Transport upscaling in heterogeneous aquifers: What physical parameters control memory functions?

M. Willmann,[1,2] J. Carrera,[3] and X. Sánchez-Vila[1]

Received 17 September 2007; revised 2 October 2008; accepted 23 October 2008; published 25 December 2008.

[1] Power law tailing is often observed in the breakthrough curves (BTCs) of tracer tests. Tailing is attributed to heterogeneity of aquifer properties and cannot be properly modeled by means of the homogeneous advection-dispersion equation. Mass transfer models (e.g., continuous time random walk method, multirate mass transfer, or fractional-order advection-dispersion equations) using memory have been widely applied for reproducing observed tails. The relationship between memory parameters obtained from BTC fitting and the parameters characterizing the heterogeneity of hydraulic properties is still unclear. Here we investigate the conditions under which heterogeneity produces the type of tailing observed in the field and how memory functions are influenced by measurable heterogeneity parameters (e.g., variance, variogram, or integral scale of the underlying transmissivity field). We find that the slope of a BTC in a log-log plot is mainly influenced by the connectivity of the underlying permeability field but is insensitive to its variance. The slope BTC reaches asymptotically 2 as connectivity increases. We conclude that an appropriate choice of the memory function allows reproducing the spreading caused by hydraulic heterogeneity but not necessarily the rate of mixing.

Citation: 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.

1. Introduction

[2] Contaminant transport has been traditionally modeled using the advection dispersion equation (ADE). Yet, transport in natural aquifers usually displays anomalous (i.e., inconsistent with the ADE) behavior. Observed deviations are numerous [*Carrera*, 1993]. They include scale dependence of dispersivity [e.g., *Lallemand-Barres and Peaudecerf*, 1978; *Neuman*, 1990], directional and time dependence of apparent porosity [*Sánchez-Vila and Carrera*, 1997; *Guimerà and Carrera*, 2000] and tailing of breakthrough curves [e.g., *Valocchi*, 1985; *Freyberg*, 1986; *Cortis and Berkowitz*, 2004]. Stochastic hydrology has succeeded in qualitatively explaining these deviations and in quantifying the scale dependence of dispersivity [e.g., *Dagan*, 1989; *Gelhar*, 1993]. In the process, tools have been developed to predict the evolution of dispersivity with scale given a stochastic description of variability of hydraulic conductivity [*Kitanidis*, 1988; *Dentz et al.*, 2000]. However, much less efforts have been devoted to understanding the causes of tailing [*Haggerty et al.*, 2000; *Shapiro*, 2001; *Cortis and Berkowitz*, 2004; *Dentz et al.*, 2004].

[3] Tailing is defined as the markedly asymmetric shape of breakthrough curves (BTCs), which cannot be reproduced by the homogeneous medium ADE. Field BTCs typically display a sharp rising limb for the early arrival but a slowly decaying limb at late time (Figure 1). More important, the decay limb often displays a power law behavior. That is, late-time concentrations decay as $t^{-m_{BTC}}$, so that they plot as a straight line on a log-log scale [*Farrell and Reinhard*, 1994; *Hadermann and Heer*, 1996; *Werth et al.*, 1997; *Becker and Shapiro*, 2000; *Shapiro*, 2001; *Meigs and Beauheim*, 2001].

[4] A proper description of this "late time" behavior is important not only for practical reasons (i.e., time for cleanup below a threshold or reactive transport modeling), but also because the apparent ubiquity of power law decay suggests it must reflect something of a fundamental nature.

[5] Different mechanisms are known to cause tailing. These include heterogeneity of permeability, diffusion or chemical heterogeneity. Tailing was initially attributed to matrix diffusion and/or sorption kinetics [*Neretnieks et al.*, 1982; *Neretnieks and Rasmuson*, 1984]. In fact, the 1.5 slope usually observed in fractured media was soon attributed to diffusion into rock matrix [*Hadermann and Heer*, 1996]. Actually, different slopes can be obtained by acknowledging that low permeability blocks exhibit a distribution of sizes and diffusion coefficients. As a result, the memory function of diffusion dominated mass transfer is predictable [*Rasmuson and Neretnieks*, 1986; *Zhang et al.*, 2007; *Gouze et al.*, 2008b]. However, here we concentrate on the role of hydraulic heterogeneity. Heterogeneity of permeability leads to conductive paths that carry most of the water and arrive early (hence the sharp rising limb of BTCs) and less conductive paths that trail behind and cause tailing. This is referred to as "slow advection."

[1]Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia, Barcelona, Spain.
[2]Institute of Environmental Engineering, ETH Zurich, Zurich, Switzerland.
[3]Institute of Environmental Analysis and Water Studies, CSIC, Barcelona, Spain.

0043-1397/08/2007WR006531

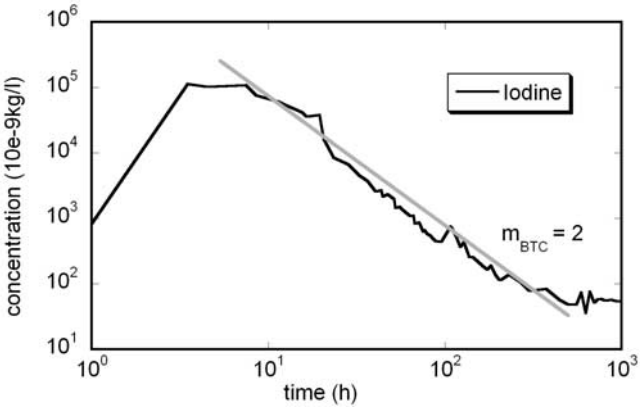

Figure 1. Breakthrough curve of iodine measured at the El Cabril site [*Gaillard et al.*, 1990]. A power law tail is observed with a slope of 2.0 (indicated as a gray line). The power law behavior spans over 1.5 orders of magnitude until the background concentration is reached. This curve cannot be reproduced by using a homogeneous ADE. Mass transfer models using a memory function can reproduce this behavior. However, the causes of the tailing need to be properly understood in order to use such memory function for predictions under different flow conditions or transport distances.

