Groundwater

Issue Paper/

Reducing Uncertainty in Calibrating Aquifer Flow Model with Multiple Scales of Heterogeneity

by Ye Zhang

Abstract

Modeling and calibration of natural aquifers with multiple scales of heterogeneity is a challenging task due to limited subsurface access. While computer modeling plays an essential role in aquifer studies, large uncertainty exists in developing a conceptual model of an aquifer and in calibrating the model for decision making. Due to uncertainties such as a lack of understanding of subsurface processes and a lack of techniques to parameterize the subsurface environment (including hydraulic conductivity, source/sink rate, and aquifer boundary conditions), existing aquifer models often suffer nonuniqueness in calibration, leading to poor predictive capability. A robust calibration methodology is needed that can address the simultaneous estimations of aquifer parameters, source/sink, and boundary conditions. In this paper, we propose a multistage and multiscale approach that addresses subsurface heterogeneity at multiple scales, while reducing uncertainty in estimating the model parameters and model boundary conditions. The key to this approach lies in the appropriate development, verification, and synthesis of existing and new techniques of static and dynamic data integration. In particular, based on a given set of observation data, new inversion techniques can be first used to estimate aquifer large-scale effective parameters and smoothed boundary conditions, based on which parameter and boundary condition estimation can be refined at increasing detail using standard or highly parameterized estimation techniques.

Introduction

While computer modeling plays an essential role in hydrogeology, due to uncertainties in describing aquifer parameters, flow and transport processes, and the associated initial and boundary conditions (BC), developing and calibrating a predictive aquifer model is challenging. This issue occurs whenever real aquifers are modeled with incomplete knowledge of system property, state, and dynamics. Reducing all sources of model uncertainty, however, is difficult to accomplish in practice, due to the cost of drilling, sampling, conducting aquifer tests, and the difficulty in portraying/forseeing future aquifer forcings (Hunt and Welter 2010). A balance between the cost of developing a model and model accuracy is clearly desired and there exists a need to (1) identify optimal complexity in representing aquifer processes and parameters for different prediction goals; (2) determine parameters and BC for the optimal model(s) accurately.

To address issues related to (1), hypotheses related to parameter (and process) resolution must be tested. For example, if we wish to predict average aquifer flow, will a model with lower hydraulic conductivity resolution be sufficient, ignoring smaller scale parameter heterogeneities? Our research on hydraulic conductivity upscaling suggests that optimal resolution in representing aquifer heterogeneity likely exists (Zhang et al. 2006). To address issues related to (2), robust and efficient model calibration (or inversion) techniques are needed, which can be facilitated by a new steady-state inverse method that simultaneously estimates model parameters, model source/sink rates, and model BC.

In this paper, a promising new direction in hydrogeological modeling is proposed in the form of a multistage and multiscale model building and calibration approach, which aims to account for subsurface heterogeneity at

Department of Geology & Geophysics, University of Wyoming, Laramie, WY 82071; 307-766-2981; fax: 307-766-6679; yzhang9@uwyo.edu

Received February 2013, accepted July 2013.

