A lithofacies approach for modeling non-Fickian solute transport in a heterogeneous alluvial aquifer

Marco Bianchi[1], Chunmiao Zheng[2,3]

[1] British Geological Survey, Environmental Science Centre, Keyworth, Nottingham, United Kingdom.

[2] School of Environmental Science and Engineering, South University of Science and Technology of China, Shenzhen, China.

[3] Department of Geological Sciences, University of Alabama, Tuscaloosa, AL, United States.

Corresponding Author:

Marco Bianchi, British Geological Survey, Environmental Science Centre, Keyworth, Nottingham, NG12 5GG, United Kingdom. marcob@bgs.ac.uk. Tel. +44 (0)155 9363136

Key points:

- Lithofacies are mapped as basis for 3D hydraulic conductivity distribution.
- Non-Fickian transport behavior emerges naturally from lithofacies distribution.
- Verifiable explanations are developed for the plume behavior at the MADE site.

Abstract. Stochastic realizations of lithofacies assemblage based on lithological data from a relatively small number of boreholes were used to simulate solute transport at the well-known Macrodispersion Experiment (MADE) site in Mississippi (USA). With sharp vertical contrasts and lateral connectivity explicitly accounted for in the corresponding hydraulic conductivity fields, experimental results from a large-scale tracer experiment were adequately reproduced with a relatively simple model based on advection and local dispersion. The geologically based model of physical heterogeneity shows that one well interconnected lithofacies, with a significantly higher hydraulic conductivity and accounting for 12% of the total aquifer volume, may be responsible for the observed non-Fickian transport behavior indicated by the asymmetric shape of the plumes and by variations of the dispersion rate in both space and time. This analysis provides a lithological basis to the hypothesis that transport at MADE site is controlled by a network of high-conductivity sediments embedded in a less permeable matrix. It also explains the calibrated value of the ratio of mobile to total porosities used in previous modelling studies based on the dual-domain mass transfer approach. The results of this study underscore the importance of geologically plausible conceptualizations of the subsurface for making accurate predictions of the fate and transport of contaminants in highly heterogeneous aquifers. These conceptualizations may be developed through integration of raw geological data with expert knowledge, interpretation and appropriate geostatistical methods.

Keywords. solute transport, heterogeneity, MADE site, lithofacies

1 Introduction and background

Despite significant theoretical, experimental and computational advances, modelling of contaminant transport in heterogeneous aquifers is still challenging and subject of continuing debate in the scientific community [e.g., *Hadley and Newell*, 2014; *Neuman*, 2014; *Molz*, 2015]. Yet, accurate simulations of the fate of contaminants are needed to address an ever growing demand for clean groundwater resources and an increasing interest in the use of the subsurface for the storage of nuclear waste, CO_2, and heat.

Transport of nonreactive solutes through porous media is traditionally modelled with the advection–dispersion equation (ADE):

$$\frac{\partial C}{\partial t} = \nabla \cdot (\mathbf{D} \nabla C) - \nabla \cdot (\mathbf{v} C) \qquad (1)$$

where C is concentration, $\mathbf{v}$ is the macroscopic advective velocity, and $\mathbf{D}$ is the hydrodynamic dispersion coefficient tensor. The latter is a function of the molecular diffusion coefficient, $\mathbf{v}$, and fixed longitudinal (α_L), horizontal transverse (α_{TH}) and vertical transverse (α_{TV}) dispersivities. Because the first term on the right hand side of Equation (1) is analogous to Fick's law of molecular diffusion, solute transport described by the ADE is referred to as Fickian. However, tracer experiments at different scales very often show "anomalous" or non-Fickian features indicated by non-Gaussian asymmetric plumes, apparent loss of mass due to sequestration in relatively immobile zones, variations of mean transport velocity, and increases in the dispersion rates (i.e., dispersivity) with mean travel distance or in time [e.g., *Silliman et al.*, 1987; *Adams and Gelhar*, 1992;

Haggerty et al., 2001; *Levy and Berkowitz*, 2003; *Cortis and Berkowitz*, 2004; *Bromly and Hinz*, 2004; *Bianchi et al.*, 2011a; *Cherubini et al.*, 2013].

For nonreactive tracers, non-Fickian transport is observed in aquifers characterized by sharp contrasts in hydraulic conductivity (*K*) and by connectivity of high-*K* regions [*Zheng and Gorelick*, 2003; *Klise et al.*, 2009; *Bianchi et al.*, 2011b; *Zhang et al.*, 2013], which are commonly found in alluvial aquifers [e.g., *Fogg, 1986; Webb and Anderson*, 1996; *Fogg et al.*, 2000; *Labolle and Fogg*, 2001; *Baratelli et al.*, 2011; *Dell'Arciprete et al.*, 2014]. The inability of the Fickian approach to describe transport in such environments is explained by the fact that the travel distance required to reach asymptotic or scale-independent conditions for macroscopic Fickian dispersion is larger than the actual scale of the observed plumes [*Eggleston and Rojstacer*, 2000; *Berkowitz et al.*, 2006; *Neuman and Tartakovsky*, 2009; *Srinivasan et al.*, 2010; *Molz*, 2015]. In fact, a scale-dependent (i.e., pre-asymptotic) behavior is observed for dispersivity, which is in contrast with the fixed macroscopic dispersivity derived from the central spatial moments of the plumes [e.g., *Adams and Gelhar*, 1992].

Field data collected at the research site in Columbus (Mississippi, USA), known as the Macrodispersion Experiment (MADE) site, have been used over the last three decades to investigate solute transport processes in alluvial aquifers. In particular, three large-scale natural gradient tracer experiments were conducted at this site in the mid '80s and in the '90s to test the applicability of the macrodispersion theory to explain solute transport in heterogeneous porous media [*Boggs et al.,*1992; *Boggs et al.*, 1993; *Boggs et al.*, 1995]. A comprehensive list of references of the numerous studies concerning the geological and hydrogeological characterization of the MADE site, as well as the results, interpretation, and modelling of the tracer experiments, is given in the review paper by *Zheng et al.* [2011]. Although the physical heterogeneity of the aquifer was initially characterized by more than 2500 flowmeter *K* measurements [*Rehfeldt et al.*, 1992], the application of the macroscopic ADE failed to explain transport behavior observed during the three large-scale experiments [*Adams and Gelhar*, 1992; *Eggleston and Rojstaczer* 1998a, 1998b; *Harvey and Gorelick*, 2000; *Feehley et al.*, 2000; *Julian et al.*, 2001].

The failure of the macroscopic ADE to accurately describe the experimental data at the MADE site has been the motivation for the application of alternative modelling methods based on two approaches. The first approach is represented by non-Fickian transport models including the dual domain mass transfer model [*Harvey and Gorelick,* 2000; *Feehley et al.*, 2000; *Guan et al.*, 2008; *Llopis-Albert and Capilla*, 2009], the fractional advective-dispersive equation [*Benson et al.* 2001; *Zhang and Benson*, 2008], and the continuous-time random walk [*Berkowitz and Scher*, 1998; *Berkovitz et al.*, 2006]. These models were able to provide a reasonable interpretation of the anomalous features of the observed plumes without an explicit representation of local-scale heterogeneity and connectivity, although their effect on transport is taken into account through mathematical formulations describing non-Fickian transport in time and space. A second approach, namely the local-ADE approach [e.g., *Fiori et al.*, 2013], considers an explicit representation of small-scale heterogeneities based on the notion that if the velocity field is sufficiently characterized, then transport can be effectively described by Equation (1) considering advection, molecular diffusion, and local dispersion [e.g., *Zheng and Gorelick*, 2003; *Salamon et al.*, 2007; *Zheng et al.*, 2011; *Fiori et al.*, 2013].

A recent application of the local-ADE approach at the MADE site is the study by *Dogan et al.* [2014], in which flowmeter measurements and additional high-resolution *K* data, collected with the direct-push injection logger [DPIL; *Liu et al.*, 2009; *Bohling et al.*, 2012], were used to generate extremely detailed representations of the *K* field [*Meerschaert et al.*, 2013] in a sector of the MADE site aquifer. This sector is about 1/6 of the total extension of the domain investigated by the three large-scale tracer experiments. Transport simulations based on nine stochastic realizations of the *K* field showed a good agreement with experimental data collected during the first tracer test (MADE-1). Results from this work are significant because they provide strong confirmation that the local ADE approach can predict solute transport in very heterogeneous porous media such as the MADE site aquifer. However, the computational effort (on a grid of 0.25 m × 0.25 m ×0.05 m, which amounts to approximately 111 million nodes for the entire MADE site domain of 120 m ×290 m × 10 m) and the amount of data used for generating the *K* field realizations (more than 5,500 measurements) were very substantial and not usually attainable.

Thus, in this work we test the hypothesis that we can explain the characteristics of the observed transport behavior at the MADE site with a much simpler local ADE-based model, without relying on exceedingly fine grid spacing or thousands of *K* data points. Differently from all the previous studies at the MADE site, we considered lithological data rather than *K* measurements (either from flowmeter or DPIL) to generate geologically consistent realizations of the spatial assemblage of five lithofacies, identified from a relatively small set of aquifer samples. These realizations were then used as basis for the *K* fields in transport simulations of the MADE-2 experiment. The agreement between simulated and experimental data provides an unprecedented lithological explanation for the observed non-Fickian transport behavior at the MADE site, while also demonstrating that this behavior can be adequately simulated by a local ADE-based model without an extraordinarily high-resolution characterization of the *K* field.

2 Data

Lithological data consist of 411 aquifer samples collected from 38 boreholes covering the total thickness of the aquifer (about 11 m on average). Location of these boreholes is shown in Figure 1, while lithological descriptions and the results of grain-size analyses performed on a subset of 214 soil samples from 29 boreholes are presented in a preliminary hydrogeological characterization study of the MADE site [*Boggs et al.,* 1990]. Aquifer sampling was conducted using a hollow stem auger and split core barrel samplers [*Boggs et al.*, 1990; 1992] and samples were generally collected at 1.5 meter intervals. The majority of the boreholes are located in the southern sector of the site, with only 6 boreholes located within the boundary of the network of multilevel sampling wells used to monitor concentrations during the tracer experiments.

