WATER RESOURCES RESEARCH, VOL. 37, NO. 10, PAGES 2503–2512, OCTOBER 2001

Predictive modeling of flow and transport in a two-dimensional intermediate-scale, heterogeneous porous medium

Gilbert R. Barth,[1,2] Mary C. Hill,[2] Tissa H. Illangasekare,[3] and Harihar Rajaram[1]

Abstract. As a first step toward understanding the role of sedimentary structures in flow and transport through porous media, this work deterministically examines how small-scale laboratory-measured values of hydraulic conductivity relate to in situ values of simple, artificial structures in an intermediate-scale (10 m long), two-dimensional, heterogeneous, laboratory experiment. Results were judged based on how well simulations using measured values of hydraulic conductivities matched measured hydraulic heads, net flow, and transport through the tank. Discrepancies were investigated using sensitivity analysis and nonlinear regression estimates of the in situ hydraulic conductivity that produce the best fit to measured hydraulic heads and net flow. Permeameter and column experiments produced laboratory measurements of hydraulic conductivity for each of the sands used in the intermediate-scale experiments. Despite explicit numerical representation of the heterogeneity the laboratory-measured values underestimated net flow by 12–14% and were distinctly smaller than the regression-estimated values. The significance of differences in measured hydraulic conductivity values was investigated by comparing variability of transport predictions using the different measurement methods to that produced by different realizations of the heterogeneous distribution. Results indicate that the variations in measured hydraulic conductivity were more important to transport than variations between realizations of the heterogeneous distribution of hydraulic conductivity.

1. Introduction

The heterogeneity of natural systems confounds attempts to achieve accurate groundwater flow and transport models. *Carrera* [1993, p. 30] comments,

> ... most of the differences between the actual behavior of solutes and that predicted by the advection-dispersion equation can be attributed to the spatial variability of hydraulic properties and, specifically, hydraulic conductivity....

In many instances hydraulic conductivity data consist of core-scale values collected at sparsely distributed locations. It is unclear how such core-scale values relate to larger scales, such as the scale of numerical grid blocks. As a result, considerable effort has been expended on model calibration techniques [e.g., *Hill*, 1998; *Hill et al.*, 1998; *Poeter and Hill*, 1997; *Sun and Yeh*, 1990; *Carrera and Neuman*, 1986; *Yeh*, 1986; *Cooley*, 1977, 1979] and determining equivalent parameters to represent the hydraulic properties of heterogeneous porous media at the grid block scale [e.g., *Renard and de Marsily*, 1997; *Rubin and Gomez-Hernandez*, 1990; *Rubin*, 1991; *Gelhar*, 1993; *Dagan*, 1989].

The focus on calibration and scaling issues has overshadowed a more fundamental issue. A rarely stated underlying assumption of calibration and scaling discussions is that heterogeneity explicitly represented, with parameter values applied at the same scale as they are measured, will produce accurate flow and transport predictions. An improved understanding of how measured, core-scale, hydraulic conductivity values relate to in situ values is needed to examine this assumption. Investigation of this assumption is not trivial and amounts to determining the predictability of measured parameter values, that is, in the absence of scaling and parameterization, how well the measured parameter values predict flow and transport. This requires controlled heterogeneous experiments with detailed observations of flow and transport and a complete description of the heterogeneous structure. Without this level of detail, scaling and parameterization issues can mask errors caused by parameter value predictability. Intermediate-scale laboratory experiments provide the only practical mechanism to evaluate this issue: Explicit knowledge of the distribution of material properties in natural heterogeneous systems is practically impossible, analytical solutions of flow and transport are limited to relatively simple configurations, and synthetic test cases can only compare results to a base simulation.

Previous laboratory investigations of flow and transport, beginning with *Darcy*'s [1856] simple column experiments, have gradually increased in complexity. Early heterogeneous experiments were primarily qualitative in nature [e.g., *Skibitzke and Robinson*, 1963], while later two- and three-dimensional evaluations quantified the impact of simple heterogeneity [e.g., *Silliman and Simpson*, 1987; *Silliman et al.*, 1987]. Two-dimensional, two-media experiments by *Wood et al.* [1994] and *Murphy et al.* [1997] compared observed and simulated results for transport through simple heterogeneities in a 1-m-long tank. Other recent efforts have focused on creating more complex heterogeneity with statistical properties similar to that found in natural systems [e.g., *Barth et al.*, 1996; *Chao et al.*, 1996; *Welty and Elsner*, 1997; *Silliman et al.*, 1998]. These laboratory investigations focused only on the observations of flow

[1]Department of Civil and Environmental Engineering, University of Colorado, Boulder, Colorado, USA.
[2]U.S. Geological Survey, Boulder, Colorado, USA.
[3]Division of Environmental Sciences and Engineering, Colorado School of Mines, Golden, Colorado, USA.

Copyright 2001 by the American Geophysical Union.