doi: 10.1111/gwat.12111

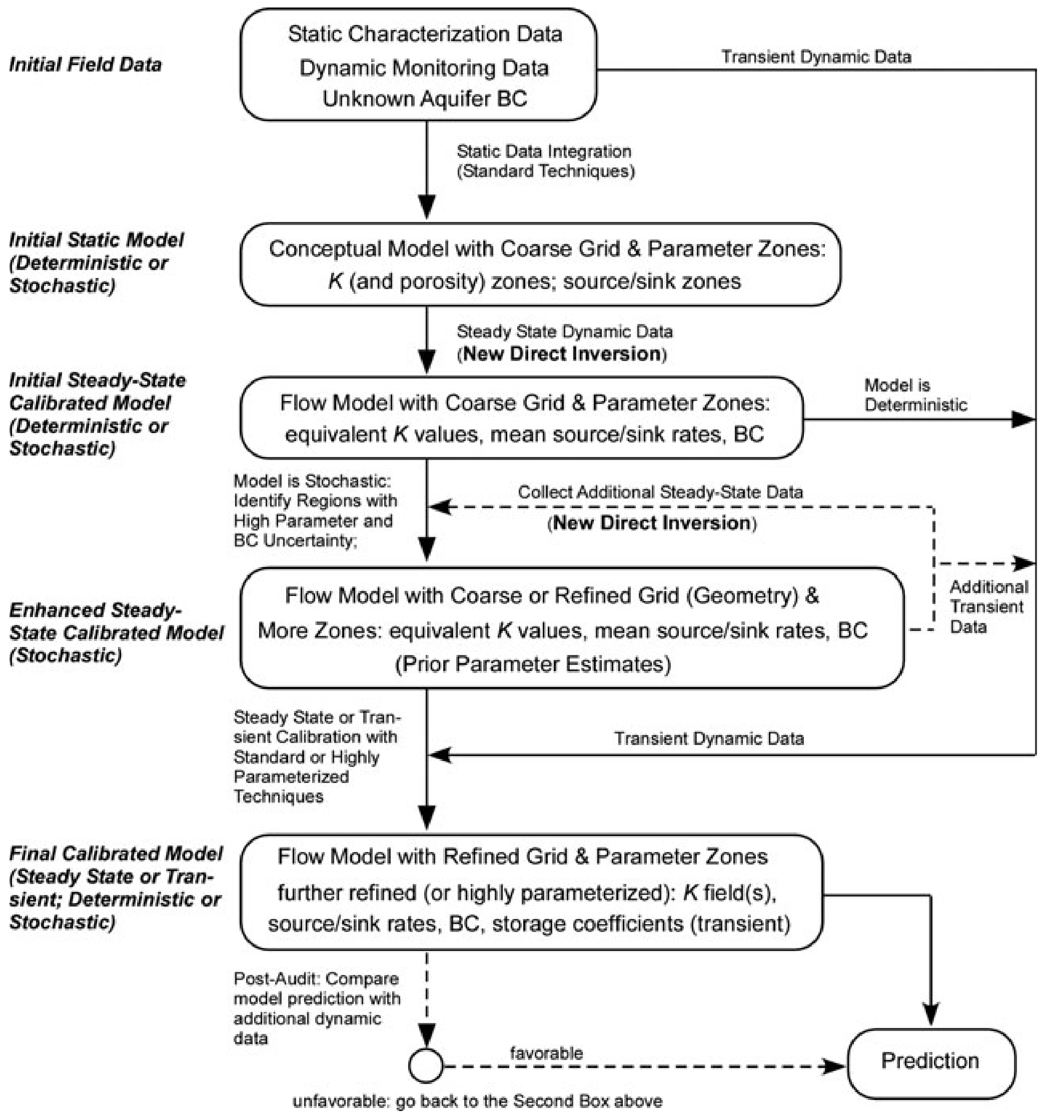

Figure 1. A proposed multistage and multiscale calibration method for groundwater flow modeling.

multiple scales, while reducing uncertainty in estimating model parameters and model BC. The key to this approach lies in the appropriate development, verification, and synthesis of the existing and new techniques of static and dynamic data integration. Specifically, for a given set of observation data, direct inversion techniques can be used to estimate long-term (steady-state) aquifer large-scale equivalent conductivities, average source/sink rates, and (smoothed) BC. Based on these initial but lower resolution estimates, grid, parameters, and BC can be refined at increasing detail using standard parameter estimation techniques (Figure 1).

In this paper, groundwater flow modeling is the focus and problems associated with process uncertainty (i.e., transport modeling) are not addressed. In the following sections, the issues related to model complexity are first presented, before the issue of nonuniqueness in model calibration is discussed. Recent research on developing a new steady state inverse method is summarized, which together with a calibration exercise estimating hydraulic conductivities for multiple conceptual models, supports the proposed multistage and multiscale calibration approach.

Modeling and Model Complexity

In developing a hydrogeological site model, initial model is almost always constructed using limited data, including static site characterization data (i.e., aquifer geometry, internal structure, porosity, hydraulic conductivity) and dynamic monitoring data (i.e., water levels, pumping rates, or recharge estimates). In general, a simple conceptual framework model, or a hydrostratigraphic model, is built first, integrating all available static data. It is then calibrated based on the available site dynamic data, either by trial-and-error, or with the aid of a parameter estimation algorithm. At this stage, the worth and limitation of the data can be explored within a formal sensitivity analysis and inversion framework, which can yield uncertainty measures of the initial set of estimated parameters, while pointing to the type and location of additional data for collection (Hill and Tiedeman 2007). Such data, often with enhanced information content that can be used to inform parameter estimation, can then lead to the reduction of various model errors and uncertainties (Carrera and Neuman 1986; Poeter and Hill 1997; Saiers et al. 2004; Tiedeman et al. 2003). During this process, in addition to the direct hydraulic data, auxiliary data such as temperature, tracer/isotope concentrations, and geophysical measurements can provide complementary or qualitative verifications. This process then leads to a more refined model, which can be subject to further testing, verification, and refinement. Ideally, as additional relevant data are collected, greater insights are gained on the system behavior, and the model will become a more accurate representation of the reality. This workflow is often utilized in analyzing aquifer problems, although room for improvement and optimization exists, particularly on the issue of parameter resolution and model complexity, as discussed below.

In aquifer modeling, the need to balance parameterization complexity with study objectives and quality of the data has been recognized by previous authors (Hunt and Zheng 1999). Using a highly parameterized inversion approach integrating many types of data, a smooth and parsimonious hydraulic conductivity (K) field is obtained by Fienen et al. (2009) in delineating aquifer flow paths. Research on K up-scaling suggests that if realistic and sufficient geological resolution is incorporated, hydrostratigraphic models with equivalent conductivities can capture the overall flow field of a heterogeneous synthetic aquifer (reference model), that is, large-scale head distribution, bulk flow velocity, and flow paths (Zhang et al. 2006, 2011). The level of complexity in resolving aquifer heterogeneity is also determined by the prediction goal, the hydraulic BC, and the level of acceptable modeling error. For example, as demonstrated by Zhang et al. (2006), under a set of BC that drives steady-state vertical flow, if the aquifer mean flow path is the prediction goal, a 2-unit hydrostratigraphic model does an equally good job of capturing the flow paths of the reference model as a 14-unit hydrostratigraphic model (see Figure 9 in Zhang et al. 2006; also, Figure 3 herein). The 2-unit model is thus of sufficient complexity for making this prediction. Moreover, if the length of the no-flow boundary is extended, both hydrostratigraphic models (2- vs. 14-unit) are nearly equally accurate in predicting the true hydraulic head field (see Figure 14 in Zhang et al. 2006). When the length of the no-flow boundary is reduced, however, mean relative error in head prediction increases to 9% for the 2-unit model, but stays low for the 14-unit model at around 3%. Therefore, if a head prediction error of 10% is considered an acceptable level of modeling error, the 2-unit model will be of sufficient complexity; if 5% is acceptable, the 14-unit model is needed.