[6] *Becker and Shapiro* [2000, 2003] performed field tracer tests such that they could neglect other causes for tailing. They found power law slopes $m_{BTC} = 2$ caused by slow advection. *Di Donato et al.* [2003] modeled slow advection in a heterogeneous system using the streamline method. They found a slope for 2.2. *Zhang et al.* [2007] modeled a heterogeneous 3D sedimentary aquifers where the dominant process in the low-conductive zones was diffusion. They found a power law tail of about 2. Most interestingly, *Gouze et al.* [2008a] and *Shapiro et al.* [2008] found for field tracer tests slopes between 2 and 2.2 at intermediate times and 1.5 at very late time.

[7] A number methods have been proposed to describe tailing. Multirate mass transfer (MRMT) [*Haggerty and Gorelick*, 1995; *Wang et al.*, 2005] and memory functions [*Carrera et al.*, 1998], which are equivalent [*Haggerty et al.*, 2000], can be viewed as dividing the medium into overlapping mobile and immobile continua (hence the generic term multicontinuum models). Each immobile zone exchanges solute mass with the mobile zone by linear mass transfer (i.e., mass exchange is proportional to concentration gradient). These models may be expressed as a function of only the mobile concentration by introducing a memory term into the ADE to account for mass transfer between mobile and immobile zones. The name ''memory'' reflects the fact that this term represents how current mobile concentrations are affected by past events. Multicontinuum models have been successfully applied to interpret anomalous transport [*McKenna et al.*, 2001; *Haggerty et al.*, 2004; *Zinn et al.*, 2004; *Medina and Carrera*, 1996; *Sánchez-Vila and Carrera*, 2004; *Zhang et al.*, 2007]. While these models were originally developed to represent diffusion into immobile regions, they can also be used to reproduce the effect of slow advection.

[8] The most widely used method for representing tailing is the continuous time random walk method (CTRW) [*Berkowitz and Scher*, 1998; *Dentz et al.*, 2004]. CTRW can be viewed as a generalization of Random Walk methods in that not only spatial displacements, but also time step lengths are random variables [*Berkowitz et al.*, 2006, section 7.1]. The effective transport equation results from ensemble averaging the transport of individual particles. As it turns out, CTRW is broader in scope than multicontinuum representations, which can be viewed as a particular case of CTRW [*Dentz and Berkowitz*, 2003]. However, the most commonly adopted form of CTRW is equivalent to the memory function approach and the corresponding memory functions can be derived from each other. CTRW has been successful not only in reproducing field BTCs, but also in reproducing appropriate scaling behavior. That is, CTRW models calibrated against BTCs measured at one scale have been successful in predicting BTCs at different scales [*Berkowitz and Scher*, 1998; *Kosakowski et al.*, 2001; *Levy et al.*, 2003; *Cortis and Berkowitz*, 2004; *Le Borgne and Gouze*, 2008]. CTRW models have also accurately reproduced the outcome of pore network models, reproducing the dependence of dispersion on molecular diffusion [*Bijeljic and Blunt*, 2006].

[9] A further method to describe tailing are the fractional-order advection-dispersion equations (fADE) [*Benson et al.*, 2000]. As with CTRW, the model is rather general, but can also be characterized by a power law memory function when only the time derivative term in the ADE is fractional. *Berkowitz et al.* [2006] showed that, in such case, fADE can be considered as a limiting case of CTRW.

[10] In summary, multicontinuum models, CTRW and fADE, all sharing a time nonlocality, must be considered as excellent representations of transport in natural media. Moreover, they allow discriminating between mixing and spreading. Mixing controls many chemical reactions [e.g., *De Simoni et al.*, 2005, 2007; *Cirpka and Valocchi*, 2007]. Therefore, a proper representation of mixing is a prerequisite for proper reactive transport modeling. Yet, the ADE (and all formulations that base dispersion solely on the spreading of plumes) equate dispersion and mixing (thus overestimating mixing and mixing-driven reactions). Nonlocal in time formulations separate these two processes. Therefore, they shed some hope on the possibility of predicting reactive transport accurately.

[11] Despite the above nice properties, non local formulations have been subject to criticism [*Neuman and Tartakovsky*, 2009]. All these formulations require specifying some (arbitrary) memory function. At present, memory functions are calibrated against tracer test data, without explicit reference to heterogeneity. This is unsatisfactory both from conceptual and practical viewpoints. Conceptually, it is generally agreed that tailing can be caused by heterogeneity. In fact, *Berkowitz and Scher* [1997, 1998] and *Berkowitz et al.* [2008] suggest deriving the memory function from the velocity pdf. However, *Le Borgne et al.* [2008a, 2008b] point out that velocity correlation along a stream tube (Lagrangian correlation) may be a key factor. Since there is no way to obtain the full statistical characterization of the velocity field, other than numerical simulation, it is clear that no explicit link is available between memory function parameters and measurable properties of the heterogeneity

in hydraulic conductivity. This casts a shadow of doubt on the predictive capabilities of nonlocal formulations for scales dramatically different than those for which they were calibrated. There are also practical implications. Tracer tests cannot be performed on the long scales often needed for transport predictions. Certainly, using nonlocal formulations for reactive transport requires ascertaining the conditions under which they are valid. On the other hand, stochastic approaches offer no clear alternative. In fact they are criticized because, with rare exceptions [e.g., *Di Donato et al.*, 2003; *Alcolea et al.*, 2008; *Luo et al.*, 2008], they usually fail to reproduce the kind of tailing observed in the field.

[12] Accepting that spatial variability of hydraulic conductivity is a frequent cause of tailing, the objective of this paper is twofold. First, we explore some conditions under which heterogeneity can explain the kind of tailing observed in tracer tests, namely power law decay. Second, we investigate the relationship between certain field properties describing heterogeneity and the parameters describing the memory function.