Paper number 2001WR000242.
0043-1397/01/2001WR000242$09.00

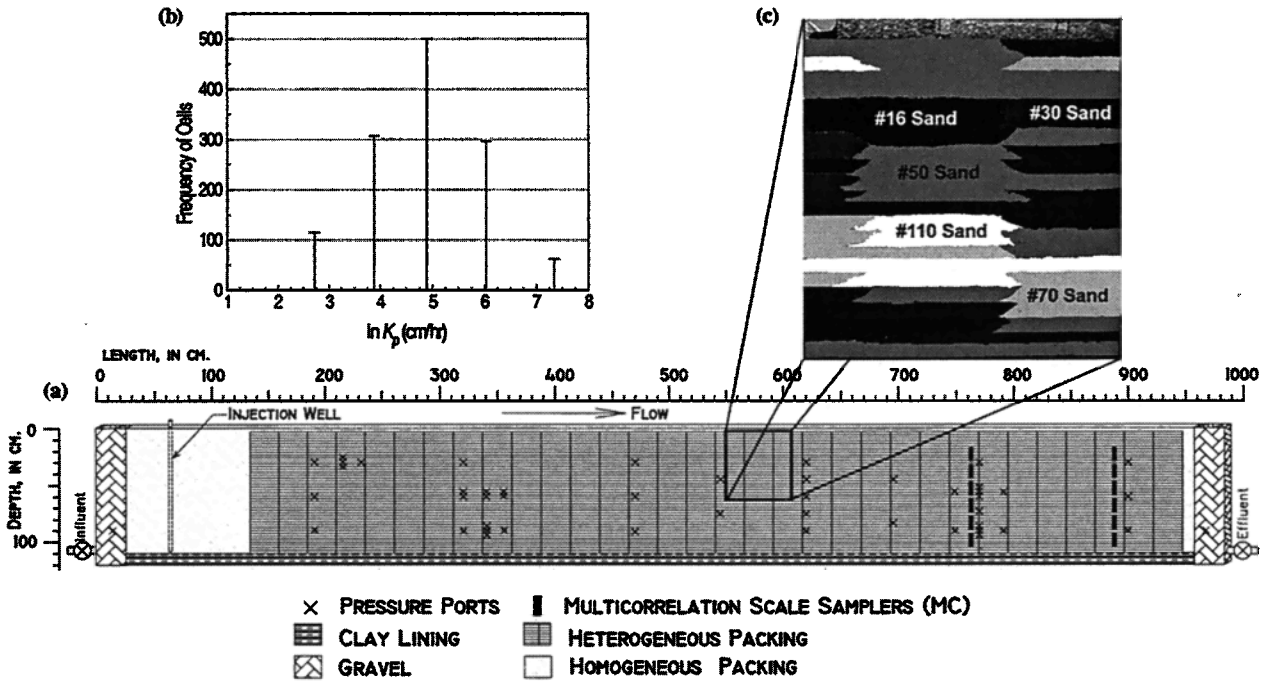

Figure 1. Intermediate-scale tank design showing (a) the dimensions of the entire tank including a simplified representation of the sand cell packing and the hydraulic head and sampling ports, (b) the frequency distribution of the five sands packed in the tank categorized using the natural log of their permeameter values, and (c) an enhanced gray-scale detail of the tank packing. Each sand lens is 2.54 cm tall and 25.4 cm long and is tapered at the ends as shown in Figure 1c.

and transport and did not include comparison of the observations to numerical predictions and were not large enough to prove that, for example, discrepancies were not a result of errors of constant head boundary conditions.

The work presented here is a critical step toward refining our understanding of how measured core-scale hydraulic conductivity values relate to in situ values and the importance of measured hydraulic conductivity variability compared to errors caused by inaccurate zonation of sedimentary features. A series of controlled, intermediate-scale tracer experiments are used to compare detailed flow, hydraulic head, and concentration observations to predictions simulated using measured hydraulic conductivity. Discrepancies are evaluated using sensitivity analysis and nonlinear regression using the methods of *Hill* [1998] and the finite difference groundwater flow model MODFLOWP [*Hill*, 1992]. The experiments were performed in a two-dimensional heterogeneous porous medium of sufficient correlation lengths to be statistically comparable to field site heterogeneity. The experiments provide a complex, explicitly characterized system that is simple enough to be controlled, definitive, and to allow explicit numerical representation yet complex enough to be relevant to field site heterogeneity. The results demonstrate significant variability in flow and transport predictions due to measurement method differences in hydraulic conductivity values and significant discrepancies between observations and predictions using the most relevant measured hydraulic conductivities. The significance of prediction variability, due to hydraulic conductivity measurement variability, is evaluated by comparing it to the variability of transport predictions from 150 realizations of the heterogeneous distribution.

2. Intermediate-Scale Experiments

This section describes the construction of the experimental facility, how hydraulic heads, flow through the tank, and concentrations were measured, the methods for measuring hydraulic conductivity, and the tracer experiments.

2.1. Porous Medium Construction

The intermediate-scale porous medium was constructed in a tank approximately 10 m long, 1.2 m tall, and 0.06 m inside width (Figure 1a). Each end of the porous medium consisted of a 20-cm section of pea gravel to provide constant head boundaries for the system. The overall gradient and saturated zone thickness were adjusted with a set of constant head tanks that controlled the water level in the pea gravel. The water table was level with the top of the sand packing at the upgradient end of the tank. At the downgradient end it was 15.7 cm below the top of the packing, producing an overall gradient of approximately 0.016. While higher than typical gradients in the field, the Reynolds number for flow within the porous medium was at least, and usually more than, an order of magnitude less than the lowest values associated with turbulent flow [*Bear*, 1972]. Deionized water was supplied to the upgradient constant head tank. The gradient and resulting flow of approximately 3.2 L h^{-1} were maintained throughout each experiment and the periods between experiments. Between experiments, NaOCl was added to the deionized water supply to produce a one pore volume pulse of 100 ppm NaOCl solution, eliminating the potential for significant microbial growth within the tank.

The packing within the tank consisted of two zones: a ho-

Table 1. Symbols Identifying Sets of Hydraulic Conductivity and Respective Values[a]

	Mesh Size (ASTM E-11)					
	#8[a]	#16[b]	#30[b]	#50[b]	#70[b]	#110[b]
K_p,[c] cm h^{-1}	NA	1550	417	133	48.6	15.1
K_{cl},[d] cm h^{-1}	NA	2148	674	111	74.2	22.8
K_c,[e] cm h^{-1}	6077	2250	708	136	84.7	23.0
K_{ch},[f] cm h^{-1}	NA	2360	780	165	92.5	23.2
K_r,[g] cm h^{-1}	NA	3170	716	156	104	45.1
d_{50},[h] mm	1.25	0.88	0.49	0.30	0.19	0.103
d_{60}/d_{10}[i]	1.56	1.72	1.50	1.94	1.86	~2.0

[a]Sand in the homogeneous zone.
[b]Sands used to create the heterogeneous zone.
[c]Measured using flexible wall permeameter.
[d]Lowest measured values using constant head column.
[e]Average measured values using constant head column.
[f]Highest measured values using constant head column.
[g]Values determined by regression.
[h]Fifty percent of grains are smaller.
[i]Uniformity coefficient (values <4.0 indicate uniform soil).

mogeneous section of coarse sand (#8 sieve) in the upstream 1.1 m of the tank followed by an 8.1-m heterogeneous section (Figure 1a). The heterogeneous zone served as the laboratory analogy of random field site sedimentary structure, was created using five different sands, and designed to support explicit representation in a numerical model. It was included to produce transport results with statistical properties similar to heterogeneous field sites. The heterogeneous zone approximated a lognormal distribution of hydraulic conductivity (K) with a mean value of 4.18 ($\mu_{\ln K}$) and a variance of 1.22 ($\sigma^2_{\ln K}$), where K has units of cm h^{-1}. Each lateral and vertical correlation scale was 50.8 and 5.08 cm, respectively. A continuous distribution with a negative exponential covariance was generated using a Fourier summation algorithm [*Shinozuka and Jan*, 1972] and then discretized into five categories. Each category was assigned a particular sieve-size sand: #16, #30, #50, #70, or #110 (Table 1). *Chao et al.* [1996] evaluated $\mu_{\ln K}$, $\sigma^2_{\ln K}$, and the correlation structure of the discretized distribution and verified that they matched the corresponding statistics of the original continuous distribution. The heterogeneous zone provided a region to inject the tracer and promote initial mixing as it exited the injection well, producing a relatively consistent vertical line source. The coarser sand, with a relatively high dispersivity, reduced the effect of microheterogeneities in the packing and the potential for variation from the injection well.

A consistent packing procedure was used for the entire tank; details of the packing procedure have been reported by *Barth* [1999] and *Barth et al.* [2001]. The sand was wet packed in the tank to minimize consolidation and air entrapment. Variations in tank thickness, due to variations in wall thickness and wall separation, precluded using consistent volumes of sand to fill each cell. A packing grid, drawn on both sides of the tank exterior, guided the packing so that each cell in the heterogeneous section measured 25.4 cm long and 2.54 cm tall. Variations in tank thickness, as discussed in section 3.1, were incorporated into the flow and transport analysis. A total of 1280 cells were packed in the heterogeneous section: 32 columns and 40 layers producing 16 lateral and 20 vertical correlation scales. Vertical interfaces were avoided to reduce the chance for preferential migration of nonaqueous phase liquids used during other experiments (Figure 1c).