The above discussion suggests that optimal complexities in describing aquifer parameters for meeting specific prediction goals likely exist, although words of caution are needed in translating these insights modeling the synthetic system to real-world practice: *To model a real aquifer, how can we obtain equivalent conductivities without using detailed measurements that are required by most upscaling methods?* Moreover, the hydrostratigraphic models reported in the previous studies are driven with the true BC of the reference model. *In the real world, even if we are able to estimate equivalent conductivities accurately, how can we obtain accurate estimation of the aquifer BC*? These questions lead to our second issue: how can optimal flow models be built with reduced uncertainties in model parameters and model BC? To address this, an appropriate integration of the existing and new inverse methodologies will likely provide a way forward. How to determine aquifer BC is an integral element of this integration, which is discussed in detail in the following section.

Model Calibration and Nonuniqueness

Most of the existing parameter estimation methods assume that aquifer BC is either known or can be determined from model calibration. However, due to limited subsurface access, BC is often unknown at most field sites, while BC calibration may lead to nonuniqueness in the estimated parameters, BC, and flow field (Irsa and Zhang 2012). The nonuniqueness issue can be explained using a simple example. Detailed mathematical proof was given in Irsa and Zhang (2012) and is not presented here. Suppose that from a two-dimensional homogeneous isotropic steady-state aquifer we sampled three hydraulic heads (dots in Figure 2a) and one flow rate along the right hand side of the model, which can be obtained from, for example, separating stream hydrographs. Given these observed data, we can write a simple weighted least-square objective function $S(\mathbf{b})$ describing the model-data mismatch (modified after Equation 3.1a in Hill and Tiedeman 2007):

$$S(\mathbf{b}) = \sum_{i=1}^{NH} \omega_{h_i} \left[y_{h_i} - y'_{h_i}(\mathbf{b}) \right]^2 + \sum_{i=1}^{NQ} \omega_{q_j} \left[y_{q_j} - y'_{q_j}(\mathbf{b}) \right]^2 \quad (1)$$

where $\mathbf{b}$ is the parameter vector to be estimated (e.g., hydraulic conductivity here for the steady-state flow problem), y_{h_i} is the ith observed hydraulic head being matched, $y'_{h_i}(\mathbf{b})$ is the model simulated hydraulic head that corresponds to the ith observed head; y_{q_j} is the jth observed flow rate being matched, $y'_{q_j}(\mathbf{b})$ is the model simulated flow rate that corresponds to the jth observed flow rate, NH is the number of hydraulic head observations, and NQ is the number of flow rate observations, ω_{h_i} is a weight assigned to the ith head observation, ω_{q_j} is a weight assigned to the jth flow rate observation. The goal of most parameter estimation methods is to minimize $S(\mathbf{b})$ using optimization techniques, which minimizes the model-data mismatch. With these techniques, parameters are updated using repeated simulations of the forward flow model, which requires the specification of the model BC. For this simple problem, say we decide on a set of

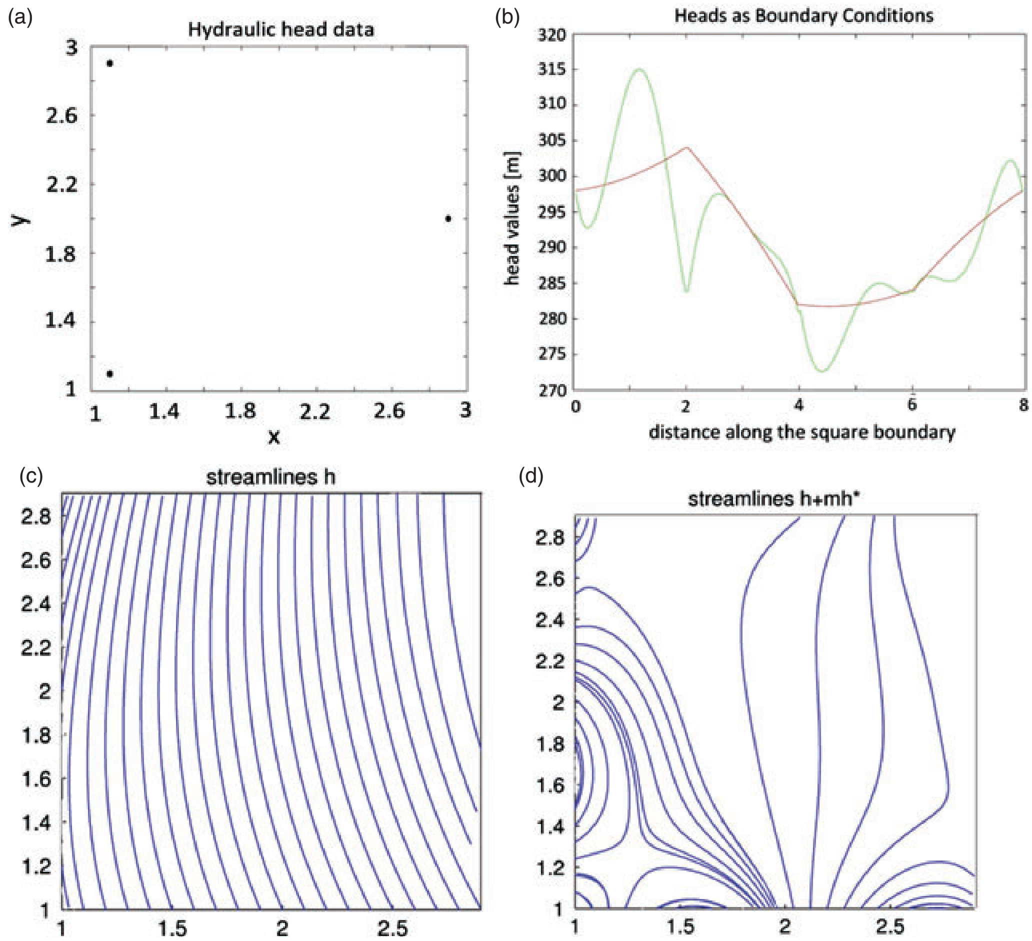

Figure 2. Nonuniqueness in fitting model BC in a homogeneous and isotropic aquifer within a square region [1,1]-[3,3] (reprinted from Irsa and Zhang 2012). (a) Observation data: three heads and one integrated flow rate along the right boundary, (b) two different Dirichlet BC, (c) and (d) streamlines driven by each BC.