2. Background

[13] We conceptualize solute transport by assuming a superposition of a (homogeneous) mobile zone and an infinite number of immobile zones. Mass is exchanged between the mobile and the immobile ones by diffusion-like processes. This results in the following transport equation:

$$\phi_m \frac{\partial c}{\partial t} = \nabla \cdot (\mathbf{D}\nabla c) - \mathbf{q} \cdot \nabla c - \Gamma \tag{1}$$

where c is solute concentration in the mobile zone, $\mathbf{D}$ is the dispersion tensor, ϕ_m is the mobile porosity and Γ is the source/sink term controlling the mass transfer between the mobile and a continuum of immobile zones. In the case of a discrete description with a finite number of immobile zones Γ can be expressed in terms of immobile concentrations resulting in a system of $n + 1$ equations for n immobile zones. *Carrera et al.* [1998] showed that Γ can be expressed in terms of mobile concentration by using a convolution product noted by (*) with a memory function g:

$$\Gamma = \phi_{im} g * \frac{\partial c}{\partial t} \tag{2}$$

$$g(t) = \sum_{n=1}^{N} \alpha_n b_n e^{-\alpha_n t} \tag{3}$$

$$\sum_{n=1}^{N} b_n = 1 \tag{4}$$

where N is the number of immobile zones, α_n are first-order rate coefficients and b_n is the fraction of total immobile porosity characterized by α_n (inverse of the characteristic time of zone n). Note, that the memory function (equation (3)) is slightly different to the one of previous authors. The constant characteristic diffusion of *Carrera et al.* [1998] is accounted for by α_n. *Haggerty et al.* [2000] define their memory function in terms of the full aquifer, while we prefer to define it in only in terms of immobile zone parameters. Thus our memory function does not depend on mobile porosity. By excluding the immobile porosity our memory function does not depend on the size of the immobile zone, but only on its geometry. In Appendix A we show how equations (1)–(4) can be easily implemented into a standard numerical code. Throughout this paper, we will only refer to conservative transport and 1D transport equation. This implies that we could solve the transport equation in Laplace space and then invert the solution. This would avoid discretization of the memory function. However, our ultimate goal is to apply these results on to reactive transport, which is simpler in the discrete version of g (equation (3)).

[14] A memory function can be of irregular shape. But as power law behavior in BTCs is observed frequently, we use memory functions that display power law behavior. *Haggerty et al.* [2000] related late time concentrations to the memory function and found that a power law behavior in a BTC is caused by a power law behavior of the memory function. Similar findings are reported by *Berkowitz and Scher* [1997, 1998] in the context of CTRW. Therefore, we choose α_n and b_n in equation (3) so as to ensure this power law behavior (Figure 2). A truncated power law (TPL) seems to represent best anomalous transport because it allows an evolution to Fickian transport [*Dentz et al.*, 2004; *Berkowitz and Scher*, 2009].

[15] This reduces the number independent unknown parameters describing a memory function to three: two characteristic times, power law behavior initial (t_1) and final (t_2) cutoff times and the power law slope m_g (Figure 3).

3. Methodology

[16] We use a numerical approach based on synthetic aquifer analysis to study the heterogeneous features that control tailing. The methodology consists of 4 major steps: (1) Generation of heterogeneous transmissivity fields, (2) transport simulation on the heterogeneous fields using the ADE at the local scale, (3) analysis of vertically integrated BTCs at selected locations and derivation of a representative memory function, and (4) simulation of transport though a homogeneous medium with the above memory function. All simulations are performed within a domain of 1024 times 512 cells of unit size. Water flow is steady state, prescribed by constant heads at the left and the right boundaries and no-flow conditions on top and bottom, resulting in mean uniform flow with a mean gradient of 0.0098.

3.1. Generation of Transmissivity Fields

[17] We generate 2D transmissivity fields with different heterogeneity characteristics. First, we need varying heterogeneity scales, as we expect in nature, evolving over a range of scales [*Neuman*, 1990]. We also need poorly connected and well connected fields, that is where high transmissivity, T zones extend over long distances. We analyze transport in individual realizations of heterogeneous aquifers, as opposed to a Monte Carlo approach, because we are interested in what happens in a given aquifer rather than in an

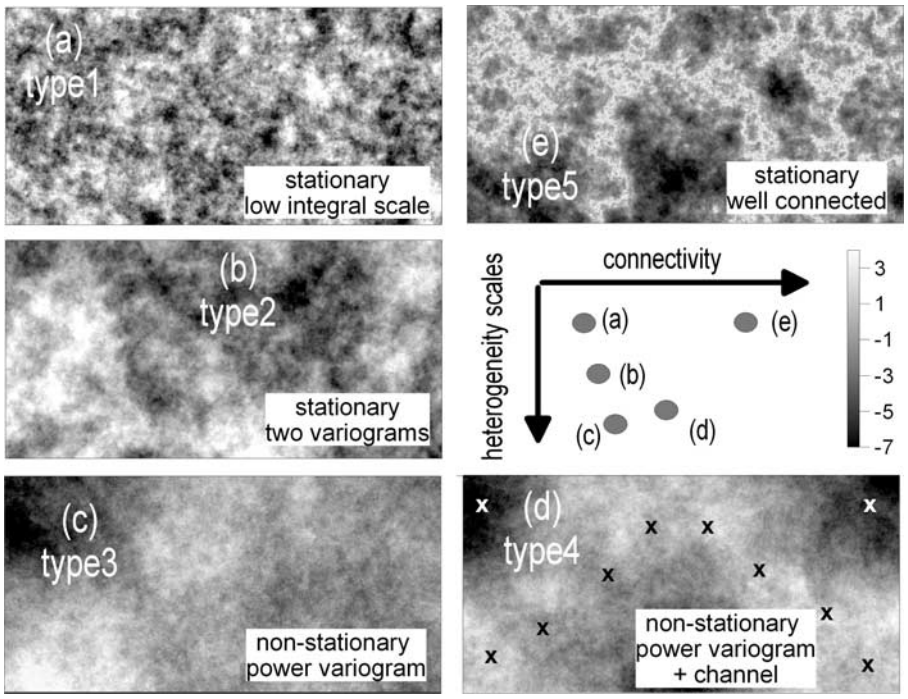

Figure 2. Five types of transmissivity fields used in this work. They show different ranges of scales and different degrees of connectivity: (a) multi-Gaussian field with a single exponential variogram and a small correlation length, (b) field comprising two nested variograms with different correlation lengths, (c) power variograms representing a continuous distribution of heterogeneity scales, (d) similar to type 3 but conditioned to leave a preferential flow path (white and black crosses indicate conditioning point with low and high T values, respectively), and (e) highly connected fields.

ensemble. The realizations are generated with the Gaussian Sequential Simulation Method [*Gómez-Hernández and Journel*, 1993]. The method uses a single or a combination of variograms, always exponential in this work, to generate simulations conditioned to field data. We used five types of heterogeneous lognormal transmissivity, lnT, fields (Figure 2) with a default variance of 6.