2.2. Injection Well and Samplers

Tracer was injected using a specially designed injection well described in detail by *Barth* [1999]. The well consisted of a 0.635-cm-diameter stainless steel tubing injector with 24 evenly spaced outlets covered by #200 stainless steel mesh inserted into a fully screened, 2.54-cm-diameter PVC casing. Packers mounted on the stainless steel tubing insured delivery of the tracer into the casing over the injection interval. The line source injected for the reported experiments was 36 cm tall with the top located at a depth of 34 cm. The vertical source interval insured that, despite relatively small vertical dispersion, the plume's vertical dimension spanned multiple correlation scales when it entered the heterogeneous section. The resulting tracer plume reflected the overall transport characteristics of the hydraulic conductivity distribution rather than a specific portion of the packed realization. Dye injections were performed in the tank to verify creation of a line source that was uniform both in the vertical direction and across the width of the tank. However, creation of a tracer plume with sufficient visual contrast to document its progress throughout the 10-m-long tank was outside of the objectives of this investigation.

Multicorrelation scale samplers (MC) were designed to reduce the potential for local hydraulic conductivity to dominate sample constitution [*Barth*, 1999]. The MCs provide a sample across approximately two vertical correlation scales but do not have the inherent dissolution and excessive dead space of a well installed from the top of the packing. This is especially important given the potential for disrupting the flow field with large-volume samples. Each sampler consists of a vertical 10-cm-long perforated 0.32-cm-diameter piece of copper tubing wrapped with #200 stainless steel mesh. One end of the tubing is sealed, and the other is bent 90° and attached to a bulkhead compression fitting, allowing sampling through the tank wall. The sampler requires extraction of only about 1.5 mL to provide a 1-mL sample. Two transects of six MC samplers each (Figure 1a), labeled MC1 (top) to MC6 (bottom) and MC7 (top) to MC12 (bottom) for transects 1 and 2, respectively, were installed. The samplers cover all but the top 10 cm and bottom 5 cm of the porous media. The saturated thickness above MC1 and MC7 is on the order of a few centimeters

Table 2. Tracer Injections

Experiment	Tank Effluent Rate, L h^{-1}	Injection Interval, hour	Injection Volume, L	Tracer	Concentration, molarity or dpm	$\Delta\rho/\rho_0$
C7	3.31	0.91	2.8	bromide	0.002	0.0001
C8	3.26	0.92	2.7	bromide	0.002	0.0001
C9	3.03	0.92	2.8	bromide	0.002	0.0001
D1	3.30	0.93	2.8	tritium	2.94×10^8	0.0000

Disintegrations per minute, dpm.

because of the location of the sampler transect along the tank (Figure 1a) in relation to the sloping water table. The sampler interval, because of its length, did not correspond to an exact integer number of sand cells. In addition, neither the top nor the bottom of each MC were positioned exactly at the interfaces between sand cells.

2.3. Hydraulic Conductivity Measurements

Table 1 summarizes the sets of hydraulic conductivity values for the six different Tyler mesh sieve-size sands used in this work. The measured values were obtained using a flexible wall permeameter (American Society for Testing and Materials (ASTM) D 5084-90) and a constant head column (ASTM D2434-68, 93) [*American Society for Testing and Materials*, 1994]. The constant head column was packed using the same method as the intermediate-scale tank. The permeameter's flexible walls eliminate the potential for wall effects and the 8.9-cm-diameter constant head column was at least 50 times the mean grain diameter, exceeding the ASTM D2434 recommended minimum column diameter by a factor of 8–12. Flexible wall permeameter samples were approximately 4–5 cm in length, while the constant head column values (K_c) are from 20- or 40-cm separation hydraulic head measurements in a vertical 90-cm column; thus K_c represents hydraulic conductivity measurements of column lengths close to the length of the lenses in the tank and under conditions of similar effective stress. Hydraulic head measurements along the 90-cm column revealed no significant trend in hydraulic conductivity as a function of depth [*Barth*, 1999].

The permeameter values (K_p) are from J. Mapa et al. (Upscaling of water and solute transport in saturated porous media, unpublished report for the U.S. Army Engineers Waterways Experiment Station, 1994) (hereinafter referred to as Mapa et al., unpublished report, 1994), who report only a single measured value for each sand. The constant head column evaluations were conducted as part of the present study and were repeated from 3 to 20 times to evaluate variability. Coefficients of variation for the column-measured values of hydraulic conductivity ranged from 0.04 to 0.11. The variability is reported using three sets of values: K_{cl}, K_c, and K_{ch}, consisting of the lowest, average, and highest constant head column-measured values, respectively. K_c was determined by taking the arithmetic average of the individual K measurements.

The values of conductivity for the different mesh-size sands span more than 2 orders of magnitude (Table 1). The sands evaluated were considered uniform because, based on the manufacturer's specifications, each sand satisfied the criteria of having a uniformity coefficient (d_{60}/d_{10}) of less than 4.0. Comparison of the column-evaluated values of hydraulic conductivity to those produced in a 1-m-long, two-dimensional tank, where flow was parallel to any potential packing-induced microheterogeneities, indicated that the individual mesh-size sands were isotropic.

The differences between K_p and K_c are attributed to large differences (~50 kPa) in the effective stress applied to the sample and possibly the difference in sample size. The variation among the column values is attributed to differences in packing despite concerted efforts to avoid such differences. The column values are measured under conditions similar to those in the intermediate-scale tank and were expected to be closest to the in situ values. The variability of the column measurements is likely to be reproduced in the tank, and it was anticipated that the average column values, K_c, would be closest to the in situ values.

2.4. Tracer Experiments

A total of four tracer injections referred to as C7, C8, C9, and D1 were performed under very similar conditions, as listed in Table 2. Aspects of some of the experiments listed were discussed by *Barth et al.* [2001], and the experiment names used here are consistent with the names used in that work. The data sets from each experiment consisted of the flow rate leaving the tank, or effluent rate, hydraulic head at 46 ports, and concentrations at the two sampling transects (Figure 1a). Head measurements from pressure scans prior to filling the tank with sand, with only a static water column, were very consistent having a coefficient of variation of 0.001 and no discernable trend. During the experiments, repeated measurements over time at a single port typically had a coefficient of variation of 0.003. Each pressure scan of the tank included a scan of a known reference hydraulic head, allowing detection of and correction for any potential drift in the pressure transducer readings. The standard deviation of effluent rate, determined from repeated samples while maintaining a constant gradient, was ~0.068 L h^{-1} or equivalent to a coefficient of variation of 0.02.

For each experiment the injection rate was ~3.0 L h^{-1}, just slightly less than the nominal tank effluent rate, to avoid flow field disruption. Samples were collected every 4 hours or approximately every 0.08 pore volumes until roughly 3.5 pore volumes had passed through the tank. Samples from the experiments using potassium bromide (KBr) were analyzed using an ion selective electrode. To verify the absence of density effects during C7, C8, and C9, the fourth tracer test (D1) was conducted using tritium. Tritium samples were analyzed with a liquid scintillation counter.