specified-head BC that vary in a nearly linear manner along the model boundary (the red curve in Figure 2b). Given these BC, and knowing the hydraulic conductivity, we can analytically generate the flow field (Figure 2c), which honors both the three observed heads and the one observed flow rate. On the other hand, we can use a parameter estimation technique to estimate K, specifying the same BC to the forward model (i.e., the flow problem of Figure 2a discretized and solved with a numerical technique), which is required for the regression iterations. Assuming no measurement errors, the parameter estimation technique will exactly recover the three observed heads, one observed flow rate, and the state variables, for example, the streamlines as shown in Figure 2c. At this point, $S(\mathbf{b})$ is optimized to be zero and the aquifer model is perfectly calibrated for the given set of BC. (A parameter estimation code, if correctly set up, and barring numerical issues that can arise during regression iterations, should ideally recover the analytical solution.) Next, the exercise is repeated by assuming a different set of specified head BC, for example, the green curve in Figure 2b. Given the new BC, for the same K value, we can generate another analytical flow solution (Figure 2d), which also honors the same three observed heads and the one observed flow rate. Again, K can be calibrated given the new BC, which will recover the analytical solution when the regression converges at $S(\mathbf{b}) = 0$. Another aquifer model is thus perfectly calibrated for the new BC. If we perturb any of the two sets of the BC and repeat this exercise ad infinitum, each time a new BC is assumed, a different flow field will be perfectly calibrated to the same observation data. It has been proven that there exists an infinite number of BCs and flow fields that all satisfy the same observed data (Irsa and Zhang 2012), leading to zero $S(\mathbf{b})$. Therefore, nonuniqueness in the inversion outcomes can arise due to the unknown model BC.

The above discussion suggests that multiple sets of calibrated parameters and BC may equally satisfy the model calibration criteria (i.e., minimization of an objective function), which may lead to a wrong model being used for prediction and management purposes. In natural aquifers, true BC are typically poorly known. If we make a wrong assumption about the BC, we may obtain a perfectly calibrated flow model using an objective-function-based parameter estimation technique, even if error-free observed data are used and even if there are no model structure errors (in the above example, inverse parameterization is given the true parameterization of the forward model). Indeed, an example of the influence of BC in model calibration is provided in Hunt et al. (1998), where for an aquifer-lake system in Wisconsin, earlier calibration under a set of assumed BC yielded 9 best-fit K zones and 94 best-fit lake-bed conductances. However, a revision of the BC by Hunt et al. (1998) yielded a better calibrated model with only 2 best-fit K zones and 18 lake-bed conductances. Although highly parameterized techniques have been developed to delineate aquifer heterogeneity at greater detail (Zhu and Yeh 2006), nonuniqueness in parameter estimation may also arise (Bohling and Butler 2010).

To address the nonuniqueness issue, a new steady-state direct inversion method was developed to simultaneously estimate hydraulic conductivity, state variables, and BC of a confined aquifer, under the condition where source/sink to the aquifer was negligible (Irsa and Zhang 2012). Unlike the objective-function based approaches, this method does not require forward flow simulations to assess the model-data mismatch, thus knowledge of the BC is not needed. Given sufficient measurement data, the method yielded well-posed systems of equations that can be solved efficiently with linear optimization. The solution was also stable when measurement errors were increased. However, only one K was estimated (the other conductivities were obtained from known conductivity ratios in the form of prior information equations), while source/sink effects were not accounted for (Irsa and Zhang 2012). As a result, groundwater flux or flow rate observation(s) must be sampled from the subsurface, which limits the applicability of the method. Recently, for both confined and unconfined aquifers, the method has been extended to simultaneously estimate a number of conductivities (Ks) and recharge (or leakage) rates, along with the unknown model BC (Zhang 2013; Zhang et al. 2013). Interestingly, a single pumping rate, in addition to hydraulic heads, suffices to provide the necessary measurement for inverting multiple Ks and multiple recharge rates.

Because of subsurface uncertainty, the inverse method must also be able to handle model structure errors, that is, when inverse parameterizations do not reflect the true conditions in the aquifer. With the new method, if the conductivity and recharge variabilities are unknown, inversion yields physically meaningful equivalent conductivities and average recharge rates. Alternatively, if the inverse parameterization contains spurious parameters, the new method can identify such parameters, while the simultaneous estimation of nonspurious parameters is not affected. The new method thus obviates the well-known issues associated with model "structure error," whereas the inverse formulation either simplifies or complexifies the true parameter fields. Moreover, to investigate uncertainty in inversion, the method can be combined with indicator geostatistics. Uncertainty in the static data (i.e., hydrofacies proportion, variogram, and correlation ranges) can be propagated into the inversion outcomes—a set of realizations of model parameters, flow fields, and BC can be created (Wang et al. 2013). These realizations center on the "true" solution created from an underlying "true" model (i.e., a forward simulation with known parameters and BC), while increased sampling of the static and dynamic data leads to reduced spread in the estimated parameters, flow fields, and BC. Therefore, not only can uncertainty in K pattern be accounted for, issues of data worth and sampling density can be addressed.