Grain size data consist of percentages of gravel (diameter of soil grains greater than 4.76 mm), sand (diameter between 0.074 mm and 4.76 mm) and fines (diameter smaller than 0.074 mm). Values of the 10th (d_{10}), 25th (d_{25}), and 60th (d_{60}) percentiles of the cumulative grain size distribution are also available. Most of the aquifer at the MADE site consists of bimodal mixtures of gravel and sand with a low percentage of fines (less than 5% on average). In general, mixtures of gravel, sand and fines are more predominant in the most superficial part of the aquifer (up to 4 m of depth). Gravel content decreases with depth (less than 25% on average), and it is particularly low toward the bottom boundary of the aquifer represented by low-permeable marine deposits of the Eutaw formation. This deeper portion of the aquifer consists mostly of well sorted sand with fines content ranging from 1% up to 22%. Additional details on the vertical variability of gravel, sand, and fines content are provided by *Boggs et al.* [1990, 1992].

3 Methods

3.1 Lithofacies identification

Aquifer samples were classified into five lithofacies on the basis of the relative content of gravel (G), sand (S) and fines (f), as well as of values of d_{10}, d_{25} and of the uniformity coefficient ($U = d_{60} / d_{10}$). The criteria used for the identification of these lithofacies and key parameters are summarized in Table 1.

Lithofacies HCG ("highly conductive gravel") and GS ("gravel with sand"), which represent the 12% and the 18% of the samples respectively, consist of poorly sorted sandy gravels (gravel content > 50%) with minor fines (< 5%). The two lithofacies are distinguishable on the basis of the d_{10} (> 0.25 mm for HCG) and d_{25} values (> 1.0 mm for HCG). The two threshold values of 0.25 mm and 1.0 mm were chosen to be corresponding to the smallest grain sizes to define "medium sands" and "very coarse sands" according to the widely used soil classification by *Krumbein* [1934]. Grain size in HCG is also relatively more uniform than in GS (U = 30 vs. 41). In particular, HCG represents coarse gravelly sediments, as shown by the values of the d_{60} with values ranging between 6.4 mm and 19.7 mm. Lithofacies SGf ("sand, gravel and fines") consists of mixtures of gravel, sand and fines in various proportions. In general, sand content is higher than that of gravel, although some samples have gravel content up to 70%. The content of fines is higher than 5% in all samples. This lithofacies is the most represented in the aquifer samples (35%). Lithofacies SG ("sand and gravel") consists of moderately sorted gravelly sands and represents the 14% of the samples. On average, SG has moderately high sand content (about 65%), minimal fines (< 3% average), and d10 values similar to those in GS, albeit with more uniformity in the grain-size distribution (U=16). Lithofacies S ("sand") consists of well sorted sand (sand content > 85%; average U = 2.6) with an average d_{10} values similar to that in SGf (0.14 mm and 0.12 mm, respectively).

3.2 Stochastic simulation of lithofacies assemblage

Spatial continuity of the identified lithofacies was initially assessed along cross-sections intercepting the boreholes to identify transition trends and estimate lateral and vertical extensions. In a second stage, transition probabilities between lithofacies were calculated and modelled with a three-dimensional Markov chain in a conditional simulation framework [*Carle*, 1999]. The transition probability approach introduced by *Carle and Fogg* [1996, 1997] has been used to produce geologically consistent representations of subsurface heterogeneity by preserving the connectivity of lithofacies and juxtapositional tendencies [e.g., *Carle et al.*, 1998; *Weissmann and Fogg*, 1999; *Ritzi*, 2000; *Ritzi et al.*, 2004; *Lee et al.*, 2007; *Dai et al.*, 2007; *Ye and Khaleel*, 2008; *Bianchi et al.*, 2011b]. Differently from traditional variogram-based geostatistical methods, with this approach the spatial structure of the data is represented by transition probabilities, which are defined in terms of the following conditional probability:

$$t_{i,k}(\mathbf{h}) = \Pr\{k(\mathbf{x}+\mathbf{h}) \mid i(\mathbf{x})\} \qquad (2)$$

where $t_{i,k}$ is the transition probability from lithofacies i to lithofacies k, and $\mathbf{x}$ and $\mathbf{h}$ are the spatial location and lag distance vectors. Because, from Equation (2), the occurrence of lithofacies k at location $\mathbf{x} + \mathbf{h}$ is only dependent on the occurrence of lithofacies i at location $\mathbf{x}$, three-dimensional continuous-lag Markov Chain models can be developed to model discrete transition probabilities observed in the data. In this work, the fitting of a 3D Markov chain to the transition probabilities measured in the borehole data was performed by adjusting embedded transition probabilities and mean length and thickness values of lithofacies (Figure 2). Because of the relatively small number of boreholes, the estimation of mean length values from the plots of auto-transition probabilities in the horizontal direction is characterized by a certain degree of uncertainty. Therefore, in order to apply a more objective criterion for the estimation of the spatial correlation of the lithofacies in the horizontal direction, we have chosen to apply an early lag data approach [*Carle and Fogg*, 1997] in which the lag-one transition probability was used to compute the Markov chain model. This fit also produces probabilistic estimates of the mean length for each lithofacies (Figure 2a). We also tested the sensitivity of the transport modelling results with respect to this choice, especially regarding variations of the mean length of lithofacies HCG. The results of this sensitivity analysis will be discussed later. The calibrated Markov chain model also assumes isotropic behavior in the horizontal plane and lithofacies SGf as the background category. Volumetric proportions of the lithofacies, represented by the sill of the transiograms in the model, are also assumed equal to the proportions exhibited by the borehole data. Modeled transition probabilities and values of mean length and thickness for each lithofacies (Table 1) are reasonable and consistent with the spatial continuity assessed in the cross-sections. The mean lengths of the lithofacies inferred from the transition probability analysis is of the order of tens of meters, while thicknesses are in the order of a meter indicating higher variability along the vertical direction as in shown by previous investigations [e.g., *Rehfeldt et al.*, 1992; *Bohling et al.*, 2012]

3.3 Flow and transport simulations

A three-dimensional stochastic flow and transport model was implemented to simulate the second large scale tracer experiment (MADE-2; *Boggs et al.*, 1993). The block-centered numerical grid, with a total size of 120 m × 290 m × 10 m (Figure 1), has a resolution of 2 m in the horizontal plane and 0.5 m in the vertical direction. The total number of cells of the numerical grid is about 1.82×10^4, which is about 18 times less than the number of cells in the model by *Dogan et al*. [2014], even though the latter considers a smaller domain.

The K fields in the numerical simulations are directly linked to the spatial distribution of the identified lithofacies. These were generated according to the following procedure. In a first step, K values for each sample were estimated with the Kozeny-Carman empirical formula [e.g., *Riva et al.*, 2010]:

$$K = 8.3 \times 10^{-3} \frac{g\theta^3}{\upsilon(1-\theta)^2} d_e^2 \qquad (3)$$

where g is gravity (9.81 m^2/s), υ is the kinematic coefficient of viscosity of water (1.002 m^2/s at 20 ºC), d_e is a representative grain diameter, and θ is porosity. Porosity was estimated according to the empirical formula of *Vucovic and Soro* (1992) :

$$\theta = 0.255\left(1 + 0.83^U\right) \qquad (4)$$

Porosity values for each lithofacies (Table 1) and the average of all the samples (0.307) are similar to measurements in collected aquifer samples (*Boggs et al.*, 1990; *Boggs et al.*, 1992). Although there are different interpretations for d_e in the literature [e.g., *Koltermann and Gorelick*, 1995], here it was assumed to be corresponding to d_{10} for lithofacies GS, SG, SGf and S. There is in fact experimental evidence showing the reliability of this assumption in medium to coarse gravelly sands [*Odong*, 2007] and in well to moderately sorted sand/gravel mixtures [e.g., *Barahona-Palomo et al.*, 2011]. The average between d_{10} and d_{25} was chosen instead for lithofacies HCG because of lower sand content and significantly coarser grain size (Table 1). With this choice, estimated K values for the HCG samples are also more comparable with previous K estimates [*Boggs et al.,* 1990; *Eggleston and Rojstaczer*, 1998b] based on a different empirical formula, which was developed specifically for gravel and sand mixtures [*Seiler,* 1973]. In the subsequent discussion, we will test the effect of this assumption on simulated transport behavior. In a second step, descriptive statistics of the log transformed *K* estimates were computed for the five lithofacies (Table 1). As expected given the coarsest grain size, statistical analysis of the estimated *K* values for each lithofacies (Figure 3) indicates that HCG is significantly the most conductive lithofacies, with a mean *K* value that is about 1.5 to 2 orders of magnitude higher than the mean values of the other lithofacies. Next, three-dimensional conditional realizations of the spatial assemblage of lithofacies were generated according to the calculated transition probabilities and fitted Markov chain model [*Carle et al.*, 1998; *Carle*, 1999]. In the transport model domain the realizations are conditioned to the lithofacies identified in the samples from 6 boreholes (Figure 1). In the final step, an appropriate *K* value was assigned to each cell of the numerical grid of transport simulations according to the simulated distribution of lithofacies. This value was randomly generated from the truncated lognormal distribution, with mean and standard deviation equal to the corresponding values for each lithofacies. One standard deviation below and above the mean were considered as truncation thresholds to avoid excessive overlapping among different lithofacies and preserve the lithological structure on the generated *K* fields.

