AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2014WR016202

# Multiscale solute transport upscaling for a three-dimensional hierarchical porous medium

**Mingkan Zhang[1] and Ye Zhang[1]**

[1]Department of Geology and Geophysics, University of Wyoming, Laramie, Wyoming, USA

**Key Points:**
- Three-dimensional solute transport
- Multiscale solute transport modeling

**Correspondence to:**
M. Zhang,
mzhang1@uwyo.edu
Y. Zhang,
yzhang9@uwyo.edu

**Citation:**
Zhang, M., and Y. Zhang (2015), Multiscale solute transport upscaling for a three-dimensional hierarchical porous medium, *Water Resour. Res.*, *51*, 1688–1709, doi:10.1002/2014WR016202.

Received 26 JUL 2014
Accepted 11 FEB 2015
Accepted article online 18 FEB 2015
Published online 23 MAR 2015
Corrected on 7 MAY 2015

This article was corrected on 7 MAY 2015. See the end of the full text for details.

**Abstract** A laboratory-generated hierarchical, fully heterogeneous aquifer model (FHM) provides a reference for developing and testing an upscaling approach that integrates large-scale connectivity mapping with flow and transport modeling. Based on the FHM, three hydrostratigraphic models (HSMs) that capture lithological (static) connectivity at different resolutions are created, each corresponding to a sedimentary hierarchy. Under increasing system ln$K$ variances (0.1, 1.0, 4.5), flow upscaling is first conducted to calculate equivalent hydraulic conductivity for individual connectivity (or unit) of the HSMs. Given the computed flow fields, an instantaneous, conservative tracer test is simulated by all models. For the HSMs, two upscaling formulations are tested based on the advection-dispersion equation (ADE), implementing space versus time-dependent macrodispersivity. Comparing flow and transport predictions of the HSMs against those of the reference model, HSMs capturing connectivity at increasing resolutions are more accurate, although upscaling errors increase with system variance. Results suggest: (1) by explicitly modeling connectivity, an enhanced degree of freedom in representing dispersion can improve the ADE-based upscaled models by capturing non-Fickian transport of the FHM; (2) when connectivity is sufficiently resolved, the type of data conditioning used to model transport becomes less critical. Data conditioning, however, is influenced by the prediction goal; (3) when aquifer is weakly-to-moderately heterogeneous, the upscaled models adequately capture the transport simulation of the FHM, despite the existence of hierarchical heterogeneity at smaller scales. When aquifer is strongly heterogeneous, the upscaled models become less accurate because lithological connectivity cannot adequately capture preferential flows; (4) three-dimensional transport connectivities of the hierarchical aquifer differ quantitatively from those analyzed for two-dimensional systems.

## 1. Introduction

As a central topic of subsurface hydrology, solute transport in aquifers has been investigated for several decades. Various theories have been proposed, including the macrodispersion approach based on the advection-dispersion equation (ADE) [e.g., *Gelhar and Axness*, 1983; *Dagan*, 1989] and alternative or nonlocal formulations, e.g., dual and multicontinuum models [*Cushman and Ginn*, 1993; *Neuman*, 1993; *Harvey and Gorelick*, 2000; *Haggerty et al*., 2001], continuous time random walk [*Berkowitz and Scher*, 1998; *Dentz et al*., 2004; *Berkowitz et al*., 2006], moment equation [*Neuman*, 1993, 2003; *Neuman and Tartakovsky*, 2009], projector formalism [*Cushman and Ginn*, 1993; *Cushman et al*., 2002], fractional ADE [*Meerschaert et al*., 1999; *Benson et al*., 2000], and the stochastic-convective approach [*Cvetkovic et al*., 1996; *Cirpka and Kitanidis*, 2000; *Ginn*, 2001]. Simultaneous with theory development, field tracer tests have been conducted at various locations with the explicit goal of testing one or more theories, e.g., Borden, Canada; Mobile, Alabama; Twin Lake, Minnesota; Cape Cod, Massachusetts, Mirror Lake, New Hampshire; and the Macrodispersion Experiment (MADE) at Columbus, Mississippi. At these sites, aquifer hydraulic conductivity ($K$) ranges from relatively homogeneous, e.g., the Cape Cod site with a $\sigma_f^2$ (variance of ln$K$) of 0.24, where $K$ varies by approximately 1 order of magnitude [*Garabedian and Leblanc*, 1991], to highly heterogeneous, e.g., the MADE site has a $\sigma_f^2$ ranging from 4.5 [*Rehfeldt et al*., 1992] to 6.0 [*Bohling et al*., 2012], where $K$ varies from 4 to 6 orders of magnitude depending on the measurement methods. As shown by the MADE tracer experiments, when $\sigma_f^2$ is high, solute transport exhibits anomalous behavior such as non-Fickian breakthrough curves (i.e., asymmetry and heavy tails) and scale dependence in the observed dispersivity.

In nearly 3 decades, several tracer experiments at the MADE site have been interpreted with theories. Although nonlocal theories can often provide a better match [*Zheng et al*., 2011], ADE was found to capture

the main solute plume characteristics if site heterogeneity is adequately resolved [*Barlebo et al*., 2004; *Salamon et al*., 2007]. When dispersivity was modeled as time-dependent [*Liu et al*., 2008], tritium plume predicted with ADE improves over that of an earlier study simulating the same experiment using a constant dispersivity [*Feehley et al*., 2000]. Assuming low-*K* inclusion within a high-*K* matrix, *Fiori et al*. [2013] capture both the arrival and tailing of the tracer tests with an ADE-based formulation. In these cases, the ability of ADE to capture transport likely improved because of the enhanced degrees of freedom in representing dispersion, as it is well known in regression that models with increased parameterization can provide improved fits to observations. In simulating transport in a synthetic aquifer, ADE was also found to capture bulk transport statistics when dispersivity was modeled as spatially or temporally variable [*Zhang and Gable*, 2008]. Field studies at Lauswiesen, Germany, further suggest that when multiple facies are explicitly accounted for, ADE can provide predictions similar to a mass-transfer model [*Riva et al*., 2008, 2010]. By comparison, nonlocal theories (including the above cited mass-transfer model) often invoke additional transport parameters besides dispersivity, while physical meaning of these parameters and their interpretation using field data can become difficult to reconcile. As an example, among the nonlocal theories, those with more parameters in their formulations can usually provide better fits, e.g., a space fractional ADE with spatially variable dispersivity improves over a constant-dispersivity variant of the same model [*Zhang et al*., 2007]. The above observations suggest that all theories with increased parameterization can potentially provide improved fits to data, although ADE is the most parsimonious model.

The previous paragraph highlights an ongoing debate over the issue of transport formulation for describing tracer movement in heterogeneous aquifers. What makes this issue challenging is the complexity and uncertainty of subsurface heterogeneity, as field and laboratory studies can suffer from data scarcity, measurement support effects, and sampling or scale limitation. (There is also the issue of process uncertainty, as transient flows due to aquifer recharge can impact dispersion—this topic will be left to another treatment.) In sedimentary aquifers, heterogeneity of hydrological properties (notably, *K*) forms in response to dynamic processes, e.g., sea level changes, basement subsidence, sediment supply variation, etc. As a result, conductivity heterogeneity exists at multiple scales, exhibiting nested structures within a lithological hierarchy, i.e., sediment laminas nested within larger scale beds that are further organized into facies or facies assemblages [e.g., *Fogg*, 1990; *Scheibe and Freyberg*, 1995; *Webb and Anderson*, 1996; *Anderson*, 1997; *Scheibe and Yabusaki*, 1998; *Labolle and Fogg*, 2001; *Klise et al*., 2008; *Milliken et al*., 2008; *Comunian et al*., 2011]. At even greater scales, facies assemblages can give rise to aquifers and aquitards that are frequent targets in field-scale contamination studies. At a given problem scale (often containing one or more sediment hierarchy), heterogeneity can also exhibit long-range correlation or "connectivity," which can give rise to preferential flows or barriers that lead to early arrival and long tail in the solute breakthrough curve [*Wen and Gómez-Hernández*, 1998; *Zinn*, 2003; *Knudby and Carrera*, 2005]. At the MADE site, sedimentological and geophysical mapping confirm the existence of long-range facies connectivity that extends beyond the tracer test domain [*Bianchi et al*., 2011; *Meerschaert et al*., 2013]. Thus, on the one hand, hierarchical deposits have been explicitly modeled by combining large-scale lithological variations with the analysis of smaller scale heterogeneities [e.g., on the other hand, by explicitly modeling connectivities at a given problem scale, factors influencing preferential flow and transport can be evaluated [*Zheng and Gorelick*, 2003; *Salamon et al*., 2007; *Bianchi et al*., 2011]. Various connectivity definitions have been proposed by prior workers [*Renard and Allard*, 2013]. In this work, we define connectivities as static (i.e., those related to lithological variation) and dynamic (i.e., those related to flow and transport characteristics).

Despite the advances made in developing theory and field characterization techniques, practical problem solutions require efficient transport simulators. Because resolving heterogeneity down to the smallest resolvable continuum scale is impractical, when problem scale is large and site data are limited, flow and transport simulators must necessarily ignore heterogeneity at one or more smaller scales (or hierarchies). This necessity motivates the development of upscaling theories of flow and transport that develop deterministic or stochastic solutions of the governing equations to derive formulations with effective or equivalent parameters in order to capture bulk flow and transport behaviors arising out of the underlying heterogeneity [*Dagan*, 1989; *Gelhar*, 1993; *Dagan and Neuman*, 1997; *Cushman et al*., 2002; *Vogel and Roth*, 2003; *Frippiat and Holeyman*, 2008; *Fiori et al*., 2013]. Two parameterizations are common in upscaling: (1) heterogeneity is populated throughout the problem domain by assigning to each grid cell a grid-effective conductivity and dispersivity that can account for flow and transport behaviors arising out of the subgrid

heterogeneities [e.g., *Efendiev et al*., 2000; *Rubin*, 2003; *Fernàndez-Garcia and Gómez-Hernández*, 2007]; (2) heterogeneity is represented by a series of homogeneous facies or aquifers/aquitards units, ignoring subunit variation (hydrostratigraphic models, or HSMs). The first parameterization is frequently associated with geostatistical modeling, e.g., creating conditional realizations that are then coarsened to create a simulation model. The second parameterization, frequently invoked in larger scale studies lacking detailed field data, is more amenable to developing hierarchical models [*Monsen et al*., 2005; *Milliken et al*., 2008; *Ramanathan et al*., 2010; *Li et al*., 2011]. Upscaling for HSMs is the focus of this work.

Based on a three-dimensional (3-D) experimental stratigraphy exhibiting heterogeneity and connectivity at different scales, this study evaluates the ADE-based macrodispersion theory for upscaling transport with a suite of HSMs that captures lithological (static) connectivity at different sedimentary hierarchy. Equivalent conductivities are first computed for individual connectivity (or unit) of the HSMs. Transport upscaling is then carried out: in the heterogeneous reference model (i.e., experimental stratigraphy), ADE is assumed applicable; in the HSMs, ADE is parameterized with macrodispersivities to represent solute spreading due to unresolved, subunit-scale heterogeneity. The macrodispersion theory is chosen because it is the most widely studied, has shown utility for capturing field transport, and has matured to an extent that its underlying assumptions are generally understood [*Rubin*, 2003]. By comparison, upscaling theory for nonlocal formulations is in its infancy, while practical applicability of many nonlocal methods still remains to be demonstrated.Because the experimental stratigraphy was a unique outcome of a physical sedimentation experiment, it is not amenable to a geostatistical (i.e., multiple-realization) treatment. Flow and transport upscaling of this study is thus deterministic and adoption of the macrodispersion theory necessitates the working assumption that ensemble predictions can be applied to evaluating single or unique experiment. The ergodicity assumption has been adopted by prior workers when addressing issues with connectivity, i.e., a large problem domain many times the ln$K$ correlation range cannot be established [*Desbarats and Srivastava*, 1991; *Adams and Gelhar*, 1992; *Fernàndez-Garcia et al*., 2004, 2005; *Zhang and Gable*, 2008]. This assumption is necessary where long-range correlation exists in the field or in the laboratory, as the nature of the transport condition cannot be changed to suit theory. In evaluating the experimental stratigraphy, this study thus follows prior research in model conceptualization while focusing on the influence of connectivity resolution on upscaling at increasing system ln$K$ variances (0.1, 1.0, 4.5; $K$ varies over 1, 3, and 6 orders of magnitude, respectively). The effect of explicit connectivity modeling on predicting non-Fickian transport by the ADE-based upscaled models is examined, yielding insights into why ADE can capture transport in some situations (e.g., site heterogeneity is moderate and/or modeled at high resolution capturing dominant connectivity), but not in others (e.g., sites with strong variability and/or models ignore or incorporate insufficient connectivity resolution). Because few hierarchical aquifer data have been analyzed for transport, this study also examines several dynamic connectivity measures in comparison to prior values reported in the literature.

In the remainder of this article, the heterogeneous reference model is introduced first, followed by a description of the procedure used to construct the HSMs. Flow and transport upscaling methods are presented next, followed by the results section which consists of: (1) flow upscaling and verification; (2) transport upscaling and verification; (3) dynamic connectivities. Insights of this study are summarized before conclusion and future research is indicated.

## 2. Methods

### 2.1. Model Creation

This study is based on stratigraphic data of an experimental deposit that exhibits hierarchical heterogeneity corresponding to physical sedimentation, i.e., fine-scale variations unique to local depositional processes are nested within larger scale, sand/clay transitions that formed in response to global forcings [*Sheets et al*., 2002; *Zhang et al*., 2010]. Compared to geostatistically generated models that are often devoid of connectivity [*Zinn*, 2003], the experimental data exhibit hierarchical structure not dissimilar to that observed at the MADE site [*Zheng et al*., 2011], e.g., sand-rich channel-like features embedded in floodplain clay, facies proportion and sizes exhibit spatial persistence, i.e., static connectivity. By scanning the sediment at high resolution and aligning the images in 3-D, a fully heterogeneous reference model (FHM) is created. This model is scaled to field dimensions ($L_x \times L_y \times L_z = 2500$ m $\times$ 2500 m $\times$ 100 m, of $N_x = 100$, $N_y = 100$, $N_z = 40$) to