The new inverse method is based on a zoned parameterization scheme, that is, either a deterministic pattern or a stochastic ensemble. The estimated parameters and boundary conditions therefore reflect large-scale spatial averages that are smoothed over local variations. While we plan to extend the method to highly parameterized inversion to account for small scale K variation and its spatial correlation, the inverse method in its current form is more suitable for inverting low-resolution conceptual hydrostratigraphic models. Also, transient data, which can provide additional information for inversion, cannot yet be incorporated into the analysis, while both standard regression techniques and a variety of data assimilation methods, for example, transient hydraulic tomography or the methods based on the Kalman filter, can utilize such data. (Extension to transient flow is currently being researched, whereas the inversion of drawdowns may be utilized to remove the need to estimate the unknown aquifer initial conditions.) Finally, objective-function based parameter estimation techniques have diversified into handling many types of flow problems while utilizing a variety of observation data, for example, steady-state or transient hydraulic data, solute concentrations, temperature, water chemistry, isotopes, etc. Spurred by the growth of computing power, improvement in numerical algorithms and solution techniques, and the increased awareness of subsurface uncertainty, these techniques can now address problems with a large number of parameters and both parameter and prediction uncertainties can be assessed with flexibility and efficiency (Liu and Kitanidis 2011; Wen et al. 2002; Zimmerman et al. 1998). Looking forward, can we benefit from the strengths of all these approaches?

A Multistage and Multiscale Calibration

Given the state-of-practice in aquifer modeling and the recent advancements in parameter scaling and model inversion, a fruitful direction for further development may lie in a multistage and multiscale integration that can address subsurface heterogeneity at multiple scales, while reducing uncertainty in estimating model parameters and model BC. As illustrated in Figure 1, direct inversion techniques can be first used to estimate long-term (steady-state) aquifer large-scale effective parameters and smoothed BC, based on which parameter and boundary condition estimation can be refined at increasing detail using the objective-function-based techniques. The proposed approach complements the prevalent modeling workflows: either deterministic or stochastic element can be incorporated, with the later accounting for uncertainties in parameter and BC estimation. This approach is also consistent with the idea, as proposed by previous workers (Anderson and Woessner 1992; Hunt et al. 1998), of developing a low-resolution "screening model" prior to building complex and highly resolved models.

The multistage and multiscale integration can be supported by the previous research. For example, as revealed by our earlier upscaling studies, as long as the model BC is close to the true BC, equivalent conductivities can lead to accurate bulk flow predictions by the hydrostratigraphic models. The new inverse method, by estimating both large-scale parameters and smoothed BC, can serve just this purpose. However, accuracy of the inversion is limited by the type, location, quality, and quantity of the measurement data. For the synthetic problems we tested so far, given sufficient data support leading to well-posed inversion problems, the estimated parameters fall within one order (often much better) of the true values. For real-world problems, because measurements will never be exhaustive nor error-free, inversion can only provide physically reasonable but nonexact estimates. One way to remediate this issue is to account for parameter, flow field, and BC uncertainties in inversion, for example, via a Monte Carlo analysis (Wang et al. 2013). For example, histogram of an estimated parameter may suggest a reasonable accuracy (small spread) or otherwise. A variance map of the inverted hydraulic head ensemble can indicate locations where additional sampling is needed to reduce the largest hydraulic head estimation uncertainty. This identification, along with the new data collected to reduce such uncertainty, should lead to a revision/update of the earlier inversion (see the loop in Figure 1). In areas of low uncertainty (of both the parameters and head fields), local analysis with a refined grid or parameterization can be carried out depending on the study objective, for example, solute transport modeling typically requires more resolved K fields.

The new inverse method, as it stands, could provide a low-cost, low-resolution aquifer characterization tool with which initial conceptual models can be built with coarsened representations of parameter heterogeneity (both of Ks and source/sink rates). Based on the same observed data, such models can be refined using either standard or highly parameterized techniques that can account for sub-hydrostratigraphic, smoothly varying, and possibly correlated heterogeneity, for example, pilot point or geostatistical inverse techniques. The estimated Ks, recharge rates, as well as the smoothed BC of the low-resolution models can provide the prior estimates for these refinement studies; their values are considered physically reasonable due to the fact that *they're inferred from the conservation of mass and flux principles as enforced by the initial inversion with the steady-state data*. With the physically reasonable BC, prior Ks, and source/sink estimates, inversion based on optimizing objective functions will become more well-posed, which then lead to physically reasonable parameter estimates and/or refinements. A multiscale calibration exercise supports this view and is presented in the following section.

An Illustrative Example

To support the view that multistage calibration can yield well-posed problems leading to better estimated parameters, a calibration exercise is presented using a suite of models with decreasing spatial K resolutions: a fully heterogeneous model and three hydrostratigraphic models with increasingly fewer K zones (Figure 3). All models employ the same high-density grid containing 845,298 finite element cells (detail on model creation can be found in Zhang et al. 2006). The heterogeneous model is considered a reference model with which steady-state groundwater flow is simulated under a set of *true* BC, that is, no-flow along the model sides and bottom and a specified sloping potentiometric head along the model top creating confined conditions. From this model, a set of simulated observation data is obtained at five monitoring wells: 13 hydraulic heads and 5 Darcy fluxes (Figure 3a). To represent measurement uncertainty, random errors with a coefficient of variation of 10% are added to the "observed" data, assuming no spatial correlation among them, that is, error covariance matrix is diagonal.