Groundwater flow was simulated in three stress periods of the duration of 2, 158 and 168 days using MODFLOW-2005 [*Harbaugh*, 2005]. The duration of the first stress period was chosen to represent the tritium injection. During the MADE-2 experiment, a total of 9.3 m^3 of a solution containing tritium was injected for approximately 48 hours through a linear array of five injection wells, spaced 1 m apart, and centered on the origin of the Cartesian coordinates system in Figure 1 (Boggs et al. 1993). The injection wells were screened at a depth interval between 57.5 m and 58.1 m a.s.l. The injection procedure in the model was simplified such that only two cells of the numerical grid were considered for the injection. However, the location of these cells and the total injected tritium mass

(0.5387 Ci) are consistent with the experimental conditions. The remaining stress periods were chosen to represent two distinct climatic periods observed over the 328 days of the experiment, which are clearly shown by significant water table fluctuations registered by the groundwater level monitoring network [*Boggs et al.*, 1993; *Stauffer et al.*, 1994; *Guan et al.*, 2008]. Accordingly, average values of groundwater levels measured at different wells during these two climatic periods were used to define specified-head boundary conditions at $\mathrm{Y = \text{-}20\ m}$ and $\mathrm{Y = 270\ m}$, while no-flow boundary conditions were imposed at $\mathrm{X = \text{-}50\ m}$, $\mathrm{X = 70\ m}$ and $\mathrm{Z = 52\ m}$. Despite the possible importance of transient flow conditions on transport at the MADE site [*Llopis-Albert and Capilla*, 2009], flow was assumed steady state in all stress periods. The ratio between vertical and horizontal *K* assumed in the model (0.13) is based on the results of a pumping test conducted at the MADE site [*Boggs et al.*, 1990].

Transport simulations based on Equation (1) were performed with MT3DMS [*Zheng*, 2010] with the advection component solved with the total-variation-diminishing (TVD) scheme to minimize numerical dispersion given the relative coarseness of the grid and avoid mass balance inconsistencies. A Courant number of 0.75 was used for all transport simulations. Porosity values were assigned to the grid according to the lithofacies distribution. These correspond to the average of the values estimated with Equation 4 for each lithofacies (Table 1). Other input parameters include a molecular diffusion coefficient for tritium of 1.16×10^{-9} m^2/s [*Salomon et al.*, 2007], α_L equal to 1 m [*Feehley et al.*, 2000; *Llopis-Albert and Capilla*, 2009], and values of α_{TH} and α_V of one and two orders of magnitude lower than α_L.

The accuracy of the implemented model was tested by comparing simulated and observed 1-D longitudinal mass distributions at 27, 132, 224, and 328 days after the injection. These times correspond to the first four "snapshots" of the MADE-2 experiment [*Boggs et al.*, 1993]. For the calculation of experimental mass distributions, the mass along each monitoring well was integrated vertically and then interpolated in 2-D over the same grid used for flow and transport simulations. Observed and simulated mass distributions for each snapshot were then obtained by integrating the fraction of total recovered mass in 30 equally spaced zones of 10 m width along the general flow direction (y axis).

The mean longitudinal displacement ($\overline{y}$) and the longitudinal variance of the observed and simulated 1-D mass profiles (σ_{YY}^2) were also calculated on the basis of the central spatial moments according to the following equations (e.g., *Adams and Gelhar*, 1992):

$$\overline{y} = M_1 / M_0 \qquad (5)$$

and

$$\sigma_{YY}^2 = M_2 / M_0 - M_1^2 / M_0 \qquad (6)$$

The generic spatial moment M_i for the observed and simulated longitudinal mass profiles was calculated with the following equation:

$$M_i = \sum_{p=1}^{N} m_p y^i \qquad (7)$$

where m_p is the fraction of recovered mass at the point p of coordinates y, and N is the total number of points. Note that since tritium mass was normalized with the total recovered mass, the zero-th moment M_0 is equal to 1 for both observed and simulated mass profiles.

4 Results and discussion

The ensemble mean and median of 1-D longitudinal mass distributions from 500 Monte Carlo realizations of the model are shown in Figure 4a-d. The interquartile range is also reported to provide a description of the variability of the simulated results. In general, the model is accurate in reproducing the mass accumulation near the injection site and the spreading to the far field. The model tends to overestimate the position of the edge of the plume in the first two snapshots, even though the mismatch is limited to fractions of recovered mass below 0.01. At later times (224 and 328 days), the model does not match the relative peak of mass observed between 160 m and 200 m from the injection site. This peak is most probably the effect of transient variations in the flow field during the experiment as suggested by fluctuations in the water table of up to 30% of the saturated thickness, which were observed during later stages of the MADE-2 test [*Stauffer et al.*, 1994; *Llopis-Albert and Capilla*, 2009]. These variations were not considered in the presented model. A better match between observed and simulated mass profiles could also be probably achieved with calibration of some of the model input parameters (e.g., porosity and K values of the lithofacies, boundary conditions). However, a calibration procedure not only is beyond the scope of the present work, but also would reduce the predictability of our lithofacies approach and compromise the insight about its transferability to other sites. Notwithstanding these simplifications, the implemented transport model is able to capture the overall characteristics of the MADE-2 plume with reasonable accuracy, especially considering the limited number of hard conditioning points used in the stochastic realizations of subsurface heterogeneity.

Reasonable accuracy is further confirmed by comparisons between observed and simulated central moments (Figure 5a-c). The percentage error between the observed longitudinal displacement and the ensemble mean of the simulated values is between 11%

and 51%. The highest discrepancy is calculated for the displacement at 132 days, because simulated plumes tend to advance too rapidly (9.2 m vs. 13.9 m). The error between observed and simulated displacement at 224 and 328 is around 25%, but this discrepancy is strongly influenced by the relative peak of mass observed at later times and by the extremely rapid movement of the center of mass observed between 132 and 224 days. One important aspect regarding the proposed model shown in Figures 5b and 5c is that the second central moment representing the longitudinal variance of the plume grows at different rates in both time and space. This characteristic and the asymmetric shape of the simulated mass distributions are indicative of non-Fickian transport behavior.

From the comparison between the spatial distributions of the identified lithofacies (Figure 6a), the corresponding K fields (Figure 6b), and the location of the plume front at different times (Figure 6c), it is evident that the asymmetric shape of the plume and the rapid movement of the edge are controlled by the location and the lateral continuity of the highly conductive lithofacies HCG. Given the dimension of the simulated domain in the longitudinal direction (145 cells) and its mean length (30 m = 15 cells), the percolation threshold for lithofacies HCG is expected to be around 0.14, according to *Harter* [2005]. Because the percolation threshold corresponds to the critical volumetric fraction for which there is occurrence of one cluster of cells spanning the entire domain, the estimated volumetric fraction of 0.12 for lithofacies HCG indicates that this lithofacies defines an interconnected network of high-K values that almost fully percolate the MADE site aquifer. This result provides a further confirmation of the hypothesis advanced by several previous studies [e.g., *Fogg*, 1986; *Fogg et al.*, 2000; *Labolle and Fogg*, 2001; *Zheng and Gorelick*, 2003; *Zheng et al.*, 2011; *Moltz*, 2015] that the "anomalous" transport behavior observed in heterogeneous alluvial aquifers is mostly the effect of connectivity of high-K sediments. This connectivity enhances fast advective transport of a fraction of mass along preferential flow-paths, while a larger fraction travels in a relatively less permeable matrix. In the matrix, the role of diffusive transport is more significant especially in directions perpendicular to the main flow. When high-*K* zones connectivity is taken into account, faster than expected breakthrough times and late-time tailing of contaminants concentrations, which are commonly observed in contaminated aquifer sites, can be successfully predicted [*Labolle and Fogg*, 2001].

The influence of lithofacies HCG on the velocity field and consequently on advective transport is also shown by the analysis of the frequency distributions of the generated *K* fields (Figure 7). These are clearly bimodal, with the majority of the $\log_{10}(K)$ values clustered around a value of about 0.75 m/d, and a smaller set of values around the average value for lithofacies HCG (Table 1). Comparisons between the distribution for the generated *K* fields and the distributions of *K* data previously collected at the MADE site with two different methods [*Rehfeldt et al.*, 1992; *Bohling et al.*, 2012] indicate similarity between the modal value of the *K* estimates for lithofacies GS, SGf, SG and S and average value of the flowmeter measurements. The *K* estimates for lithofacies HCG are also comparable to the upper tails of the distributions of both the flowmeter and the DPIL data. However, the three *K* data sets differ in terms of sample variances, and the correlation between corresponding values at different depths in boreholes located within a 3.5 m radius is generally poor. A discussion of the possible causes for the mismatch between the flowmeter data and the DPIL data is presented by *Bohling et al.* [2012], while mismatches between the *K* estimates based on grain-size analysis and flowmeter data have been also observed in other alluvial aquifers [*Barahona-Palomo et al.*, 2011; *Guting et al.*, 2015]. As for these other aquifers, the lack of correlation between types of *K* data for the MADE site aquifer is most likely explained by the difference in the support scale associated with each method, which ranges from a few centimeters for DPIL, to about 1.5 decimeters for the flowmeter measurements, up to several decimeters for the grain-size estimates.

Our interpretation may also provide a geological explanation for the success of the dual-domain mass transfer rate approach (DDM) in reproducing the experimental data at this site [*Harvey and Gorelick*, 2000; *Feehley et al.*, 2000; *Guan et al.*, 2008; *Bianchi et al.*, 2011a]. This approach simulates transport in two distinct but overlapping mobile and immobile domains, each characterized by a certain porosity value, and the total porosity of the system is given by the sum of the mobile and immobile porosities. A mass transfer rate coefficient controls the exchange of solute mass between the two domains. According to the dual-domain conceptualization, pore space in the mobile domain is filled with water that can actually move through the porous structure and solute transport is mainly due to advection. On the other hand, pores in the immobile domain are filled with stagnant water and molecular diffusion is the main transport process. This separation into two mobile and immobile domains is therefore particularly suitable for reproducing transport when interconnected high-*K* sediments (i.e., the mobile domain) are embedded in a relatively lower permeable matrix (i.e., the immobile domain).