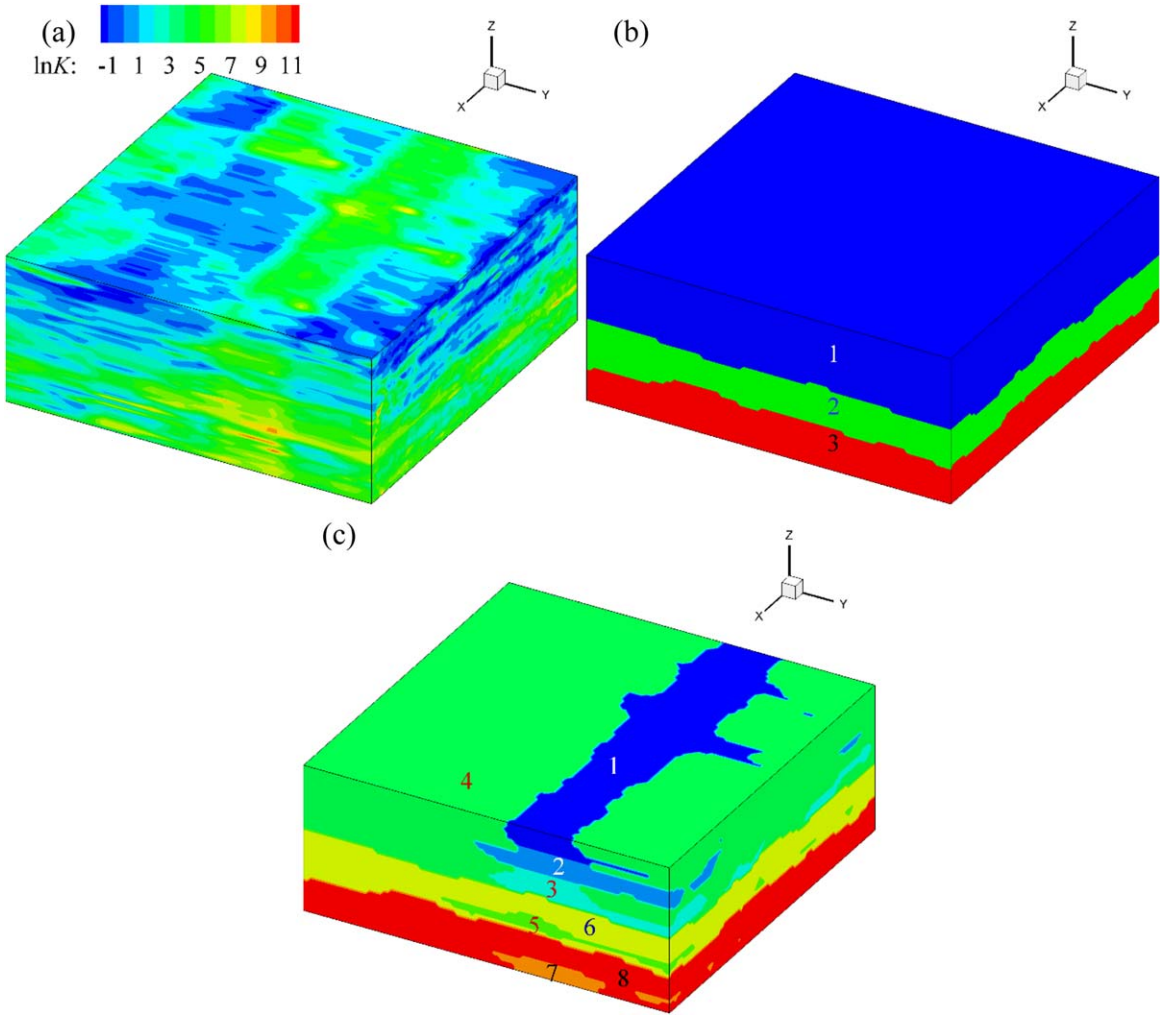


**Figure 1.** (a) The reference fully heterogeneous model withlocal *K* variation in natural log scale (system $\sigma_f^2$ is 4.5). *K* varies over 6 orders of magnitude in this model. Connectivity mapping at two scales: (b) 3-unit model, and (c) 8-unit model, while units of the latter model are nested within units of the former model.

create a synthetic aquifer with *K* variation mirroring sedimentation (Figure 1a). Each local *K* is assumed isotropic to facilitate transport upscaling with a well-known theory (see section 2.4).

To create the HSMs, lithology of the FHM was mapped to capture distinct facies (8-unit model; Figure 1c) before the facies were grouped into depositional environments (3-unit model; Figure 1b), which were further grouped into a single aquifer unit (1-unit model) (see detail in *Zhang et al.* [2011]). Because this procedure was carried out honoring lithological principles, the mapped connectivity is static. In practice, such models can be built using logging or geophysical data that are sensitive to formation lithology. The static connectivity is captured at different resolutions, in effect reflecting different levels of field characterization that can be carried out in practice. For each connectivity (or unit) of the HSMs, a geostatistical analysis estimated subunit ln*K* mean, $\sigma_f^2$, and integral scales (Table 1). This analysis was then repeated after increasing system ln*K* variance while fixing the mean ln*K*. At a given system variance, as connectivity resolution increases, subunit $\sigma_f^2$ becomes

**Table 1.** Integral Scales of ln*K* (m) for the HSM Units

| Models | ID | $\lambda_x$ | $\lambda_y$ | $\lambda_z$ |
|---|---|---|---|---|
| 1-unit | | 155.0 | 690.0 | 5.0 |
| 3-unit | 1 | 155.0 | 610.0 | 3.8 |
| | 2 | 140.0 | 900.0 | 12.0 |
| | 3 | 200.0 | 900.0 | 12.0 |
| 8-unit | 1 | 155.0 | 610.0 | 3.5 |
| | 2 | 135.0 | 320.0 | 3.0 |
| | 3 | 135.0 | 700.0 | 3.4 |
| | 4 | 155.0 | 250.0 | 3.0 |
| | 5 | 155.0 | 275.0 | 7.5 |
| | 6 | 155.0 | 720.0 | 15.0 |
| | 7 | 230.0 | 300.0 | 3.0 |
| | 8 | 330.0 | 350.0 | 16.0 |

**Table 2.** Equivalent Hydraulic Conductivity Principal Components (m/yr) Computed With Flow-Based Upscaling[a]

| | | $\sigma_f^2 = 0.1$ | | | | | $\sigma_f^2 = 1.0$ | | | | | $\sigma_f^2 = 4.5$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Models | ID | Var(lnK) | $K_{xx}$ | $K_{yy}$ | $K_{zz}$ | $\alpha_L^M$ | Var(lnK) | $K_{xx}$ | $K_{yy}$ | $K_{zz}$ | $\alpha_L^M$ | Var(lnK) | $K_{xx}$ | $K_{yy}$ | $K_{zz}$ | $\alpha_L^M$ |
| 1-unit | | 0.1 | 40.97 | 40.10 | 38.15 | 15.5 | 1.0 | 60.74 | 50.32 | 33.43 | 155.00 | 4.5 | 230.41 | 113.82 | 30.64 | 697.50 |
| 3-unit | 1 | 0.078 | 34.13 | 33.30 | 32.63 | 12.09 | 0.784 | 32.56 | 26.03 | 22.54 | 121.52 | 3.528 | 58.63 | 24.81 | 17.35 | 546.84 |
| | 2 | 0.066 | 48.97 | 47.89 | 46.47 | 9.24 | 0.662 | 96.54 | 78.90 | 66.40 | 92.68 | 2.979 | 475.42 | 216.91 | 172.62 | 417.06 |
| | 3 | 0.063 | 46.26 | 45.46 | 44.83 | 12.60 | 0.629 | 80.15 | 68.17 | 60.43 | 125.80 | 2.831 | 330.03 | 178.35 | 121.00 | 566.06 |
| 8-unit | 1 | 0.046 | 42.10 | 41.85 | 40.90 | 7.13 | 0.458 | 55.66 | 18.81 | 24.57 | 70.99 | 2.061 | 123.78 | 102.95 | 53.39 | 319.46 |
| | 2 | 0.048 | 43.65 | 43.20 | 42.63 | 6.48 | 0.482 | 63.88 | 26.22 | 47.88 | 65.07 | 2.169 | 179.05 | 129.85 | 72.34 | 292.82 |
| | 3 | 0.037 | 44.04 | 43.63 | 42.63 | 5.00 | 0.373 | 62.67 | 28.86 | 61.92 | 50.36 | 1.679 | 148.15 | 116.01 | 72.97 | 226.62 |
| | 4 | 0.046 | 29.59 | 29.55 | 28.48 | 7.13 | 0.464 | 18.66 | 26.33 | 46.51 | 71.92 | 2.088 | 13.46 | 14.17 | 3.78 | 323.64 |
| | 5 | 0.027 | 65.34 | 64.97 | 61.98 | 4.19 | 0.266 | 212.08 | 96.26 | 19.76 | 41.23 | 1.197 | 1715.69 | 1401.05 | 676.38 | 185.54 |
| | 6 | 0.050 | 45.55 | 44.91 | 43.64 | 7.75 | 0.504 | 72.99 | 71.82 | 23.40 | 78.12 | 2.268 | 227.64 | 141.98 | 84.52 | 351.54 |
| | 7 | 0.044 | 59.96 | 60.12 | 56.89 | 10.12 | 0.437 | 171.45 | 80.88 | 17.99 | 100.51 | 1.967 | 1288.16 | 1530.46 | 304.19 | 452.30 |
| | 8 | 0.055 | 44.49 | 43.96 | 43.35 | 18.15 | 0.546 | 68.33 | 69.90 | 18.32 | 180.18 | 1.967 | 208.41 | 135.81 | 95.29 | 810.81 |

[a]For all HSM units, system variance ($\sigma_f^2$) and unit-specific $\alpha_L^M$ (m) predicted by a first-order theory are also listed.

progressively smaller (Table 1). Finally, besides lithology, other static measures can also be used to map connectivity, e.g., percolation thresholds [*Hunt*, 2001; *Zhang et al*., 2010] and geostatistics-based ranking metrics [*Deutsch*, 1998].

### 2.2. Flow Upscaling

Flow upscaling was conducted to compute equivalent conductivity (**K***) for each unit of the HSMs [*Zhang et al*., 2011]: (1) the FHM was simulated solving the steady state incompressible flow equation under different global flow boundary conditions (BC), (2) for each unit, mean flux and hydraulic gradient were computed for each BC; (3) by imposing Darcy's Law using mean fluxes and gradients from all BC, **K*** was computed with a least-squares method. Given the system variances investigated ($\sigma_f^2 = 0.1, 1.0, 4.5$), three sets of upscaled **K*** were obtained for each HSM (Table 2). To eliminate numerical errors associated with grid coarsening, the HSMs employ the same grid as the FHM. Accuracy of the upscaled **K*** can be assessed by comparing flow predictions of the HSMs against those of the reference model under the same global flow BC (see 3.1).

### 2.3. Transport Modeling

One steady state flow experiment imposing a lateral hydraulic gradient along the *x* axis is selected for the transport analysis. Hydraulic heads of the left and right sides of the model are assigned constant values of 100 m and 10 m, respectively, and all other boundaries are no-flow. Flow is driven from the inflow boundary ($x$=0 m) to the outflow boundary ($x$=2500 m). Within the flow field, the release of a dilute, conservative tracer is simulated by all models (the computational domain and simulation grid are the same as the flow model). Given the *K* range (which changes with system variance), the simulated lateral transport is advection-dominated, and diffusion is ignored. In all models, advective-dispersive transport is simulated: local dispersion is assumed for the FHM; macrodispersion for the HSMs. Longitudinal macrodispersivity is estimated for the latter, as this parameter can be linked to geostatistical parameters of the underlying heterogeneity and because, for flow parallel to stratification, longitudinal spreading is significant compared to transverse spreading. Below, the transport formulation is briefly explained.

The continuum-scale ADE that describes transport of a dilute, conservative solute in groundwater is:

$$\frac{\partial c}{\partial t} = \nabla \cdot (\mathbf{D}\nabla c) - \nabla \cdot \left(\frac{\mathbf{q}}{\theta} c\right) \tag{1}$$

where $c$ is solute concentration, **D** is dispersion tensor, **q** is Darcy flux, and $\theta$ is effective porosity. Groundwater velocity (**v**) is calculated as $\mathbf{q}/\theta$. In this work, $\theta$ is given a uniform value of 0.25, typical for sand-clay systems [*Freeze and Cherry*, 1979]. **D** is expressed as [*Bear*, 1988]:

$$\mathbf{D} = (\alpha_T V + D_d)\mathbf{I} + (\alpha_L - \alpha_T)\mathbf{vv}/V \tag{2}$$

where $D_d$ is effective porous medium diffusion coefficient ($D_d = 0$ in this work), **I** is the identity matrix, $\alpha_L$ and $\alpha_T$ are longitudinal and transverse dispersivities, respectively, and $V$ is magnitude of the groundwater velocity.

Equation (1) can be discretized to create a random walk particle tracking (RWPT) equation of the form [*Tompson and Gelhar*, 1990; *Kitanidis*, 1994]:

$$\mathbf{x}_p(t_n)=\mathbf{x}_p(t_{n-1})+\left[\mathbf{v}(\mathbf{x}_p(t_{n-1}),t_{n-1})+\nabla\cdot\mathbf{D}(\mathbf{x}_p(t_{n-1}),t_{n-1})\right]\Delta t+\mathbf{B}(\mathbf{x}_p(t_{n-1}),t_{n-1})\cdot\mathbf{z}_n\sqrt{\Delta t} \tag{3}$$

where $\mathbf{x}_p(t_n)$ is particle position at time $t_n$, $\mathbf{v}(\mathbf{x}_p(t_n), t_n)$ is groundwater velocity at $\mathbf{x}_p(t_n)$ and $t_n$, $\mathbf{D}(\mathbf{x}_p(t_n), t_n)$ is the dispersion tensor at $\mathbf{x}_p(t_n)$ and $t_n$, $\Delta t$ is time step, and $\mathbf{z}_n$ is a vector of independent random variables of standard normal distribution. $\mathbf{B}(\mathbf{x}(t_n), t_n)$ is related to $\mathbf{D}(\mathbf{x}_p(t_n), t_n)$ as follows:

$$\mathbf{D}(\mathbf{x}(t_n),t_n)=\frac{1}{2}\mathbf{B}(\mathbf{x}_p(t_n),t_n)\cdot\mathbf{B}^{\mathrm{T}}(\mathbf{x}_p(t_n),t_n) \tag{4}$$

At the subgrid level, $\mathbf{v}(\mathbf{x}_p(t_n), t_n)$ is sampled from a set of semi-analytic functions that accurately delineate streamlines using bilinear interpolation of the interface Darcy fluxes [*Pollock*, 1988]. Under higher ln$K$ variances, divergence of dispersion could be significant in regions with high velocity gradients, equation (3) is thus solved for all models with $\nabla\cdot\mathbf{D}$ evaluated at every time step using finite differences.

For a given system variance, at the start of a RWPT simulation, 100,000 particles are released at $x = 300$ m, their positions uniformly distributed within a $60 \times 775$ m$^2$ area ($z \in [25, 85]$ m and $y \in [900, 1675]$ m). To compare results of the FHM and the upscaled models (where groundwater velocity is increasingly smoothed as connectivity resolution is reduced), the same tracer test is simulated by all models. To conserve mass in computing the tracer moments (next), particle tracking is terminated when the fastest particle reaches the outflow boundary. The remaining no-flow boundaries are set to reflect the particles back if any attempts to cross them. The time step size is chosen following this rule: at least 10 time steps are required for a particle to move through the fastest-flowing cell, thus particle motion is resolved at high temporal resolution throughout the simulation time [*Prickett et al.*, 1981; *Tompson et al.*, 1987]. Because of the explicit need to compute $\nabla\cdot\mathbf{D}$ at every time step and the fact that a large number of particles is needed to avoid artificial fluctuations in the computed breakthrough curve [*Salamon et al.*, 2006], a parallel RWPT code based on Message Passing Interface is developed, significantly reducing computation time. For a 3-D test problem simulating 100,000 particles, for example, the parallel code achieves 14× speedup using 32 processors. All simulations were run on the Yellowstone Supercomputer at the National Center for Atmospheric Research (Computational & Information Systems Lab).

For each tracer simulation, moment analysis is conducted to investigate the time evolution of solute mass centroid, average plume velocity, and spreading of the solute around its mass centroid. These moments are defined as:

$$M=\iiint_{\Omega}(\theta c)dxdydz \tag{5}$$

$$\mathbf{I}_p=\frac{1}{M}\iiint_{\Omega}(\mathbf{x}_p\theta c)dxdydz \tag{6}$$

$$\mathbf{s}^2=\frac{1}{M}\iiint_{\Omega}(\mathbf{x}_p-\mathbf{I}_p)(\mathbf{x}_p-\mathbf{I}_p)\theta c dxdydz \tag{7}$$

where $\Omega$ is the computational domain, $M$ is total solute mass, $\mathbf{I}_p$ is mean displacement of solute mass centroid, and $\mathbf{s}^2$ is spatial covariance describing spreading of the solute about its centroid. The mean velocity of the plume ($\mathbf{v}_p$) is computed as [*Burr et al.*, 1994]:

$$\mathbf{v}_p=\frac{d\mathbf{I}_p}{dt} \tag{8}$$

The above central moments cannot reveal plume's higher moments such as skewness and early and late-time behavior. Solute breakthrough curve (BTC) is thus calculated at two downstream control planes at $x = 500$ m and $x = 1000$ m. Over time, the number of particles crossing each plane is counted to create a set of vertically integrated BTC.

