Uncertainty in simulated groundwater-quality trends in transient flow

J. Jeffrey Starn · Amvrossios C. Bagtzoglou ·
Gary A. Robbins

Abstract In numerical modeling of groundwater flow, the result of a given solution method is affected by the way in which transient flow conditions and geologic heterogeneity are simulated. An algorithm is demonstrated that simulates breakthrough curves at a pumping well by convolution-based particle tracking in a transient flow field for several synthetic basin-scale aquifers. In comparison to grid-based (Eulerian) methods, the particle (Lagrangian) method is better able to capture multimodal breakthrough caused by changes in pumping at the well, although the particle method may be apparently nonlinear because of the discrete nature of particle arrival times. Trial-and-error choice of number of particles and release times can perhaps overcome the apparent nonlinearity. Heterogeneous aquifer properties tend to smooth the effects of transient pumping, making it difficult to separate their effects in parameter estimation. Porosity, a new parameter added for advective transport, can be accurately estimated using both grid-based and particle-based methods, but predictions can be highly uncertain, even in the simple, nonreactive case.

J. J. Starn (✉)
US Geological Survey,
101 Pitkin Street, East Hartford, CT 06108, USA
e-mail: jjstarn@usgs.gov
Tel.: +1-860-2916746
Fax: +1-860-2916799

J. J. Starn · A. C. Bagtzoglou
Department of Civil and Environmental Engineering,
University of Connecticut,
261 Glenbrook Road, U 2037, Storrs, CT 06269-2037, USA

G. A. Robbins
Department of Natural Resources and the Environment,
University of Connecticut,
1376 Storrs Road, U-4087, Storrs, CT 06269-4087, USA

Introduction

Groundwater tracers are sometimes used in the inverse calibration of basin-scale (as opposed to plume scale) groundwater-flow models (Sanford et al. 2004; Manning and Solomon 2005; Michaels and Voss 2009; Sanford 2011). However, there are outstanding concerns with regard to simulation methods and parameter estimation (Konikow 2011; Voss 2011a, b). New insights into aspects of these issues are provided in this paper. Numerical dispersion, and its effect on parameter estimates, is one of those concerns. At the basin-scale, real dispersion often is neglected, in part because larger sources of uncertainty exist than in the sub-grid-scale heterogeneity represented by the dispersion coefficient (Sanford 2011). For example, Zinn and Konikow (2007a, b) simulate complex solute distributions that arise from wells represented in several model layers. Even if real dispersion is not simulated, inverse model estimates of parameters may be affected by numerical dispersion caused by grid resolution, transient flow, solution method, and the interaction of all these factors. An alternative to more complex advection-dispersion models and solution techniques is to use computationally simpler models to explore parameter space and the relations among parameters.

Grid resolution affects simulation accuracy in several ways. Velocity gradients near features such as pumping wells (Starn et al. 2012) or streams (Haitjema et al. 2001) are not accurately represented by velocities at grid nodes unless appropriately small grid spacing is used. Linear interpolation of velocity, commonly used in particle-based solute transport (Pollock 1988; Goode 1990; Schafer-Perini and Wilson 1991), often is referred to as "exact," but it is exact only in its reproduction of cell mass balance, not in its estimation of velocity (Pokrajac and Lazic 2002). Apart from numerical effects of inaccurate velocity interpolation (which can be rectified by using a fine enough model discretization), spatial averaging of geologic heterogeneity (Bower et al. 2005) affects groundwater flow and transport simulation results. Mehl and Hill (2001) tested the effect of numerical solution methods on inverse modeling parameter estimates and found that numerical dispersion in the solution methods propagated to estimates of dispersion that were larger than real dispersion in order to compensate for error in the numerical method. Similarly, Zyvoloski and Vesselinov (2006) reported that error in estimated parameters could

be reduced by as much as several hundred percent solely by grid refinement. Although some guidelines are offered in these references, there is no formal relation between grid resolution and model accuracy (for advective flow).

Transient groundwater flow also can lead to apparent dispersion (Goode and Konikow 1990) because of changing flow directions; however, if the periodicity of a cyclical stress that causes changing flow directions (such as pump on/off periods) is less than the average groundwater residence time, the effects on overall transport may be minimal (Reilly and Pollock 1996). Starn (1994) investigated a similar phenomenon where apparent dispersion was caused by periodic stream stage changes at a model outflow boundary. Even in a flow system with spatially stationary sources and sinks, apparent dispersion is produced by the complex interaction of aquifer heterogeneity and transient flow (Elfeki et al. 2012).

Model parameter values determined through inverse modeling are products of their relation to the input data, and different sets of calibration data will yield different parameter estimates, although each may fit the calibration data equally well. Although a simultaneous inverse simulation involving groundwater flow and solute transport can be used (for example, Fienen et al. 2009), Voss (2011a, b) points out that hydraulic head disturbance propagates through an aquifer as a diffusive process, whereas tracer concentrations propagate through an aquifer by advection and dispersion. He concludes that tracer and head data should not be used in the same inverse problem as they would each likely produce different parameter estimates. On the other hand, Hill and Tiedeman (2007) present successful studies that use multiple types of data in inverse simulation. In one example, Barth and Hill (2005) simulated breakthrough curves (BTCs) of virus transport and found that artifacts of the numerical solution methods (flow and transport) influenced the simulated BTC and that error-based weighting potentially leads to a few observations dominating the regression. Saiers et al. (2004) compared simulations calibrated to heads alone, heads and flows, and heads, flows, and concentrations, and found that simulated head and flow predictions varied little among the simulations, even though estimated parameter values were different for each simulation. Anderman and Hill (1999) decoupled flow and transport models in a three-step approach to inverse modeling, thus separating the effects of different types of data on parameter estimates.

The unifying idea in the preceding studies is that inverse modeling will seek to match whatever data are available, and although constrained by the physics in the model, will produce parameter estimates that compensate for a variety of conceptual, numerical, and structural defects in the model (which are ubiquitous). Konikow (2011) and Voss (2011a, b), considering these difficulties (among others), exhort modelers to construct simple well-understood models and to base model predictions on major trends and locally averaged values.

Head and flow data can be used in inverse modeling to identify groundwater flow paths, and volumetric flow information produced by groundwater flow models can be transformed to average linear velocity and groundwater residence time (Konikow 2011). Uncertainty in porosity estimates, which are important in overall residence time estimates (Zhu et al. 2010; Konikow 2011; Zhu 2012), can be propagated to uncertainty in tracer concentrations by convolution for any steady-state period and arbitrary tracer input function. The probability density function of groundwater residence times can be used to derive a solute BTC using convolution-based particle tracking (CBPT) (Robinson et al. 2010). CBPT was extended to simulate samples collected from pumping wells by Starn et al. (2012) in steady flow and to simulate plumes in transient flow by Srinivasan et al. (2011).

In calibrating advective transport models to atmospheric tracer concentration data, questions arise as to the appropriate level of discretization needed to achieve meaningful results and to the effect of the solution method on model parameter estimates. This paper looks at the effects of grid resolution and numerical method on inverse-simulation estimates of advective-transport model parameters. The problem is framed using a typical situation: a calibrated groundwater-flow model exists and at some later time, advective transport is added. Porosity is the principal parameter that is added and, in the work presented here, is the only parameter estimated in inverse simulation. Advective transport is governed by groundwater velocity, which is the ratio of hydraulic conductivity to porosity, both of which can be highly correlated in typical problems (Barth and Hill 2005). In the limited test cases presented here, porosity represents a linear modification of the velocity field that has already been determined through flow modeling. The flow model fixes flow paths in space, and the porosity parameter encapsulates uncertainty in travel times along those flow paths.

The simulation of breakthrough curves (BTCs) by convolution-based particle tracking (CBPT) in a transient flow field at a pumping well is introduced in this paper. In CBPT, the residence time distribution of groundwater calculated using particle tracking is convolved with the solute source input to produce BTCs. The situation investigated here is basin-scale advective transport where predictions of solute BTC are desired. Calibration data in this case could consist of sparse observations of an atmospheric tracer, such as tritium, chlorofluorocarbons, or sulfur hexafluoride. The calibration data are vertically integrated tracer observations such as might be obtained from a pumping well. A groundwater-flow model constructed on a relatively coarse grid is available such as might be done to simulate regional flow. This scenario can be considered worst-case, but this paper shows the merit in undertaking such simulations if appropriate regard is given to choosing simulation methods. The paper presents new understanding based on a simple set of simulations that may help a modeler understand larger questions of basin-scale groundwater transport.

Overview of methods

The effects of grid resolution and numerical methods on parameter estimates are tested using a synthetic groundwater model of transient groundwater flow. Three distributions of aquifer properties on two grid resolutions (a third resolution added for one case) are simulated: (1) homogeneous properties, (2) layered heterogeneous properties, and (3) distributed heterogeneous properties. BTCs are simulated by CBPT and then are compared to grid-based (Eulerian) solutions of advective transport. Porosity of the synthetic model cases is estimated with an inverse model. Predicted BTCs resulting from the introduction of a second hypothetical tracer (other than the one used for calibration) are made using CBPT for selected cases. Predictive uncertainty is estimated based on model calibration and a calibration-constrained Monte Carlo simulation. The effects of inverse model linearity are discussed in relation to predictions.