Because our results suggest that the lithofacies HCG can be considered the mobile domain through which fast advective transport occurs, it is very noteworthy that the volumetric fraction estimated from the borehole data (0.12) corresponds to the calibrated value of the ratio between mobile and total porosities (1/8 = 0.125) of dual-domain models, which were able to fit the observed plume spreading at the MADE site [*Zheng et al.*, 2011 and references therein]. As a confirmation, we implemented a DDM model (single-rate) based on a homogenous field with *K* equal to the ensemble mean of the equivalent *K* values for a subset of realizations of the *K* field. For each realization, the equivalent *K* was estimated by applying Darcy's law between the two specified-head boundaries of the simulated domain in Figure 1, and by assuming a preservation of the total discharge. This approach is similar to that used by *Liu et al*. [2007] to test the applicability of the DDM to represent transport in binary *K* fields characterized by decimeter-scale highly conductive channels. The model also assumes a mobile to total porosity ratio equal to the volumetric fraction of HCG. Comparisons between observed and simulated plumes show that we can match the observed the transport behavior with adequate accuracy by a simple calibration of the mass transfer rate coefficient (Figure 8 and Figure 5d). As in the model proposed by *Guan et al.* [2008], calibrated values for this parameter indicate that the single-rate mass transfer coefficient is scale-dependent and decreases with time.

Results shown in Figure 4a-d are based on the input parameters of Table 1. Because of the uncertainty associated with some of these parameters and the dominant influence of lithofacies HCG on the simulated transport behavior, we also analyzed the sensitivity of the results with respect to changes of *K* and mean length for this lithofacies. The results for the snapshot at 328 days are shown in Figure 4e. When lithofacies HCG is ignored in the generation of the *K* fields and its *K* value is assumed equal to that of lithofacies GS, the mass distribution showed very limited spreading and a symmetric shape. A similar result was obtained in a scenario in which the *K* of lithofacies HCG is estimated by considering the d_{10} as the value for d_e in Equation 2. The model is also sensitive with respect to changes of the mean length of lithofacies HCG. However, even when the mean length is assumed to be one half of the value in Table 1, we still observe a significantly asymmetric mass distribution although the leading edge of the plume is about 40 m shorter. This result indicates that even if a small range of mean length values would fit the estimated auto-transition probabilities equally well for lithofacies HCG (Figure 2a), the main conclusion regarding its role on controlling non-Fickian transport is still valid.

5 Conclusions

Site-scale transport behavior observed during one of the MADE site experiments (MADE-2) was effectively reproduced with a relatively simple, local ADE-based model. The physical aquifer heterogeneity in the transport model was conceptualized and represented by 3-D realizations of the spatial distribution of lithofacies identified from aquifer samples collected from 39 boreholes, mostly located outside the domain used for transport simulations. The lithofacies approach appears to have provided an unprecedented explanation to "anomalous" plume-scale behavior at the MADE site that has motivated a long line of studies over the past 30 years. Furthermore, results suggest that such behavior can be reproduced with a model based on a much smaller set of aquifer property data than previously thought possible.

In particular, this analysis shows that some of the non-Fickian features of the observed plume can be explained by a highly permeable lithofacies with limited (less than 1 m) vertical extent and moderate (>10 m) horizontal correlation. The presence of a network of well interconnected highly permeable sediments embedded in a less permeable matrix has been previously suggested for the MADE site [*Harvey and Gorelick*, 2000; *Zheng and Gorelick*, 2003] and tested in small sectors of the investigated domain [*Liu et al.*, 2010; *Ronayne et al.*, 2011; *Bianchi et al.*, 2011a, 2011b], but never assessed at the scale of the large scale tracer experiments. In the context of about three decades of research work at the MADE site, the identification of the most conductive lithofacies (HCG) from borehole lithological data is a significant result providing a previously elusive, simple explanation for the observed non-Fickian transport behavior from a geological perspective.

The proposed model of physical heterogeneity for the MADE site aquifer seems also to provide a lithological basis for the success of dual-domain mass transfer rate approach in reproducing non-Fickian transport behavior at this site [*Zheng et al.*, 2011]. In this respect, this work can also be seen as a first successful attempt to infer the ratio between mobile to total porosities, which is at the basis of dual-domain conceptualization, from grain-size analysis data and volumetric fractions of lithofacies.

Even though this study is focused on a particular alluvial aquifer, the impact of the results is broader because they show that if the geological structure – here represented by the spatial distribution of the lithofacies – is properly represented in the 3-D hydraulic conductivity field, then solute transport in heterogeneous aquifers can be accurately simulated with local ADE-based models without relying on exceedingly fine grid spacing or high-resolution *K* data. The incorporation of the geological structure in the physical model of heterogeneity also provides verifiable explanations for the observed plume behavior. Therefore, this work underscores the importance of geologically based representations of the subsurface, which can be developed through integration of raw geological data (e.g., borehole logs, aquifer analog descriptions, geophysical surveys) with expert knowledge, interpretation and appropriate geostatistical methods.

Acknowledgments. This work was undertaken as part of the "Research Fellowship Programme" funded by the British Geological Survey (Natural Environment Research Council). Additional support was provided by the National Natural Science Foundation of China (Grant No. 41330632). We are grateful to Graham Fogg, Daniel Feinstein, and Erica Siirila-Woodburn for their constructive review comments , which have improved the final manuscript. Co-author Bianchi publishes with the permission of the Executive Director of the British Geological Survey. Lithological data used in this study can be

found in the referenced report *Boggs et al.* [1990]. Tracer data used in this study are available in the referenced report *Boggs et al.* [1993].

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.

Barahona-Palomo, M., M. Riva, X. Sanchez-Vila, E. Vazquez-Sune, and A. Guadagnini (2011), Quantitative comparison of impeller flowmeter and particle-size distribution techniques for the characterization of hydraulic conductivity variability, *Hydrogeol. J.,* 19(3), 603–612, doi:10.1007/s10040-011-0706-5.

Baratelli, F., M. Giudici, and C. Vassena, C. (2011), Single and dual domain models to evaluate the effects of preferential flow paths in alluvial porous sediments, *Transp. Porous Med.,* 87, 465–484, doi:10.1007/s11242-010-9695-4.

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

Benson, D. A., R. Schumer, M. M. Meerschaert, and S. W. Wheatcraft (2001), Fractional dispersion, Lévy motion, and the MADE tracer tests, *Transport in Porous Media*, 42, 211-240, doi: 10.1023/A:1006733002131.

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.

Bianchi, M., C. Zheng, G.R. Tick, and S.M. Gorelick (2011a), Investigation of Small-Scale Preferential Flow with a Forced-Gradient Tracer Test, *Groundwater,* 49, 503–514, doi: 10.1111/j.1745-6584.2010.00746.

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

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

Boggs, J. M., S. C. Young, D. J. Benton, and Y. C. Chung (1990). Hydrogeologic characterization of the MADE site, Interim Rep. EN-6915, Electr. Power Res. Inst., Palo Alto, Calif.

Boggs, J.M., S.C. Young, and L.M. Beard (1992), Field study of dispersion in a heterogeneous aquifer 1. Overview and site description, *Water Resour. Res.,* 28(12), 3281–3291.

Boggs, J.M., L.M. Beard, S.E. Long, M.P. McGee, W.G. MacIntyre, C.P. Antworth, and T.B. Stauffer (1993), Database for the Second Macrodispersion Experiment (MADE-2), Tech. Rep. TR-102072, Electric Power Res. Inst., Palo Alto, California.

Boggs, J.M., J.A. Schroeder, and S.C. Young (1995), Data to support model development for natural attenuation study. Report No. WR28-2-520-197. TVA Engineering Laboratory, Tennessee Valley Authority, Norris, Tennessee

Bromly, M., and C. Hinz (2004), Non-Fickian transport in homogeneous unsaturated repacked sand, *Water Resour. Res.*, 40, W07402, doi:10.1029/2003WR002579.

Carle, S.F. (1999), T-PROGS: Transition Probability Geostatistical Software, version 2.1. Davis, California: University of California.

Carle, S.F., and G.E. Fogg (1996), Transition probability-based indicator geostatistics, *Math. Geol.*, 28(4), 453–476.

Carle S.F., and G.E. Fogg (1997), Modeling spatial variability with one and multidimensional continuous-lag Markov chains, *Math. Geol.*, 29(7), 891–918.

Carle S.F., E.M. LaBolle, G.S. Weissmann, D. VanBrocklin, and G.E. Fogg (1998), Geostatistical simulation of hydrostratigraphic architecture: a transition probability / Markov approach, *in* Concepts in Hydrogeology and Environmental Geology No. 2, *SEPM Special Publication*, p. 147–170

Cherubini, C., C.I. Giasi, and N. Pastore (2013), Evidence of non-Darcy flow and non-Fickian transport in fractured media at laboratory scale, *Hydrol. Earth Syst. Sci.,* 17, 2599-2611, doi:10.5194/hess-17-2599-2013.

Cortis, A., and B. Berkowitz (2004), Anomalous transport in ‘‘classical’’ soil and sand columns, *Soil Science Society of America Journal,* 68, no. 5, 1539-1548.

Dai, Z., A. Wolfsberg, Z. Lu, and R. Ritzi Jr. (2007), Representing aquifer architecture in macrodispersivity models with an analytical solution of the transition probability matrix, *Geophys. Res. Lett.,* 34, L20406, doi:10.1029/2007GL031608.

Dell'Arciprete, D., C. Vassena, F. Baratelli, M. Giudici, R. Bersezio, and F. Felletti (2014), Connectivity and single/dual domain transport models: tests on a point-bar/channel aquifer analogue, *Hydrogeology Journal,* 22(4), 761-778. doi:10.1007/s10040-014-1105-5 ·

Dogan, M., R. L. Van Dam, G. Liu, M. M. Meerschaert, J. J. Butler Jr., G. C. Bohling, D. A. Benson, and D. W. Hyndman (2014), Predicting flow and transport in highly heterogeneous alluvial aquifers, *Geophys. Res. Lett.*, 41, 7560–7565, doi:10.1002/2014GL061800.

Eggleston, J. R., and S. Rojstaczer (1998a), Identification of large-scale hydraulic conductivity trends and the influence of trends on contaminant transport, *Water Resour. Res.*, 34(9), 2155–2168, doi:10.1029/98WR01475.

Eggleston, J. R., and S. Rojstaczer (1998b), Inferring spatial correlation of hydraulic conductivity from sediment cores and outcrops, *Geophys. Res. Lett.,* 25, 2317–2320, 1998.