3. Numerical Simulations

3.1. Numerical Modeling of Flow: Forward Simulations and Regression Analysis

The finite difference groundwater flow model MODFLOWP [*Hill*, 1992] was used to simulate steady state hydraulic head and flow in the tank. The free surface in the tank was represented with a no-flow boundary that approximated the free surface elevation; testing using a calculated free surface indicated little error from the approximation, as expected given the steady state flow field. To simplify data input, the finite difference grid was oriented vertically so that depth in the single layer corresponded to thickness of the two-dimensional packing. This made it possible to represent the two-dimensional tank as a single layer of 40 rows and 150 columns, for a total of 6000 finite difference cells, without any loss in accuracy of the numerical simulation. Each finite difference cell was ~2.5 cm tall and 6.4 cm long so that each 2.5-cm-tall, 25.4-cm-long sand cell was represented by four finite difference cells. Variations in the tank width, from 5.02 to 6.60 cm, were mapped by measuring tank wall deflection at 306 points and kriging the values to produce a detailed mapping of tank width variations. Repeated measurements on a reference cross section indicated that the deflection measurements had a coefficient of variation of <0.01. The tapered ends of the sand cells shown in Figure 1b were not represented, but simulations with the grid refinement increased by a factor of 50 [*Mehl*, 1998] indicated the resulting error in flow through the tank is less than 1.5%. The upstream and downstream ends of the tank were represented as constant heads. Large manual perturbations of the upgradient and downgradient constant head boundaries by +0.17 cm and −0.17 cm, respectively, which is more than 2 times the typical hydraulic head observation standard deviation, increased the simulated effluent by only 1%. Hydraulic conductivity values that produced the best fit to measured heads and effluent from the tank were determined using MODFLOWP [*Hill*, 1992].

Inverse flow modeling was performed because, as noted in section 4.1, simulations using laboratory-measured hydraulic conductivities did not reproduce flow and hydraulic head, and as a result, concentration observations, as well as expected. Inverse modeling and associated sensitivity analyses were performed as described by *Hill* [1998]. Simultaneous regression of the four hydraulic head and effluent data sets produced optimal values of hydraulic conductivity for the five sands in the heterogeneous packing reported as K_r in Table 1. Hydraulic conductivity of the homogeneous section, the #8 sieve-size sand, was not estimated because of insensitivity. Observation weights are based on the variability about mean observation values quantified as a standard deviation of 0.08 cm for hydraulic head and 21 $cm^3\ h^{-1}$ for effluent. The values of weights used, since they are proportional to the variance-covariance matrix of the observation measurement errors, are the most appropriate values for indicating the importance of each observation used in the regression and producing parameter estimates with the smallest possible variance [*Hill*, 1998, p. 45].

Composite scaled sensitivities, correlation coefficients, calculated error variance, and linear confidence intervals on the parameter estimates [*Hill et al.*, 1998; *Hill*, 1998] were used to appraise parameter estimates. Composite scaled sensitivities indicate the amount of information provided by the observations for each parameter. Two dimensionless parameters, the percent discrepancy and the flow-relevant scaled discrepancy, are used to quantify the difference between measured and regression values of hydraulic conductivity. The percent discrepancy is calculated as in (1):

$$\varepsilon = \frac{K_m^j - K_r^j}{K_r^j}(100). \tag{1}$$

The m indicates measurement type (Table 1), and j refers to the sand sieve number. Linear 95% confidence intervals for the percent discrepancy are determined by using the ±2 standard deviation values of K_r^j in (1) to produce the corresponding high and low values of percent discrepancy. Equation (2) defines flow-relevant scaled discrepancies. It indicates the potential effect of differences between measured and regression values of hydraulic conductivity on flow predictions as a flow-relevant scaled discrepancy:

$$\varepsilon^* = \frac{(K_m^j - K_r^j)}{\bar{Q}} \frac{\partial Q}{\partial K_m^j}(100), \tag{2}$$

where $\bar{Q}$ is the mean observed effluent for the four experiments and $\partial Q/\partial K_m^j$ is the sensitivity of calculated flow through the tank to K_m^j evaluated at the measured value of K using MODFLOWP. Nonlinearity of the term $\partial Q/\partial K_m^j$ will affect the accuracy of (2); however, as mentioned in section 4.1, this term did not vary markedly for the different sets of K values.

3.2. Numerical Modeling of Transport

Transport was simulated using the modular three-dimensional transport model MT3DMS [*Zheng*, 1998] which uses the flows generated by MODFLOWP. The third-order, total variation diminishing, solver was used. Single values of porosity and dispersivity for each sand were reported by Mapa et al. (unpublished report, 1994) and *Szlag* [1995], respectively. As expected for the granular, silica sands used, porosity was very consistent across the five mesh sizes, and dispersivity increased with increasing grain size. Tracer injection was represented as an initial concentration in two adjacent columns of finite difference cells which corresponded to the height of the injection interval and the width of the source immediately following the injection period.

Simulated concentrations from finite difference cells approximating the location of each MC sampler were integrated to provide simulated breakthrough curves (BTCs). Both simulated and observed solute transport BTCs were integrated by combining the flux-weighted concentration from each sampler in each of the two transects. Simulated concentrations for each set of K values were weighted using the respective simulated flux. Observed concentrations from the physical experiments were weighted with the flux values from the K_r flow simulations. Discrepancies between observed and predicted transport were evaluated by analyzing the temporal moments of the integrated BTC from each transect. The nth absolute temporal moment (M_n) is defined as

$$M_n = \int_0^\infty t^n C(x, t)\, dt, \tag{3}$$

where t is time and $C(x, t)$ is concentration as a function of space and time. The normalized absolute nth moment (m_n) is obtained by dividing M_n by M_0,

$$m_n = M_n/M_0, \tag{4}$$

and μ_n represents the nth normalized central moment:

$$\mu_n = \frac{\int_0^\infty (t - m_1)^n C(x, t)\, dt}{M_0}. \quad (5)$$

For this paper, m_1 and μ_2 were evaluated for each BTC and are referred to simply as the first and second moments, respectively. The first and second moments provide summaries of the mean arrival time and the amount of tracer plume spreading, respectively, for the measured and regression values of hydraulic conductivity. These summaries do not capture all the subtleties of the tracer BTC but provide an efficient method of quantifying differences in transport results.

3.3. Simulating Flow and Transport for the 150 Realizations of the Heterogeneous Distribution

The regression-estimated K values closely reproduce the measured transport. Lack of significant discrepancy meant that a sensitivity analysis and regression were not needed. However, discrepancies between the measured and regression-estimated values of hydraulic conductivity produce significant differences in simulated transport. Of concern, is the importance of the variability produced by measured values of K, relative to other common types of variability. The variability of transport results due to different realizations of the heterogeneous distribution is used as a measure for that induced by the different hydraulic conductivity measurement methods. Differences in transport results between realizations can be used as an analogy for the differences expected from repeated experiments at various locations in a stationary, heterogeneous aquifer or the errors associated with improper zonation of heterogeneous, sedimentary features. Packing more than one realization of the heterogeneous distribution in the intermediate-scale tank was not practical, instead forward flow and transport was simulated in different realizations of the hydraulic conductivity distribution. Variability of simulated flow and transport across the different realizations represents the variability expected for a given distribution of materials and provides a baseline against which the variability in transport predictions due to the different hydraulic conductivity measurement methods is compared.