### 2.4. Transport Upscaling

For the reference heterogeneous model, hydrodynamic dispersivities are assigned to equation (3) to represent local dispersion. As $\alpha_T$ is usually much less than $\alpha_L$ [*Bear*, 1988], $\alpha_L$ is set to $10^{-2}$ m and $\alpha_T$ is

ignored. Alternatively, local dispersion can be ignored entirely because its effect on large-scale solute spreading is usually negligible [*Rubin*, 2003; *Zhang and Gable*, 2008; *Fiori et al*., 2013]. To simulate transport in the HSMs, two upscaling formulations are tested: (1) unit-specific macrodispersivity: $\alpha_L^M$ is estimated for each unit of the HSMs to represent solute spreading due to unresolved, subunit-scale velocity variation, (2) time-dependent macrodispersivity: $\alpha_L^M$ is obtained from moment analysis of the tracer plume simulated by the FHM. For the upscaled models, because $\alpha_L^M$ is spatially or temporally variable, an enhanced degree of freedom in representing dispersion (i.e., as compared to using a single dispersivity) is employed. Moreover, because classic theories predict that macrodispersion scales with ln$K$ variance, transport upscaling is carried out at each system variance. Below, both approaches are briefly described.

In the first approach, a first-order stochastic theory is implemented to estimate spatially varying macrodispersivity [*Gelhar and Axness*, 1983; *Rubin*, 2003; *Fiori et al*., 2013]:

$$\alpha_L^M = \sigma_f^2 \lambda \tag{9}$$

$$\alpha_T^M = 0 \tag{10}$$

where $\lambda$ is ln$K$ integral scale along mean flow direction (i.e., $x$ axis), $\sigma_f^2$ is local ln$K$ variance describing subunit variability. As shown in Table 1, most HSM units are weakly to moderately heterogeneous, and, as connectivity resolution is increased (i.e., from 1 to 3 to the 8-unit models), *local* ln$K$ variances decrease. In equation (9), ln$K$ integral scales transverse to flow are not used, as statistical anisotropy is found to have minor effects on predicting BTC under similar transport conditions [*Zarlenga et al*., 2012; *Fiori et al*., 2013]. Tracer's initial vertical dimension (60 m) is large compared to vertical ln$K$ integral scales of the various connectivities (3∼16 m), satisfying theory requirement of a large plume in direction orthogonal to flow [*Dagan*, 1989; *Rubin*, 2003]. Additional simulations were carried out in the FHM using a larger initial plume size (70×1025 $m^2$), with results suggesting that tracer moments are not significantly affected. However, it must be pointed out that ergodicity as a working assumption cannot be strictly proven. Moreover, for the lateral transport regimes analyzed, transverse macrodispersivity is assigned zero. At all system variances, transverse expansion of the plume, as simulated by the FHM, is found negligible over time (not shown).

In the second approach, an apparent longitudinal macrodispersivity is obtained by analyzing tracer simulation in the FHM [*Dagan*, 1989], which can be traced to Einstein's original time-dependent dispersion concept [*Einstein*, 1905]:

$$\alpha_L^M(t) = \frac{1}{2}\frac{s_{xx}^2(t) - s_{xx}^2(0)}{\mathbf{I}_p(t) - \mathbf{I}_p(0)} \tag{11}$$

With equation (11), a time-dependent $\alpha_L^M$ can be obtained by fitting a function to the apparent $\alpha_L^M$ of the FHM experiment (the apparent $\alpha_T^M$ is again set to zero). This approach is analogous to situations where solute plume is monitored at high resolution with detailed measurements. Instead of evaluating $\alpha_L^M$ for every particle at every time step (as is done with the unit-specific approach, where $\alpha_L^M$ is evaluated depending on the position of each particle), the apparent $\alpha_L^M$ is assigned to all particles regardless of their positions. For a given system variance, the apparent $\alpha_L^M$ obtained from analyzing the FHM tracer test is assigned to all HSMs, regardless of their differences in connectivity resolution.

Because velocity simulated by the FHM becomes more strongly variable under increasing system variance, tracer experiment is simulated for 200 years, 40 years, and 4 years by all models, corresponding to system $\sigma_f^2 = 0.1$, 1.0 and 4.5, respectively.

### 2.5. Flow and Transport Connectivity

At each system variance, select flow and transport connectivity measures are computed for the FHM and each of the upscaled models. Unlike the static connectivity (i.e., lithology mapping) that defines the HSMs, dynamic connectivities can be strongly influenced by $\sigma_f^2$ and thus the variability of the velocity field. For fluid flow connectivity, the definition of *Knudby and Carrera* [2005] is chosen because it can potentially represent preferential flows:

$$CI = \frac{K_{xx}^{eff}}{K_G} \tag{12}$$

where $CI$ is a flow connectivity index, $K_G$ is geometric mean of the local conductivity, and $K_{xx}^{eff}$ is an *effective* conductivity of the bulk flow field:

$$K_{xx}^{eff} = \frac{QL_x}{(h_1 - h_2)A} \tag{13}$$

where $Q$ is bulk flow rate across the outflow boundary over $L_y \times L_z$, and $h_1$ and $h_2$ are the hydraulic heads applied at the inflow and outflow boundaries, respectively. For all models, $K_{xx}^{eff}$ is computed from the reference or the upscaled flow field. For example, for the 8-unit model, only one $K_{xx}^{eff}$ is computed, reflecting a global effective flow parameter.

For every model, two transport connectivity metrics are computed. The first metric is the power-law slope fitted to the BTC's tail in the log-log space [*Willmann et al*., 2008; *Renard and Allard*, 2013]. The second metric is defined by *Knudby and Carrera* [2005] as:

$$CT = \frac{t_{\text{peak,hom}}}{t_{\text{peak,het}}} \tag{14}$$

where $t_{\text{peak,het}}$ is peak breakthrough arrival time predicted by a heterogeneous model (i.e., FHM, 8-unit, and 3-unit models) at a given capture plane, and $t_{\text{peak,hom}}$ is peak breakthrough arrival time predicted by a homogeneous model (i.e., 1-unit model) at the same capture plane.

## 3. Results

### 3.1. Hydraulic Conductivity Upscaling and Flow Verification

For a given system variance, steady state flow is simulated in the FHM under different global BC [*Zhang et al*., 2011]. For each HSM unit, average hydraulic gradients and Darcy fluxes can be computed with which an equivalent **K*** can be estimated. For all units, equivalent **K***passed the Cholesky decomposition test [*Bhatia*, 2007], and is diagonally dominant because model coordinate is either parallel or orthogonal to stratification (Table 1 lists the principal components of **K***). To evaluate the accuracy of the equivalent **K***, the same flow experiments conducted for upscaling are repeated in the HSMs. Using the FHM as a reference, deviations in the average outflow rate and hydraulic head can be computed, which are found to vary with flow direction and system variance [*Zhang et al*., 2011]. Upscaling result of the lateral flow experiment is analyzed here with two error metrics of prediction, i.e., relative error in flow rate (*err*) and mean relative error (*MRE*) in hydraulic head:

$$\text{err} = \frac{|q_{\text{HSM}} - q_{\text{ref}}|}{q_{\text{ref}}} \times 100\%, \tag{15}$$

$$MRE = \frac{1}{I}\sum_{i=1}^{I} \frac{|h_{i,\text{HSM}} - h_{i,\text{ref}}|}{h_{\text{i,ref}}} \times 100\% \tag{16}$$

where $q_{\text{HSM}}$ and $q_{\text{ref}}$ represent average Darcy flux at the outflow boundary computed by the HSM and the FHM, respectively, $I$ is number of grid cells within a HSM unit, and $h_{i,\text{HSM}}$ and $h_{i,\text{ref}}$ represent head computed by HSM and FHM at each cell location, respectively.

Figure 2a presents the *err* of the HSMs. For a given HSM, higher system ln$K$ variance leads to less accurate flow rate prediction. When variance is fixed, *err* of the 8-unit model is consistently the smallest, that of the 1-unit model consistently the largest. Clearly, along with variance, connectivity resolution plays a significant role: models with higher connectivity resolutions can better capture flow. Figure 2b compares the *MRE* of hydraulic head among the HSMs. Again, *MRE* increases with system variance: when $\sigma_f^2 = 0.1$, *MRE* is less than 2% for all models, but increases to 3.5–4.4% and 4.9–6.5% when $\sigma_f^2 = 1.0$ and 4.5, respectively. Hydraulic head prediction becomes less accurate with increasing $\sigma_f^2$. For a given system variance, *MRE* of the 8-unit model is the smallest, that of the 3-unit model the highest. To summarize, both system variance and connectivity resolution influence the accuracy of the upscaled flow fields. Increasing variance is also found to

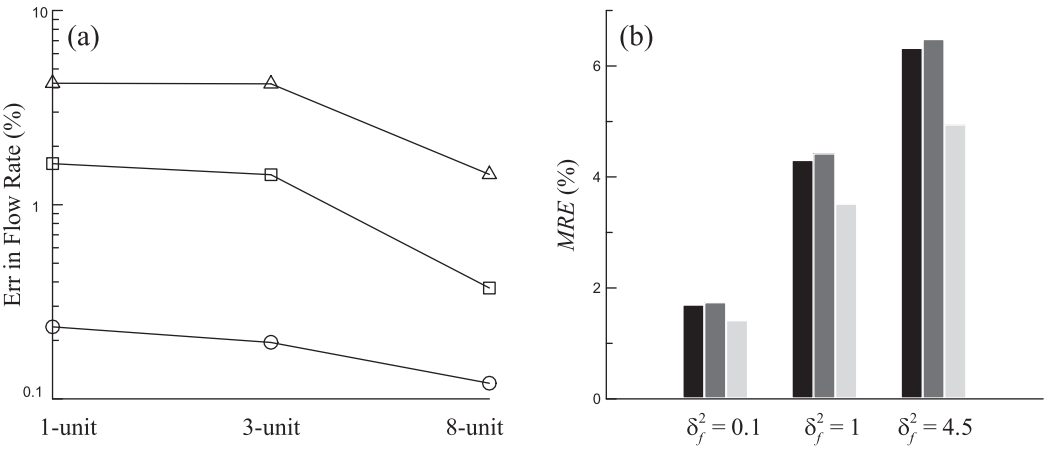


**Figure 2.** (a) Error in flow rate prediction by the HSMs at increasing system variances: $\sigma_f^2 = 0.1$ (circles), 1 (squares),and 4.5 (triangles). (b) *MRE* in hydraulic head prediction: 1-unit model (black), 3-unit model (dark gray), and 8-unit model (gray).

result in increasing horizontal-to-vertical anisotropy ratio ($K_{xx}/K_{zz}$ or $K_{xx}/K_{zz}$) of **K***, as preferential flows start to form in the reference model [*Zhang et al*., 2010]. Figure 3 presents lateral Darcy flux computed by all models when system variance is 4.5. As connectivity resolution is reduced, velocity becomes increasingly smoothed, although the HSMs generally capture the mean velocity of the reference flow field.

### 3.2. Transport Simulations and Upscaling

Assuming identical initial concentration distribution, a conservative tracer is released into the flow fields of all models. The upscaled $K_{xx}$, computed for each HSM unit using flow properties (see above), are compared against an apparent $K_{xx}$ inferred from the mean solute plume velocity (i.e., $v_{Px}$) when solute is traveling through the same unit: $K_{xx} = \theta v_{Px}/(\partial h/\partial x)$. To calculate the apparent $K_{xx}$, tracer initial dimensions were reduced to fit into each unit, so velocities from the other units are not sampled by solute particles. Apparent $K_{xx}$ should approach the upscaled $K_{xx}$ if solute has sampled most of the subunit velocities. (Using tracer moment to infer equivalent conductivity is a common practice, although it is not used here to replace flow upscaling because of the general difficulty of tracking a large number of particles inside each irregularly shaped HSM unit, and because such an approach can only lead to equivalent parameters along the mean flow direction). Table 3 presents this comparison for the 3-unit model. For all units, the two conductivities agree well with each other when $\sigma_f^2 = 0.1$ and 1.0, thus static connectivity mapping corresponds to dynamic flow connectivity. When $\sigma_f^2 = 4.5$, however, $K_{xx}$ derived from $v_{Px}$ overestimates the upscaled $K_{xx}$ by 17% to 42%, which is attributed to the appearance of lateral preferential flows: particles channel through fast-flowing streamtubes within a HSM unit without sampling the slower velocities in the same unit, which leads to an overestimated $K_{xx}$ compared to the upscaled $K_{xx}$. The latter is computed using *all* velocities within the unit, thus the effect of fast channeling is averaged out. At higher system variance, only a portion of the mapped static connectivity is dynamically connected.

#### 3.2.1. Transport Simulation in the FHM

Time evolution of the mass centroid and longitudinal plume covariance ($s_{xx}^2$) of the reference model is first examined. At all variances, mass centroids over time are near straight lines (solid curves in Figures 4a–4c), suggesting that plume centroids move at constant speeds, consistent with the observation of *Dagan* [1989]. The centroid velocity (i.e., slope of a line fit) increases with variance, e.g., $v_{Px} = 7.5$, 16, and 138 m/yr for $\sigma_f^2 = 0.1$, 1.0, and 4.5, respectively. As preferential flows are developed under higher variance, mean lateral groundwater velocity is increasing, as is the equivalent $K_{xx}$ (Table 1). At a fixed variance, $s_{xx}^2$ increases with time parabolically (solid curves in Figures 4d–4f), thus tracer not only expands continuously along mean flow, but rate of this expansion (i.e., the apparent $\alpha_L^M$) increases with time. The apparent $\alpha_L^M$, obtained with equation (11), also increases with $\sigma_f^2$ (Figure 5): when $\sigma_f^2 = 0.1$, it grows from 0 to 35 m over 200 years, but reaches as high as 150 m (in 48 years) and 250 m (in 4 years) when $\sigma_f^2 = 1.0$ and 4.5, respectively. At each variance, a third-order polynomial function of time is fitted, with a goodness-of-fit ($R^2$) greater than 99%. These functions will be used by the HSMs employing time-dependent macrodispersion.