Eggleston, J.R., and S.A. Rojstaczer (2000), Can we predict subsurface mass transport? *Environmental Science and Technology* 34, 18, 4010–4017, doi:10.1021/es000903s.

Feehley C.E., C. Zheng, and F.J. Molz (2000), A dual-domain mass transfer approach for modeling solute transport in heterogeneous porous media, application to the MADE site, *Water Resour. Res.,* 36(9), 2501–2515, doi:10.1029/2000WR900148.

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., S.F. Carle, and C. Green (2000), Connected-network paradigm for the alluvial aquifer system. In: Zhang, D., C.L. Winter, eds., Theory, modeling, and field investigation in hydrogeology: A special volume in honor of Shlomo P. Neuman's 60th birthday. Geological Society of America, Special paper 348, p. 25–42.

Guan, J., F. J. Molz, Q. Zhou, H. H. Liu, and C. Zheng (2008), Behavior of the mass transfer coefficient during the MADE-2 experiment: New insights, *Water Resour. Res.*, 44, W02423, doi:10.1029/2007WR006120.

Guting, N., A. Klotzsche, J. Van Der Kruk, J. Vanderborght, H. Vereecken, and A. Englert (2015), Spatially highly resolved mapping of aquifer heterogeneities using ground penetrating radar full-waveform tomography, AQUA 2015, 42nd IAH Congress, 13-18 September 2015, Rome, Italy.

Hadley, P. W., and C. Newell (2014), The New Potential for Understanding Groundwater Contaminant Transport, *Groundwater*, 52, 174–186, doi: 10.1111/gwat.12135

Haggerty, R., S. W. Fleming, L. C. Meigs, and S. A. McKenna (2001), Tracer tests in a fractured dolomite: 2. Analysis of mass transfer in single-well injection-withdrawal tests, *Water Resour. Res.,* 37(5), 1129–1142, doi:10.1029/2000WR900334.

Harbaugh, A.W. (2005), MODFLOW-2005, The U.S. Geological Survey modular ground-water model—the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16, variously p.

Harter, T. (2005), Finite-size scaling analysis of percolation in three-dimensional correlated binary Markov chain random fields. *Phys Rev E*, 72(2), 26120. doi:10.1103/PhysRevE.72.026120.

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, doi:10.1029/1999WR900247.

Julian, H.E., J.M. Boggs, C. Zheng, and C.E. Feehley (2001) Numerical simulation of a natural gradient tracer experiment for the Natural Attenuation Study: flow and physical transport, *Groundwater*, 39(4), 534-545.

Klise, K. A., G. S. Weissmann, S. A. McKenna, E. M. Nichols, J. D. Frechette, T. F. Wawrzyniec, and V. C. Tidwell (2009), Exploring solute transport and streamline connectivity using lidar-based outcrop images and geostatistical representations of heterogeneity, *Water Resour. Res.*, 45, W05413, doi:10.1029/2008WR007500.

Koltermann, C. E., and S. M. Gorelick (1995), Fractional packing model for hydraulic conductivity derived from sediment mixtures, *Water Resour. Res.,* 31(12), 3283–3297, doi:10.1029/95WR02020.

Krumbein W.C. (1934), Size frequency distributions of sediments*, Journal of Sedimentary Petrology*, 4, 65–77.

LaBolle E.M., and G.E. Fogg (2001), Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, *Transport Porous Media,* 42 (1-2), 155–179.

Levy, M., and B. Berkowitz (2003), Measurement and analysis of non-Fickian dispersion in heterogeneous porous media, *J. Contam. Hydrol.,* 64, 203–226.

Lee S-Y., S.F. Carle, and G.E. Fogg (2007), Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation, *Adv. Water Resour.*, 30, 1914–1932.

Liu, G., C. Zheng, and S. M. Gorelick (2004), Limits of applicability of the advection-dispersion model in aquifers containing connected high-conductivity channels, *Water Resour. Res.,* 40, W08308, doi:10.1029/2003WR002735.

Liu, G., C. Zheng, and S. M. Gorelick (2007), Evaluation of the applicability of the dual-domain mass transfer model in porous media containing connected high-conductivity channels, Water Resour. Res., 43, W12407, doi:10.1029/2007WR005965.

Liu, G., J.J. Butler Jr., G.C. Bohling, E. Reboulet, S. Knobbe and D.W. Hyndman (2009), A new method for high-resolution characterization of hydraulic conductivity, *Water Resour. Res.,* 45, W08202, doi:10.1029/2009WR008319.

Liu, G., C. Zheng, G. R. Tick, J. J. Butler Jr., and S. M. Gorelick (2010), Relative importance of dispersion and rate-limited mass transfer in highly heterogeneous porous media: Analysis of a new tracer test at the Macrodispersion Experiment (MADE) site, *Water Resour. Res.*, 46, W03524, doi:10.1029/2009WR008430.

Llopis-Albert, C. and J.E. Capilla (2009), Gradual conditioning of non-Gaussian transmissivity fields to flow and mass transport data: 3. Application to the Macrodispersion Experiment (MADE-2) site, on Columbus Air Force Base in Mississippi (USA). *J. Hydrol.*, 371, 1-4, 75–84.

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.

Molz, F.J. (2015), Advection, dispersion, confusion, *Groundwater*, 53: 348–353. doi: 10.1111/gwat.12338.

Neuman, S. P. (2014), The New Potential for Understanding Groundwater Contaminant Transport, *Groundwater*, 52, 653–656. doi: 10.1111/gwat.122

Neuman, S., Tartakovsky, D. (2009), Perspective on theories of non-fickian transport in heterogeneous media, *Adv. Water Resour.*, 32, 670–680 .

Odong J. (2007) Evaluation of empirical formulae for determination of hydraulic conductivity based on grain-size analysis, *J. Am. Sci.,* 3, 54–60.

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.

Ritzi, R.W. (2000), Behavior of indicator variograms and transition probabilities in relation to the variance in lengths of hydrofacies, *Water Resour. Res.*, 36(11), 3375-3381.

Ritzi, R. W., Z. Dai, D.F. Dominic, and Y.N Rubin (2004), Spatial correlation of permeability in cross-stratified sediment with hierarchical architecture, *Water Resour. Res.*, 40(3), W03513, doi: 10.1029/2003WR002420.

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, 955–970, doi:10.1007/s00477-010-0399-7.

Ronayne, M. J., S. M. Gorelick, and C. Zheng (2010), Geological modeling of submeter scale heterogeneity and its influence on tracer transport in a fluvial aquifer, *Water Resour. Res.,* 46, W10519, doi:10.1029/2010WR009348.

Salamon, P., D. Fernandez-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.

Seiler, K.-P. (1973), Durchlässigkeit, Porosität und Kornverteilung quartärer Kies-Sand-Ablagerungen des bayrischen Alpenvorlandes, Gas- und Wasserfach - Wasser, Abwasser – 114. Jahrgang, Heft 8, S. 353-358.

Silliman, S. E., L. F. Konikow, and C. I. Voss (1987), Laboratory investigation of longitudinal dispersion in anisotropic porous media, *Water Resour. Res.*, 23(11), 2145–2151, doi:10.1029/WR023i011p02145.

Srinivasan, G., D.M. Tartakovsky, M. Dentz, H. Viswanathan, B. Berkowitz, and B.A. Robinson (2010), Random walk particle tracking simulations of non-Fickian transport in heterogeneous media, *Journal of Computational Physics,* 229(11), 4304-4314.

Stauffer, T. B., C. P. Antworth, R. G. Young, W. G. MacIntyre, J. M. Boggs, and L. M. Beard (1994), Degradation of aromatic hydrocarbons in an aquifer during a field experiment demonstrating the feasibility of remediation by natural attenuation, Rep. AL/EQ TR 1993-0007, Armstrong Lab., Tyndall Air Force Base, Fla.

Vukovic´ M, Soro A (1992) Hydraulics and water wells: theory and application. Water Resources Publications, Highlands Ranch, CO, USA. 1143 Hydrogeology.

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.

Weissmann, G.S., S.F. Carle and G.E. Fogg, 1999, Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics, *Water Resour. Res.*, 35(6), 1761-1770.

Ye, M., and R. Khaleel (2008), A Markov chain model for characterizing medium heterogeneity and sediment layering structure, *Water Resour. Res.*, 44, W09427, doi:10.1029/2008WR006924.

Zhang, Y., C. Green, and G. Fogg (2013), The impact of medium architecture of alluvial settings on non-fickian transport, *Adv. Water Resour.,* 54, 78–99.

Zhang, Y., and D. A. Benson (2008), Lagrangian simulation of multidimensional anomalous transport at the MADE site, *Geophys. Res. Lett*., 35, L07403, doi:10.1029/2008GL033222.

Zheng, C., and S. M. Gorelick (2003), Analysis of the effect of decimeter scale preferential flow paths on solute transport, *Groundwater,* 41(2), 142– 155.

Zheng, C. (2010), MT3DMS v5.3 Supplemental User's Guide, Technical Report, Department of Geological Sciences, University of Alabama, Tuscaloosa, Alabama.

Zheng, C., M. Bianchi, and S.M. Gorelick (2011), Lessons learned from 25 years of research at the MADE site, *Groundwater,* 49, 649–662, doi:10.1111/j.1745–6584.2010.00753.x.

LIST OF TABLES

Table1. Criteria used for lithofacies identification and representative parameters.