Description of synthetic models

Synthetic models used in this analysis are modified slightly from those used by Starn et al. (2012) (Fig. 1 and Table 1). A standard finite-difference groundwater flow model (MODFLOW; Harbaugh 2005) and particle tracker (MODPATH; Pollock 2012) with refinements by Starn et al. (2012) are used to calculate nodal velocities. The

Table 1 Model grid dimensions and boundary conditions

Model domain dimensions		
x	5,400 m	
y	5,400 m	
z	540 m	
Number of layers	9	
Boundary conditions		
Row 1 specified flux	87,480 m^3/d	
Last row constant head	540 m	
Flow and transport parameters (hydrofacies shown on Fig. 10)		
	K	Porosity
Homogeneous	30 m/d	0.30
Clay	5 m/d	0.05
Muddy sand	20 m/d	0.20
Sand	30 m/d	0.30
Gravel	40 m/d	0.40
Specific Storage	0.00001/m	
Sink strength		
Monitoring well	0 m^3/d	
Pumping well (intermittent)	50,000 m^3/d	
Grid cell dimensions		
9-grid ($x = y$)	600 m	
27-grid ($x = y$)	200 m	
243-grid ($x = y$)	22.22 m	
Layer thickness (z)	60 m	

K hydraulic conductivity

simulations are formulated on increasingly refined grids by dividing the domain into 9 × 9, 27 × 27, or 81 × 81 cells in the horizontal dimension (abbreviated as 9-grid etc.). The reference simulation is based on CBPT on a 243 × 243 cell grid. Each grid is divided vertically into 9 evenly spaced model layers, and the number of cells between adjacent levels of refinement differs by a factor of 9. Inflow of water is through a specified-flux (Neumann) boundary and outflow is through a specified-head (Dirichlet) boundary. Water inflow at the upstream boundary is specified in proportion to hydraulic conductivity (K) such that the total inflow is identical to the homogeneous aquifer (described in the following). Boundaries not otherwise specified are zero-flux (including no recharge). Two wells are simulated: a pumping well in layers 4 and 5 and a monitoring well in layers 5 and 6 that is not pumped. Pumping history is varied to create transient flow (Fig. 2). Flow into the pumping well from each layer is calculated using the Multi-Node Well (MNW2)
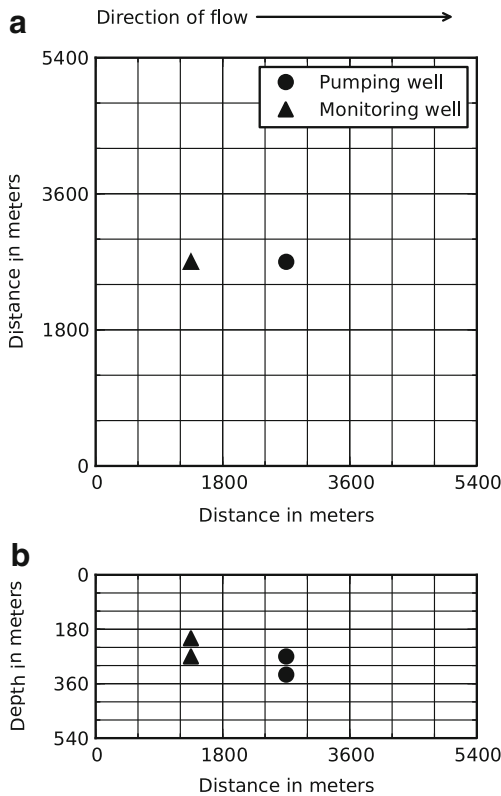

Fig. 1 **a** Map view and **b** section view of a 27 × 27 synthetic simulation grid. *Left boundary* is specified solute and water flux. *Right boundary* is constant head. All boundaries not otherwise specified are zero-flux boundaries. All grid dimensions in meters
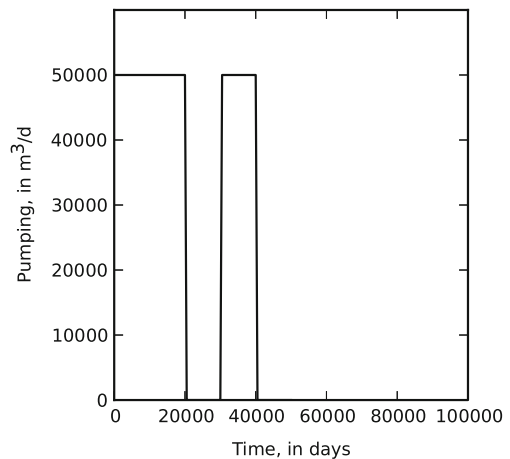

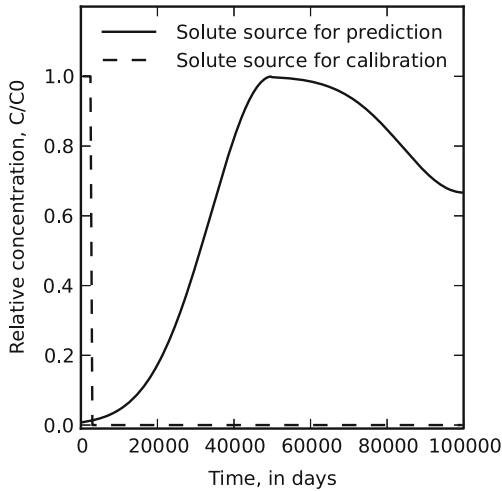**Fig. 2** Pumping history for transient simulation

Fig. 3 Solute input functions used to compute breakthrough curves for parameter estimation

package (Konikow et al. 2009). The specific storage used in these simulations (Table 1) is representative of a confined to semi-confined aquifer, which results in stress periods that very quickly reach steady state. CBPT also was tested with a storage representative of an unconfined aquifer (0.30) with similar results (not presented in this paper).

Solute is introduced to the flow system at the upstream boundary, and BTCs are calculated at the two simulated

wells. The solute source is a 2,500-day pulse of 1.0 mg/L (Fig. 3). CBPT is performed by backward tracking the release of 100 particles per time step for the simulated pumping well and monitoring well. A 0.01 m spatial error is specified for the movement of each particle in the adaptive time-step Runge-Kunge solution (Starn et al. 2012). Particles are released at evenly spaced time intervals, in this case every 500 days for 50,000 days, and tracked backward from the well screen to the inflow boundary. Particles are assigned such that the number of particles in each layer is proportional to the flow into the pumping well from that layer. An analytical solution describing velocity distribution around the well is used to provide divergence within the finite-difference model cell (Starn et al. 2012). Truncated Gaussian density estimation (Starn et al. 2012) is used to estimate complete residence time density at each time step, and the concentration at each time step is selected from the convolution of the input function with the time-step density function. The final BTC is minimally smoothed by applying a centered moving-mean window with a width of 3 time steps. This minor smoothing resulted in significantly less high-frequency noise in the BTC without adding discernible numerical dispersion.

For comparison, BTCs are produced on the same grids using a grid-based (Eulerian) solution (MT3DMS; Zheng and Wang 1999). The total variation diminishing (TVD)
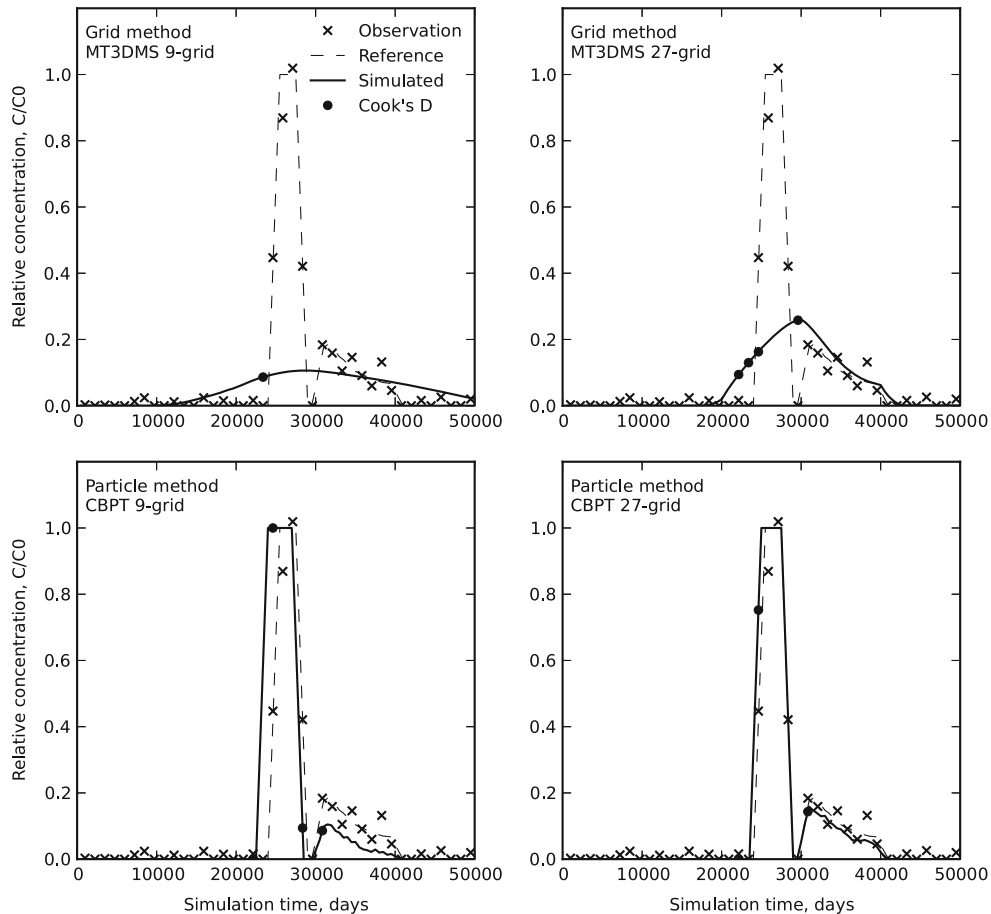

Fig. 4 Pumping well breakthrough curves for a homogeneous aquifer

solver in MT3DMS is used. TVD is explicit and requires a maximum time-step constraint, which in this case was set to 0.5 to ensure stability. Results are comparable to the CBPT, but execution time can be longer because concentration must be calculated over the entire grid for potentially small time steps and not just where BTCs are desired. As the grid is refined, solutions computed using CBPT and MT3DMS converge (Starn et al. 2012).