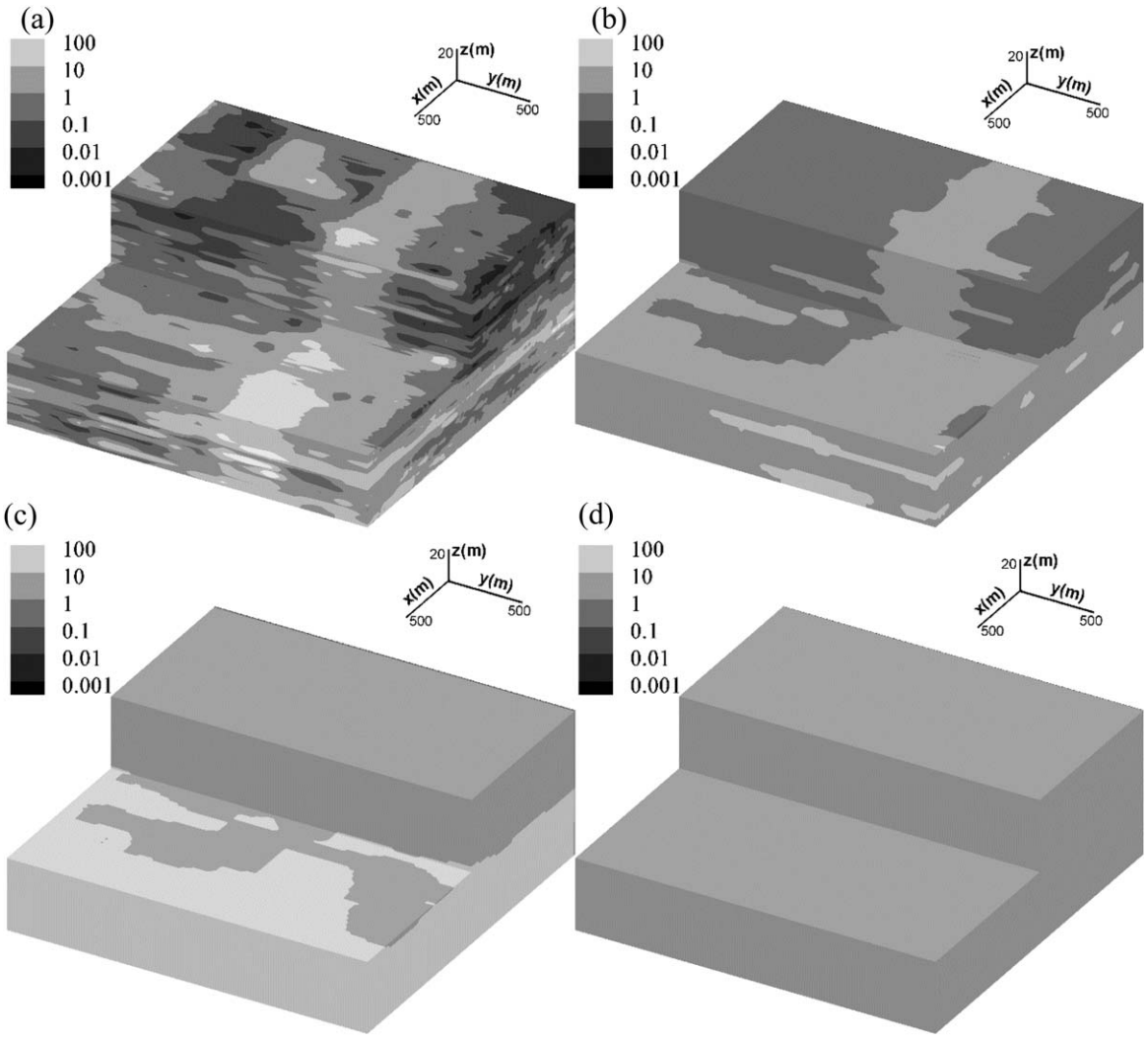


**Figure 3.** Dracy flux(*qx*) distribution for system variance of 4.5, with (a) FHM, (b) 8-unit model, (c) 3-unit model and, (d) 1-unit model. Note that *qx*of the FHM varies about 6 orders of magnitude.

Tracer simulated by the FHM is visualized next (Figure 6, first column). Of the total 100,000 particles, a random subset of 1,000 particles, capable of capturing the full extent of the plume, is plotted. When $\sigma_f^2 = 0.1$, the flow field is weakly heterogeneous and tracer expands slowly (Figure 5a). When $\sigma_f^2 = 1.0$, due to increased flow variability, tracer at 38.4 years is more laterally extensive than it is at 180 years under $\sigma_f^2 = 0.1$. When $\sigma_f^2 = 4.5$, tracer is the most laterally expansive: at 3.84 years, while tracer's leading edge nearly reaches the outflow boundary, most of the solute mass is still found near the source release plane. This phenomenon was observed at the MADE site with a similar level of *K* variability, although both downstream and upstream plume spreading was observed there at early times, possibly due to injection artifacts and transient flow effects [*Boggs et al*., 1992].

Breakthrough curve (BTC) can reveal tracer's higher moments including arrival, tailing, and symmetry. When $\sigma_f^2 = 0.1$, BTC at the first control plane ($x = 500$ m) is narrow and reaches a peak particle count (proxy for concentration) at $t = 30$ years (Figure 7a, solid curve). Due to the relatively homogeneous flow field and short travel distance, tracer lateral spreading is limited and BTC exhibits a slight asymmetry. At $x = 1000$ m, a stronger asymmetry is observed (Figure 7d, solid curve): over the longer travel distance, tracer has experienced more variable flow, resulting in a lower peak concentration (at $t = 110$ years) and more lateral spreading. This is similar to the behavior observed in *Fiori et al*. [2013]. When $\sigma_f^2 = 1.0$, due to increased flow variability, BTC at $x = 500$ m is more laterally spread out with a stronger asymmetry (Figure 7b, solid curve)

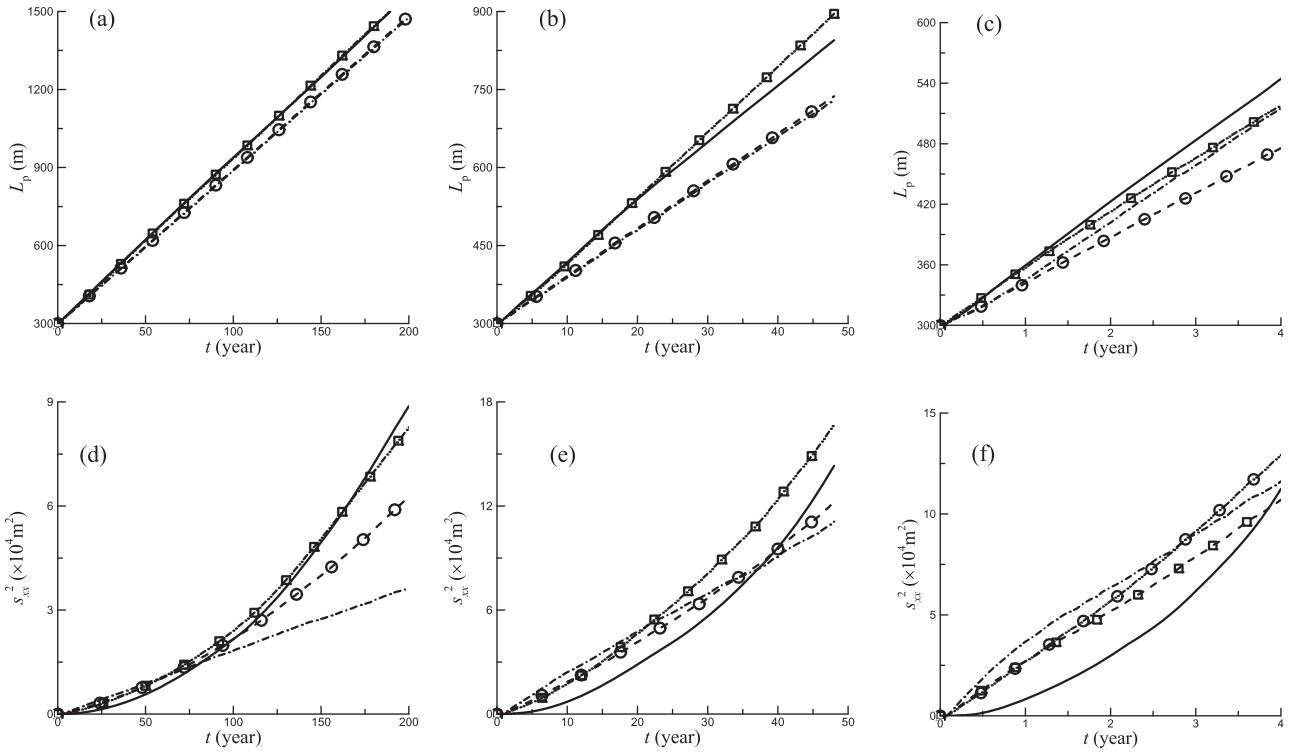


**Figure 4.** Time evolution of tracer moments using unit-specific macrodispersivity: mean plume displacement for $\sigma_f^2 =$ (a) 0.1, (b) 1.0, and (c) 4.5, respectively; longitudinal plume covariance for $\sigma_f^2 =$ (d) 0.1, (e) 1.0, and (f) 4.5, respectively. The solid, dash-dot, dashed-circle, and dash-square curves represent FHM, 1-unit, 3-unit,and 8-unit models, respectively.

compared to that observed when $\sigma_f^2 = 0.1$. As the result of a higher mean velocity, breakthrough peak also occurs earlier at $t = 15$ years. At $x = 1000$ m, BTC peaks at $t = 30$ years (Figure 7e, solid curve), while tracer has also expanded significantly compared to that observed at $x = 500$ m. When the leading edge of the plume reaches the outflow boundary, a large number of particles have not yet reached the second control plane. When $\sigma_f^2 = 4.5$, velocity variability is the strongest, i.e., $q_x$ varies close to 6 orders of magnitude (Figure 3). BTC at $x = 500$ m is even more spread out (Figure 7c, solid curve) compared to those observed at lower system variances. BTC at $x = 1000$ m (Figure 7f, solid curve) reaches a peak concentration at 2.5 years, decreases briefly, before increasing again. While a significant number of particles have reached this control plane via fast-flowing pathways (i.e., the first peak), the remaining particles lag behind and are advecting slowly in the lower velocity zones.

### 3.2.2. Transport Upscaling (Unit-Specific Macrodispersion)

Using unit-specific macrodispersivities, tracer moments of the HSMs, computed with the same grid, initial particle distribution, and number of particles, are compared to the FHM (Figure 4). In visualizing the plumes, the same random subset of particles is used. At all system variances, plume centroid displacement is nearly linear with time, while that simulated by the 8-unit model is closest to that of the FHM. Because centroid displacement is determined by mean velocity, the 8-unit model yields the best bulk velocity field, allowing it to accurately capture the reference plume's position over time for all system variances. The 1 and 3-unit models, because their flow errors do not differ significantly, exhibit similar deviations in predicting the centroid velocity (Figure 2). Moreover, at all variances, $s_{xx}^2$ of the HSMs increases with time, indicating plume expansion. At all the times examined, because macrodispersion enhances mixing *within* individual units, the HSM plumes are better mixed compared to the reference plume. Below, tracer moments predicted by all models are compared in greater detail.

When $\sigma_f^2 = 0.1$, $s_{xx}^2$ of the FHM, 8-unit, and 3-unit models are parabolic (rate of plume expansion increases with time), while that of the 1-unit model is linear (constant rate). At $t$=60 years, all models display a similar plume size (Figure 6a, first row), corresponding to a similar $s_{xx}^2$ value at the same time (Figure 4d). At $t$=180 years (Figure 6a, second row), the 8-unit model accurately captures the lateral extent of the reference

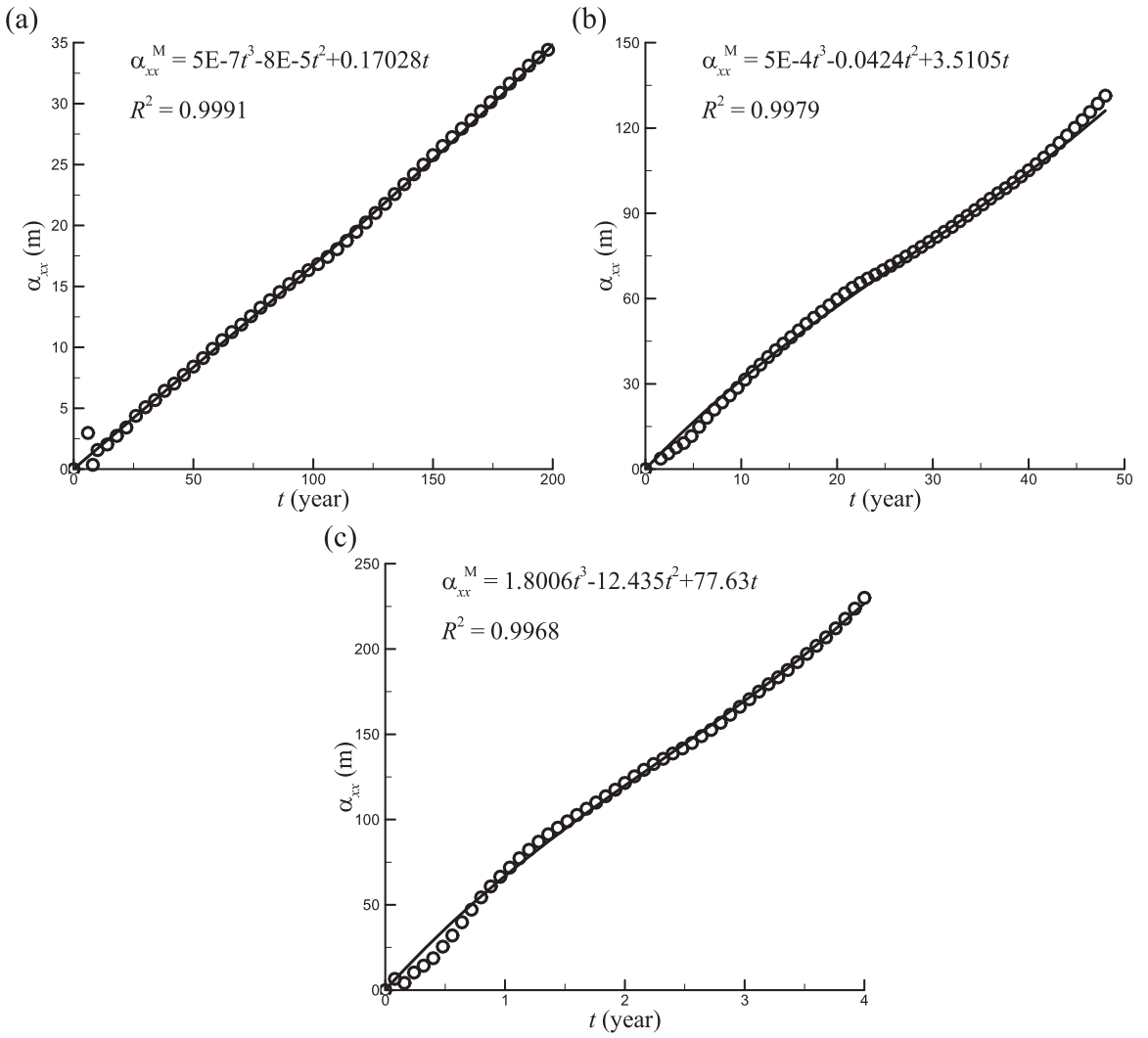


**Figure 5.** Apparent longitudinal macrodispersivity (circles) of the tracer simulated by the FHM under (a) $\sigma_f^2 = 0.1$, (b) $\sigma_f^2 = 1.0$, and (c) $\sigma_f^2 = 4.5$. A polynomial function (solid curve) is fitted with the goodness-of-fit shown.

plume, thus a HSM with high connectivity resolution can capture longitudinal spreading. When $\sigma_f^2 = 1.0$, $s_{xx}^2$ of the 8-unit model is (correctly) parabolic, while those of the 3 and 1-unit models are nearly linear (Figure 4e). At a fixed time, comparing $s_{xx}^2$ of a HSM with that of the same model under $\sigma_f^2 = 0.1$, lateral plume dimension is greater. For all except the 1-unit model, when lateral flow field is more variable, advection-controlled longitudinal spreading is enhanced. Because macrodispersivities increase with variance (Table 1), dispersion-controlled longitudinal spreading is also enhanced. All HSMs predict upstream dispersion, which is absent in the FHM. When $\sigma_f^2 = 4.5$, HSMs are not able to capture the parabolic behavior of the FHM (Figure 4f), although the 8-unit model is the least inaccurate. At higher system variance, upstream dispersion is also significant in the HSMs due to the greatly increased macrodispersivities.

