systems of equations.

To test the proposed integration technique, a reference forward model provides both the facies characterization and dynamic measurements at increasing sampling densities (i.e., data quantity) and measurement errors (i.e., data quality). Because flow field can be uniquely determined if conductivities and BC are known, the estimated BC ensemble is examined in lieu of the flow fields. Uncertainty in the estimated conductivities and BC is quantified against the reference model to evaluate the quality of inversion. Results suggest that for the ranges of tested variation in data quantity, quality, and inverse parameterization, (1) data quantity has the greatest impact on both inversion accuracy and inversion precision; (2) data quality impacts inversion accuracy; (3) inverse parameterization has the weakest influence on inversion as long as the overall facies pattern is captured (i.e., sufficient data quantity). A balance can thus be achieved between parameterization, computational efficiency, and inversion performance. Moreover, for the heterogeneity investigated herein, by defining an acceptable margin of uncertainty for either conductivity or flow field estimation, optimal well spacing in relation to the characteristic length of heterogeneity can be determined under unknown aquifer BC. Finally, results of this study suggest that inversion domain should be closely defined by the measurement locations in order to minimize extrapolation errors.

Results and insights of this work can lead to field implementation of the proposed stochastic integration technique for problems with unknown information about aquifer's boundary conditions. Along with parameter estimation, the integration technique can explicitly quantify boundary conditions including their mean values and their uncertainty, thus providing a powerful tool for characterizing deep, data-poor environments which are increasingly being exploited for waste disposal (e.g., geological carbon storage and sequestration) as well as for energy production (e.g., hydraulic fracking, geothermal, oil/gas) operations. Using the integration technique, by assuming a reasonable level of measurement errors in the hydrogeological data, well spacing ranging from 1/3 to 1/6 of the lateral facies correlation length can lead to adequate inversion outcomes with acceptable estimation uncertainties. Future work will extend the methods of this study to three-dimensional problems, while static data integration will incorporate not only direct borehole information but other auxiliary data (e.g., seismic facies analysis) that can yield additional information of subsurface heterogeneity in between wells.

**Acknowledgements** This work is supported by the University of Wyoming School of Energy Resources Center for Fundamentals of Subsurface Flow (WYDEQ49811ZHNG), and NSF CI-WATER (Cyberinfrastructure to Advance High Performance Water Resource Modeling), and NSF EPSCoR (EPS 1208909).

## References

1. Richard L. Chambers, Jeffrey M. Yarus, and Kirk B. Hird. Petroleum Geostatistics for Nongeostatisticians. *The Leading Edge*, 19(5):474–479, 2000.
2. C. Deutsch and P. Cockerham. Practical considerations in the application of simulated annealing to stochastic simulation. *Mathematical Geology*, 26(1):67–82, 1994.