Three cases, corresponding to different conceptual models of aquifer hydraulic property distribution, are used to demonstrate the CBPT method. The first case simulates homogeneous aquifer properties. Effective porosity is 0.30 and K is 30 m/d.

The second case simulates exponential decrease of porosity with depth, similar to the systems described by Sanford (2011). Two porosity parameters are needed; here, 0.30 at the surface decreasing exponentially to 0.05 at the base of the aquifer are used. Because hydraulic conductivity (K) and effective porosity are weakly correlated; K is simulated as 100 times porosity. K and porosity have opposing effects on velocity, and their ratio results in a single effective parameter that governs the velocity field; however, in inverse modeling here, K is fixed and porosity is varied.

The third case uses a transitional probability model (Fogg 1996) to generate one realization of cell-by-cell heterogeneous aquifer properties. The synthetic geology used here was developed using well log data from the Salt Lake Valley aquifer in Utah, USA (C.T. Green, US Geological Survey, personal communication, 2011). Synthetic geology of this type leads to improved simulation of long, non-Gaussian BTC tails. The model used here is typical of an alluvial fan (Weissmann et al. 2002) and comprises four hydrofacies: clay, muddy sand, sand, and gravel; thus, four porosity parameters are needed. These values of K and effective porosity vary over a much smaller range than those used by other investigators; however, the purpose is not to simulate real deposits, but to simulate synthetic deposits that have variability similar to textural variation observed in well logs (Table 1). The original hydrofacies values are on a 100 m (length and width) by 1 m (thickness) grid. The first step to map them onto the synthetic grids is to interpolate the values onto the 243-grid. A median filter is applied in each vertical stack of cells to get the median value for each layer in the synthetic models. A median filter is then applied horizontally to map hydrofacies from the finer to the coarser grids.

For each case and each method/grid combination, inverse simulation was applied to estimate the known distribution of porosity using a set of 40 equally spaced simulated observations, 20 each in the pumping and monitoring wells. In the case of real atmospheric tracer data, only a sparse subset of these simulated observations may be available. The sparseness of typical data sets in
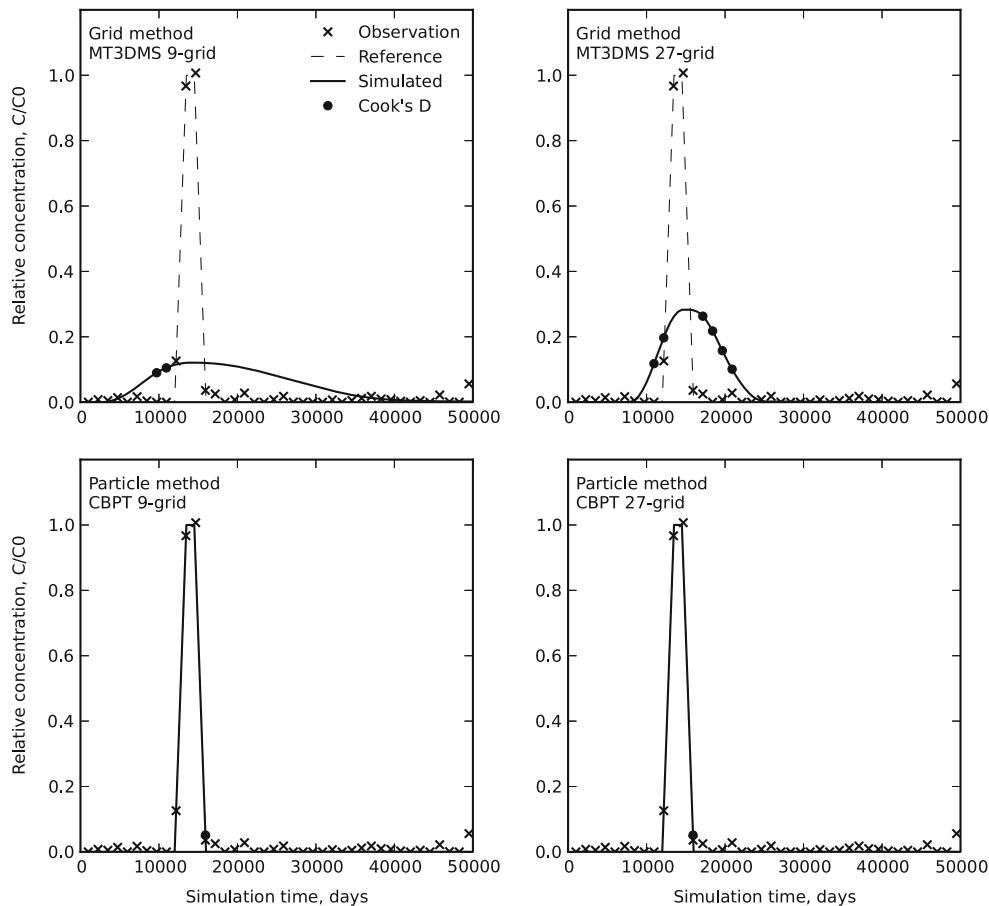

Fig. 5 Monitoring well breakthrough curves for a homogeneous aquifer

simulating atmospheric tracers, coupled with irregular BTCs produced by time-varying pumping, make it problematic to use the first temporal moment (arrival time of center of solute mass) as done by Barth and Hill (2005). The objective of the inverse simulation is to minimize the sum of squared weighted residuals of tracer concentrations. The simulated observations (hereafter termed "observations") were drawn from the reference simulation (243-grid) and perturbed by adding normally distributed noise to approximate measurement error. The variance of the noise was assigned using a coefficient of variation of 0.1, from which the variance of the measurement was calculated. The weight used in the inverse model is the inverse of the variance. To prevent very small measurements (which, if this were real data, might result from values near the detection limit of the analytical method) from dominating the regression (Hill and Tiedeman 2007), a minimal variance of 0.0004 (standard deviation 0.02 mg/L) was used. If the perturbed measurement was less than zero, it was reset to zero. This weighting scheme corresponds to the "observation-based weights" of Barth and Hill (2005). In general, "simulation-based weights" described in Barth and Hill (2005) are a good alternative because they yield unbiased parameter estimates. The computer program UCODE (Poeter et al. 2005) was used to perform the inverse modeling. UCODE employs a modified Gauss-Newton gradient-based nonlinear regression to estimate model parameters. A perturbation method with a user-specified range of perturbation is used to calculate the parameter sensitivity (Jacobian) matrix.

In order to see which measurements affect model calibration, Cook's D statistic (Yager 1998) is plotted on the BTCs. Cook's D is a measure of influence of each measurement on the regression by calculating how much the calibration would change if the measurement were removed. Critical values calculated using methods described by Hill and Tiedeman (2007) are used to show influential measurements. Cook's D is strictly applicable for linear or nearly linear models and may not apply to all the examples. It is shown to point out, through its consistent pattern in all models, which parts of the breakthrough curve have greater importance to model calibration. Cook's D does not relate to model predictions, only to calibration. Although there is no analogous statistic to Cook's D for predictions, the OPR statistic is an appropriate surrogate measure of influence (Hill and Tiedeman 2007). OPR was not used here because, in these simple models where porosity is the only parameter, measurements that affect calibration are likely to affect predictions. A measure of model linearity (Beale's measure; Hill and Tiedeman 2007) also is discussed with the inverse model results.
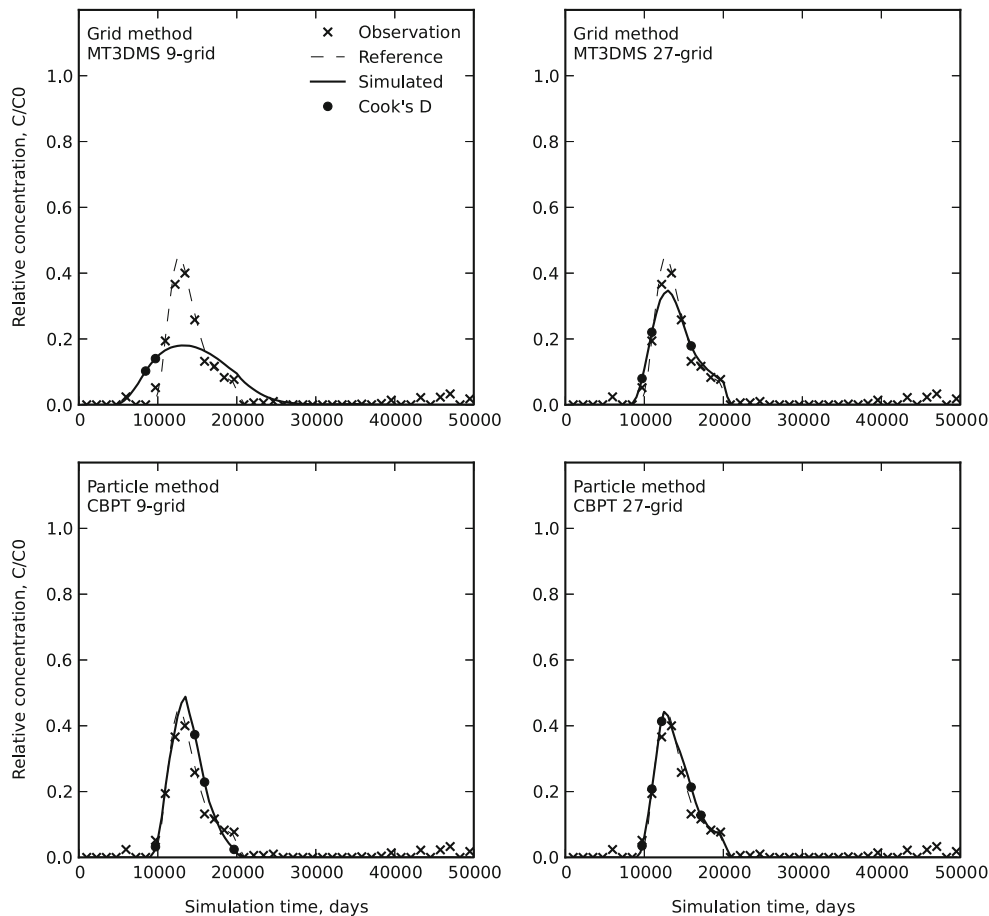