Under the condition that the true BC is known, the three hydrostratigraphic models are calibrated based on the same dataset using nonlinear regression as implemented in UCODE and PEST (both of the standard version and PEST SVD) (Doherty 2005; Poeter et al. 2005). During regression iterations, all forward simulations with these models are therefore driven by the true BC. A weighted least-squares objective function is used, which consists solely of data misfit terms without imposing regularization nor prior information equations. The optimization algorithms, as implemented in UCODE and PEST, are gradient-based local methods. The horizontal conductivity (K_x) of each unit of the hydrostratigraphic models is set as the calibration target. Prior to calibration, a parameter sensitivity analysis is conducted to identify the units for which K_x can be estimated for the given set of the observation data, following the guidelines of Hill and Tiedeman (2007). For the starting parameter guess in inversion, equivalent K_x computed for the same unit is used (see detail on its calculation in Zhang et al. 2006). Both UCODE and PEST are compiled into 64-bit executables on a Linux computer cluster. To enhance computational efficiency, a serial iterative solver (Bramley and Wang 1995) is implemented in the forward flow models, its results are verified by a direct Gaussian Elimination solver. The iterative solver sped up the simulation time of the forward model by approximately 60 times. For most inversion runs, both UCODE and PEST converged overnight or within a day.

Results of this exercise are presented in Table 1, where both the calibrated K_x and the equivalent K_x are listed for each hydrostratigraphic model. Several observations can be made:

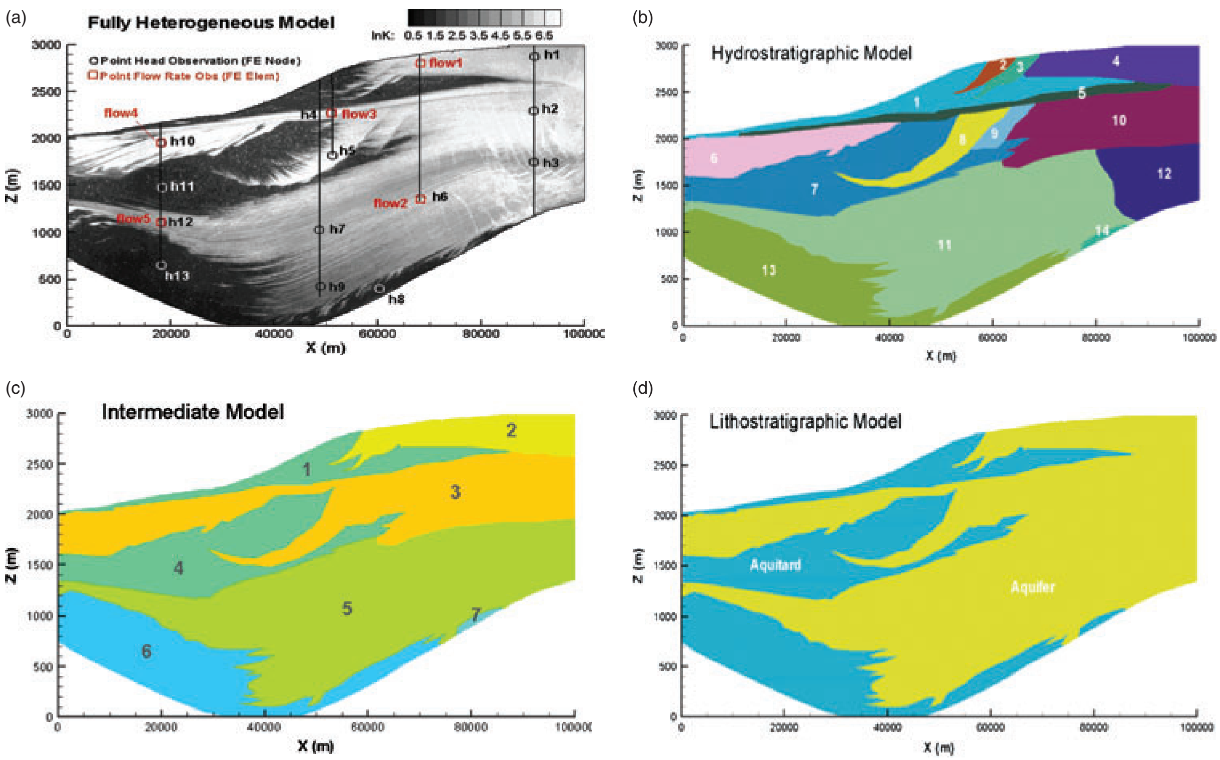

Figure 3. (a) A fully heterogeneous groundtruth model with *lnK* variation in m/year. Steady-state groundwater simulation is conducted under a BC of no-flow along the model sides and bottom and a specified sloping potentiometric head along the model top. Observation data (13 head and 5 Darcy flux measurements) are sampled at five well locations. (b) A 14-unit model. Color represents the unit ID. (c) A 7-unit model. (d) A 2-unit model. Models b, c, and d are examples of alternative conceptual models developed with decreasing amount of site characterization data. They are calibrated using nonlinear regression based on the simulated observations of the groundtruth model.

Table 1
Steady-State Calibration Outcomes of K_x for the Units of the Three Hydrostratigraphic Models

Model ID	Unit ID	Equivalent K_x	Calibrated K_x (UCODE)	Calibrated K_x (PEST classic)	Calibrated K_x (PEST SVD)
Model b	4	330.5	345.1	349.5	300.2
	5	554.6	758.0	712.1	601.3
	6	871.0	942.8	909.2	801.0
	8	583.8	626.2	625.4	501.8
	10	591.6	639.8	653.2	501.9
	11	140.4	133.2	130.3	199.8
Model c	2	305.0	308.7	308.4	305.0
	3	632.4	157.9	158.0	632.4
	5	162.7	127.8	127.6	162.7
Model d	Aquifer	324.6	121.0	120.9	NA

The equivalent K_x of the same units, previously computed in Zhang et al. (2006), are listed. Conductivity unit is m/year. Location of these units and their IDs are shown in Figure 3. Before calibration, a parameter sensitivity analysis was conducted to identify the units for which K_x can be estimated for the given set of the observation data (Figure 3a).