	Highly conductive gravel (HCG)	Gravel with sand (GS)	Sand gravel and fines (SGf)	Sand and gravel (SG)	Well sorted sand (S)
Identification criteria	$G > 50\%$ $f < 5\%$ $d_{10} > 0.25$ mm $d_{25} > 1$ mm	$G > 50\%$ $f < 5\%$	$f > 5\%$	$S>50\%$ $f<5\%$	$S > 85\%$ $U < 3$
G^* [%]	64.6	56.2	40.8	32.2	3.1
S^* [%]	32.0	40.7	51.7	64.9	90.2
f^* [%]	3.4	3.1	7.5	2.9	6.7
d_{10}^* [mm]	0.62	0.22	0.14	0.21	0.12
d_{25}^* [mm]	2.7	0.72	0.45	0.36	0.16
d_{60}^* [mm]	12.4	8.73	5.56	3.3	0.28
U*	30.4	41.0	38.3	15.6	2.6
Proportions [%]	12	18	35	14	21
Mean length [m]	30	31	39	25	35
Mean thickness [m]	1.0	0.5	0.9	0.4	1.7
Mean $Log_{10}(K)$ [m/d]	2.482	0.830	0.402	0.889	0.752
Variance $Log_{10}(K)$	0.589	0.210	0.343	0.228	0.165
Mean θ	0.265	0.257	0.259	0.298	0.415

G: gravel content;

S: sand content;

f: fines content

*average value

LIST OF FIGURES

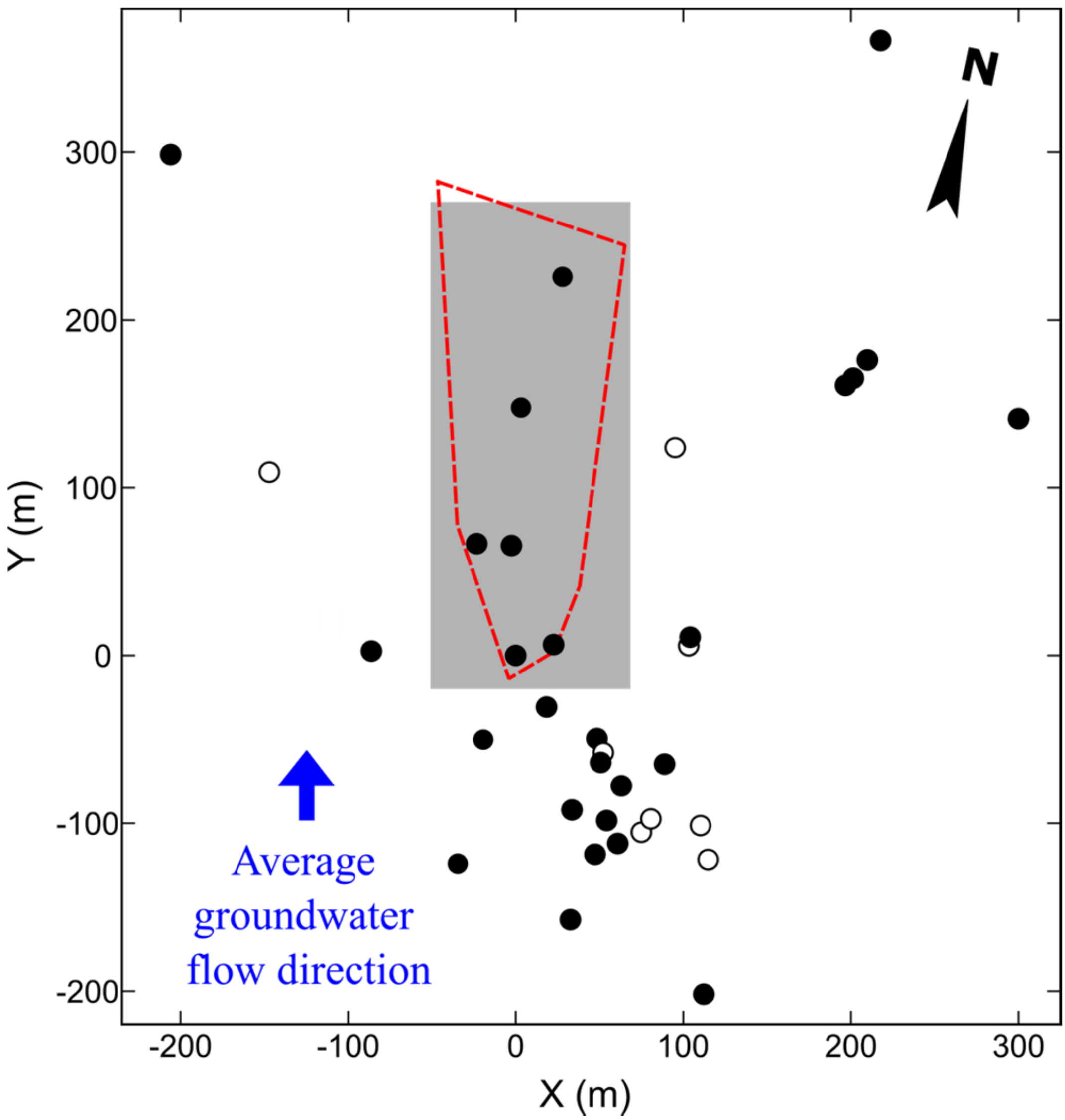

Figure 1. Map of boreholes used for lithological characterization of the MADE site. Black circles indicate boreholes with grain size data in Appendix A in *Boggs et al.* [1990]. Boreholes with only lithological description are indicated by open circles. The grey shaded area indicates the extension of the domain used for flow and transport modelling. The red dashed line indicates the boundary of the network of multilevel sampling wells used during the large-scale tracer tests.

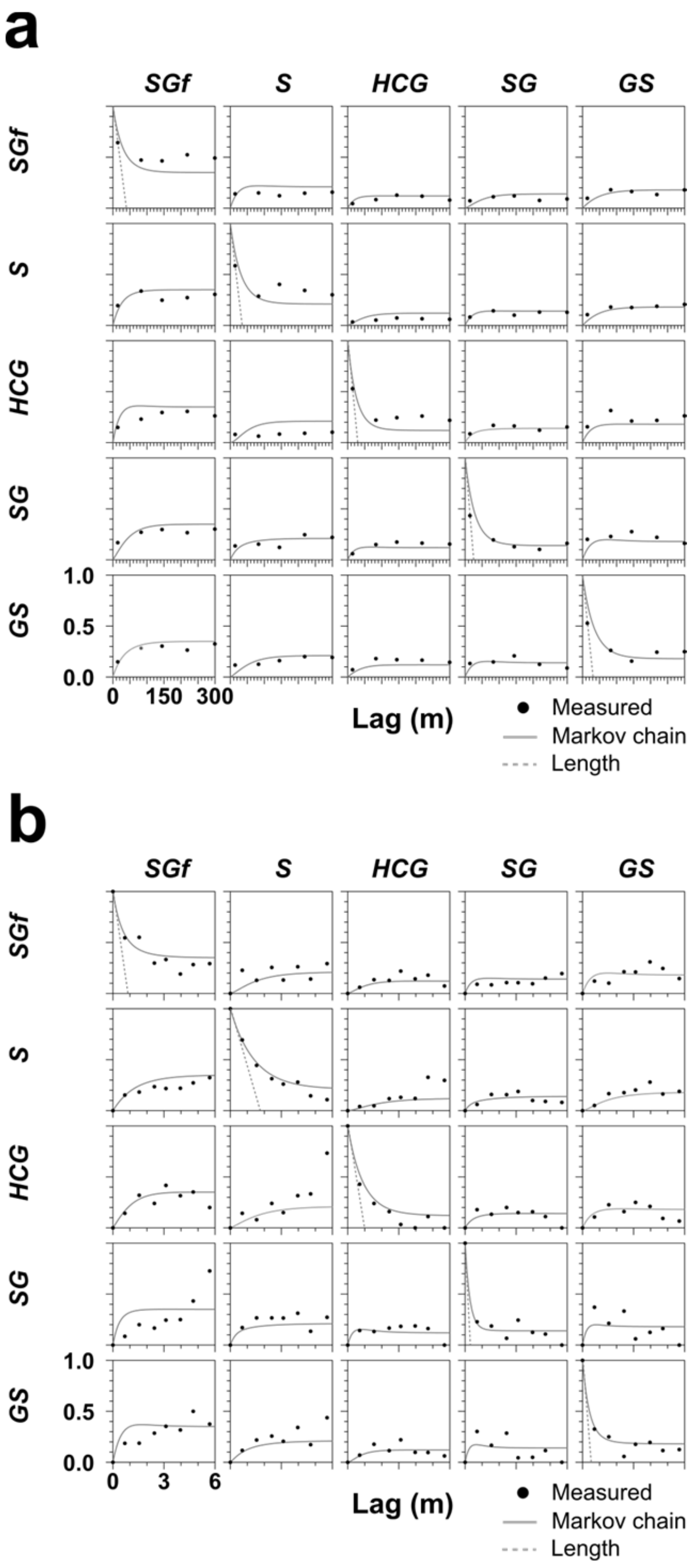

Figure 2. Lateral (a) and vertical (b) transition probabilities and fitted Markov chain model.