Fig. 6 Pumping well breakthrough curves for a layered aquifer

Results of synthetic simulations

Homogeneous properties

An interesting artifact of the synthetic simulations is that the pumping history (Fig. 2) causes a bimodal split of the input pulse. During the first pump-off period, the solute pulse migrates past the pumping well. When the pump is turned back on, some of the solute reverses direction and is pulled back into the well. The grid solution cannot resolve the bimodal BTC and partially smoothes over the variations (circles in Fig. 4), and in doing so under-predicts porosity. Increasing the number of cells by a factor of 9 (from 9-grid to 27-grid) increases the accuracy of the grid method, but also increases the simulation time by a factor of 4. CBPT (Fig. 4) had no difficulty resolving the peaks, and grid refinement had little effect on the match. The correct porosity was estimated by the inverse model for the particle method for both grids. Critical measurements (as defined in Hill and Tiedeman 2007) are likely to occur just before and at the early part of solute breakthrough, and at the low point in between the two peaks. The monitoring well (Fig. 5), which is much less affected by a mixture of groundwater residence times than the pumping well, has the characteristic BTC of a pulse input. The grid method cannot resolve this pulse due to numerical dispersion, but the particle methods are able to do so. Beale's measure indicates that the two grid methods

are effectively linear, but the two particle models were highly nonlinear.

Layered properties

The particle method tended to produce better matches than the grid method, although improved grid resolution improved the match for both methods (Fig. 6). The same pattern of influential measurements is evident as for the homogeneous case. Compared to the homogeneous aquifer simulation, the BTC at the pumping well had a lower, wider peak. The monitoring well again had a BTC characteristic of a pulse source (Fig. 7). The grid method matched the center of the peak accurately, but overestimated the width and yielded biased estimates of porosity (Fig. 8). The particle method was able to resolve the center and width of the peak and yielded unbiased estimates of porosity. The correlation matrix generated by UCODE shows that the two porosity parameters are weakly correlated ($r \sim 0.5$–0.7) for all synthetic models. Beale's measure indicated, as in the homogeneous case, the grid method produced linear models and the particle method produced nonlinear models. The sum-of-squared weighted residuals objective function for both porosity parameters (Fig. 9) shows some of the reason for difficulty estimating the correct parameters: low sensitivity to surface porosity and local minima.
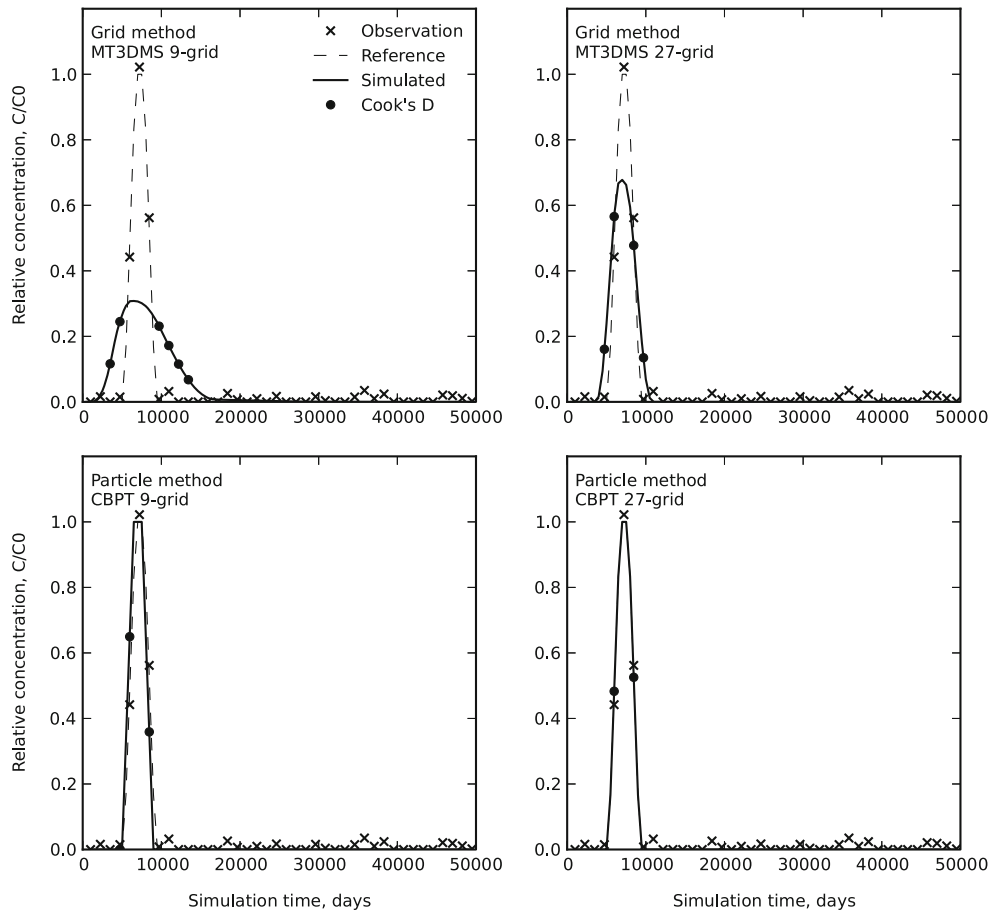

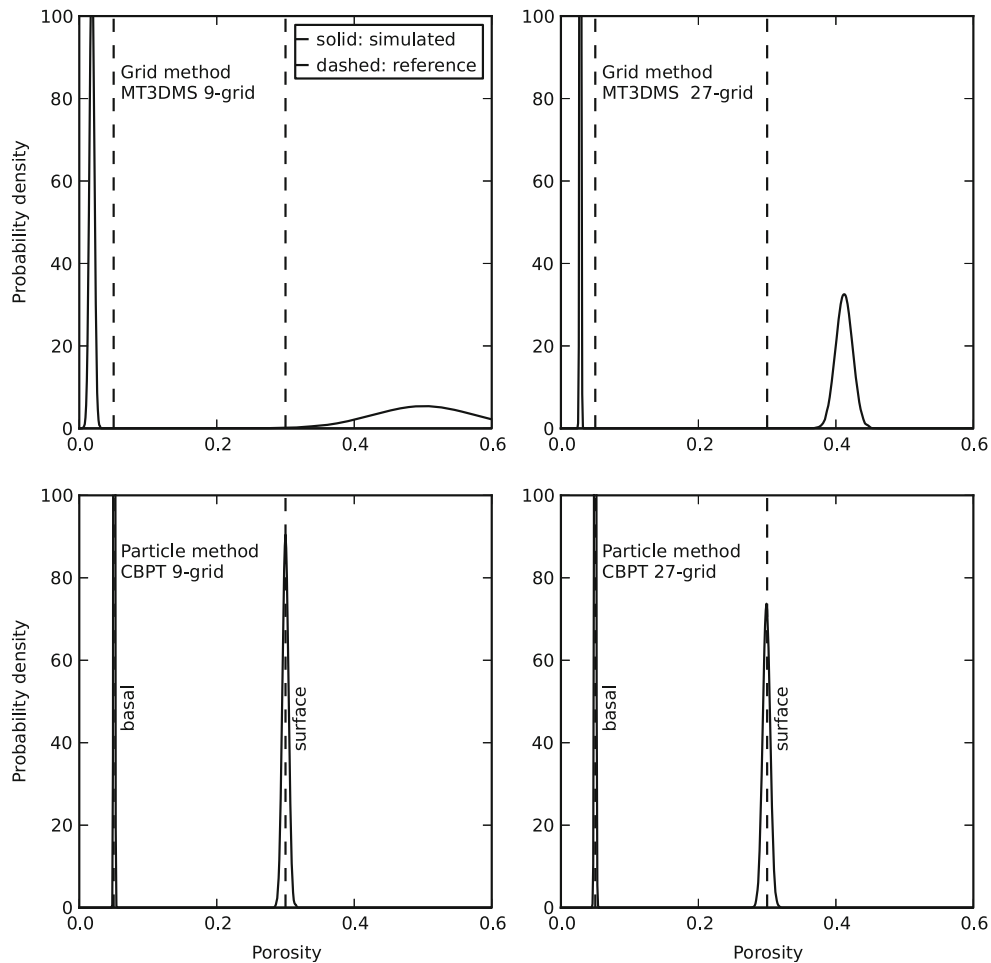**Fig. 7** Monitoring well breakthrough curves for a layered aquifer