One hundred fifty realizations of the heterogeneous packing were generated using a Fourier summation algorithm and then discretized using the K_r values to produce 150 discretized realizations with $\mu_{\ln K} = 5.33$ and $\sigma^2_{\ln K} = 1.07$, where K is in cm h^{-1}. Transport was simulated in the 150 realizations to generate BTCs, and the results were used to estimate the ensemble average transport and its 95% linear confidence intervals. Variability of the results as a function of the hydraulic conductivity measurement method was compared to the 95% confidence intervals for transport produced by the different realizations.

4. Results

4.1. Comparison of Measured and Predicted Hydraulic Head, Flow, and Transport

The results presented focus on the discrepancy between simulated and observed flow rates and the resulting implications for transport simulations. As would be expected for a two-dimensional system constrained by constant head boundaries, head observations were only a minor influence on the regression. The contribution to the total sum-of-squared weighted residuals ($SSWR_{total}$) from the error in head predictions amounted to only 2.3%, 4.6%, and 8.0% of $SSWR_{total}$ using the K_p, K_{cl}, and K_{ch} values, respectively. The weighted head residuals became a significant influence on the regression only when the K values were close to the K_r values. Changing from the K_p values to the K_r values reduced the $SSWR_{head}$ by a factor of <2, while improving $SSWR_{flow}$ by a factor of >300.

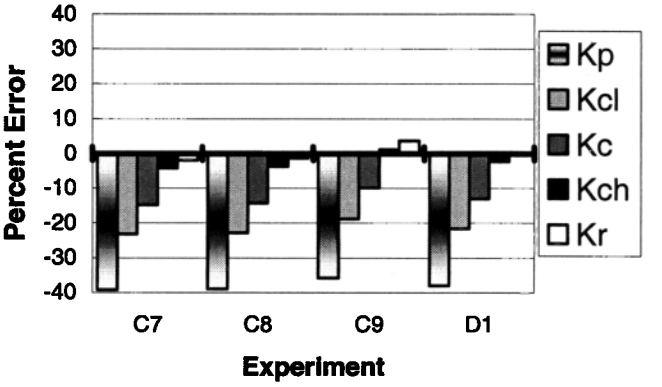

Figure 2. Percent error of simulated tank effluent compared to that observed for data sets C7, C8, C9, and D1 for permeameter values (K_p); low (K_{cl}), mean (K_c), and high (K_{ch}) values of column measurements; and values estimated by the regression (K_r).

The percent errors in predicted effluent, compared to measured rates, are shown in Figure 2. Despite the controlled nature of these experiments, effluent and transport predictions using the hydraulic conductivity values measured by permeameter and the mean of the column-measured values were considerably different from observations. Surprisingly, only the highest column-measured values closely reproduced the measured flow. Predictive simulations using K_p and K_{cl} values underpredicted the effluent rate by about 40% and 20%, respectively. Using K_c values, flow was underpredicted by 9.9–14.7%, with a mean value of 13.0%. Predictions using the K_{ch} values had a mean error of −2.2% (Figure 2).

For the controlled laboratory conditions presented in this work, the discrepancy in flow using K_c was expected to be much smaller than the mean value of 13.0%. The relatively large discrepancy prompted the following question: What values of hydraulic conductivity would reproduce the observations? It was not possible to measure directly hydraulic conductivity of the individual lenses in situ; however, sensitivity analysis and nonlinear regression provide an indirect method of obtaining estimates of in situ values. The sensitivity analysis can identify parameters that can be estimated given the observations, in this case heads and flows; the regression hydraulic conductivity values (K_r) reproduce the observed heads and effluent. Not unexpectedly, they also reproduced the observed concentrations.

Composite scaled sensitivities from the sensitivity analysis (Figure 3) are based on simultaneous consideration of the hydraulic head and effluent data from all four data sets. The fact that sensitivities are all within an order of magnitude of each other and all correlation coefficients are less than 0.88 indicates that optimal values of hydraulic conductivity for each sand can be estimated [*Hill*, 1998, p. 38]. Regression results confirm this; regression values (K_r) are listed in Table 1.

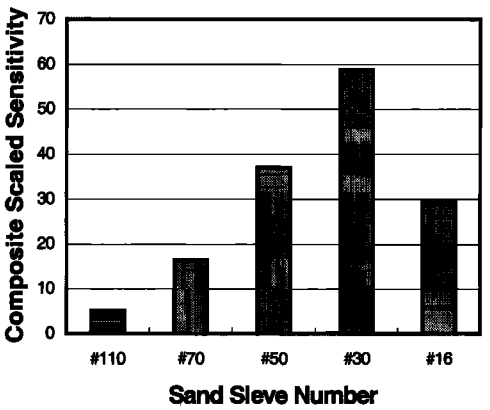

Figure 3. Composite scaled sensitivities of the five different sieve-size sands based on simultaneous optimization of hydraulic conductivity for C7, C8, C9, and D1.

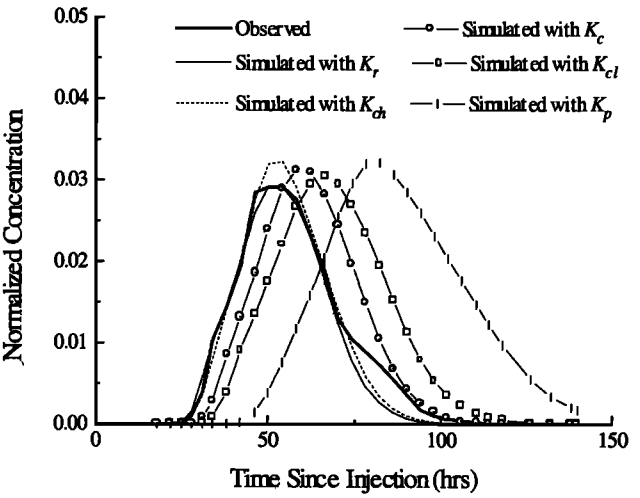

Figure 5. Observed and predicted, using K_r, K_{ch}, K_c, K_{cl}, and K_p, second transect breakthrough curves (BTCs) from experiment C7.

The discrepancies between measured and regression hydraulic conductivities are investigated to determine whether a single method of measurement or sieve-size sand dominated the differences between predictions and observations. Figure 4a shows the percent discrepancy, (1), between regression and measured values of hydraulic conductivity. Linear 95% confidence intervals on the discrepancies are presented. For all permeameter-measured values of hydraulic conductivity the discrepancies are significant. For all measurement methods the percent discrepancies for the #110 and #16 sands were significant. For the #70, #50, and #30 sands the confidence intervals on the methods producing larger values, such as K_{ch}, tend to include zero, so that the corresponding regression values are not significantly different from the measured values.