BTCs of the HSMs at the control planes reveal that (Figure 7): (1) when variance is low, HSMs provide fairly accurate predictions of solute arrival, tailing, and overall breakthrough asymmetry. The observed tailing in these models can be attributed to the explicit modeling of connectivity, as groups of particles move at different mean velocities. Again, the 8-unit model is the most accurate at capturing the BTC. (2) when variance is higher, HSMs predict earlier arrival times, but capture the overall asymmetry at $x = 500$ m; at $x = 1000$ m, HSMs correctly predict the arrival time but cannot capture the BTC. The earlier ($x = 500$ m) underestimation of the arrival time is due to the fact that macrodispersivities predicted by theory reflect asymptotic values, which are usually achieved after solute has migrated over significant distances sampling a number of

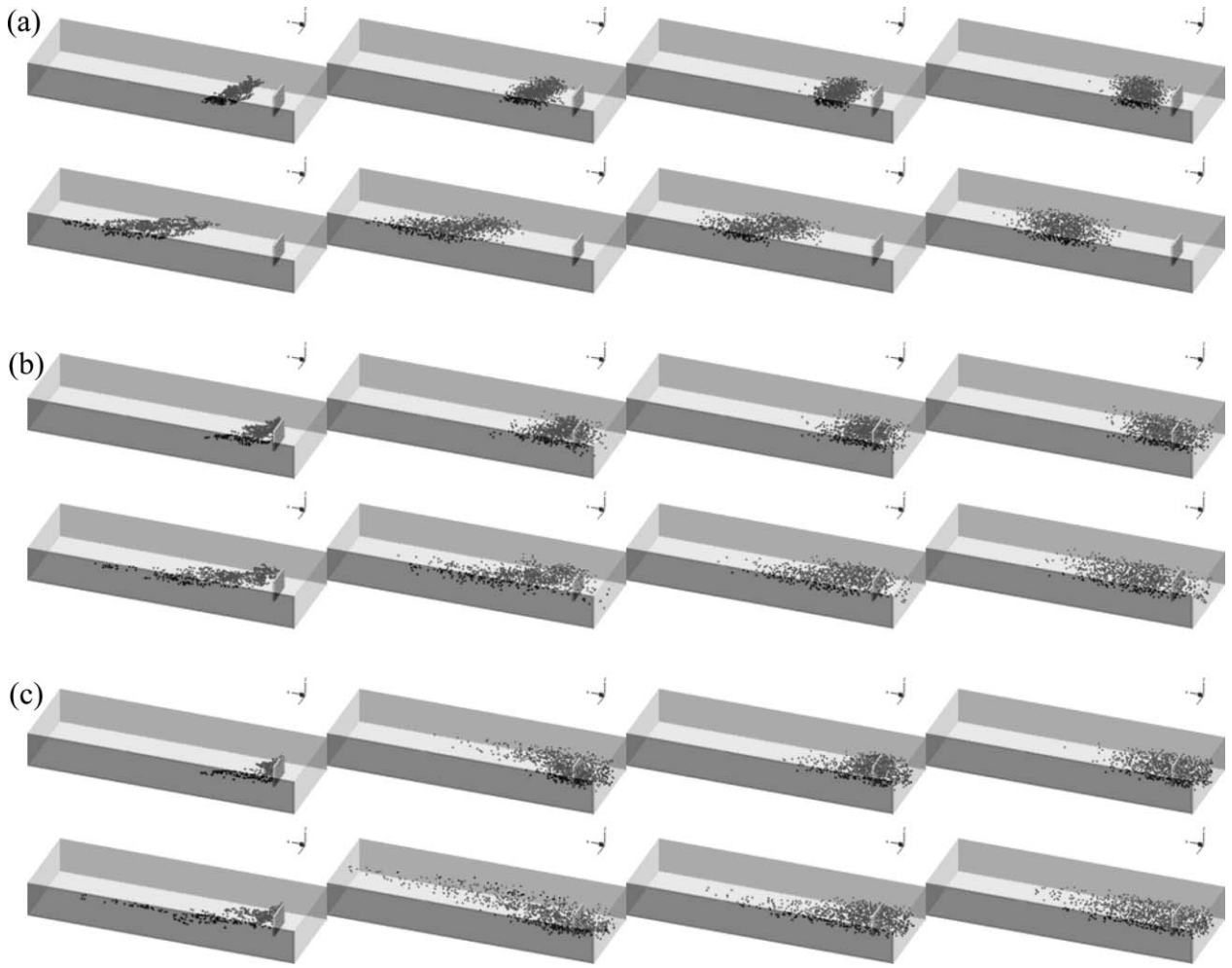


**Figure 6.** Particle locations simulated under (a) $\sigma_f^2 = 0.1$,(b) $\sigma_f^2 = 1.0$, and (c) $\sigma_f^2 = 4.5$. The first, second, third, and fourth columns represent tracer plumes simulated by the FHM, 8-unit, 3-unit, and 1-unit models, respectively. All HSMs employ unit-specific macrodispersion. The first row of each plot represents particle positions at (a) $t = 60$ years, (b) 12.8 years, and (c)1.28 years. The second row of each plot represents particle positions at (a)$t = 180$ years, (b)38.4 years, and (c) 3.84 years. Only a selectedsubset of 1000 particles is shown in each plot. Particle initial positions (light gray) are also shown.

heterogeneities. The later ($x = 1000$ m), more accurate, arrival prediction can be attributed to the fact that tracer has migrated over several lateral ln*K* correlation scales (Table 2). (3) when variance is the highest, all HSMs predict poor BTC. Lithological mapping, though sufficient to capture flow and transport connectivity when variance is low, fails to do so when variance is high. This confirms the earlier observation when the apparent *Kxx* inferred from tracer velocity is higher than the upscaled *Kxx*.

### 3.2.3. Transport Upscaling (Time-Dependent Macrodispersion)

Time-dependent macrodispersion, when assigned to the HSMs, mimics the observed spreading in the FHM (Figure 5). When plume centroids are examined first (Figures 4 and 8), time-dependent macrodispersivity provides almost identical results at low-to-medium variances as those simulated previously with unit-specific macrodispersivities. When variance is higher, greater differences are observed, but mainly in results of the 8-unit model. Plume centroid averages the positions of all particles and appears less influenced by how macrodispersion is formulated. When plume size (i.e., $s_{xx}^2$) is examined next, results of the HSMs are no longer similar to those using unit-specific $\alpha_L^M$. At all variances, $s_{xx}^2$ of the HSMs exhibits the characteristic parabolic behavior of the reference plume (Figure 8), thus the increasing rate of tracer expansion is captured. The time-dependent functions assigned to them (i.e., approximately linear growth of the apparent $\alpha_L^M$ with time) ensure this. At a given variance, the 8-unit model is the most accurate, despite the fact that the same time-dependent $\alpha_L^M$ is assigned to all HSMs. When variance increases, prediction of the HSMs degrades: at

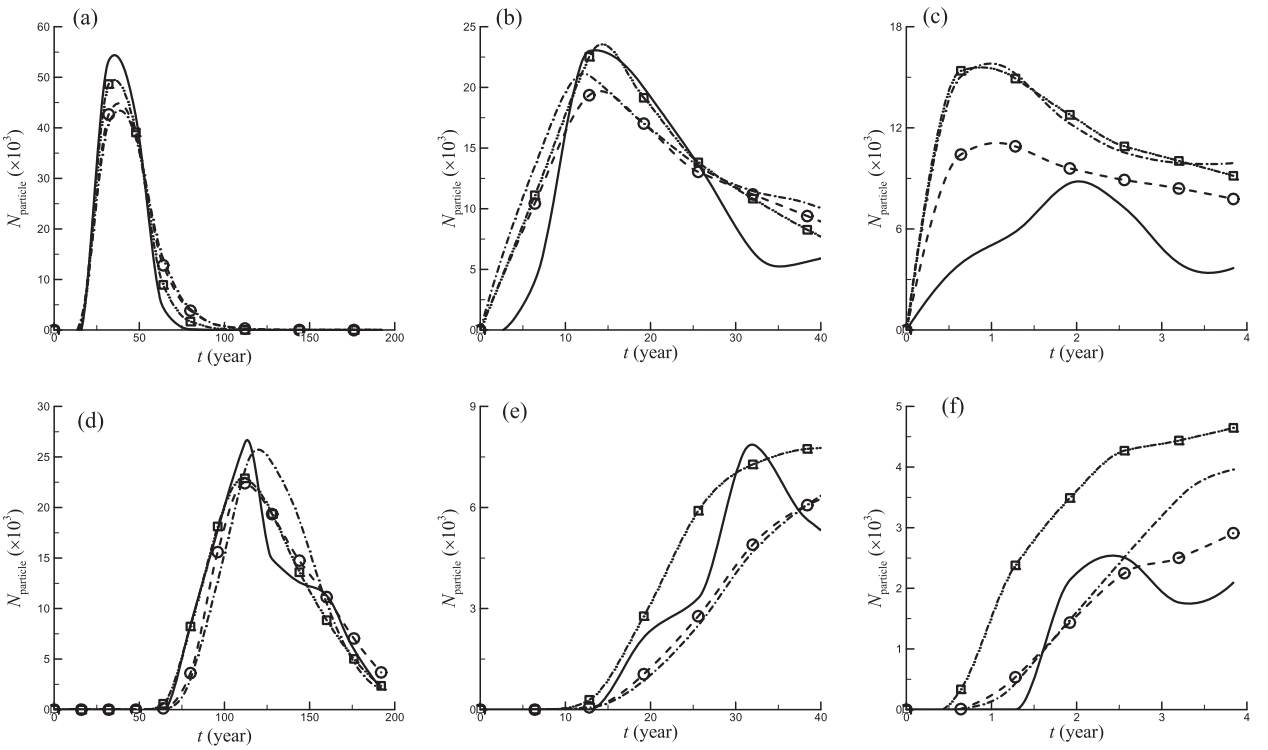


**Figure 7.** Breakthrough curves predicted by all models, with the HSMs using unit-specific macrodispersivity: at $x = 500$ m for $\sigma_f^2 =$ (a) 0.1, (b) 1.0, and (c) 4.5, respectively; at $x = 1000$ m for $\sigma_f^2 =$ (d) 0.1, (e) 1.0, and (f) 4.5, respectively. The solid, dash-dot, dash-circle,and dash-square curves represent FHM, 1-unit, 3-unit,and 8-unit models, respectively.

$\sigma_f^2$ of 4.5, all HSM predictions are quite poor (Figure 8f). Plumes visualized in Figure 9 reveal that the 8-unit model is still the best, as it approximately captures the leading edge of the reference plume even when variance is high. The time-dependent formulation also leads to reduced upstream dispersion near the tracer release plane (Figure 9). The apparent $\alpha_L^M$ is zero at $t = 0$, regardless of system variance (Figure 5).

BTCs of the HSMs at the control planes reveal that (Figure 10): (1) when variance is low, all HSMs except the 1-unit model predict the BTC well, and their predictions are similar to those of the unit-specific models (Figures 7a and 7d). (2) when variance is higher, HSMs become less accurate, although compared to the unit-specific models (Figure 7b), better arrival time is predicted at $x = 500$ m. This is likely a result of more accurately capturing the initially small plume with the apparent $\alpha_L^M$. At $x = 1000$ m, however, compared to Figure 7e, there is no significant arrival time improvement. Due to the longer travel distance, asymptotic $\alpha_L^M$ adopted by the unit-specific models are not significantly inferior. (3) when variance is the highest, performance of the HSMs degrades further, although the 8-unit model can still capture the arrival times of the FHM at both control planes. The arrival time is significantly influenced by the existence of preferential flows in the FHM, which can be captured to some degree by the 8-unit model, but not by the others.

### 3.2.4. Sensitivity Analysis

Tracer simulations in the upscaled models generally yield a poor fit when system variance is high. Could these be numerical artifacts from insufficient discretization? Because a more refined flow grid can lead to more resolved (but not necessarily more accurate) subgrid as well as global streamlines, for $\sigma_f^2 = 4.5$, velocity field of the 8-unit model is recalculated with an 8× refinement, doubling each of *Nx, Ny*, and *Nz*. With the refined flow field, tracer simulation is repeated with unit-specific macrodispersivities. As shown in Figure 11, both mean plume displacement and longitudinal plume covariance are very close to those obtained with the original flow field. Moreover, increasing number of particles in the RWPT experiments can potentially improve the smoothness and accuracy of the computed tracer moments as well as BTC. However, results of the 8-unit model using 1,000,000 particles nearly coincide with the original results (Figure 11), thus the original number of particles are sufficient for the transport analysis of this study. Finally, for all system variances, when local dispersion is ignored, breakthrough curves computed by the FHM are virtually

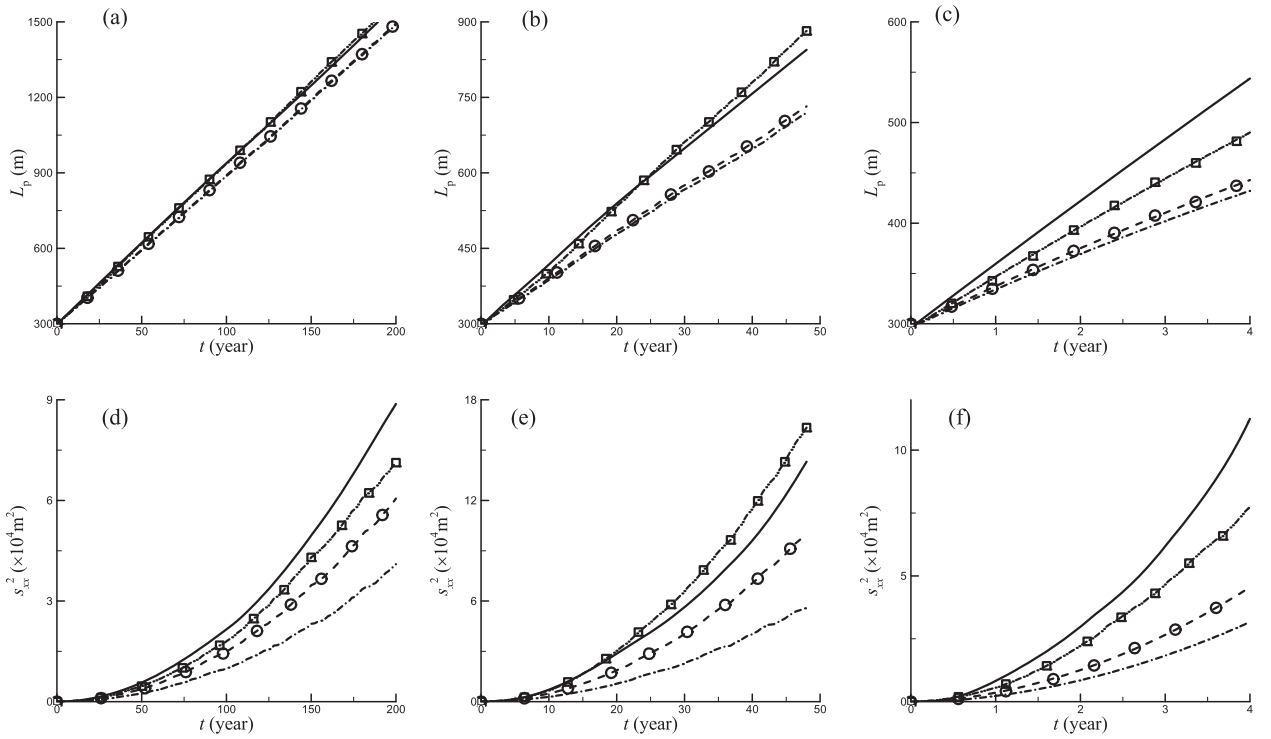


**Figure 8.** Evolution of plume moments predicted by all models, with the HSMs using time-dependent macrodispersivity: mean plume displacement for $\sigma_f^2 =$ (a) 0.1, (b) 1.0, and (c) 4.5, respectively; longitudinal plume covariance for $\sigma_f^2 =$ (d) 0.1, (e) 1.0,and (f) 4.5, respectively. The solid, dash-dot, dash-circle,and dash-square curves represent the FHM, 1-unit, 3-unit,and 8-unit models, respectively.

identical (not shown) to those presented earlier with nonnegligible local dispersion. For anisotropic heterogeneity where transport is advection-dominated, local dispersion generally has no discernable impact on large-scale solute spreading [*Fiori*, 1996].