[18] Type 1 fields are unconditioned multi-Gaussian fields generated with a correlation length of 20, which is small compared to the domain size (1024). The variance of the lnT fields was set to 2 and 6. The resulting fields display neither multiple scales nor preferential flow paths. Therefore, no anomalous transport behavior is expected for long travel distances, compared to the domain size.

[19] Type 2 fields are obtained with a nested variogram consisting of two exponential variograms, one describing small-scale heterogeneity with a correlation length between 8 and 128, and one with a large correlation length between 128 and 1024. The sill of the two different variograms are varied between 1 and 5, but the total variance always adds up to 6. The resulting fields display two distinct scales of heterogeneity scales but no apparent preferential flow paths.

[20] Type 3 heterogeneous fields show an evolving range of scales. We use a power law variogram:

$$\gamma(s) = C_0 s^{2H} \tag{5}$$

where s is distance, C_0 is a constant and H is the Hurst coefficient. The power law variogram can be seen as an infinite series of nested variograms [*Neuman and Di Federico*, 2003]. We generated fields with H values of either 0.1, 0.25 or 0.4 and scale them for comparison with the other fields to a variance of 6. Such a variogram was postulated by *Neuman* [1990] to account for increasing variance and correlation length of aquifers as their size increases. Whether such a variogram would be still valid at very large scales can be neither proven nor discarded. The

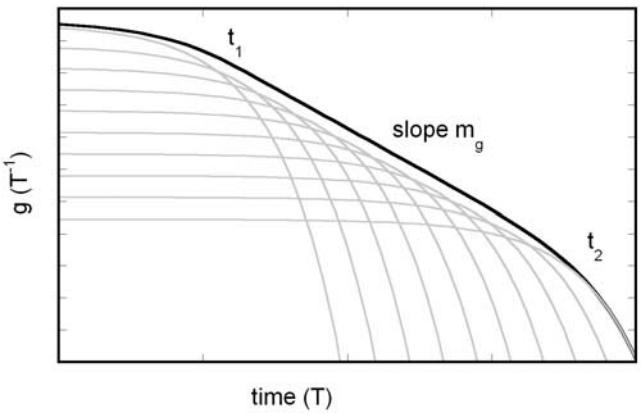

Figure 3. Representation of a memory function (black). It can be viewed as the superposition of 10 individual memory functions each corresponding to an immobile zone (gray) (equation (3)). The memory function exhibits three independent variables: characteristic time t_1 when power law decay starts, characteristic time t_2 when power law behavior ends, and the power law slope m_g.

fields display a continuous range of heterogeneity scales but the existence of interconnected zones of high T depends on the individual realization.

[21] Type 4 fields are a modification of type 3 fields, where the simulations are now conditioned to 14 points to ensure interconnected zones of high T (Figure 2). These fields still display a continuous range of heterogeneity scales, but also a well defined preferential flow path. Still, the results may be affected by the conditioning process.

[22] The type 5 fields are built with the methodology of *Zinn and Harvey* [2003] to ensure good connectivity. The original multi-Gaussian fields display connectivity only for intermediate T values. These are transformed to high T values to enhance hydraulic connectivity. To ensure consistency we use here an exponential variogram with correlation lengths between 20 and 200.

[23] We quantify here connectivity using the indexes of *Knudby and Carrera* [2005]:

$$CF = \frac{T_{eq}}{T_G}. \tag{6}$$

where T_{eq} is the equivalent transmissivity (i.e., the transmissivity of a homogeneous medium allowing the same flux as the actual heterogeneous medium) and T_G is the geometric mean of point transmissivity values. CF is considered a measure of connectivity because it reflects the increase in flow caused by highly transmissive channels transversing the domain. The transport connectivity index is defined as

$$CT = \frac{t_{peak_{hom}}}{t_{peak_{het}}} \tag{7}$$

where $t_{peak_{het}}$ is the peak arrival time for transport across the heterogeneous field and $t_{peak_{hom}}$ is the peak arrival time for the equivalent homogeneous field. CT is considered a connectivity measure because it reflects the early arrival caused by channels. *Cortis and Knudby* [2006] use a CTRW formulation for representing flow through heterogeneous media. We expect their memory function to contain information on connectivity, but it has not yet been quantified as an index.

[24] Figure 2 displays an schematic plot comparing the 5 types of lnT fields in terms of regional connectivity and scales of heterogeneity. This is a qualitative plot aimed at indicating the features that are explored in each of the realizations.

3.2. Fine-Scale Transport Simulations Within the Heterogeneous Fields

[25] Fine-scale conservative transport simulations are performed on the heterogeneous fields following the steady state flow conditions described above. Transport is simulated at the local scale with the ADE using the finite element code FAITH [*Sánchez-Vila et al.*, 1993]. Local dispersivities are assumed to be 10 units in longitudinal and 1 unit in transverse direction. To test the effect of local dispersivities, these values are varied for one specific setup between 0.1 and 10. The porosity is set to 0.3 in all cases. The time step starts with 100, increases up to 5000, and then remains constant until the end of the simulation period. The simulation is stopped when only 10^{-4}% of the initial mass remains in the domain. Breakthrough curves are measured as integrated values at selected sections, perpendicular to the mean flow and at several distances from the source.

[26] The above mentioned assumption (validity of the ADE at the local scale) has been questioned by some authors [e.g., *Berkowitz et al.*, 2006]. To test this we add some additional runs with a local-scale equation based on the mass transfer scheme presented in Appendix A. We used the same input value as before but add an additional porosity of 0.05 be accessible to mass transfer. To keep the analogy of intragranular diffusion we use $m_g = 0.5$, the value of matrix diffusion. We use different characteristic diffusion times (t_1/t_2) to define three local memory functions: 0.017/17, 0.017/1.7, and 0.17/17 for memory functions 1, 2 and 3 respectively.

[27] Concentration is initially zero throughout the domain. Solute mass input into the system can be expressed either as resident or as flux averaged. Most of the runs were performed by imposing a flux dependent mass of 1 at the left boundary. That is, mass inflow at every node is proportional to the water inflow. The boundary condition was applied during an initial time interval of 1000 which is small compared to the characteristic transport times. After this time interval, the boundary condition is changed to let clean water enter the domain. To test the effect of using initial resident concentration, one test case was performed by evenly distributing the unit mass over the first column.