3. C. V. Deutsch and A. G. Journel. *GSLIB: Geostatistical Software Library and User's Guide, Applied Geostatistics Series*. Oxford University Press, New York, NY 10016, 2 edition, 1997.
4. Banta E.R. Hill M.C. Harbaugh, A.W. and M.G. McDonald. *MODFLOW-2000, the U.S. Geological Survey modular ground-water model – User guide to modularization concepts and the Ground-Water Flow Process*. U.S. Geological Survey, Reston, Virginia, open-file report 00-92, 121 p. edition, 2000.
5. M. C. Hill and C. R. Tiedeman. *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*. Wiley-Interscience, Hoboken, N. J., 2007.
6. J. Irsa and A. N. Galybin. Stress Trajectories Element Method for Stress Determination from Discrete Data on Principal Directions. *Eng. Anal. Boundary Elem.*, 34:423432, 2010.
7. Juraj Irsa and Ye Zhang. A Direct Method of Parameter Estimation for Steady State Flow in Heterogeneous Aquifers with Unknown Boundary Conditions. *Water Resour. Res.*, 48:W09526, 2012.
8. J. Jiao and Y. Zhang. A method based on Local Approximate Solution (LAS) for Inverting Transient Flow in Heterogeneous Aquifers. *Journal of Hydrology*, 514:145–149, 2014.
9. J. Jiao and Y. Zhang. Two-Dimensional Inversion of Confined and Unconfined Aquifers under Unknown Boundary Conditions. *Advances in Water Resources*, 65:43–57, 2014.
10. D. Kleinecke. Use of Linear Programming for Estimating Geohydrologic Parameters of Groundwater Basins. *Water Resour. Res.*, 7(2):367374, 1971.
11. R. W. Nelson. In Place Measurement of Permeability in Heterogeneous Media: 1. Theory of a Proposed Method. *J. Geophys. Res.*, 65:17531760, 1960.
12. R. W. Nelson. In Place Measurement of Permeability in Heterogeneous Media: 2. Experimental and Computational Considerations. *J. Geophys. Res.*, 66:24692478, 1961.
13. R. W. Nelson. In Place Determination of Permeability Distribution of Heterogeneous Porous Media through Analysis of Energy Dissipation. *Soc. Pet. Eng. J.*, 3:3342, 1968.
14. S. P. Neuman, A. Blattstein, M. Riva, D. M. Tartakovsky, A. Gaudagnini, and T. Ptak. Type curve interpretation of late-time pumping test data in randomly heterogeneous aquifers. *Water Resources Research*, 43:W10421, 2007.
15. C. C. Paige and M. A. Saunders. Lsq: An Algorithm for Sparse Linear Equations and Sparse Least Squares. *Trans. Math. Software*, 8(1):43–71, 1982.
16. V. E. A. Post and J. R. von Asmuth. Review: hydraulic head measurements – new technologies, classic pitfalls. *Hydrogeology Journal*, 21:737–750, 2013.
17. A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis: The Primer*. Wiley.
18. X. Sanchez-Vila, J. Carrera, and J. Girardi. Scale effects in transmissivity. *Journal of Hydrology*, 183:1–22, 1996.
19. V. Shevnin, O. Delgado-Rodriguez, A. Mousatov, and A. Ryjov. Estimation of Hydraulic Conductivity on Clay Content in Soil Determined from Resistivity Data. *Geofisica Internacional*, 45(3):195–207, 2006.
20. G. W. Stewart and J. G. Sun. *Matrix Perturbation Theory*. Academic Press, 374 Pages, 1990.
21. N.-Z. Sun. *Inverse Problems in Groundwater Modeling, Theory Appl. Transp. Porous Media, vol. 6*. Kluwer Acad., Dordrecht, Netherlands, 1994.
22. D. Wang. Physically based stochastic inversion and parameter uncertainty assessment on a confined aquifer with a highly scalable parallel solver. Master's thesis, University of Wyoming, 2014.
23. W. W.-G. Yeh, Y. S. Yoon, and K. S. Lee. Aquifer Parameter Identification with Kriging and Optimum Parametrization. *Water Resour. Res.*, 19(1):225233, 1983.
24. Y. Zhang. Nonlinear Inversion of an Unconfined Aquifer: Simultaneous Estimation of Heterogeneous Hydraulic Conductivities, Recharge Rates, and Boundary Conditions. *Transport in Porous Media*, 102:275–299, 2014.
25. Y. Zhang, J. Irsa, and J. Jiao. Three-Dimensional Aquifer Inversion under Unknown Boundary Conditions. *Journal of Hydrology*, 509:416–429, 2014.
26. Y. Zhang, B. Liu, and C. W. Gable. Homogenization of Hydraulic Conductivity for Hierarchical Sedimentary Deposits at Multiple Scales. *Transport in Porous Media*, 87(3):717–737, 2011.
27. Y. Zhang, M. Person, and C. W. Gable. Representative Hydraulic Conductivity of Hydrogeologic Units: Insights from an Experimental Stratigraphy. *Journal of Hydrology*, 339:65–78, 2007.

28. V. A. Zlotnik, B. R. Zurbuchen, T. Ptak, and G. Teutsch. Support volume and scale effect in hydraulic conductivity: experimental aspects. *Geological Society of American Special Paper*, 348:215–231, 2000.