Fig. 8 Optimal estimated porosity distribution for a layered aquifer

Heterogeneous properties

As the aquifer properties are upscaled toward coarser grid resolutions, the estimated properties tend toward homogeneity with median property values (Fig. 10). However, coarse features in the synthetic geology remain. For example, in the 9-grid, the pattern is much simpler than the 243-grid, but the
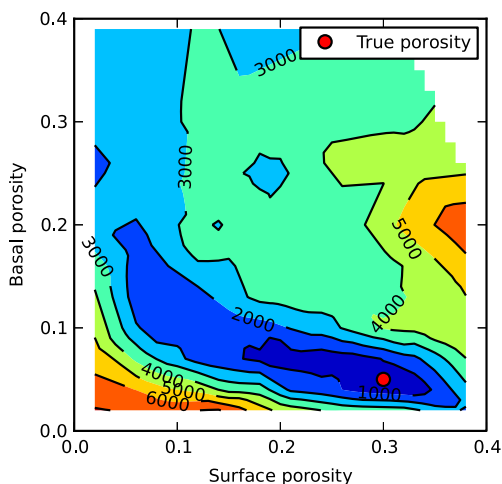

Fig. 9 Sum-of-squares objective function surface for a CBPT layered aquifer on 27-grid

preferential path ways remain. Upscaling produces a more uniform (but not completely so) solute inflow distribution. Only the particle method was tried in this case, but an extra level of grid refinement (the 81-grid) was added. The purpose is to focus on the effects of upscaling and to assess how well the particle solution performs in this more realistic case. The particle method matches the pulse relatively accurately, and there is a slight improvement with increased grid refinement (Fig. 11). Similar influential observations are evident and their distribution follows the pattern as in previous cases. In the monitoring well, and to a lesser extent the pumping well, the peak arrival is estimated accurately in the 9-grid and 81-grid, but not in the 27-grid (Fig. 12). The pulse retains its shape as in previous cases, but it is apparent that some solute bypasses the well due to preferential pathways in the aquifer properties. All 3 grids can resolve the correct porosity parameters (Fig. 13), with increasing reliability with increasing grid refinement, except that the 9-grid overestimates porosity in clay. Estimated parameters show weak correlation, as in the layered case, except for the 27-grid, in which the muddy sand and sand hydrofacies were highly correlated ($r \sim 0.96$). The dilution rate (the pumping rate divided by the volume of water in the grid cell) is an indication of grid accuracy for particle methods (Starn et al. 2012). For a pumping well with time-varying pumping, the dilution factor
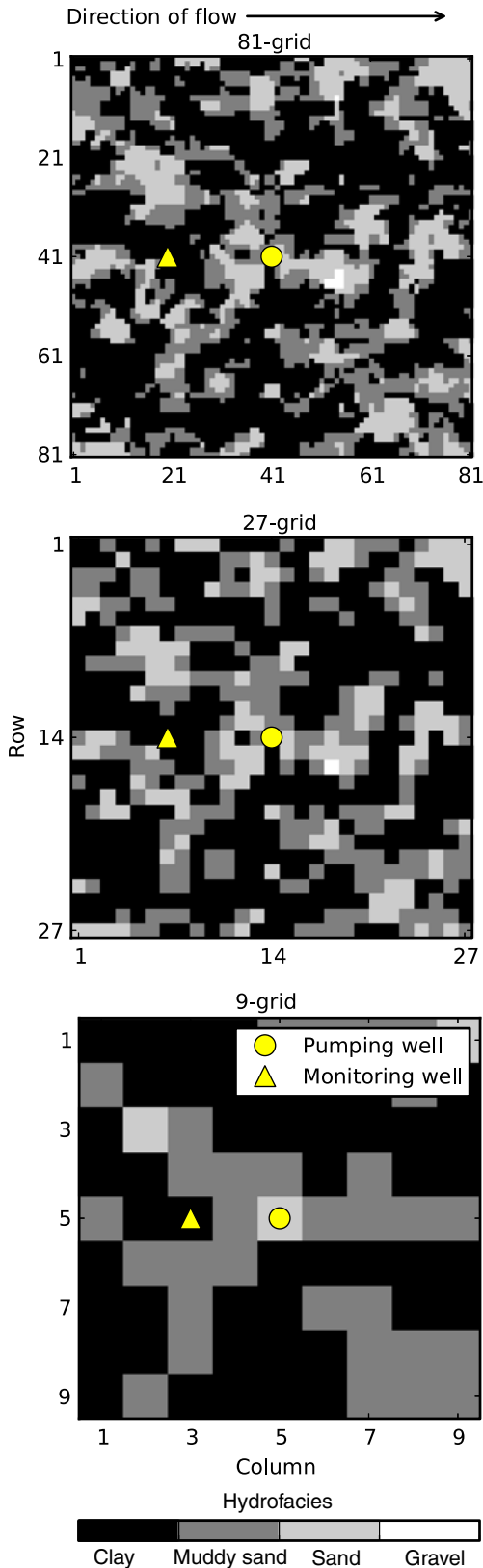
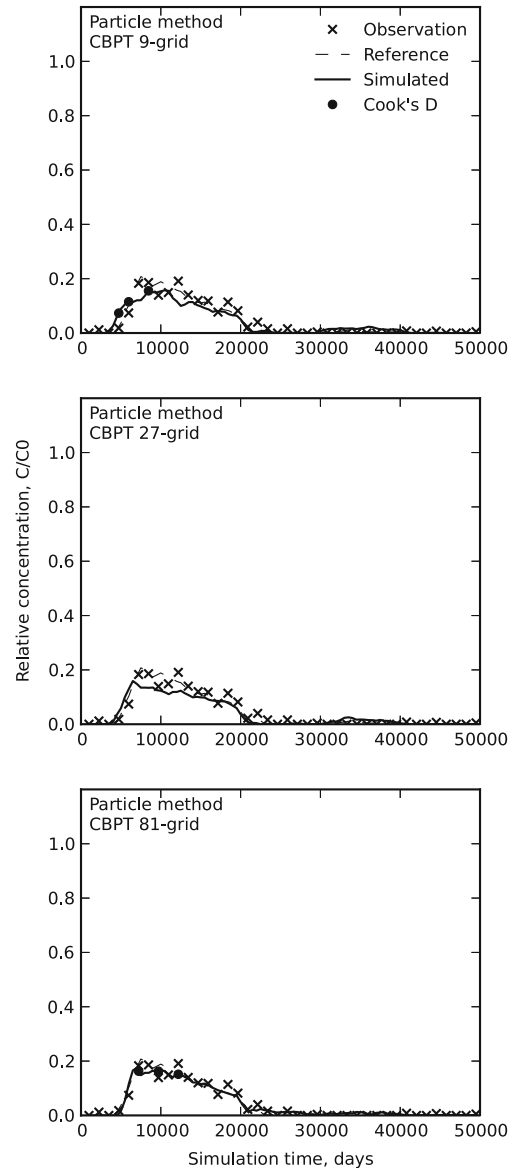Fig. 10 Map view showing hydrofacies (defined in Table 1) in layer 4

Fig. 11 Pumping well breakthrough curves using the CBPT solution for a heterogeneous aquifer

differently for each layer (Fig. 14), depending on grid resolution. Likewise, the sink strength factor (the pumping rate divided by the total rate of water flow into the cell) changes with time in a transient flow system (Fig. 15).

Simulation of predictive uncertainty

In order to see the effects of parameter uncertainty on predicted trends, a new solute (Fig. 3) was introduced into the synthetic models. This new solute can be thought of as a slowly increasing atmospheric tracer such as chlorofluoro-carbons or sulfur hexafluoride, but it is not meant to represent a specific input. These simulations were done using 9-grid and 27-grid CBPT for the layered and heterogeneous cases. These were chosen because they illustrate several key points. Variance-covariance matrices
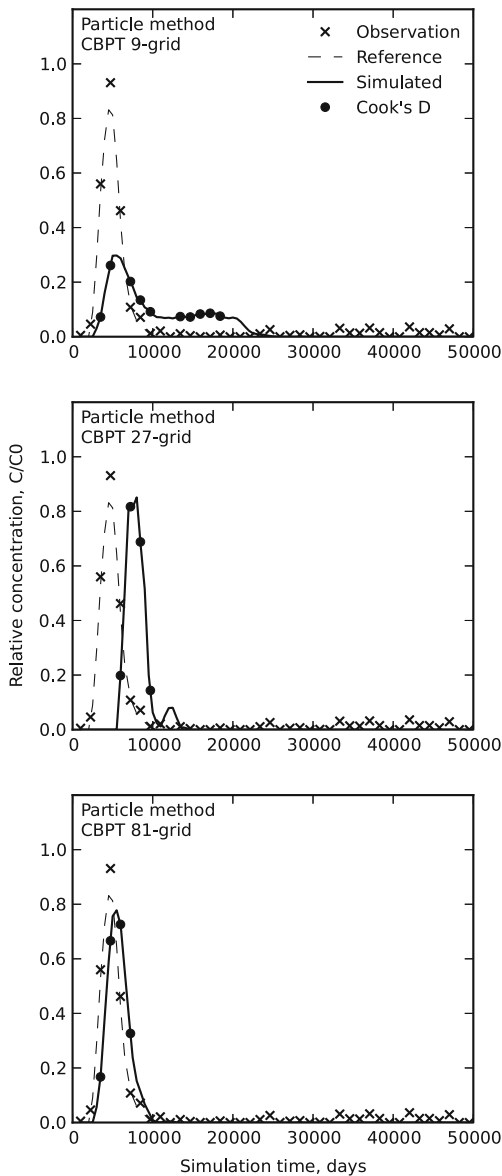
changes with time also, and if the well is represented in more than one model layer, the dilution factor can change