[28] Integrated breakthrough curves are sampled along transects at several distances. To avoid boundary effects these measurements are taken always at least 24 cells away from the aquifer boundaries. The distinction between resident and flux concentrations [*Kreft and Zuber*, 1978; *Zhang et al.*, 2006] is relevant for our work. While resident concentrations represent the mass within a certain domain (e.g., numerical grid cell) at a certain time, flux concentrations provide the mass passing a cross section during a time interval. Resident concentrations are relevant for chemical reactions (i.e., reactive transport). Flux concentrations are the ones to be used when all mass flowing across a boundary has to be taken into account, which is the case in tracer tests. The relationship between both concentrations is based on the definition of flux concentration at control plains as the vertically averaged solute flux component along the mean flow direction normalized by the vertically averaged fluid flux in mean flow direction:

$$c_f = \frac{\int_0^{Ly} j_x dy}{\int_0^{Ly} q_x dy} \tag{8}$$

$$j_x = q_x c_r - D_{xx}\frac{\partial c_r}{\partial x} - D_{xy}\frac{\partial c_r}{\partial y} \tag{9}$$

where c_f is the flux concentration, c_r is the resident concentration, j_x is the mass flux in mean flow direction, q_x is the fluid flux in mean flow direction, D_{xx} and D_{xy} are the corresponding components of the dispersion tensor, and Ly is the length of the cross section.

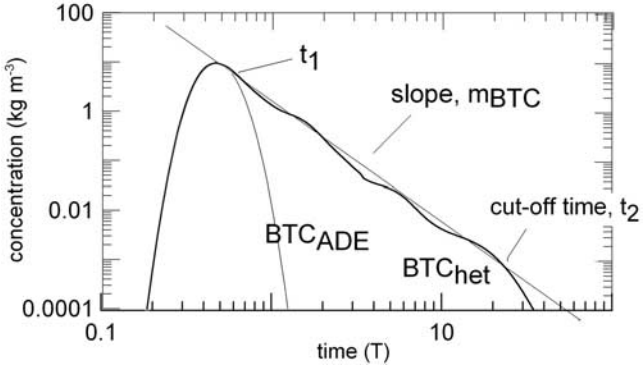

Figure 4. Derivation of the memory function of breakthrough curves taken in a heterogeneous aquifer BTC_{het}. Early time data are fitted against the corresponding BTC derived by means of the ADE (BTC_{ADE}); t_1 is read directly from the point where the two curves deviate. Starting from t_1 a straight line is fitted to the tail to get m_{BTC}. Finally, the cutoff t_2 is taken as the point where the BTC starts to decrease exponentially.

3.3. Obtaining Memory Functions From Breakthrough Curves

[29] Figure 3 shows the memory function defined in equation (3) for 10 immobile zones ($N = 10$). *Haggerty et al.* [2000] related the late time slope of BTCs (m_{BTC}) with the slope of memory functions depending on the initial conditions, provided that the longest characteristic times of the memory function are much longer than the advection time. A BTC with initial concentration specified throughout the mobile and immobile zones is linear on the memory function, g. A BTC resulting from an initial mass pulse into the mobile zone is linear on the time derivative of the memory function, $\partial g/\partial t$. For a tracer test BTC with the initial pulse and detection mode c_f, the relation between the slopes m_g and m_{BTC} is $m_g = m_{BTC} - 1$. For example, m_{BTC} for a matrix diffusion tracer test is 1.5 and the slope of the corresponding memory function, m_g, is 0.5.

[30] We take advantage of the above properties to identify the memory function from the BTC obtained from the heterogeneous fields simulations. It should be noticed that the objective here is just to fit the BTC, which could be done by any conventional fitting procedure. In fact, we use TRANSIN [*Medina and Carrera*, 1996], which is an inverse problem code, for fitting runs, which should facilitate estimation. However, direct inversion of the memory function shown in equation (3) is extremely ill posed (a potentially large number of α_n and b_n coefficients just to fit the BTC tail). Therefore, g must be parameterized. As mentioned in section 2, we have used a standard parameterization in terms of m_g, t_1 (time for which the slope m_g starts) and t_2 (time for which power law behavior ends), as shown in Figure 4. Time t_2 represents the time for which the immobile zone equilibrates with the mobile zone and corresponds roughly with the time for which the BTC departs from power law behavior and drops to zero exponentially, hence the name cutoff time. Time t_1 is the characteristic time of the smallest heterogeneity scale encountered [*Bijeljic and Blunt*, 2006; *Berkowitz et al.*, 2008]. For the case of slow advection, both characteristic times scale with flow rate [*Berkowitz et al.*, 2008] while the slope remains constant. Additionally, if t_1 is smaller than t_{peak}, t_1 can be chosen larger and the contribution of the missed part can be modeled using an upscaled dispersivity. TRANSIN does not handle this parameterization. Moreover, we preferred a generic methodology which could be easily applied with any code. Therefore, we use the following steps: First, we fit an ADE to the first arrival and the peak of the BTC by applying the same boundary conditions and using the T_{eq} as upscaled T (T_{up}). The fitting parameters are initial mass, porosity and dispersivity. Second, overlaying the resulting BTC (BTC_{ADE}) with the measured one, we define t_1 as the time when the two curves start departing from each other (again, this is somewhat arbitrary, shorter times would work as well). Second, a straight line is fitted starting at the point where t_1 crosses the original BTC, which renders the slope m_{BTC}. And third, t_2 is read from the BTC. Sometimes, the BTC does not display a well defined cutoff time t_2. In such cases, t_2 is defined by extrapolating the power law behavior until the full mass is recovered. Finally, the immobile porosity controls the size (height) of the tail. Therefore, some trial runs may be needed to obtain ϕ_{im} (this we did automatically with TRANSIN). Ideally, ϕ_{im} should be equal to the total minus the mobile porosity [*Sánchez-Vila and Carrera*, 2004]. But t_2 may have to be adjusted for this condition to be met (increasing t_2 lowers the tail). Once t_1, t_2 and m_g are known, coefficients α_n and b_n are adjusted as explained below.