### 3.3. Analysis of Flow and Transport Connectivity

For all models, $K_{xx}^{eff}$ computed by equation (13) is greater than $K_G$, with a *CI* value ranging from 2.6 to 15.2 (Table 4). *CI* of the FHM is consistent with stratification, which amplifies horizontal flow. It also increases with system variance because of emerging preferential flow [*Zhang et al*., 2010]. For the HSMs, their *CI* values reflect the combination of explicit connectivity mapping and upscaling (i.e., equivalent ***K**** of each connectivity is generally anisotropic), both contributing to horizontal flows. When variance is low, all HSMs have a similar *CI* as that of the FHM. When variance is higher, greater differences emerge. For the 1-unit model, *CI* does not change with system variance. This suggests that when connectivity is not accounted for, *CI* is not a useful index of flow connectivity, because the emergence of preferential flow under high variance cannot be established. For the other HSMs, *CI* increases with $\sigma_f^2$ because connectivity is captured. When variance is high, the 8-unit model is best able to represent the flow connectivity of the FHM, as expected.

When comparing transport, HSMs under lower system variance and/or with higher connectivity resolution are better able to capture transport connectivities of the FHM. For the BTCs at the first control plane ($x$=500 m), results of power-law fitting and *CT* are shown in Table 5 and 6, respectively. For all variances, *CT* computed for the 3 and 8-unit models are close to each other, suggesting that this metric cannot accurately distinguish the difference in their transport connectivity. On the other hand, power-law slope can provide better quantification: at a given variance, with a few exceptions, it tends to increase with connectivity resolution (Table 5). Figure 12 compares *CT* versus power-law slope between this study (at the first control plane) and 2-D results from *Willmann et al*. [2008] for their "Type 2" (two heterogeneity scales but no preferential flow) and "Type 5" (one scale of heterogeneity but high-*K* cells are strongly connected) fields. When transport is 3-D, *CT* falls into a narrow range from 1.0 to 1.6, while in 2-D it ranges from 1.3 to 4.0. In 2-D, power law slope appears to decrease with *CT*; in 3-D, it varies independently of *CT*. Transport connectivity

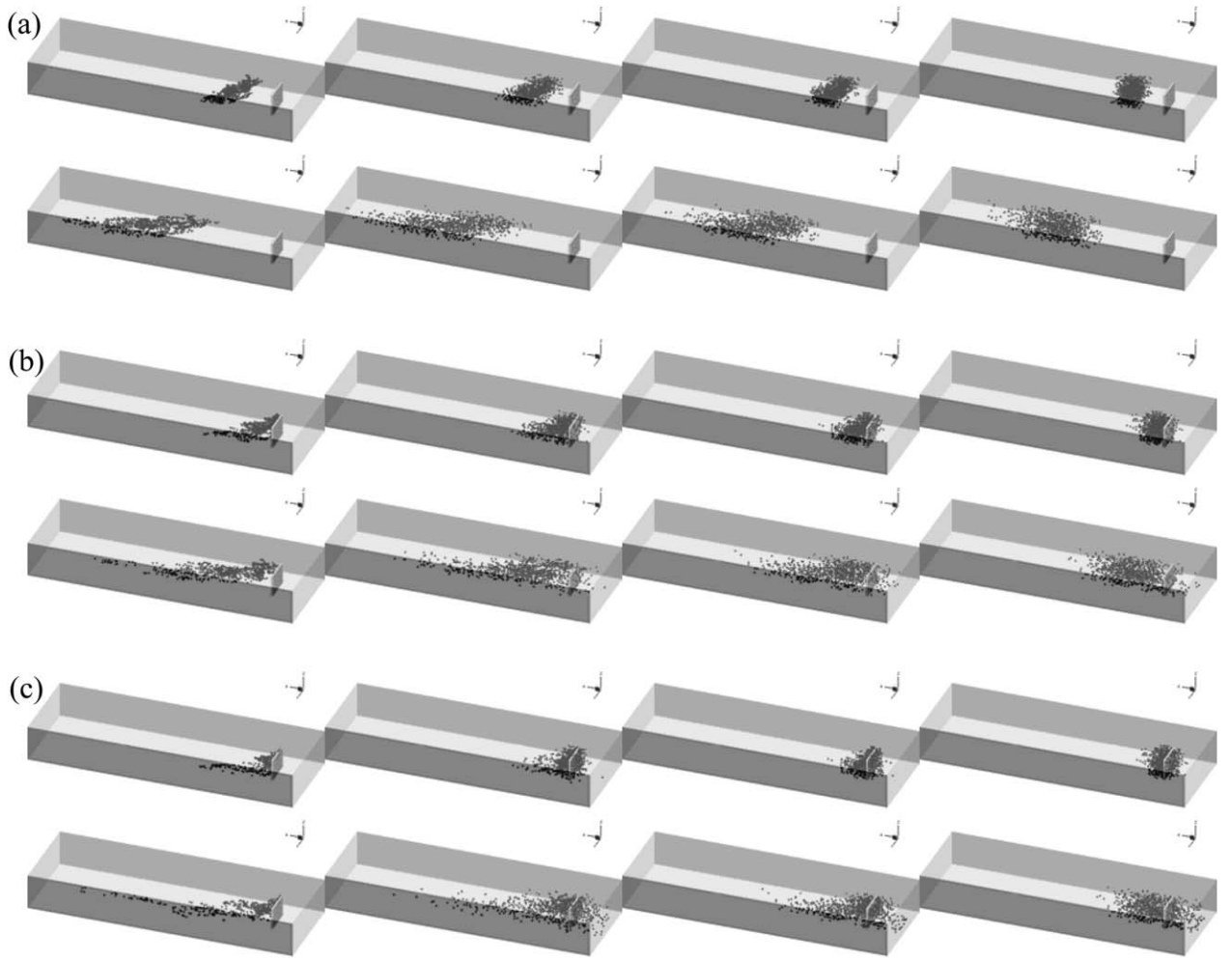


**Figure 9.** Particle locations simulated under(a) $\sigma_f^2 = 0.1$, (b) $\sigma_f^2 = 1.0$, and (c) $\sigma_f^2 = 4.5$. The first, second, third, and fourth columns represent tracer plumes simulated by the FHM, 8-unit, 3-unit and 1-unit models, respectively. All HSMs employ time-dependent macrodispersion. The first row of each plot represents particle positions at (a) $t = 60$ years, (b) 12.8 years, (c) 1.28 years. The second row of each plot represents particle positions at (a) $t = 180$ years,(b) 38.4 years, and (c) 3.84 years. Only a selected subset of 1000 particles is shown in each plot. Particle initial positions (light gray) are also shown.

behaves differently in 3-D because solute can more easily migrate around low-*K* features[*Fiori and Jankovic*, 2012].

## 4. Discussion

For a hierarchical aquifer model, ADE is assumed applicable to describing transport at the continuum scale: local dispersion is assigned to the FHM and macrodispersion to the HSMs. A 1-unit model ignoring all scales of heterogeneity provides a fair representation of the transport behavior of the FHM when system variance is low. When variance is higher, it becomes a poor choice: transport predictions using time-dependent $\alpha_L^M$ do not improve significantly over those using a constant $\alpha_L^M$, despite the enhanced degree of freedom in representing dispersion. This is consistent with the interpretation of MADE tracer tests using a homogenous model [*Zheng et al*., 2011]. In comparison, the 3 and 8-unit models provide better predictions at all variances, thus an enhanced degree of freedom in representing dispersion can improve transport upscaling. Between them, the 8-unit model is more accurate, thus a higher connectivity resolution can further improve upscaling. Because solute travels through both high and low velocity zones due to the explicit modeling of connectivities, these models can capture non-Fickian transport such as breakthrough tailing and asymmetry. This effect is also illustrated in *Fiori et al*. [2013] who used statistically populated inclusions to represent

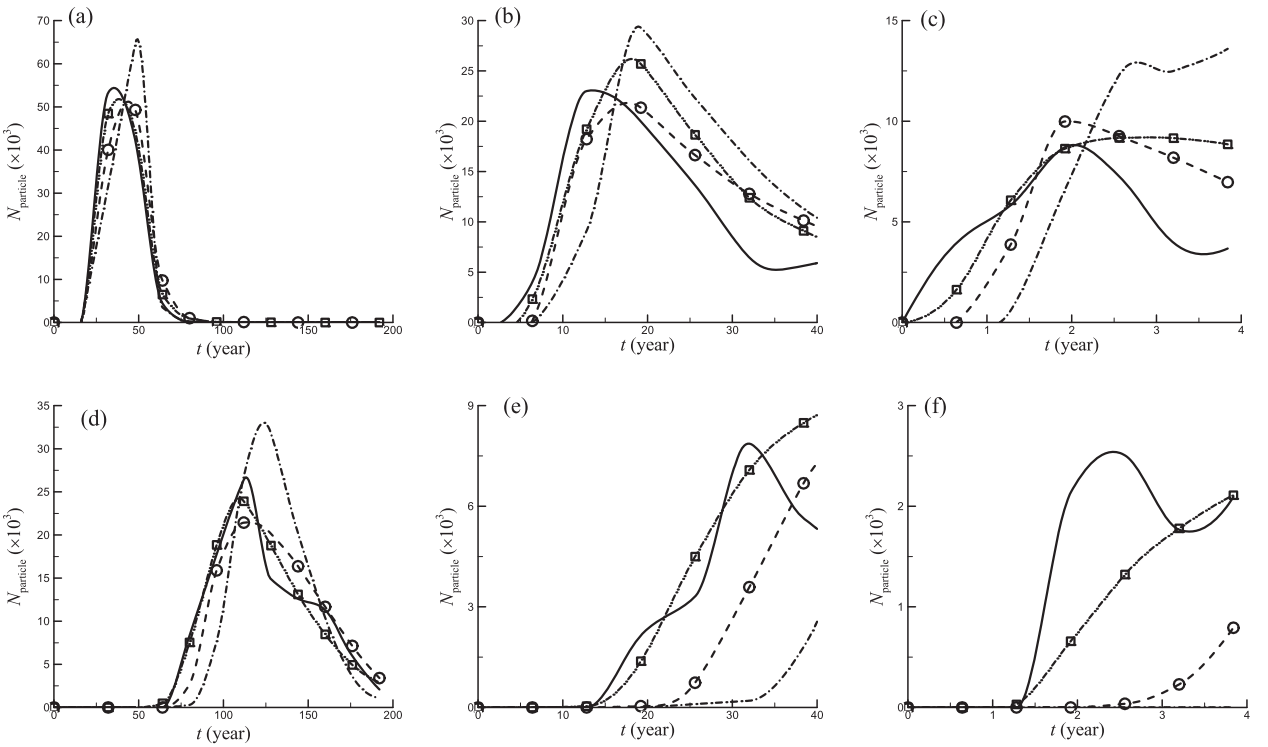


**Figure 10.** Breakthrough curves predicted by all models, with the HSMs using time-dependent macrodispersivity: at $x$= 500 m for $\sigma_f^2$= (a) 0.1, (b) 1.0, and (c) 4.5, respectively; at $x$= 1000 m for $\sigma_f^2$= (d) 0.1, (e) 1.0, and (f) 4.5, respectively. The solid, dash-dot, dash-circle, and dash-square curves represent the FHM, 1-unit, 3-unit, and 8-unit models, respectively.

low-velocity zones within a high-velocity matrix. At the Lauswiesen site in Germany where system ln$K$ variance is around 3.0, when heterogeneity was modeled at two scales (stochastic facies and subfacies correlated $K$ distributions) that were conditioned to site static data, ADE as well as a purely advective model can capture depth-averaged BTC that exhibits heavy tails [*Riva et al.*, 2008]. In this case, inclusion of subgrid mass-transfer processes using a dual-porosity formulation did not lead to significant improvement. Together, the above studies suggest that ADE with an enhanced degree of freedom in representing dispersion can capture non-Gaussian transport due to the underlying (hierarchical) heterogeneity.

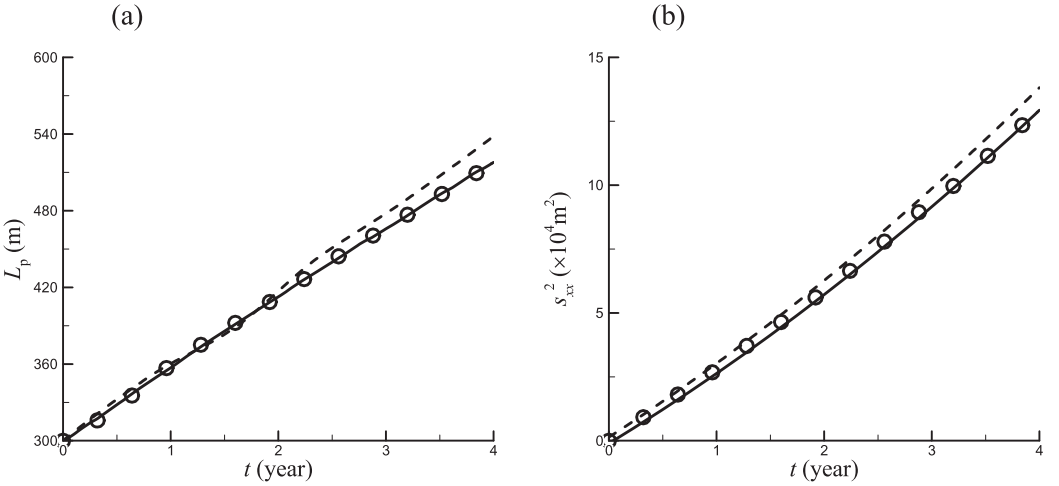


**Figure 11.** Time evolution of (a) mean plume displacement and (b) longitudinal plume covariance under $\sigma_f^2$= 4.5. The solid curve, dashed curve, and circle represent the simulation results with the 8-unit model using the original flow and tracer discretization, using a refined flow grid, and using a refined RWPT simulation with 1,000,000 particles, respectively.

**Table 3.** Upscaled $K_{xx}$ (m/yr) for Units of the 3-Unit Model Compared to $K_{xx}$ of the Same Units Inferred From Mean Plume Velocity

| | Unit 1 | | Unit 2 | | Unit 3 | |
|---|---|---|---|---|---|---|
| System Variance | Upscaled $K_{xx}$ | $K_{xx}$ From Mean Plume Velocity | Upscaled $K_{xx}$ | $K_{xx}$ From Mean Plume Velocity | Upscaled $K_{xx}$ | $K_{xx}$ From Mean Plume Velocity |
| $\sigma_f^2 = 0.1$ | 34.24 | 34.13 | 49.11 | 48.97 | 46.20 | 46.26 |
| $\sigma_f^2 = 1.0$ | 32.56 | 32.80 | 96.54 | 96.92 | 80.15 | 81.21 |
| $\sigma_f^2 = 4.5$ | 58.63 | 68.05 | 475.42 | 588.54 | 330.03 | 467.51 |

For the transport problems tested, predictive ability of the alternative macrodispersion formulations does not differ significantly. Both approaches are fairly accurate when implemented with the 8-unit model and/or when system variance is low. With the 3 and 1-unit models, both perform fairly well in capturing breakthrough tailing and asymmetry at low-to-medium $\sigma_f^2$. Differences between the two do exist. For example, time-dependent $\alpha_L^M$ is able to capture the rate of lateral plume expansion over time, even with the 1-unit model. Such an effect is not well captured with the unit-specific $\alpha_L^M$. During early migration, time-dependent $\alpha_L^M$ predicts more accurate solute arrival time while reducing upstream dispersion, althoughthis approach requires detailed tracer measurements while unit-specific models only need static (geostatistical) parameters. For all variances, the 8-unit model is the best predictor, regardless of the formulation used: when connectivity is sufficiently resolved, the type of data conditioning for modeling transport appears less critical. However, what resolution is "sufficient" is likely problem dependent, and future work needs to test more heterogeneities. Furthermore, how upscaled models should be conditioned by data is influenced by prediction goal. If the location/speed of plume centroid is of interest, data type used to condition transport does not matter significantly, especially when system variance is low to moderate. If early prediction is of interest, detailed measurements as required by the time-dependent models are more appropriate. If late-time behavior is of interest, the type of data again becomes less important.