Fig. 12 Monitoring well breakthrough curves using the CBPT solution for a heterogeneous aquifer

Fig. 13 Optimal estimated porosity distribution for heterogeneous aquifer

from the inverse model were used to generate Latin hypercube samples (LHS) of porosity (Starn et al. 2010; Starn and Bagtzoglou 2011). Sampling in this way preserves correlations among parameters. LHS guarantees that the parameter density functions (Figs. 8 and 13) are sampled evenly. In this case, 20 equally probable regions of the density functions were sampled, and the sampling was repeated 50 times for a total of 1,000 parameter sets. Monte Carlo simulations were run using each of the 1,000 parameters sets, and the standard error in relation to the calibration data was recorded for each simulation. For the final analysis, parameter sets were eliminated in which the standard error exceeded plus or minus one standard deviation of the standard error. The number of valid runs (out of the original 1,000) is presented in Figs. 16 and 17. To judge whether the simulations were converging, mean concentrations were calculated for each time step, and the
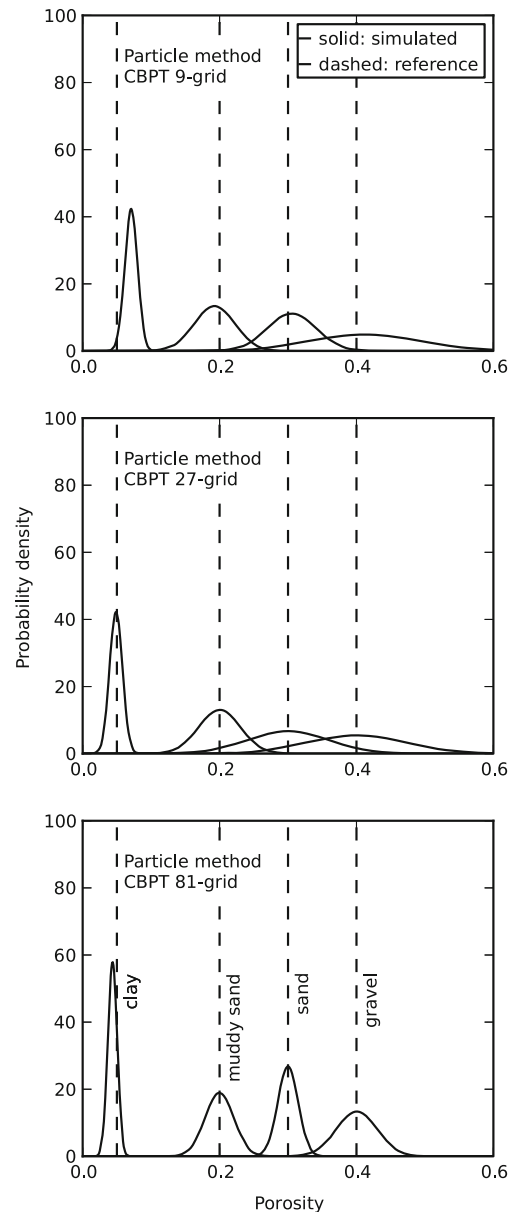
maximum difference between simulations was plotted against simulation number. In each case, the maximum change in mean concentration after all simulations was less than 0.001 mg/L. This test confirms the stability of the mean concentration, but does not address the extremes. To produce specific confidence intervals would require much more testing than was done here. Instead, the envelope of minimum and maximum around the mean predicted curve (Figs. 16 and 17) are plotted using only valid runs.

Discussion

Similar to the results described by Barth and Hill (2005), discrepancies are noted among methods and
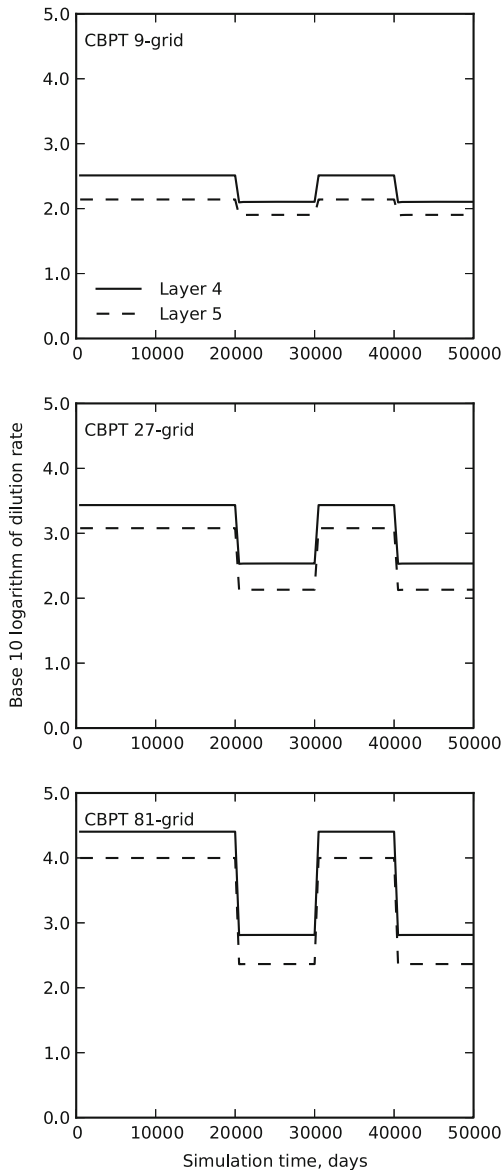
Fig. 14 Pumping well dilution factor for a heterogeneous aquifer

Fig. 15 Pumping well weak sink factor for a heterogeneous aquifer

grid resolutions in their abilities to correctly estimate the true porosity distribution. The homogeneous case illustrates the strengths of the particle method in resolving sharp concentration fronts. This is beneficial in the inverse model for estimating parameter mean and variance; however, the nonlinearity of these models may make their use in prediction uncertainty problematic in that most widely used techniques rely on linear theory. It is possible that their nonlinearity is "apparent," not real, caused by a locally rough objective-function surface. Although larger sets of particles were tried without improving the situation, it is possible that by judicious manipulation of the number of particles and their release times that a smoother objective-function surface could result. As noted by Barth and Hill (2005), the time-step size can influence parameter sensitivity and estimates. In general, it seems the more heterogeneous the flow
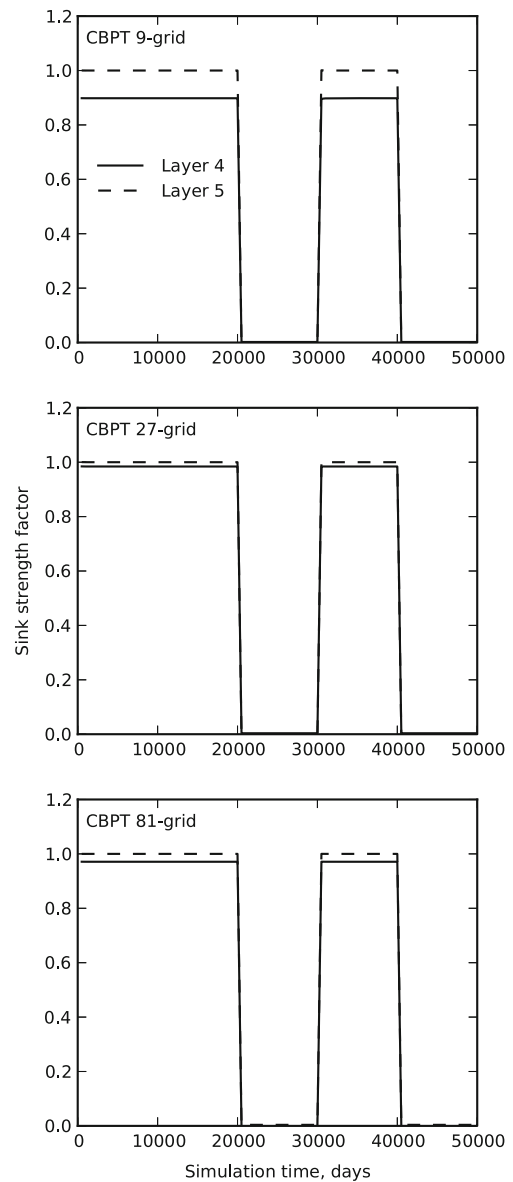
system, the more particles are required. The effective linearity of the grid method for the same models could be because the grid method promotes smoothness in the BTC, which in turn helps smooth the objective function surface. Yager (2004) points out that the amount by which sensitivities are perturbed in UCODE can affect parameter estimation convergence, and this could be manipulated to achieve a well-posed linear problem with particles. Point estimates of objective function gradient such as by exact sensitivity equations or adjoint methods, might be more susceptible to the ill effects of non-smooth objective function surfaces (Yager 2004).