3.4. Modeling of BTCs Using the Memory Term

[31] As a last step we take the memory function derived above and implement it in a homogeneous 1D mass transfer model. Again, the finite element code TRANSIN [*Medina and Carrera*, 1996] was used for solving the ADE with general mass transfer (Appendix A). Dispersivity values and mobile porosity are estimated from early time data. When possible, dispersivity is set to its local value. The total porosity is kept at its true value of 0.3, except for the cases in which BTCs were obtained with a local mass transfer scheme, when total porosity is 0.35. The immobile porosity is calculated from the mobile one. $N = 20$ was found sufficient in most cases. The α_n values are bounded by (the inverse of) cutoff times, within which the power law behavior takes place ($\alpha_1 = t_2^{-1}$ and $\alpha_N = t_1^{-1}$). If we chose the rate coefficients, α_n, evenly distributed on a logarithmic scale, the corresponding b_N can be easily calculated using the slope of the memory function and equation (1) together with

$$b_n^* = \alpha_n^{m-1} \tag{10}$$

$$b_{tot}^* = \sum b_n^* \tag{11}$$

$$b_n = \frac{b_n^*}{b_{tot}^*} \tag{12}$$

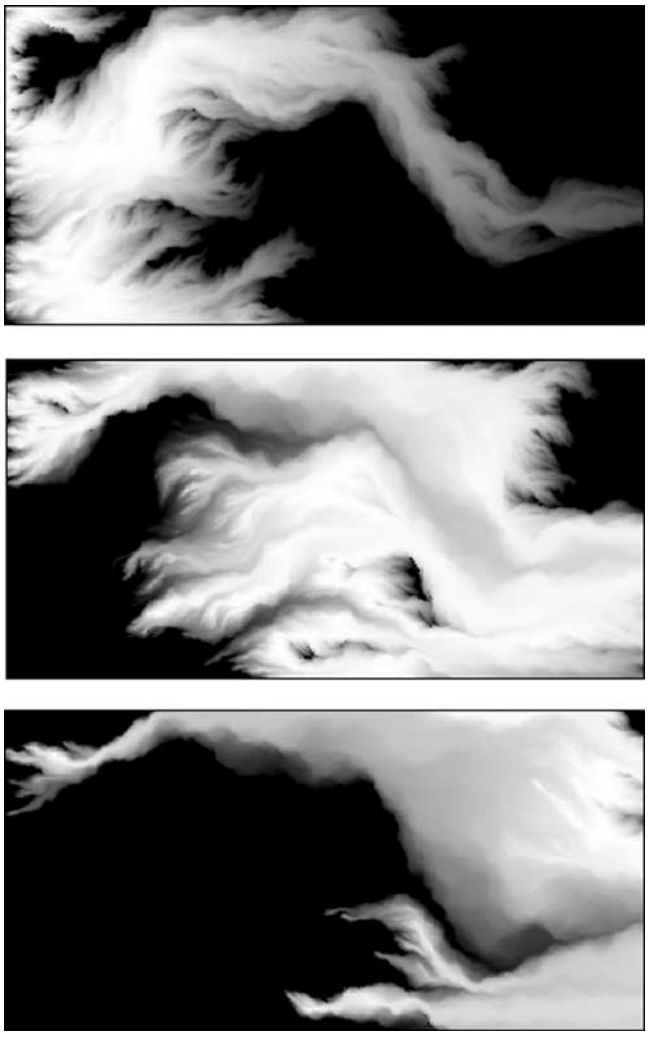

Figure 5. Snapshots of concentrations for a type 3 field at three different time steps. The shape of the plume clearly indicates non-Fickian behavior (marked asymmetry) that cannot be described by an ADE-like equation with upscaled parameters, whether constant or time-dependent. Actually, two different preferential flow paths can be observed, where the upper one is more conductive than the lower one.

where b_n^* and b_{tot}^* are introduced to normalize b_n. Note, that $b_n\phi_{im_n} = \phi_{im_n}$ where ϕ_{im_n} is the corresponding immobile porosity fraction characterized by α_n.

4. Results

[32] Plume snapshots (Figure 5) show that transport is non-Fickian at this scale. Spatial distributions of solute mass are highly asymmetric, with peak values trailing behind early arrivals. It is clear that properly reproducing the geometry of these plumes would require a thorough knowledge of the medium, which is unrealistic. Instead, the main features of this and other simulations are captured with a homogeneous transport equation with mass transfer term, which motivates the type of upscaling proposed here.

4.1. Effects of Mass Input and Detection Mode: Scaling

[33] Figure 6 shows eight BTCs investigating solute mass input conditions and detection (sampling) mode. They are all obtained with the same heterogeneous field (type 3 field in Figure 3). This particular field displays a relatively homogeneous right boundary, but a highly heterogeneous left one. The homogeneous boundary is characterized by an almost uniform distribution of velocities while the heterogeneous boundary displays a high-velocity variance along the cross section (low values on top and high values on bottom). Four tests are performed with flow from left to right and four from right to left with all possible combinations of detection mode (flux or resident concentration) and input conditions (flux-averaged input pulse or fixed initial concentration). In our numerical setup, the input conditions are affected by the upstream boundary, while the detection mode is influenced by the downstream boundary (where measurements are taken). BTCs separate all tests into two distinct groups. The individual behavior depends only on the condition applied at the heterogeneous boundary. When a heterogeneous boundary is used as inflow, the curves corresponding to constant initial concentration display a smaller slope than those of an input pulse proportional the local velocity at each point. When the heterogeneous boundary corresponds to the outflow, the same effect is displayed for resident concentrations detection mode with respect to flux concentration. The difference between the slopes in the log-log plots from the two groups of curves is 0.9, which agrees quite well with a difference of 1.0 derived analytically by *Haggerty et al.* [2000] for different initial conditions. This result agrees also with *Di Donato et al.* [2003] where they found a difference of 0.83. It is also

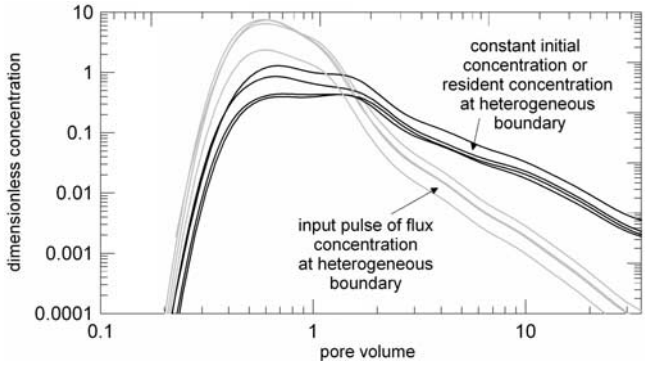

Figure 6. Comparison of breakthrough curves for two different initial conditions (uniform initial concentration or flux-averaged mass inflow) and two detection modes (resident or flux-averaged concentration). All eight tests are performed in the type 3 field example of Figures 3 and 5, where the left boundary is highly heterogeneous and the right boundary is relatively homogeneous. In four tests, the solute moves from left to right, and in the remaining four, it moves from right to left with all the possible combinations of detection mode (flux or resident concentration) and initial conditions (uniform or flux averaged). Test results do not depend on the condition applied at the homogeneous boundary, which consistently display a slope around 2.2. The slope decreases by about 0.9 for the four cases where either resident concentrations are measured or fixed concentrations are applied at the heterogeneous boundary. Notice that the effect of the two preferential flow paths of Figure 5 is reflected as two, more or less smoothed, humps in the BTC.