Also of interest is how much these discrepancies contributed to accurately reproducing net flow. This is evaluated using the flow-relevant scaled discrepancies, (2). Results are shown in Figure 4b. The values of $\partial Q/\partial K_c^j$, calculated using K_c^j, were 3.5, 4.6, 6.1, 1.5, and 0.1 cm^2 for the #110, #70, #50, #30, and #16 sands, respectively. Sensitivities calculated using the other hydraulic conductivity measurement methods were similar.

Although magnitudes of (1) for the #110 and #16 sands are significant for all measurement methods and are large relative to values for other size sands (Figure 4a), examination of the flow-relevant scaled discrepancies (Figure 4b) reveals that the contribution to the flow prediction error is fairly consistent across all sieve-size sands for a particular method of measurement. This analysis suggests that no one K change dominated the improved fit. Regression simulations optimizing only one parameter at a time supported this conclusion. The fit between simulated and observed values for single-parameter regression was significantly worse than for the simultaneous optimization. Each of the individually optimized values, which are not reported here, were significantly different from their respective K_r values.

4.2. Transport Variability Caused by Variability of the Measured Hydraulic Conductivity

The BTCs in Figure 5, from simulations of the second transect during C7, illustrate the typical impact of different hydraulic conductivity measurement methods on simulated transport. The K_r values produced using head and flow observations provide the best match between simulated and observed BTCs. None of the simulated BTCs, however, reproduce the shoulder exhibited in the tail of the observed BTC

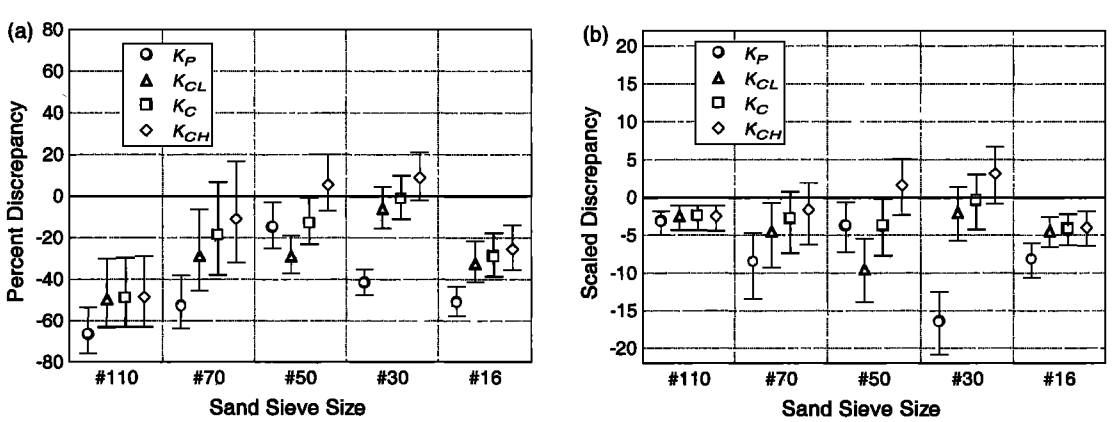

Figure 4. Differences, expressed as (a) percent (equation (1)) and (b) flow-relevant scaled discrepancies (equation (2)) between optimized (K_r) and measured (K_p, K_{cl}, K_c, and K_{ch}) hydraulic conductivity values for each of the five sands. Error bars reflect the linear 95% confidence intervals of K_r values.

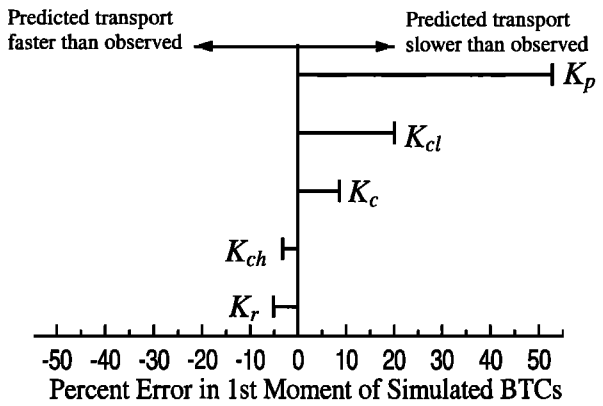

Figure 6. Discrepancies between observed and predicted first moment of second transect in experiment C7. Predictions are based on measured hydraulic conductivity (K_p, K_{cl}, K_c, and K_{ch}) and the values of hydraulic conductivity that best fit flow and hydraulic head observations (K_r).

(Figure 5). Adjustment of other parameters also failed to improve the match between simulated and observed BTCs. For example, increasing local dispersivity simply increased spreading and smoothing of the BTC, decreasing the possibility of simulating details such as the shoulder in the BTC tail. It is hypothesized that, even though the resolution was sufficient to explicitly represent the heterogeneity, it was not enough to reproduce some of the more subtle features that may affect solute transport. For example, additional refinement of the finite difference grid would improve representation of the samplers and allow incorporation of the sloped interfaces between horizontally adjacent sand cells. Refining representation of these features may enable the simulation to capture features such as the shoulder in the BTC tail.

Applying (4) to the BTCs in Figure 5 produces a range of simulated m_1 that reflect the variation in measured hydraulic conductivity. Figure 6 summarizes the error in m_1: the difference between the m_1 values from the BTCs simulated using different hydraulic conductivity measurement methods versus the value from the observed BTC. As mentioned in section 3.2, summary statistics such as m_1 and μ_2 do not capture all the subtleties; in Figure 5 the BTC simulated using K_r matches the observed BTC best, but because of the unreproduced shoulder the K_{ch}-predicted m_1 is closest to the observed m_1 (Figure 6).

4.3. Transport Variability Between Realizations of the Hydraulic Conductivity Distribution

Simulation of transport using K_r values in 150 realizations of the heterogeneous distribution provided perspective for the large variations in transport predictions caused by the differences in measured hydraulic conductivity. Figure 7 plots the simulated BTCs from Figure 5 reflecting the variation caused by the differences in measured values of hydraulic conductivity. Also shown are BTCs from 150 realizations of the correlated random field using K_r. The variability of simulated BTCs is of the same order for both sources of variation in transport.