Estimating porosity from BTC data can be a challenge because there are multiple points that can be matched by the simulation. For example, the inverse model could adjust porosity to match a
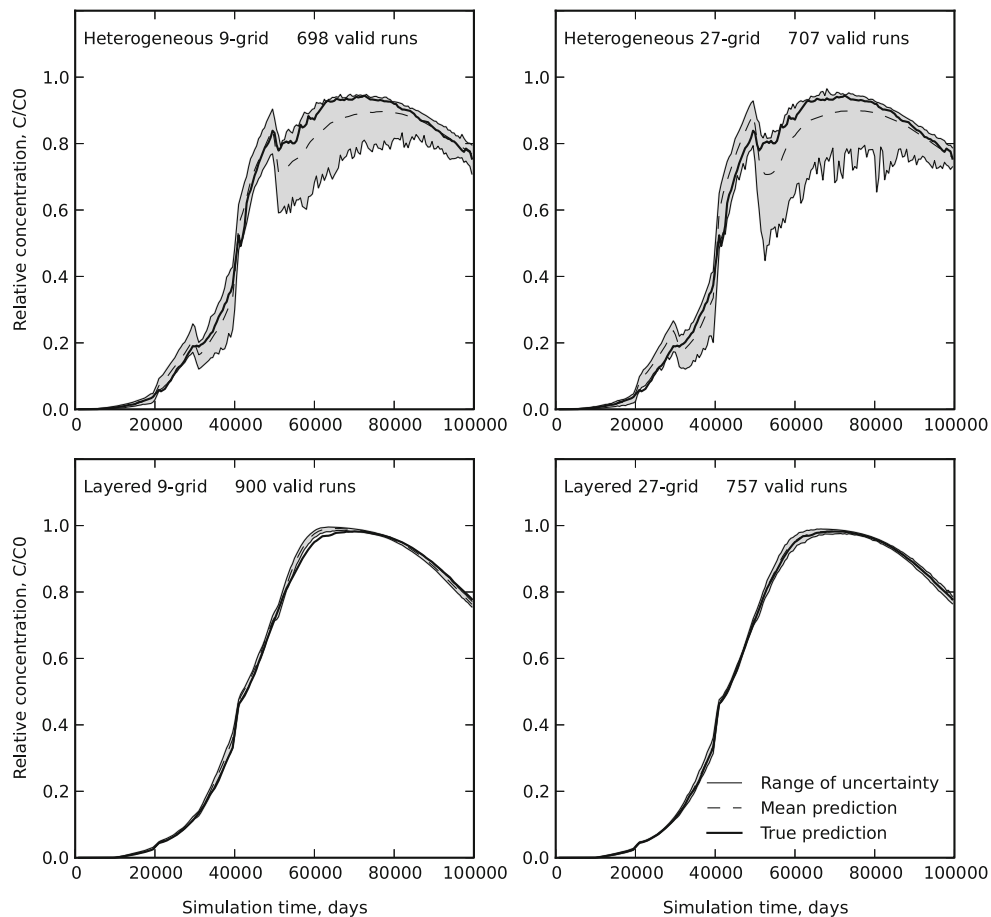
Fig. 16 Monte Carlo predictions for a pumping well

measured value as if it were on either the rising limb or falling limb of the BTC. This effect was seen in the objective function surface (Fig. 9). Because of its irregular shape, judicious choice of parameter starting values is necessary. Having observations at multiple locations within single porosity zones can help achieve the correct minima. Highly parameterized models, in which more parameters can be adjusted, may have more difficulty in this regard.

Although a rigorous analysis of data worth was not undertaken, the simulations suggest data that describe the first arrival of solute are influential, as well as data that document any large changes in direction of slope of the BTC. The first arrival of solute fixes the BTC in its place in time and space, and the physics of the model determine the rest of its shape. Yager (1998) points out, however, that even the least influential measurements have worth in increasing confidence in predictions. Also, Barth and Hill (2005) found that weights based on simulated values rather than observed values produced more accurate sensitivities. In this study, the observed and simulated BTCs overlie each other, so the final weights are the same whether using observed or simulated values as the basis for calculating weights.

The degree of upscaling of aquifer properties affects parameter correlation and accuracy of estimates. As preferential pathways change with grid refinement, the residence time distribution changes in relation to the truth, and depending on local juxtaposition of hydrofacies, can lead to inaccurate results. In the heterogeneous case, refining the grid from the 9-grid to the 27-grid actually resulted in a less accurate simulation for the monitoring well (Fig. 12). This refinement caused a dramatic increase in parameter correlation between two of the hydrofacies.

One characteristic of groundwater tracer data is that there often are multiple low-level detections. Data precision can mask real-but-low parts of BTCs and make the inverse model insensitive to them. Non-normality of residuals caused by clustering of concentrations near zero can lead to incorrect conclusions about the applicability on confidence intervals that are based on linear theory and normality of residuals.

Conclusions

Estimating complete BTCs, rather than point data, as is typically done in inverse modeling, is helpful in
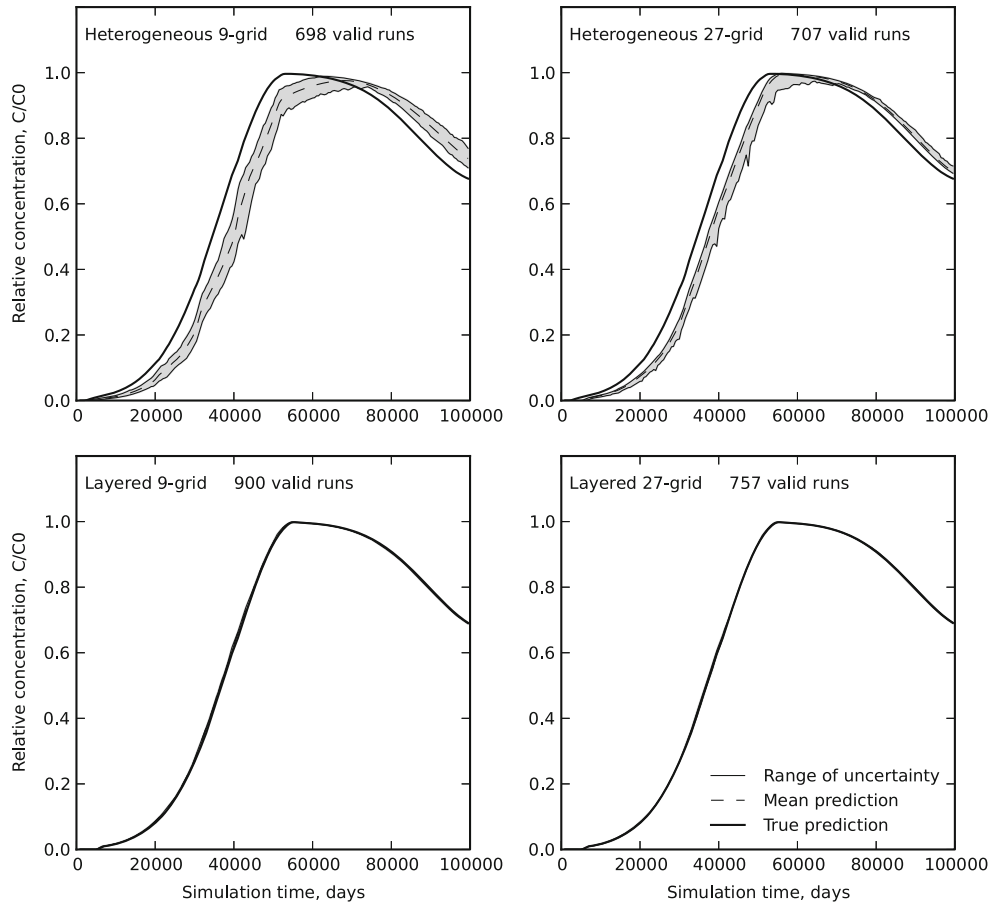
Fig. 17 Monte Carlo predictions for a monitoring well

diagnosing non-uniqueness, incorrect parameter values, and lack of fit. Transient particle methods (CBPT here) and grid-based methods produce BTCs differently depending on grid resolution. Unique solutions to models based on solute BTCs may be difficult to obtain, even with rich data sets. This study shows that simple transient pumping history and simple solute input functions can produce complex multi-modal BTCs. Real pumping histories, even in the absence of complicated geology, can be expected to produce complicated BTCs.

Both particle and grid methods were able to resolve porosity parameters for increasingly complex aquifer property conditions on relatively coarse model grids. In these simple synthetic models, a model calibrated on one set of solute data was used to successfully predict a second set of data. For the pumping well, the range of Monte Carlo outcomes contained the "true" prediction (from the high resolution 243-grid). By contrast, the range of Monte Carlo outcomes at the monitoring well for the heterogeneous aquifer did not contain the true prediction. This demonstrates how incorrect model structure caused by upscaling can lead to incorrect predictions, even though the correct parameters are obtained from inverse modeling. Ranges of plausible outcomes are greater for the heterogeneous case than for the layered case. This is expected because of the existence of preferential pathways in the heterogeneous case. Pumping history strongly affects the BTC, and numerical error from discretization can cause reversals in trend direction even though the input function is monotonically increasing. Effects of pumping on the BTC are still evident in a monitoring well far from the pumping well, but the variability of outcomes is less than those at the pumping well.