When aquifer variability is weak (e.g., $\sigma_f^2 \leq 1$), all upscaled models are fairly robust and increased accuracy can be achieved with higher connectivity resolutions. This is consistent with the observations at the Cape Cod and Borden sites, where aquifers are weakly heterogeneous and longitudinal spreading of the measured tracers is in fair agreement with theory. When system exhibits strong variability, however, all macrodispersion models perform poorly, which is attributed to their general failure to capture preferential flows in the FHM. In these cases, lithological variation cannot accurately capture dynamic connectivity. In this study, connectivity was delineated by kriging [*Zhang et al*., 2011] which is a smooth interpolator that can lead to classification errors that then contribute to upscaling errors, e.g., high-*K* cells assigned to low-*K* facies or connected thin lenses become separated. Such issues cannot be remedied easily using macrodispersion, although a finer connectivity resolution (i.e., further division of the 8-unit model) may improve prediction. However, techniques that can distinguish dynamic connectivity are needed, while for problems where such connectivity cannot be practically identified, pre-asymptotic or higher order theories, or one or more nonlocal formulations, should be explored. Upscaling will again aim to map dynamic transport behaviors to parameters of these models.

Models of this study reveal insights into flow and transport behavior in hierarchical media, based on which we propose an upscaling strategy with 3 steps: (1) large-scale connectivity, preferably corresponding to dominant transport pathways, is first identified [*Monsen et al*., 2005]. (2) after capturing such connectivity, effective conductivity can be estimated *without* conducting flow upscaling using, e.g., (a) tracer test; (b) analytic-stochastic techniques subject to similar variability constraints [*Zhang et al*., 2007]; (c) parameter estimation techniques that can infer effective parameters from the observed state variables [*Irsa and Zhang*, 2012; *Zhang*, 2014]. (3) after the flow field has been upscaled, transport upscaling is conducted by conditioning to either spatial statistics assuming that a local stationary variogram is applicable, or to tracer moments using time-dependent formulation. Note that a similar approach was suggested by *Anderson* [1997], as explicit modeling of connected features is believed to exert a stronger control on predicting transport than smaller scale heterogeneities within such features [e.g., *Fogg*, 1986; *Jussel and Stauffer*, 1994; *Poeter and Townsend*, 1994]. This study suggests that accuracy of such an approach will

**Table 4.** Flow Connectivity (*CI*) Computed for All Models Under Increasing System Variances

| System Variance | 1-Unit Model | 3-Unit Model | 8-Unit Model | FHM |
|---|---|---|---|---|
| $\sigma_f^2 = 0.1$ | 2.59 | 2.62 | 2.65 | 2.71 |
| $\sigma_f^2 = 1.0$ | 2.59 | 2.91 | 3.25 | 4.01 |
| $\sigma_f^2 = 4.5$ | 2.59 | 3.96 | 6.00 | 15.16 |

**Table 5.** Power-Law Slopes Fitted to Breakthrough Tailing at $x=500$ m for All Models at Increasing System Variances

| $\sigma_f^2$ | 0.1 | 1.0 | 4.5 |
|---|---|---|---|
| 1-unit | 1.57 | 0.72 | 0.60 |
| 3-unit | 1.57 | 0.90 | 0.33 |
| 8-unit | 2.08 | 1.33 | 0.46 |
| FHM | 2.91 | 2.73 | 3.06 |

depend on whether dynamic connectivity is captured at sufficient detail as well as the underlying system variability, i.e., higher variability will likely require a finer resolution. However, these insights may change when a different heterogeneity is evaluated depending on, e.g., how realistic it is to assume that a local stationary variogram can describe heterogeneity within individual connectivities. To verify this, not only will we need to seek a different connectivity mapping scheme robust to variance, we also need to address the local stationarity assumption. For a given problem, when connectivity resolution (i.e., spatial division) is increased, local variances of the individual connectivities will decrease. Can a threshold be reached such that potential local nonstationarity can be ignored? In studies that tested upscaling theories on nonstationary media with bimodal, channeled, and hierarchical heterogeneities [*Desbarats*, 1990; *Desbarats and Srivastava*, 1991; *Zhang and Gable*, 2008], theory provides reasonable predictions when (local) ln$K$ variances are relatively low.

## 5. Conclusion

Solute transport in hierarchical porous media is examined by developing hydrostratigraphic models (HSMs) integrating large-scale connectivity mapping with flow and transport upscaling. A laboratory-based synthetic aquifer, which exhibits hierarchical heterogeneity (FHM), provides a reference hydraulic conductivity model. Within the sedimentary hierarchy, HSMs with different lithological (static) connectivity resolutions are created: an 8-unit facies model, a 3-unit depositional model, and a 1-unit homogeneous model. For each connectivity (or unit) of the HSMs, equivalent hydraulic conductivity is first calculated using flow-based upscaling. In the computed flow field that is parallel to stratification, an instantaneous, conservative tracer is simulated by all models. While local dispersion is modeled in the FHM, transport is upscaled for the HSMs using alternative approaches: (1) macrodispersivities, conditioned to geostatistical parameters of subunit heterogeneity, are assigned to the HSMs using a first-order theory; (2) time-dependent macrodispersivities, conditioned to detailed tracer measurements from the FHM, are assigned to the same models. In both flow and transport upscaling, increasing system ln$K$ variance (0.1, 1.0, 4.5) reflecting field-scale variability is tested.

For all variances tested, HSMs provide fair to good bulk flow predictions, with relative errors in flow rate and hydraulic head less than 5% and 7%, respectively. For a given model, upscaling errors of flow increase with variance. When variance is fixed, HSMs with increased connectivity resolution are more accurate. Among them, the 8-unit model provides the best flow predictions at all variances. In transport modeling, tracer migration (i.e., mass centroid, longitudinal plume covariance, breakthrough curve, peak concentration) of the FHM can be captured well by the HSMs with higher connectivity resolutions, but upscaling errors of transport increase with variance as well. Importantly, when connectivity is explicitly resolved, an enhanced degree of freedom in representing dispersion can improve the ADE-based models by capturing non-Fickian transport of the FHM. This suggests that nonlocal theories which use more parameters than dispersivity may have similarly benefited. When connectivity is sufficiently resolved, the type of data used to condition transport also becomes less critical. Data conditioning, however, is influenced by the prediction goal. Moreover, 3-D transport connectivities of the hierarchical aquifer analyzed in this work differ quantitatively from those analyzed for 2-D systems, consistent with prior observations.

When aquifer is weak-to-moderately heterogeneous ($\sigma_f^2 \leq 1$), upscaled ADE can adequately capture transport, despite the existence of hierarchical heterogeneity at smaller scales. For such media, accurate transport upscaling can be accomplished with low connectivity resolutions. At the weakly heterogeneous Cape Cod and Borden sites, spreading of the observed plumes is in fair agreement with theories. When aquifer exhibits stronger variability, the upscaled models perform poorly, which is attributed to the failure of static connectivity to capture preferential flows (i.e., dynamic connectivity) in the FHM. This observation is consistent with

**Table 6.** Transport Connectivity Computed for All Models at Increasing System Variances Using Peak Times in the Simulated BTC at $x=500$ m

| $\sigma_f^2$ | 0.1 | 1.0 | 4.5 |
|---|---|---|---|
| 3-unit | 1.03 | 1.06 | 1.10 |
| 8-unit | 1.05 | 1.06 | 1.10 |
| FHM | 1.16 | 1.23 | 1.59 |

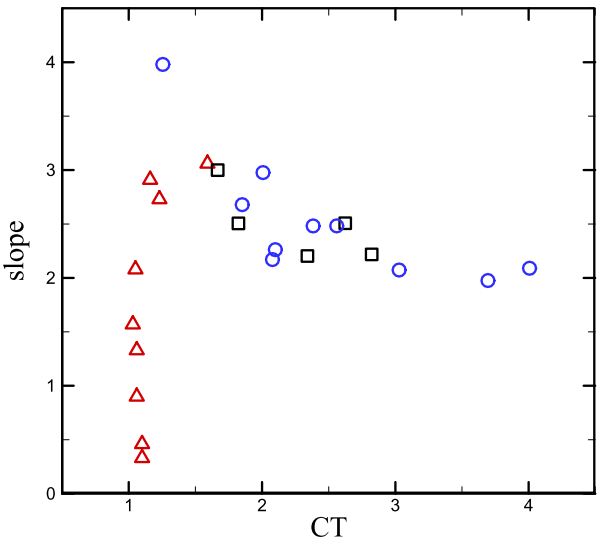


**Figure 12.** *CT* versuspower-law fitted slope. Triangles, squares, and circles represent result from the present study (FHM, 8 and 3-unit models), "Type 2"fields data [*Willmann et al*., 2008], and "Type 5" fields data [*Willmann et al*., 2008], respectively.

tracer interpretations at the MADE site. *Renard and Allard* [2013] provide a review of flow and transport connectivities, which will be examined along with a new inversion approach to develop improved connectivity mapping schemes that are robust to system variance.

The experimental stratigraphy offers one example of a hierarchical heterogeneity. Results and insights of this study thus pertain to this one heterogeneity and cannot yet be generalized to other systems. Because aquifer heterogeneity tends to be site specific, additional studies analyzing different heterogeneities are needed to help reveal insights that are more universal in nature. However, the upscaling method of this study is general and the equivalent conductivities are tensor properties that can be used to capture bulk flow for complex geometries. The study method can thus be applied to analyzing problems with any connectivity features. Future work can further assess the ergodicity condition as a working assumption, for which a stochastic approach is likely needed, i.e., instead of using one reference FHM for upscaling, an ensemble of such reference models will be created and analyzed. Such an approach will increase the computational burden but is straightforward to implement.

**Acknowledgments**

The data for this paper are available at Community Surface Dynamics Modeling System (Data set: Hydrography; Data set name: 3-D ADE simulation). The research is supported by the U.S. Department of Energy, Office of Fossil Energy, grant DE-FE-0009238. We would like to acknowledge the use of computational resources (ark:/85065/d7wd3xhc) at the NCAR-Wyoming Supercomputing Center, which are provided by the National Science Foundation and the State of Wyoming and are supported by NCAR's Computational and Information Systems Laboratory.

## References

Adams, E. E., and L. W. Gelhar (1992), Field study of dispersion in a heterogeneous aquifer: 2. Spatial moments analysis, *Water Resour. Res.*, *28*(12), 3293–3307, doi:10.1029/92WR01757.

Anderson, M. P. (1997), Characterization of geological heterogeneity, in *Subsurface Flow and Transport: A Stochastic Approach*, edited by G. Dagan, and P. Neuman, pp 23–43, Cambridge Univ. Press, Cambridge, U. K.

Barlebo, H. C., M. C. Hill, and D. Rosbjerg (2004), Investigating the Macrodispersion Experiment (MADE) site in Columbus, Mississippi, using a three-dimensional inverse flow and transport model, *Water Resour. Res.*, *40*, W04211, doi:10.1029/2002WR001935.

Bear, J. (1988), *Dynamics of Fluids in Porous Media*, Dover Civ. Mech. Eng. Ser., Dover, Mineola, N. Y.

Benson, D. A., S. W. Wheatcraft, and M. M. Meerschaert (2000), Application of a fractional advection-dispersion equation, *Water Resour. Res.*, *36*(6), 1403–1412, doi:10.1029/2000WR900031.

Berkowitz, B., and H. Scher (1998), Theory of anomalous chemical transport in random fracture networks, *Phys. Rev. E*, *57*(5), 5858–5869, doi:10.1103/PhysRevE.57.5858.

Berkowitz, B., A. Cortis, M. Dentz, and H. Scher (2006), Modeling non-Fickian transport in geological formations as a continuous time random walk, *Rev. Geophys.*, *44*, RG2003, doi:10.1029/2005RG000178.

Bhatia, R. (2007), *Positive Definite Matrices*, Princeton Univ. Press, Princeton, N. J.

Bianchi, M., C. Zheng, C. Wilson, G. R. Tick, G. Liu, and S. M. Gorelick (2011), Spatial connectivity in a highly heterogeneous aquifer: From cores to preferential flow paths, *Water Resour. Res.*, *47*, W05524, doi:10.1029/2009WR008966.

Boggs, J. M., S. C. Young, L. M. Beard, L. W. Gelhar, K. R. Rehfeldt, and E. E. Adams (1992), Field study of dispersion in a heterogeneous aquifer, 1. Overview and site description, *Water Resour. Res.*, *28*(12), 3281–3291, doi:10.1029/92WR01756.

Bohling, G. C., G. Liu, S. J. Knobbe, E. C. Reboulet, D. W. Hyndman, P. Dietrich, and J. J. Butler (2012), Geostatistical analysis of centimeter-scale hydraulic conductivity variations at the MADE site, *Water Resour. Res.*, *48*, W02525, doi:10.1029/2011WR010791.

Burr, D. T., E. A. Sudicky, and R. L. Naff (1994), Nonreactive and reactive solute transport in three-dimensional heterogeneous porous media: Mean displacement, plume spreading, and uncertainty, *Water Resour. Res.*, *30*(3), 791–815, doi:10.1029/93WR02946.

Cirpka, O. A., and P. K. Kitanidis (2000), An advective-dispersive stream tube approach for the transfer of conservative-tracer data to reactive transport, *Water Resour. Res.*, *36*(5), 1209–1220, doi:10.1029/1999WR900355.

Comunian, A., P. Renard, J. Straubhaar, and P. Bayer (2011), Three-dimensional high resolution fluvio-glacial aquifer analog—Part 2: Geostatistical modeling, *J. Hydrol.*, *405*(1-2), 10–23, doi:10.1016/j.jhydrol.2011.03.037.

Cushman, J. H., and T. R. Ginn (1993), Nonlocal dispersion in media with continuously evolving scales of heterogeneity, *Transp. Porous Media*, *13*(1), 123–138.

Cushman, J. H., L. S. Bennethum, and B. X. Hu (2002), A primer on upscaling tools for porous media, *Adv. Water Resour.*, *25*(8-12), 1043–1067, doi:10.1016/S0309-1708(02)00047-7.

Cvetkovic, V., H. Cheng, and X. Wen (1996), Analysis of nonlinear effects on tracer migration in heterogeneous aquifers using Lagrangian travel time statistics, *Water Resour. Res.*, *32*(6), 1671–1680.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer, Berlin.

Dagan, G., and S. P. Neuman (1997), *Subsurface Flow and Transport: A Stochastic Approach*, Cambridge Univ. Press, N. Y.