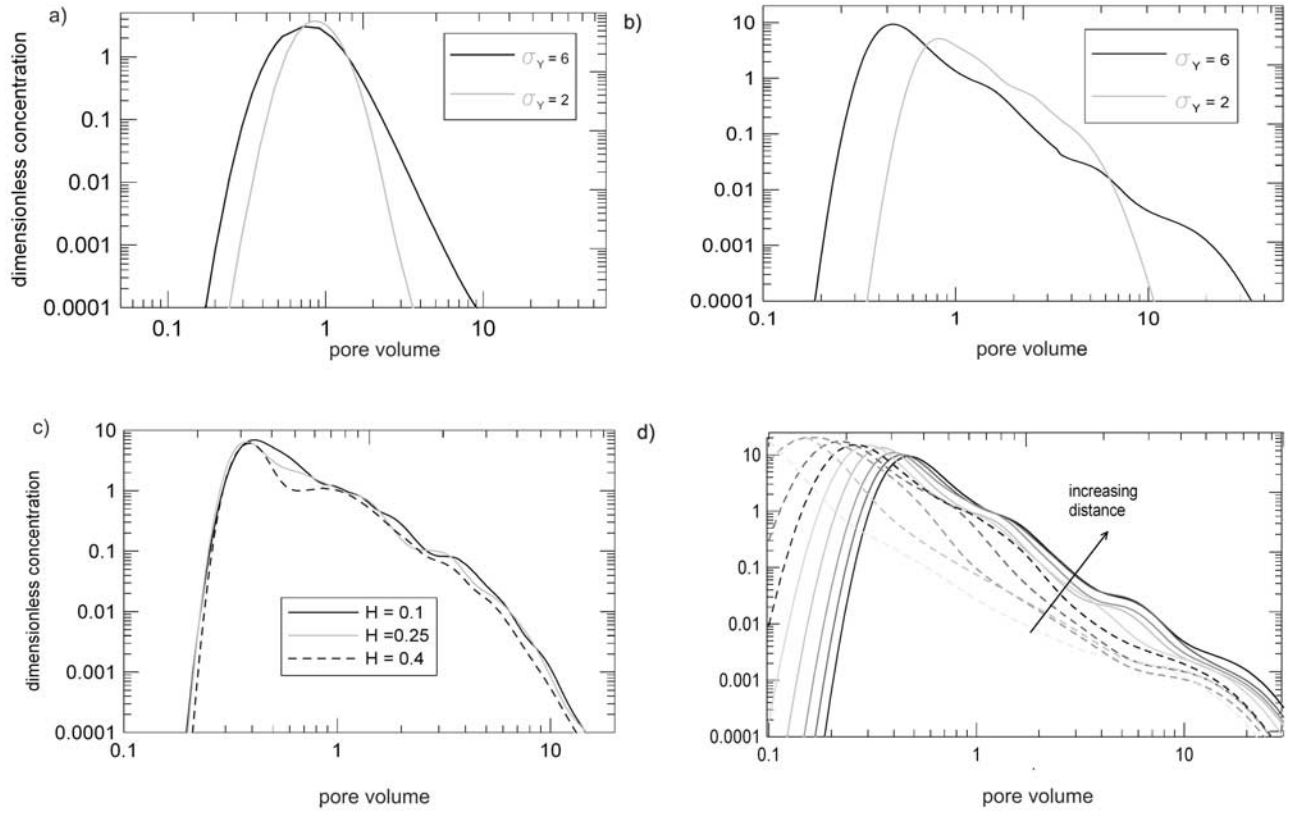

Figure 7. Breakthrough curves for different types of heterogeneity. (a) Multi-Gaussian fields (type 1). The curves do not show power law tail, and the curves can be reproduced using an ADE with upscaled parameters. (b) A type 3 field with varying variance. When variance is decreased, we observe no change of slope m_{BTC} but a significant delay in first arrival and reduction of late time cutoff t_2. (c) A type 3 field with different values for the Hurst coefficients. It indicates that small-scale heterogeneity has little effect on any of the parameters of g. (d) Type 3 field with BTC measured at uniformly increasing distances between 100 and 1000. No change in m_{BTC} is observed with increasing sampling volume.

worth pointing out that the peak would have been poorly reproduced with a single ADE for the mobile zone. All BTCs display two, more or less explicit, humps (arguably corresponding to the two main flow paths that can be discerned in Figure 5). Clearly, an accurate fit would require the superposition of at least two homogeneous solutions [*Luo et al.*, 2008]. This implies that indeed, the BTC spread is reflecting the variability in travel times of different stream tubes (only to some extent, lateral mass transfer between adjacent stream tubes smooths out part of the spread). It also implies that, as opposed to memory functions representing actual diffusion into mobile zones, the portion of memory functions representing reversible slow advection [*Gouze et al.*, 2008a], must be scaled by advection.

4.2. Effects of Different Heterogeneous Fields

[34] We now compare BTCs corresponding to the simulations performed in the different field types. BTCs of type 1 fields (multi-Gaussian) do not show anomalous features (Figure 7a). Some tailing is observed, for a *lnT* variance of 6 there, but no power law slope is developed. These curves could be well reproduced by the ADE with an upscaled dispersivity. As our BTC samples about 50 correlation lengths, we can consider transport here to be ergodic, so that macrodispersive behavior is observed.

[35] The type 2 fields (nested variograms, BTCs not shown) allow us to investigate the influence of two clearly separate heterogeneity scales. The breakthrough curves display anomalous (i.e., non ADE-like) behavior, some with a clear tail. Still, the shape of the log-log scale tail is irregular. This might be due to the fact that heterogeneity only exists at two specific scales. In any case, no clear connected paths are observed in these fields (Figure 3).