Multiple local minima and local roughness on the objective-function surface can lead to non-uniqueness and convergence problems. The first issue can be addressed by starting inverse modeling at different starting values, but even several starting points at what might be reasonable values can lead to incorrect solutions. Additionally, if starting values of porosity place the simulated BTC at a time when most data are non-detects, there will be little movement of the simulated BTC for small perturbations of porosity. Conversely, large perturbation may help hide some of the noise on the objective function surface, but also lead to inaccurate gradients that could lead to a correct minima being missed (overshoot). A careful analysis based on prior knowledge of likely parameter values and the ability to simulate multiple BTCs can give confidence in modeling results. The second issue is related, in part, to solution method. Methods that generate smooth solutions such as grid methods

reduce roughness of the objective function but cause numerical dispersion that can lead to inaccurate parameter values, especially if the BTC is multi-modal. Accuracy for grid methods is improved through grid refinement, but the execution time increases greatly because more calculations are required and because high gradients near pumping wells require small time steps. Also, using a coarse grid causes more uncertainty in parameter estimates, which in turn can increase variance on predictions that depend on those parameters. Truncated Gaussian-kernel density estimation regularizes the particle method by estimating the complete residence-time density function with relatively few particles, thus reducing noise in the computed BTC. Although fewer particles are required, the number of particles and release times can be problem-dependent and to offer general guidelines is beyond the scope of this work.

As Elfeki et al. (2012) pointed out, there is a complex interaction between transient flow and hetero-geneity. This is especially true when using concentra-tion data. For example, the vertical gradient imposed by the parameter values in the "layered aquifer" case caused a tension in the inverse model between matching concentrations in the monitoring well and the pumping well. Transient velocity fields at these wells affect parameter estimates in opposing ways. In the heterogeneous case, the upscaling that happens as a necessary consequence of model construction obscures preferential pathways, which, because of the transient flow, play greater or lesser roles depending on the location of the solute mass in relation to the sampling point at the time of sampling.

The effects of solution method, flow transients, data uncertainty, and geologic heterogeneity are all con-volved together to produce BTCs, and separating them in a model is not generally possible. Concentrations measured at pumping wells tend to integrate the effects of upgradient factors. In the end, all these factors tend to act together to produce smooth BTCs, at least at the scale where the effects of multiple confounding factors are integrated, such as in a pumping well. There can be enough information content in samples from such wells to estimate parameters, but the results should be interpreted with caution. When faced with the problem of using the information in an existing groundwater flow model to help explain groundwater trends, the choices might be to use a grid method, if the execution time is not prohibitive, for inverse modeling. The results obtained using a particle method, which has minimal numerical dispersion, can be compared to results obtained using a grid-based to highlight where numerical dispersion might be a problem. Once model parameters are estimated, more accurate solutions with less computation time can be obtained using reverse particle tracking. In the future, efforts should be made to fully understand the interactions among solution method, heterogeneity, and predictive accuracy. In particular, it would be helpful to have more concrete guidelines for choices in that regard. The methods presented in this work can be used to help pursue those guidelines.

References

Anderman ER, Hill MC (1999) A new multistage groundwater transport inverse method: presentation, evaluation, and impli-cations. Water Resour Res 35:1053–1063

Barth G, Hill MC (2005) Numerical methods for improving sensitivity analysis and parameter estimation of virus transport simulated using sorptive-reactive processes. J Contam Hydrol 76:251–277

Bower KM, Gable CW, Zyvoloski GA (2005) Grid resolution study of ground water flow and transport. Ground Water 43:122–132

Elfeki AMM, Uffink G, Lebreton S (2012) Influence of temporal fluctuations and spatial heterogeneity on pollution transport in porous media. Hydrogeol J 20:283–297

Fienen M, Hunt R, Krabbenhoft D, Clemo T (2009) Obtaining parsimonious hydraulic conductivity fields using head and transport observations: a Bayesian geostatistical parameter estimation approach. Water Resour Res 45, W08405

Fogg GE (1996) Transition probability-based indicator geostatistics. Math Geol 28:453–476

Goode DJ (1990) Particle velocity interpolation in block-centered finite difference groundwater flow models. Water Resour Res 26:925–940

Goode DJ, Konikow LF (1990) Apparent dispersion in transient groundwater flow. Water Resour Res 26:2339–2351

Haitjema H, Kelson V, De Lange W (2001) Selecting MODFLOW cell sizes for accurate flow fields. Ground Water 39:931–938

Harbaugh AW (2005) MODFLOW-2005, the U.S. Geological Survey modular ground-water model: the ground-water flow process. US Geological Survey Techniques and Methods 6-A16, USGS, Reston, VA, 253pp

Hill MC, Tiedeman CR (2007) Effective groundwater model calibration with analysis of data, sensitivities, and uncertainty. Wiley, Englewood Cliffs, NJ

Konikow LF (2011) The secret to successful solute-transport modeling. Ground Water 49:144–159

Konikow LF, Hornberger GZ, Halford KJ, Hanson RT (2009) Revised multi-node (MNW2) package for MODFLOW ground-water flow model. US Geological Survey Techniques and Methods 6-A30, USGS, Reston, VA, 67pp

Manning AH, Solomon DK (2005) An integrated environmental tracer approach to characterizing groundwater circulation in a mountain block. Water Resour Res 41, W12412

Mehl S, Hill MC (2001) A comparison of solute-transport solution techniques and their effect on sensitivity analysis and inverse modeling results. Ground Water 39:300–307

Michaels HA, Voss CI (2009) Estimation of regional-scale groundwater flow properties in the Bengal Basin of India and Bangladesh. Hydrogeol J 17:1329–1346

Poeter EP, Hill MC, Banta ER, Mehl S, Christensen S (2005) UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation. US Geological Survey Techniques and Methods 6-A11, USGS, Reston, VA

Pokrajac D, Lazic R (2002) An efficient algorithm for high accuracy particle tracking in finite elements. Adv Water Resour 25:353–369

Pollock DW (1988) Semianalytical computation of path lines for finite-difference models. Ground Water 26:743–750

Pollock DW (2012) User guide for MODPATH version 6: a particle tracking model for MODFLOW. US Geological Survey Techniques and Methods 6-A41, USGS, Reston, VA, 58pp

Reilly TE, Pollock DW (1996) Sources of water to wells for transient cyclic systems. Ground Water 34:979–988

Robinson BA, Dash ZV, Srinivasan G (2010) A particle tracking transport method for the simulation of resident and flux-averaged concentration of solute plumes in groundwater models. Comput Geosci 14:779–792

Saiers JE, Genereux DP, Bolster CH (2004) Influence of calibration methodology on ground water flow predictions. Ground Water 42:32–44

Sanford WE (2011) Calibration of models using groundwater age. Hydrogeol J 19:13–16

Sanford WE, Plummer LN, McAda DP, Bexfield LM, Anderholm SK (2004) Hydrochemical tracers in the middle Rio Grande Basin, USA: 2. calibration of a groundwater-flow model. Hydrogeol J 12:389–407

Schafer-Perini AL, Wilson JL (1991) Efficient and accurate front tracking for two-dimensional groundwater flow models. Water Resour Res 27:1471–1485

Srinivasan G, Keating E, Moulton JD, Dash ZV, Robinson BA (2011) Convolution-based particle tracking method for transient flow. Comput Geosci doi:10.1007/s10596-011-9265-z

Starn JJ (1994) Apparent solute dispersion in an alluvial aquifer/stream system. In: Morganwalp DW, Aronson DA (eds) U.S. Geological Survey Toxic Substances Hydrology Program: proceedings of the technical meeting, Colorado Springs, Colorado, September 20–24, 1993. US Geol Surv Sci Invest Rep 94–4015, pp 935–940

Starn JJ, Bagtzoglou AC (2011) Programs for calibration-based Monte Carlo simulation of recharge areas. Ground Water. doi:10.1111/j.1745-6584.2011.00868.x

Starn JJ, Bagtzoglou AC, Robbins GA (2010) Using atmospheric tracers to reduce uncertainty in groundwater recharge areas. Ground Water 48:858–868. doi:10.1111/j.1745-6584.2010.00674.x

Starn JJ, Bagtzoglou AC, Robbins GA (2012) Methods for simulating solute breakthrough curves in pumping groundwater wells. Comput Geosci 48:244–255. doi:10.1016/j.cageo.2012.01.011

Voss CI (2011a) Editor's message: groundwater modeling fantasies–part 1, adrift in the details. Hydrogeol J 19:1281–1284

Voss CI (2011b) Editor's message: groundwater modeling fantasies–part 2, down to earth. Hydrogeol J 19:1455–1458

Weissmann GS, Zhang Y, LaBolle EM, Fogg GE (2002) Dispersion of groundwater age in an alluvial aquifer system. Water Resour Res 38:161–1613

Yager RM (1998) Detecting influential observations in nonlinear regression modeling of groundwater flow. Water Resour Res 34:1623–1633

Yager RM (2004) Effects of model sensitivity and nonlinearity on nonlinear regression of ground water flow. Ground Water 42:390–400

Zheng C, Wang PP (1999) MT3DMS: a modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems; documentation and user's guide. US Army Corps of Engineers Contract Report SERDP-99-1, US Army Corps of Engineers, Washington, DC, 220pp

Zhu J (2012) Sensitivity of advective travel time of contaminants to correlated formation porosities. Hydrogeol J 20:135–141

Zhu J, Pohlmann KF, Chapman JB, Russell CE, Carroll RWH, Shafer DS (2010) Sensitivity of solute advective travel time to porosities of hydrogeologic units. Ground Water 48:442–447

Zinn BA, Konikow LF (2007a) Effects of intraborehole flow on groundwater age distribution. Hydrogeol J 15:633–643

Zinn BA, Konikow LF (2007b) Potential effects of regional pumpage on groundwater age distribution. Water Resour Res 43, W06418

Zyvoloski GA, Vesselinov VV (2006) An investigation of numerical grid effects in parameter estimation. Ground Water 44:814–825