Dentz, M., A. Cortis, H. Scher, and B. Berkowitz (2004), Time behavior of solute transport in heterogeneous media: Transition from anomalous to normal transport, *Adv. Water Resour.*, *27*(2), 155–173, doi:10.1016/j.advwatres.2003.11.002.
Desbarats, J. (1990), Macrodispersion in sand-shale sequences, *Water Resour. Res.*, *26*(l), 153–163.
Desbarats, J., and R. M. Srivastava (1991), Geostatistical characterization of groundwater flow parameters in a simulated aquifer, *Water Resour. Res.*, *27*(5), 687–698, doi:10.1029/90WR02705.
Deutsch, C. V. (1998), Fortran programs for calculating connectivity of three-dimensional numerical models and for ranking multiple realizations, *Comput. Geosci.*, *24*(1), 69–76, doi:10.1016/S0098–3004(97)00085-X.
Efendiev, Y., L. J. Durlofsky, and S. H. Lee (2000), Modeling of subgrid effects in coarse-scale simulations of transport in heterogeneous porous media, *Water Resour. Res.*, *36*(8), 2031–2041, doi:10.1029/2000WR900141.
Einstein, A. (1905), Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen, *Ann. Phys.*, *322*(8), 549–560.
Feehley, C. E., C. Zheng, F. J. Molz, S. Carolina, and A. Abstract (2000), A dual-domain mass transfer approach for modeling solute transport in heterogeneous aquifers: Application to the Macrodispersion Experiment (MADE) site, *Water Resour. Res.*, *36*(9), 2501–2515.
Fernàndez-Garcia, D., and J. J. Gómez-Hernández (2007), Impact of upscaling on solute transport: Traveltimes, scale dependence of dispersivity, and propagation of uncertainty, *Water Resour. Res.*, *43*, W02423, doi:10.1029/2005WR004727.
Fernàndez-Garcia, D., T. H. Illangasekare, and H. Rajaram (2004), Conservative and sorptive forced-gradient and uniform flow tracer tests in a three-dimensional laboratory test aquifer, *Water Resour. Res.*, *40*, W10103, doi:10.1029/2004WR003112.
Fernàndez-Garcia, D., T. H. Illangasekare, and H. Rajaram (2005), Differences in the scale dependence of dispersivity and retardation factors estimated from forced-gradient and uniform flow tracer tests in three-dimensional physically and chemically heterogeneous porous media, *Water Resour. Res.*, *41*, W03012, doi:10.1029/2004WR003125.
Fiori, A. (1996), Finite Peclet extensions of Dagan's solutions to transport in anisotropic heterogeneous formations, *Water Resour. Res.*, *32*(1), 193–198.
Fiori, A., and I. Jankovic (2012), On preferential flow, channeling and connectivity in heterogeneous porous formations, *Math. Geosci.*, *44*, 133–145, doi:10.1007/s11004-011-9365-2.
Fiori, A., G. Dagan, I. Jankovic, and A. Zarlenga (2013), The plume spreading in the MADE transport experiment: Could it be predicted by stochastic models?, *Water Resour. Res.*, *49*, 2497–2507, doi:10.1002/wrcr.20128.
Fogg, G. E. (1986), Groundwater flow and sand body interconnectedness in a thick, multiple-aquifer system, *Water Resour. Res.*, *22*(5), 679–694, doi:10.1029/WR022i005p00679.
Fogg, G. E. (1990), Architecture and interconnectedness of geologic media: Role of the low permeability facies in flow and transpo, in *Hydro-Geology of Low Permeability Environments*, edited by S. P. Neuman and I. Neretnieks, pp. 19–40, Heinz Heis, Hannover, Germany.
Freeze, R. A., and J. A. Cherry (1979), *Groundwater*, Prentice Hall, Englewood Cliffs, N. J.
Frippiat, C. C., and A. E. Holeyman (2008), A comparative review of upscaling methods for solute transport in heterogeneous porous media, *J. Hydrol.*, *362*(1-2), 150–176, doi:10.1016/j.jhydrol.2008.08.015.
Garabedian, S. P., and D. R. Leblanc (1991), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts 2. Analysis of spatial moments for a nonreactive tracer, *Water Resour. Res.*, *27*(5), 911–924.
Gelhar, L. W. (1993), *Stochastic Subsurface Hydrology*, Prentice Hall, Upper Saddle River, N. J.
Gelhar, L. W., and C. L. Axness (1983), Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, *19*(1), 161–180.
Ginn, T. R. (2001), Stochastic–convective transport with nonlinear reactions and mixing: Finite streamtube ensemble formulation for multicomponent reaction systems with intra-streamtube dispersion, *J. Contam. Hydrol.*, *47*(1), 1–28, doi:10.1016/S0169-7722(00)00167-4.
Haggerty, R., S. A. Mckenna, and L. C. Meigs (2001), On the late-time behavior of tracer test breakthrough curves, *Water Resour. Res.*, *36*(12), 3467–3479.
Harvey, C., and S. M. Gorelick (2000), Rate-limited mass transfer or macrodispersion: Which dominates plume evolution at the macrodispersion experiment (MADE) site?, *Water Resour. Res.*, *36*(3), 637–650.
Hunt, A. G. (2001), Applications of percolation theory to porous media with distributed local conductances, *Adv. Water Resour.*, *24*(3-4), 279–307, doi:10.1016/S0309-1708(00)00058-0.
Irsa, J., and Y. Zhang (2012), A direct method of parameter estimation for steady state flow in heterogeneous aquifers with unknown boundary conditions, *Water Resour. Res.*, *48*, W09526, doi:10.1029/2011WR011756.
Jussel, P., and F. Stauffer (1994), Transport modeling in heterogeneous aquifers : 1. Statistical description and numerical generation of gravel deposits of the aquifer, *Water Resour. Res.*, *30*(6), 1803–1817.
Kitanidis, P. K. (1994), Particle-tracking equations for the solution of the advection-dispersion equation with variable coefficients, *Water Resour. Res.*, *30*(11), 3225–3227.
Klise, K. A., V. C. Tidwell, and S. A. McKenna (2008), Comparison of laboratory-scale solute transport visualization experiments with numerical simulation using cross-bedded sandstone, *Adv. Water Resour.*, *31*(12), 1731–1741, doi:10.1016/j.advwatres.2008.08.013.
Knudby, C., and J. Carrera (2005), On the relationship between indicators of geostatistical, flow and transport connectivity, *Adv. Water Resour.*, *28*(4), 405–421, doi:10.1016/j.advwatres.2004.09.001.
Labolle, E. M., and G. E. Fogg (2001), Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, *Transp. Porous Media*, *42*(1-2), 155–179.
Li, S., Y. Zhang, and X. Zhang (2011), A study of conceptual model uncertainty in large-scale CO2 storage simulation, *Water Resour. Res.*, *47*, W05534, doi:10.1029/2010WR009707.
Liu, G., Y. Chen, and D. Zhang (2008), Investigation of flow and transport processes at the MADE site using ensemble Kalman filter, *Adv. Water Resour.*, *31*(7), 975–986, doi:10.1016/j.advwatres.2008.03.006.
Meerschaert, M. M., D. A. Benson, and B. Ba (1999), Multidimensional advection and fractional dispersion, *Phys. Rev. E*, *59*(5), 5026–5028.
Meerschaert, M. M., M. Dogan, R. L. Van Dam, D. W. Hyndman, and D. A. Benson (2013), Hydraulic conductivity fields: Gaussian or not?, *Water Resour. Res.*, *49*, 4730–4737, doi:10.1002/wrcr.20376.
Milliken, W. J., M. Levy, S. Strebelle, C. Energy, T. Company, and Y. Zhang (2008), The effect of geologic parameters and uncertainties on subsurface flow: Deepwater depositional systems, in *SPE Western Regional and Pacific Section AAPG Joint Meeting, SPE 114099*, Society of Petroleum Engineers, pp. 1–16.
Monsen, E., T. Randen, L. Sonneland, and J. E. Odegard (2005), Geological model building: A hierarchical segmentation approach, in *Mathematical Methods and Modeling in Hydrocarbon Exploration and Production*, edited by A. Iske and T. Randen, pp. 213–245, Springer, N. Y.
Neuman, S. P. (1993), Eulerian-Lagrangian Theory of transport in space-time nonstationary velocity fields: Exact nonlocal formalism by conditional moments and weak approximation, *Water Resour. Res.*, *29*(3), 633–645.
Neuman, S. P. (2003), Multifaceted nature of hydrogeologic scaling and its interpretation, *Rev. Geophys.*, *41*(3), 1014, doi:10.1029/2003RG000130.

Neuman, S. P., and D. M. Tartakovsky (2009), Perspective on theories of non-Fickian transport in heterogeneous media, *Adv. Water Resour.*, *32*(5), 670–680, doi:10.1016/j.advwatres.2008.08.005.
Poeter, E., and P. Townsend (1994), Assessment of critical flow path for improved remediation management, *Ground Water*, *32*(3), 439–447.
Pollock, D. W. (1988), Semianalytical computation of path lines for finite-difference models, *Ground Water*, *26*(6), 743–750, doi:10.1111/j.1745–6584.1988.tb00425.x.
Prickett, T. A., T. G. Naymik, and C. G. Longquist (1981), *A "Random-Walk" Solute Transport Model for Selected Groundwater Quality Evaluations*, Ill. State Water Surv., Champaign.
Ramanathan, R., A. Guin, R. W. Ritzi, D. F. Dominic, V. L. Freedman, T. D. Scheibe, and I. Lunt (2010), Simulating the heterogeneity in braided channel belt deposits: 1. A geometric-based methodology and code, *Water Resour. Res.*, *46*, W04515, doi:10.1029/2009WR008111.
Rehfeldt, K. R., J. M. Boggs, and L. W. Gelhar (1992), Field study of dispersion in a heterogeneous aquifer 3. Geostatistical analysis of hydraulic conductivity, *Water Resour. Res.*, *28*(12), 3309–3324.
Renard, P., and D. Allard (2013), Connectivity metrics for subsurface flow and transport, *Adv. Water Resour.*, *51*, 168–196, doi:10.1016/j.advwatres.2011.12.001.
Riva, M., A. Guadagnini, D. Fernandez-Garcia, X. Sanchez-Vila, and T. Ptak (2008), Relative importance of geostatistical and transport models in describing heavily tailed breakthrough curves at the Lauswiesen site., *J. Contam. Hydrol.*, *101*(1-4), 1–13, doi:10.1016/j.jconhyd.2008.07.004.
Riva, M., L. Guadagnini, and A. Guadagnini (2010), Effects of uncertainty of lithofacies, conductivity and porosity distributions on stochastic interpretations of a field scale tracer test, *Stochastic Environ. Res. Risk Assess.*, *24*(7), 955–970, doi:10.1007/s00477-010-0399-7.
Rubin, Y. (2003), *Applied Stochastic Hydrogeology*, Oxford Univ. Press, N. Y.
Salamon, P., D. Fernàndez-Garcia, and J. J. Gómez-Hernández (2006), A review and numerical assessment of the random walk particle tracking method., *J. Contam. Hydrol.*, *87*(3-4), 277–305, doi:10.1016/j.jconhyd.2006.05.005.
Salamon, P., D. Fernàndez-Garcia, and J. J. Gómez-Hernández (2007), Modeling tracer transport at the MADE site: The importance of heterogeneity, *Water Resour. Res.*, *43*, W08404, doi:10.1029/2006WR005522.
Scheibe, T., and S. Yabusaki (1998), Scaling of flow and transport behavior in heterogeneous groundwater systems, *Adv. Water Resour.*, *22*(3), 223–238, doi:10.1016/S0309-1708(98)00014-1.
Scheibe, T. D., and D. L. Freyberg (1995), Use of sedimentological information for geometric simulation of natural porous media structure, *Water Resour. Res.*, *31*(12), 3259–3270, doi:10.1029/95WR02570.
Sheets, B. A., T. A. Hickson, and C. Paola (2002), Assembling the stratigraphic record: Depositional patterns and time-scales in an experimental alluvial basin, *Basin Res.*, *14*(3), 287–301, doi:10.1046/j.1365-2117.2002.00185.x.
Tompson, A. F. B., and L. W. Gelhar (1990), Numerical Simulation of Solute transport in three-dimensional, randomly heterogeneous porous media, *Water Resour. Res.*, *26*(10), 2541–2562.
Tompson, A. F. B., E. G. Vomvoris, and L. W. Gelhar (1987), Numerical simulation of solute transport in randomly heterogeneous porous media: Motivation, model development, and application, *Rep. UCID-21281*, Lawrence Livermore Natl. Lab., Livermore, Calif.
Vogel, H. J., and K. Roth (2003), Moving through scales of flow and transport in soil, *J. Hydrol.*, *272*(1-4), 95–106, doi:10.1016/S0022-1694(02)00257-3.
Webb, E. K., and M. P. Anderson (1996), Simulation of preferential flow in three-dimensional, heterogeneous conductivity fields with realistic internal architecture, *Water Resour. Res.*, *32*(3), 533–545, doi:10.1029/95WR03399.
Wen, X. H., and J. J. Gómez-Hernández (1998), Numerical modeling of macrodispersion in heterogeneous media: A comparison of multi-Gaussian and non-multi-Gaussian models, *J. Contam. Hydrol.*, *30*(1-2), 129–156, doi:10.1016/S0169-7722(97)00035-1.
Willmann, M., J. Carrera, and X. Sánchez-Vila (2008), Transport upscaling in heterogeneous aquifers: What physical parameters control memory functions?, *Water Resour. Res.*, *44*, W12437, doi:10.1029/2007WR006531.
Zarlenga, A., I. Janković, and A. Fiori (2012), Advective transport in heterogeneous formations: The impact of spatial anisotropy on the breakthrough curve, *Transp. Porous Media*, *96*(2), 295–304, doi:10.1007/s11242-012-0088-8.
Zhang, Y. (2014), Nonlinear inversion of an unconfined aquifer: Simultaneous estimation of heterogeneous hydraulic conductivities, recharge rates, and boundary conditions, *Transp. Porous Media*, *102*(2), 275–299, doi:10.1007/s11242-014-0275-x.
Zhang, Y., and C. W. Gable (2008), Two-scale modeling of solute transport in an experimental stratigraphy, *J. Hydrol.*, *348*(3-4), 395–411, doi:10.1016/j.jhydrol.2007.10.017.
Zhang, Y., D. a. Benson, M. M. Meerschaert, and E. M. LaBolle (2007), Space-fractional advection-dispersion equations with variable parameters: Diverse formulas, numerical solutions, and application to the macrodispersion experiment site data, *Water Resour. Res.*, *43*, W05439, doi:10.1029/2006WR004912.
Zhang, Y., C. W. Gable, and B. Sheets (2010), Equivalent hydraulic conductivity of three-dimensional heterogeneous porous media: An upscaling study based on an experimental stratigraphy, *J. Hydrol.*, *388*(3-4), 304–320, doi:10.1016/j.jhydrol.2010.05.009.
Zhang, Y., B. Liu, and C. W. Gable (2011), Homogenization of hydraulic conductivity for hierarchical sedimentary deposits at multiple scales, *Transp. Porous Media*, *87*(3), 717–737, doi:10.1007/s11242-010-9711-8.
Zheng, C., and S. M. Gorelick (2003), Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale, *Ground Water*, *41*(2), 142–155, doi:10.1111/j.1745-6584.2003.tb02578.x.
Zheng, C., M. Bianchi, and S. M. Gorelick (2011), Lessons learned from 25 years of research at the MADE site., *Ground Water*, *49*(5), 649–62, doi:10.1111/j.1745-6584.2010.00753.x.
Zinn, B. (2003), When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields, *Water Resour. Res.*, *39*(3), 1051, doi:10.1029/2001WR001146.

## Erratum

In the second paragraph of the Introduction, the "Zhang et al." citation is incorrect: "Although nonlocal theories can often provide a better match [Zhang et al., 2011], ADE was found to capture [...]" The correct citation is Zheng et al., 2011. This version may be considered the authoritative version of record.