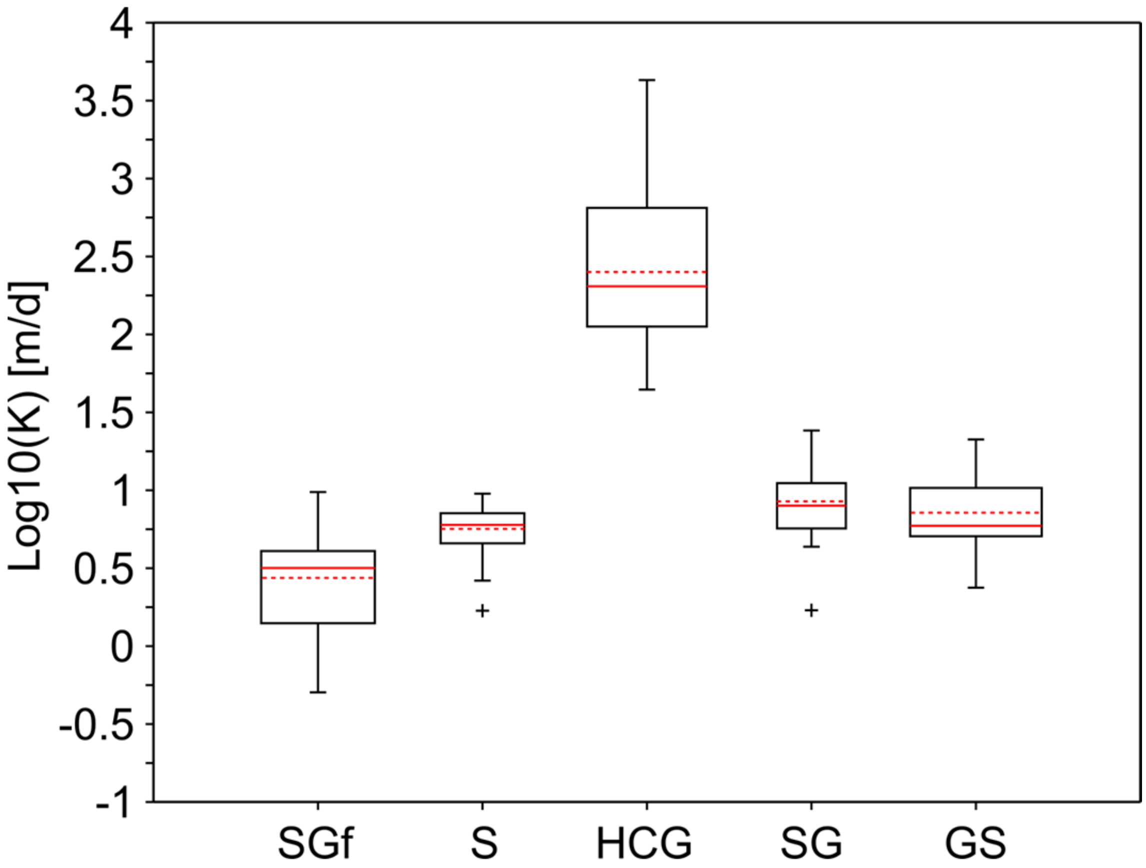

Figure 3. Box plots of the estimated log-transformed hydraulic conductivity (K) values for each lithofacies showing median, interquartile range and extreme values (crosses). Red dashed lines indicate mean values.

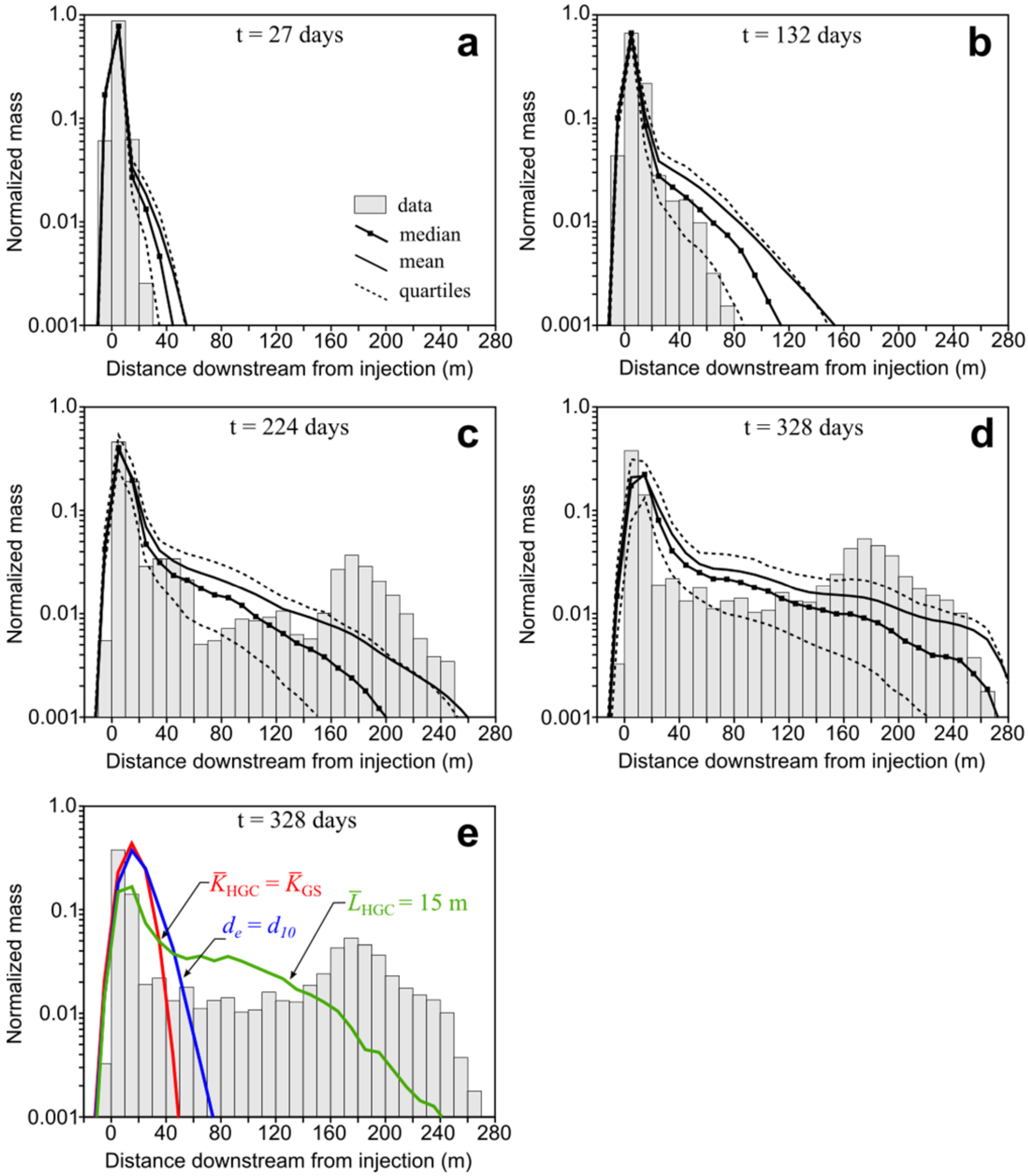

Figure 4. (a-d) Observed and simulated longitudinal mass distributions of the tritium plume. Simulated distributions were obtained with input parameters in Table 1. (e) Mass distribution at 328 days for simulations considering different mean *K* and mean length for lithofacies HCG. The scenario assuming a mean *K* value for HGC equal to that for

lithofacies GS is shown red. The scenario assuming d_e as d_{10} for K estimations is shown in blue. The scenario assuming a mean length ($\bar{L}$) of 15 m is shown in green.

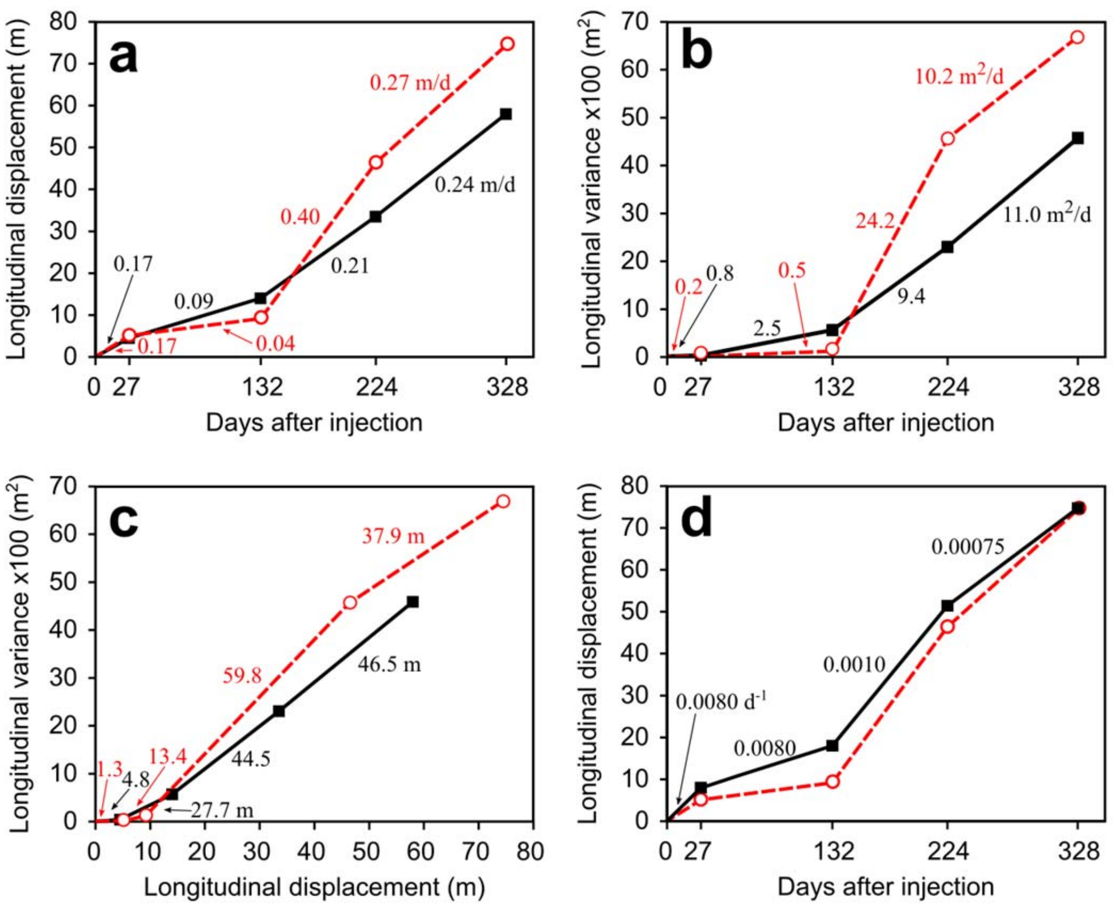

Figure 5. First and second central spatial moments evolution for the observed (in red) and simulated (in black) plumes. Simulated points in a-c represent mean values of the Monte Carlo realizations. (a) Values indicate the estimated mean plume velocity. (b) Values indicate one half of the growth rate of the longitudinal variance with time (c) Values indicate one half of growth rate of the longitudinal variance with the mean travel distance. Under the assumption of a uniform flow field these values correspond to the macroscopic longitudinal dispersivity. (d) Longitudinal displacement of a dual-domain single rate mass transfer model (DDM) in which the ratio of mobile to total porosity is equal to the volumetric fraction of HCG. Values indicate calibrated values for the mass transfer rate coefficient (see text for explanation).