The variability depicted in Figure 7 can be summarized and quantified by plotting the average m_1 (Figure 8a) and μ_2 (Figure 8b) from simulated BTCs for all four experiments. The top five sets of values depicted in each graph reflect the difference between the moments calculated from the observed BTCs and the simulated BTCs based on K_p, K_{cl}, K_c, K_{ch}, and K_r. Error bars signify the range of values over the four experiments. The bottom set of values in each graph depict the mean moments of simulated transport using K_r in 150 realizations of the distribution. The error bars for these values indicate the 95% confidence intervals. The range of first and second moments due to measurement-method variability is on the same order as

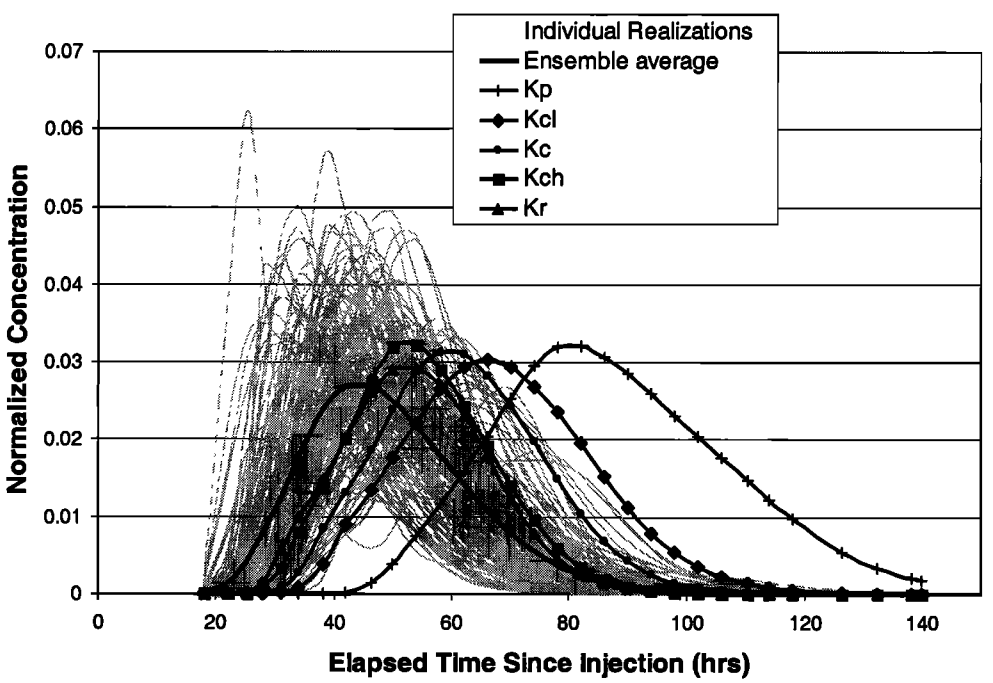

Figure 7. Simulated BTCs at second transect using experiment C7 boundary conditions. Impact of realization variability is compared to differences in measured value of hydraulic conductivity. Unlabeled shaded lines represent BTCs for individual realizations.

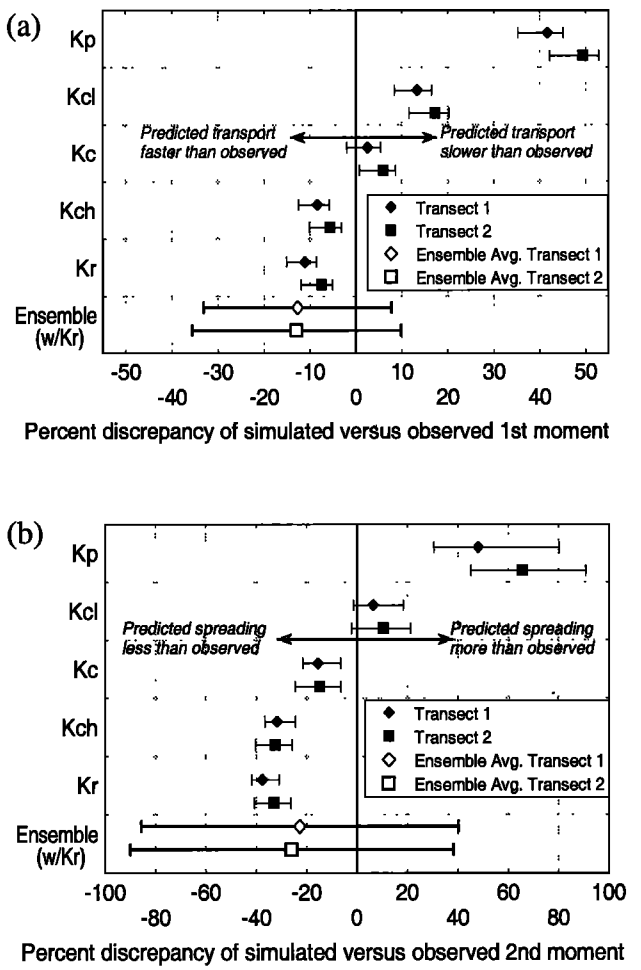

Figure 8. Discrepancies between observed and predicted BTC (a) first moments and (b) second moments for each set of measured hydraulic conductivity values and the ensemble average of the distribution using K_r. Error bars on the ensemble average indicate 95% confidence intervals based on the 150 realizations. All other error bars indicate the low and high values derived using measured moments for the four experiments.

that produced by between-realization variations in sedimentary structure.

5. Discussion/Conclusions

This investigation, especially because of its size and complexity, provides unique insight into processes that cannot be controlled or evaluated at field sites. The experiments, simulations, and analysis produced a unique perspective on our understanding of sedimentary structures, measured hydraulic conductivities, and their role in controlling flow and transport through porous media.

The experiments presented were designed to provide results free from the effects of scaling and parameterization. The data presented illustrate limitations on the application of laboratory-measured hydraulic conductivity values to predictive modeling of heterogeneous systems. Despite careful construction of the porous media and the detailed information available, tank effluent predictions using mean column-measured hydraulic conductivity were 13% less than the observed. Surprisingly, the highest column-measured values most closely reproduced the observed effluent rate and solute transport. For the mean constant head column values the scaled discrepancies were comparable for all sands. These results indicate that the measured values of each sand made similar contributions to the underprediction of flow. This indication is also supported by the fact that residuals from regression runs optimizing individual sieve-size sands were significantly greater than those from the simultaneous regression of all sands. The consistent underprediction of tank effluent suggests that, in addition to scaling issues, measured hydraulic conductivity predictability can contribute significantly to the problem of determining equivalent parameter values.

Typically, the reported mean and variance of a heterogeneous distribution represent the magnitude and variability of sedimentary structures but do not represent the variations that occur for either repeated measurements or use of alternate measurement methods on a single sedimentary structure. Even under the ideal conditions of the reported experiments the variability of flow and transport predictions, as a function of the hydraulic conductivity measurement methods, was significant compared with that produced between realizations of a heterogeneous distribution of hydraulic conductivity. The results show that the variability in measured values of hydraulic conductivity contributes as much or more to the uncertainty in groundwater model simulations as the random variations between realizations of the heterogeneous distribution. This suggests that the statistical parameters summarizing a heterogeneous distribution should be reported with confidence intervals that reflect the variability of hydraulic conductivity measurements.

The experiments and simulations illustrate the importance of regression-estimated parameter values. Flow and hydraulic head calibrated values of hydraulic conductivity provided transport predictions superior to those produced using either of the measurement methods. Between-realization variability of flow and transport, used as an analogy for errors because of inaccurate zonation of sedimentary features, was of the same order as that resulting from measured hydraulic conductivity value variability. The results suggest that because of the variability of measured values the accuracy of hydraulic conductivity measurements is less important than determining the zonation of sedimentary structures. With sufficient boundary condition information, observations, and proper zonation of the heterogeneous sedimentary structures, regression-estimated values of hydraulic conductivity will produce more accurate predictions than those based on measured hydraulic conductivity values.