1. UCODE and PEST (classic) give consistent results in the estimated K_x values;
2. For the chosen observation data, most of the insensitive parameters belong to low-K units and are not estimated by regression, indicating potential difficulty in estimating K for aquitards;
3. For the 14-unit model (Figure 3b), regression yields realistic K_x that fluctuates around the equivalent K_x;
4. For the 7-unit model (Figure 3c), K_x estimated by UCODE and PEST (classic) are less accurate, although PEST SVD converges at the starting parameter values—the equivalent K_x;
5. For the 2-unit model (Figure 3d), UCODE and PEST (classic) underestimates the aquifer K_x, while PEST SVD fails to converge.

Clearly, parameter estimation is more satisfactory for the 14-unit model than for the 7- and 2-unit models. In this case, an enhanced level of heterogeneity resolution is likely key to the success of inversion. It is of interest to note that PEST SVD failed to calibrate the 2-unit model, which suggests that this particular technique may be able to identify model structure deficiency that the others fail to reveal. However, without a systematic analysis testing different observation data density and location, and possibly, flow directions (i.e., different true BC), this finding could be coincidental.

Importantly, the above example illustrates that even with limited measurement data that are corrupted by noise (Figure 3a), *if the correct BC and good initial* K_x *(i.e., equivalent* K_x*) are provided*, standard parameter estimation techniques can yield physically reasonable K_x estimates for the hydrostratigraphic model, as long as the model incorporates a sufficient level of K heterogeneity (e.g., the 7-unit model). As demonstrated by the earlier upscaling study (Zhang et al. 2006), these simpler models can create reasonable to good representations of the flow field of the heterogeneous reference model. Moreover, success with the standard methods here suggests that highly parameterized techniques, or those that employ transient data (i.e., data fusion), may also build upon these prior models with coarse parameter resolutions. Such methods can populate finer scale correlated parameters within individual hydrostratigraphic units, smaller scale parameter variation can therefore be captured by analyzing the same data. This problem will be explored in the future.

In modeling real aquifers, a similar multiscale calibration can be carried out. Based on limited field data, the direct inversion method can provide a set of large-scale parameter and BC estimates for an initial model with low resolutions, for example, the 2-unit model in the above example. This model (or a suite of models if the direct method is combined with geostatistical simulations), can be subject to additional calibration with UCODE, PEST, or other techniques. In the new calibration, model or data errors may be revealed using fit-dependent statistics, whereas transient data, if available, can be calibrated to help reveal additional parameterization details. This combined analysis may point to deficiency in the model structure (e.g., inversion fails to converge), or if such deficiency is absent, point out locations where the estimation uncertainty is large and where additional static or dynamic data should be collected should resources become available. If additional sampling is carried out, the above workflow can be repeated, which then can result in increased parameterization detail, for example, the 7- or 14-unit model. Guided by this multistage calibration, model will evolve with increasing parameterization complexity that is supported by the data that are collected specifically to reduce estimation uncertainty. In this process, whenever appropriate, a post-audit analysis should be carried out testing whether an optimal parameterization, including the resolution of model BC, has been achieved for a given set of (post-audit) prediction goals and acceptable modeling errors.

Conclusion

Because of limited subsurface access, modeling and calibration of natural aquifers with multiple scales of heterogeneity is a challenging task and large uncertainty exists in developing a conceptual aquifer model and in uniquely calibrating this model for decision making. Because of uncertainties such as a lack of understanding of subsurface processes and a lack of techniques to parameterize the subsurface environment (including hydraulic conductivity, source/sink rates, and BC), existing aquifer models can suffer nonuniqueness in calibration, leading to poor predictive capability. A robust calibration methodology is needed that can address the simultaneous estimation of aquifer parameters, processes, and boundary conditions.

In this paper, we propose a multistage strategy that also addresses subsurface heterogeneity at multiple scales, while reducing uncertainty in estimating model parameters and model BC. The key of this strategy lies in the appropriate development, verification, and synthesis of the existing and new techniques of static and dynamic data integration. In particular, direct inversion techniques can be first used to estimate aquifer large-scale effective parameters and (smoothed) boundary conditions, based on which parameter and boundary condition estimation can be refined at increasing detail using standard or highly parameterized techniques. Both the initial analysis and the refinement stage(s) fit with the existing modeling workflows, whereas simpler models are constructed first, upon which complexity or refinement is added next, with or without collecting additional observation data.

Furthermore, because exhaustive subsurface sampling is impractical, an important question in developing hydrogeological site models is whether or not there exists one or more optimal level(s) of parameterization and process complexity that is sufficient for making accurate predictions. Research suggests that such optimality likely exists in groundwater flow modeling, as hydrostratigraphic models with reduced K heterogeneity resolutions can suffice for making certain flow predictions. On the other hand, the importance of heterogeneity on solute transport modeling has been increasingly recognized in hydrogeology and related fields. Could a hydrostratigraphic model of sufficient heterogeneity resolution also be capable of making accurate transport predictions? For certain bulk transport performance metrics, our work upscaling dispersivity suggests so (Zhang and Gable 2008). Transport issues, however, will be addressed later in tandem with the issues of process complexity, that is, coupling flow and transport. For both flow and transport modeling, our long term goal is to find an optimal modeling approach in terms of process representation, parameter scale, grid resolution, and inversion methodology.