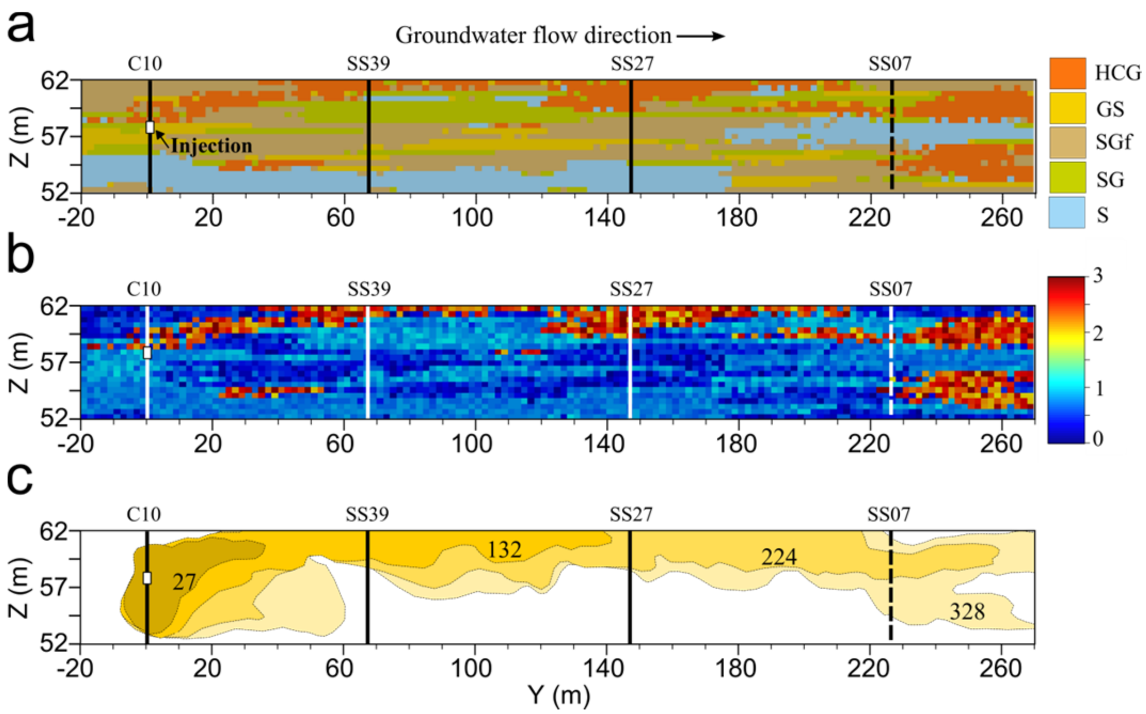

Figure 6. (a) One equally probable realization of the simulated spatial distribution of lithofacies shown in a cross section oriented parallel to the main flow direction and crossing through the injection site and three boreholes. Location of borehole SS07 is projected. (b) Corresponding $\log_{10}(K)$ field [m/d]. (c) Evolution of the simulated plume front (C = 2pCi/ml) with time.

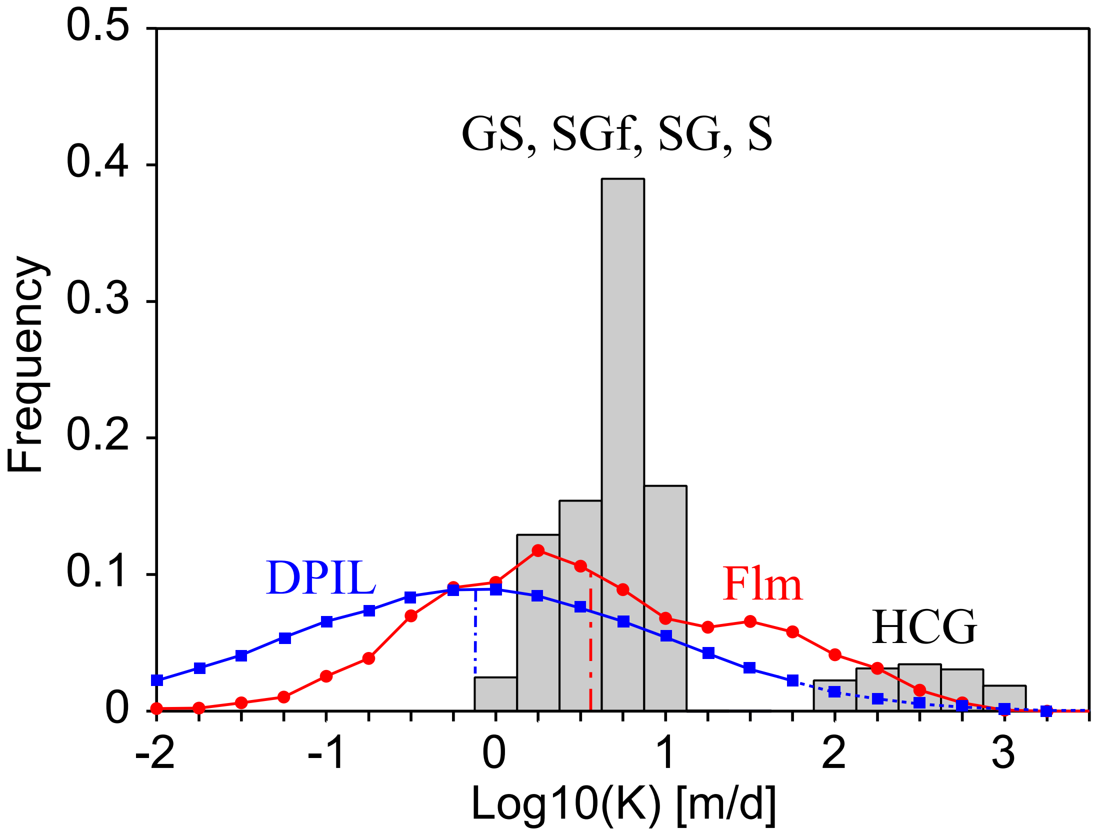

Figure 7. Example of frequency distribution of the generated K fields. The distributions of the K measurements using the impeller flowmeter (Flm) and direct-push injection logger (DPIL) are also shown. The vertical dash-dot lines indicate the mean of the two distributions. Flowmeter measurements data from *Rehfeldt et al.* [1992]. The DPIL data distribution was estimated by assuming a lognormal distribution with a geometric mean of 8.9×10^{-6} m/s and a variance of natural log-transformed K values of 6.6 [Table 1in *Bowling et al.*, 2012]. The upper limit of the DPIL instrument is about 60 m/d [*Bowling et al.*, 2012; *Dogan et al.*, 2014].

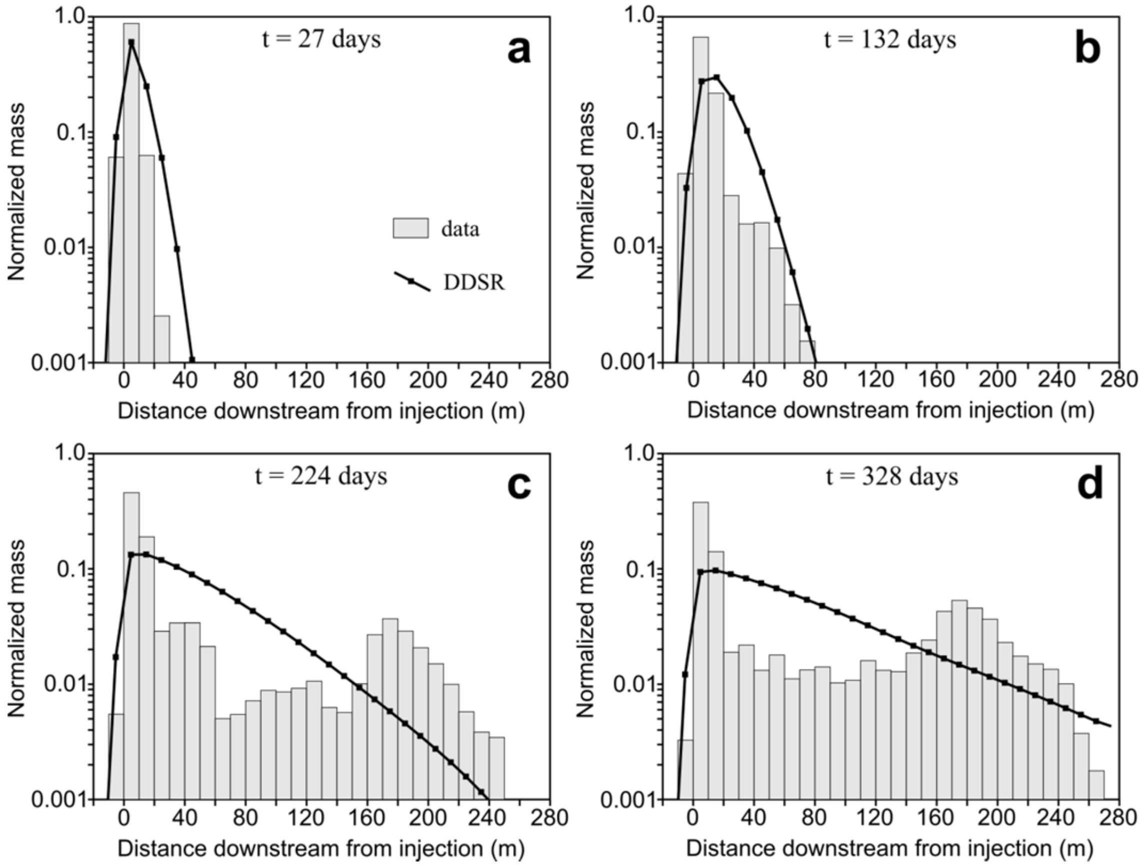

Figure 8. Observed and simulated longitudinal mass distributions of the tritium plume. Simulated profiles were calculated with a dual-domain single rate mass transfer model in which the ratio of mobile to total porosity is equal to the volumetric fraction of HCG. Values for the mass transfer rate coefficient (see Figure 5d) were estimated by calibration with a trial-and-error approach.