Acknowledgments. Financial support was provided by the U.S. Geological Survey, U.S. EPA Great Plains/Rocky Mountain Hazardous Substance Research Center, and the U.S. Army Research Office at Research Triangle Park, North Carolina. Comments from Vincent Tidwell, Steffen Mehl, Tarek Saba, and William Yeh were greatly appreciated.

References

American Society for Testing and Materials, *1994 Annual Book of ASTM Standards*, vol. 04.08, West Conshohocken, Pa., 1994.

Barth, G. R., Intermediate-scale tracer tests in heterogeneous porous media: Investigation of density, predictability and characterization of NAPL entrapment, Ph.D. dissertation, 360 pp., Dep. of Civ. Eng., Univ. of Colo., Boulder, 1999.

Barth, G. R., T. H. Illangasekare, and H. Rajaram, Use of tracer techniques to characterize nonaqueous phase liquid entrapment in heterogeneous aquifers (abstract), *Eos Trans. AGU*, *77*(46), Fall Meet. Suppl., F243, 1996.

Barth, G. R., T. H. Illangasekare, M. C. Hill, and H. Rajaram, Analysis of intermediate-scale tracer experiments for the development of tracer density guidelines in heterogeneous porous media, *Water Resour. Res.*, *37*(1), 21–31, 2001.

Bear, J., *Dynamics of Fluids in Porous Media*, Elsevier Sci., New York, 1972.

Carrera, J., An overview of uncertainties in modelling groundwater solute transport, *J. Contam. Hydrol.*, *13*(1–4), 23–48, 1993.

Carrera, J., and S. P. Neuman, Estimation of aquifer parameters under transient and steady state conditions, 3, Application to synthetic and field data, *Water Resour. Res.*, *22*(2), 228–242, 1986.

Chao, H.-C., H. Rajaram, and T. H. Illangasekare, Experimental and computational investigation of radial flow tracer tests in heterogeneous formations (abstract), *Eos Trans. AGU*, *77*(46), Fall Meet. Suppl., F275, 1996.

Cooley, R. L., A method of estimating parameters and assessing reliability for models of steady state groundwater flow, 1, Theory and numerical properties, *Water Resour. Res.*, *13*(2), 318–324, 1977.

Cooley, R. L., A method of estimating parameters and assessing reliability for models of steady state groundwater flow, 2, Application and statistical analysis, *Water Resour. Res.*, *15*(3), 603–617, 1979.

Dagan, G., *Flow and Transport in Porous Formations*, Springer Verlag, New York, 1989.

Darcy, H., *Les Fontaines Publiques de la Ville de Dijon*, Dalmont, Paris, 1856.

Gelhar, L. W., *Stochastic Subsurface Hydrology*, Prentice-Hall, Englewood Cliffs, N. J., 1993.

Hill, M. C., A computer program (MODFLOWP) for estimating parameters of a transient, three-dimensional groundwater flow model using nonlinear regression, *U.S. Geol. Surv. Open File Rep.*, *91-484*, 358 pp., 1992.

Hill, M. C., Methods and guidelines for effective model calibration, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, *98-4005*, 90 pp., 1998.

Hill, M. C., R. L. Cooley, and D. W. Pollock, A controlled experiment in ground water flow model calibration, *Ground Water*, *36*(3), 520–535, 1998.

Mehl, S., Experimental and numerical investigations of intermediate-scale experiments, M.S. thesis, Univ. of Colo., Boulder, 1998.

Murphy, E. M., T. R. Ginn, A. Chilakapati, C. T. Resch, J. L. Phillips, T. W. Weitsma, and C. M. Spadoni, The influence of physical heterogeneity on microbial degradation and distribution in porous media, *Water Resour. Res.*, *33*(5), 1087–1103, 1997.

Poeter, E. P., and M. C. Hill, Inverse models: A necessary next step in ground-water modeling, *Ground Water*, *35*(2), 250–260, 1997.

Renard, P., and G. De Marsily, Calculating equivalent permeability: A review, *Adv. Water Resour.*, *20*(5–6), 253–278, 1997.

Rubin, Y., Transport in heterogeneous porous media: Prediction and uncertainty, *Water Resour. Res.*, *27*(7), 1723–1738, 1991.

Rubin, Y., and J. J. Gomez-Hernandez, Stochastic approach to the problem of upscaling of conductivity in disordered media: Theory and unconditional numerical simulations, *Water Resour. Res.*, *26*(4), 691–701, 1990.

Shinozuka, M., and C.-M. Jan, Digital simulation of random processes and its applications, *J. Sound Vib.*, *25*(1), 111–128, 1972.

Silliman, S. E., and E. S. Simpson, Laboratory evidence of the scale effect in dispersion of solutes in porous media, *Water Resour. Res.*, *23*(8), 1667–1673, 1987.

Silliman, S. E., L. F. Konikow, and C. I. Voss, Laboratory investigation of longitudinal dispersion in anisotropic porous media, *Water Resour. Res.*, *23*(11), 2145–2151, 1987.

Silliman, S. E., L. Zheng, and P. Conwell, The use of laboratory experiments for the study of conservative solute transport in heterogeneous porous media, *Hydrogeol. J.*, *6*(1), 166–177, 1998.

Skibitzke, H. E., and G. M. Robinson, Dispersion in ground water flowing through heterogeneous materials, *U.S. Geol. Surv. Prof. Pap.*, *386-B*, 1963.

Sun, N. Z., and W. W. G. Yeh, Coupled inverse problems in groundwater modeling, 1, Sensitivity analysis and parameter identification, *Water Resour. Res.*, *26*(10), 2507–2525, 1990.

Szlag, D. C., Dissolution of nonaqueous phase liquids in sandy aquifer materials, Ph.D. dissertation, 215 pp., Univ. of Colo., Boulder, 1997.

Welty, C., and M. M. Elsner, Constructing correlated random fields in the laboratory for observations of fluid flow and mass transport, *J. Hydrol.*, *202*(1–4), 192–211, 1997.

Wood, B. D., C. N. Dawson, J. E. Szecsody, and G. P. Streile, Modeling contaminant transport and biodegradation in a layered porous medium system, *Water Resour. Res.*, *30*(6), 1833–1845, 1994.

Yeh, W. W.-G., Review of parameter identification procedures in groundwater hydrology: The inverse problem, *Water Resour. Res.*, *22*(2), 95–108, 1986.

Zheng, C. M., A modular three-dimensional multispecies transport (MT3DMS) model, *Release DoD_3.00.A*, Waterw. Exp. Stn., U.S. Army Corps of Eng., Vicksburg, Miss., 1998.

G. R. Barth and H. Rajaram, Department of Civil and Environmental Engineering, University of Colorado, Boulder, CO 80309-0428, USA. (grbarth@usgs.gov; hari@colorado.edu)

M. C. Hill, U.S. Geological Survey, 3215 Marine Street, Boulder, CO 80303-1066, USA. (mchill@usgs.gov)

T. H. Illangasekare, Division of Environmental Sciences and Engineering, Colorado School of Mines, Golden, CO 80401-1887, USA. (tillanga@mines.edu)

(Received August 1, 2000; revised February 7, 2001; accepted July 9, 2001.)