Acknowledgments

The author gratefully acknowledges helpful communications with Dr. Mary Hill in the discussion and interpretation of the nonlinear regression outcomes. The author also acknowledges the support from the Center of Fundamental of Subsurface Flow, the School of Energy Resources at the University of Wyoming, NSF CI-WATER (Cyberinfrastructure to Advance High Performance Water Resource Modeling), and NSF EPSCoR (EPS 1208909).

References

Anderson, M.P., and W.W. Woessner. 1992. *Applied Groundwater Modeling*, 381. San Diego, California: Academic Press.

Bohling, G.C., and J.J. Butler, Jr. Inherent limitations of hydraulic tomography. *Ground Water* 2010, 48: 809–824. DOI: 10.1111/j.1745-6584.2010.00757.x

Bramley, R., X. Wang. 1995. Splib: A library of iterative methods for sparse linear systems. www.cs.indiana.edu/pub/techreports/TR454.pdf.

Carrera, J., and S.P. Neuman. 1986. Estimation of aquifer parameters under transient and steady state conditions: 3. Application to synthetic and field data. *Water Resources Research* 22: 228–242.

Doherty, J. 2005. PEST: Software for model-independent parameter estimation. www.pesthomepage.org.

Fienen, M., R. Hunt, D. Krabbenhoft, and T. Clemo. 2009. Obtaining parsimonious hydraulic conductivity fields using head and transport observations: A Bayesian geostatistical parameter estimation approach. *Water Resources Research* 45: W08405. DOI:10.1029/2008WR007431

Hill, M.C., and C.R. Tiedeman. 2007. *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*, 480. New York: Wiley-Interscience.

Hunt, R.J., and D.E. Welter. 2010. Taking account of "unknown unknowns." *Ground Water* 48: 477.

Hunt, R., and C. Zheng. 1999. Debating complexity in modeling. *Eos, Transactions American Geophysical Union* 80: 29.

Hunt, R., M. Anderson, and V.A. Kelson. 1998. Improving a complex finite-difference ground water flow modeling through the use of an analytical element screening model. *Ground Water* 36: 1011–1017.

Irsa, J., and Y. Zhang. 2012. A new direct method of parameter estimation for steady state flow in heterogeneous aquifers with unknown boundary conditions. *Water Resources Research* 48: W09526. DOI:10.1029/2011WR011756

Liu, X., and P.K. Kitanidis. 2011. Large-scale inverse modeling with an application in hydraulic tomography. *Water Resources Research* 47: 1373–1413. DOI:10.1029/2010WR009144

Poeter, E.P., M.C. Hill, E.R. Banta, S. Mehl, S. Christensen. 2005. UCODE 2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation. USGS Techniques and Methods, 6-All.

Poeter, E., and M.C. Hill. 1997. Inverse methods: A necessary next step in groundwater modeling. *Ground Water* 35: 250–260.

Saiers, J.E., D.P. Genereux, and C.H. Bolster. 2004. Influence of calibration methodology on ground water flow predictions. *Ground Water* 42: 32–44.

Tiedeman, C.R., M.C. Hill, F.A. D'Agnese, and C.C. Faunt. 2003. Methods for using groundwater model predictions to guide hydrogeologic data collection, with application to the Death Valley regional groundwater flow system. *Water Resources Research* 39: 1010. DOI:10.1029/2001WR001255

Wang, D., Y. Zhang, and J. Irsa. 2013. Dynamic data integration and stochastic inversion of a two-dimensional confined aquifer. *Proceeding of the 2013 AGU Hydrology Days*. http://hydrologydays.colostate.edu/Papers_13/Dongdong_paper.pdf (accessed August 16, 2013).

Wen, X.-H., C.V. Deutschb, and A.S. Cullickc. 2002. Construction of geostatistical aquifer models integrating dynamic flow and tracer data using inverse technique. *Journal of Hydrology* 255: 151–168.

Zhang, Y. 2013. Nonlinear inversion of an unconfined aquifer: Simultaneous estimation of heterogeneous hydraulic conductivities, recharge rates, and boundary conditions. Transport in Porous Media, under review.

Zhang, Y., J. Irsa, J. Jiao, D. Wang. 2013. Direct method of parameter estimation for steady state flow in heterogeneous aquifers with unknown boundary conditions. *Proceeding of the 2013 MODFLOW and More*, Golden, Colorado, 2–5 June.

Zhang, Y., B. Liu, and C.W. Gable. 2011. Homogenization of hydraulic conductivity for hierarchical sedimentary deposits at multiple scales. *Transport in Porous Media* 87: 717–737. DOI:10.1007/s11242-010-9711-8

Zhang, Y., and C.W. Gable. 2008. Two-scale modeling of solute transport in an experimental stratigraphy. *Journal of Hydrology* 348: 395–411. DOI:10.1016/j.jhydrol.2007.10.017

Zhang, Y., C.W. Gable, and M. Person. 2006. Equivalent hydraulic conductivity of an experimental stratigraphy-implications for basin-scale flow simulations. *Water Resources Research* 42: W05404. DOI:10.1029/2005WR004720

Zhu, J., and T.C.J. Yeh. 2006. Analysis of hydraulic tomography using temporal moments of drawdown recovery data. *Water Resources Research* 42: W02403. DOI:10.1029/2005WR004309

Zimmerman, D.A., G. de Marsily, C.A. Gotway, M.G. Marietta, C.L. Axness, R.L. Beauheim, R.L. Bras, J. Carrera, G. Dagan, P.B. Davies, D.P. Gallegos, A. Galli, J. Gómez-Hernández, P. Grindrod, A.L. Gutjahr, P.K. Kitanidis, A.M. Lavenue, D. McLaughlin, S.P. Neuman, B.S. RamaRao, C. Ravenne, and Y. Rubin. 1998. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resources Research* 34, no. 6: 1373–1413.