[36] Transport in type 3 fields (power law variogram) is always anomalous. Whether strong tailing develops in a given realization depends on how high T values align to form a certain connected path, which in these realizations happens randomly. The slopes of BTCs are smoother than those obtained with type 2 fields and look similar to those often found in field tests, such as that presented in Figure 1. This supports the conjecture that aquifers are heterogeneous over a range of scales [*Neuman*, 1990] and that the power law tails result from this range of scales [*Berkowitz and Scher*, 1997, 1998].

[37] The influence of the overall variance of transmissivity can be seen in Figure 7b. The slope is not sensitive to variance of *lnT*, but both early arrival and cutoff times change dramatically. As the variance decreases, early arrival is delayed and late time cutoff, t_2, is reduced. This is consistent with the view of the memory function reflecting

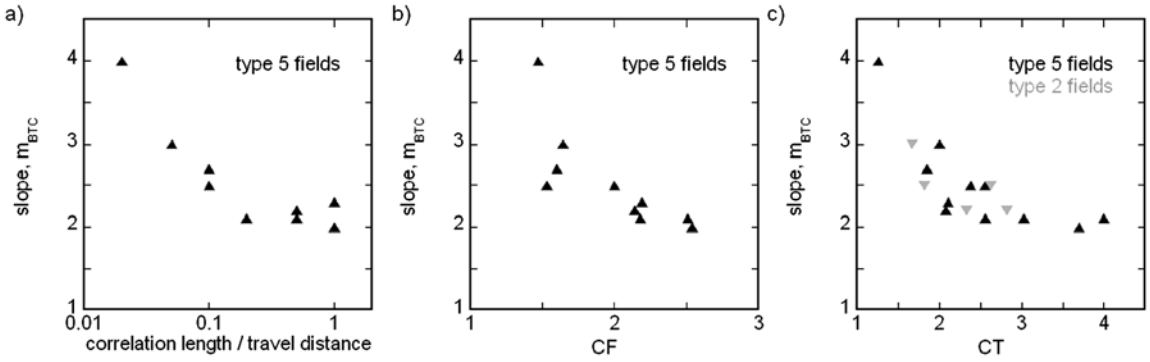

Figure 8. Slopes for different types of transmissivity fields versus (a) the correlation lengths of the underlying multi-Gaussian field in type 5 fields (recall that increasing this correlation length results in broader, better connected channels), (b) flow connectivity index, and (c) transport connectivity index. They all show that the slope m_{BTC} decreases (i.e., tailing becomes increasingly marked) as connectivity increases. The slope appears to tend $m_{BTC} = 2$ but does not decrease further.

the spread of travel times. In fact, in the limiting case of zero variance (homogeneous medium), t_2 becomes so small that a memory term is not needed. These results suggest that t_2 scales up with σ_y^2. That is, increasing variance with a factor f implies increasing t_2 by the same factor while reducing immobile porosity.

[38] To assess the effect of Hurst coefficient, independently of *lnT* variance, we changed *H* while total variance and geometrical patterns (i.e., location of high and low *T* regions) remained unchanged. This way, *H* reflects the relative importance of small-scale variability. With increasing Hurst coefficient (i.e., reduced small-scale variability) we get a slightly delayed first arrival but also a slightly delayed cutoff (Figure 7c). Again, the slope remains constant independently of *H*. This indicates that small-scale structures do not affect the large-scale slope, but only the characteristic times. In any case, the dependence is weak, which we find surprising.

[39] Transport in type 4 fields is always anomalous. Because of the conditioning process, all fields exhibit a preferential flow path that leads to a well defined power law tail in the BTC (not shown). Results are similar to those of type 3 curves. Figure 7d displays a series of BTCs taken at varying distances from the source within the same field. Tail slopes do not change with distance (or time). However, the late time cutoff, t_2, increases more or less linearly with travel distance. This confirms the earlier assertion about the scaling of t_2 with advection.

[40] Type 5 fields allow us to investigate the influence of connectivity. Recall that these fields are obtained by imposing connectivity on a multi-Gaussian field characterized by a single correlation length. When the correlation length of the original field is small, the modified field displays thin high-conductivity channels within a matrix of small low-conductivity blocks. Increasing the correlation length of the original field leads to an enlargement of both channels and low-conductivity zones, which causes an increase in connectivity. Figure 8a shows that the slope decreases with increasing correlation length of the original underlying multi-Gaussian field. The slope appears to tend asymptotically to a value of 2. The same can be said about the dependence of the slope on connectivity indicators (equations (6) and (7)). A noisy, but well defined, relationship exists between slope and flow connectivity (Figure 8b) or, even better, transport connectivity (Figure 8c). For values close to 1 (ADE limit), increasing connectivity indicators causes a decrease in slope, until it reaches values close to 2. From there, the slope remains constant despite of further increases of connectivity.

4.3. Effects of the Local-Scale Equation

[41] In all the cases discussed up to here we assumed that a local ADE exists. Subgrid heterogeneity is modeled by a local-scale dispersivity. We discuss now the effect of local-scale transport assumptions on large-scale BTCs. The effect of local transverse dispersivity (α_T) is displayed in Figure 9a. Increasing α_T, causes a small delay in first arrival and a small reduction in cutoff time, t_2, but has no impact upon the slope of the BTC. Longitudinal dispersivity (α_L, not shown) has little effect on either the cutoff time or the slope. Only the first arrival time is slightly delayed with decreasing dispersivity. The explanation of this effect is quite apparent form earlier discussions. Transverse dispersion tends to smooth away both from the leading (fast) and the trailing (slow) flow tubes by transferring mass to adjacent flow tubes. This is highlighted by the strong dependence of resident concentrations (not shown) on transverse dispersion. While this is consistent with long-standing understanding [*Taylor*, 1953], it sets a warning on the use of this formulation for reactive transport. Mixing is greatly enhanced by transverse dispersion, yet the memory function is only marginally affected.

[42] To investigate the influence of the ADE assumption on the local-scale transport equation, we study now the effect of using a local-scale equation with memory. Figure 9b displays the behavior of the BTCs for different memory functions. For characteristic times much smaller than the observation time we find that BTC is virtually identical to the one observed with the ADE, only delayed by the total porosity, which was increased from 0.30 to 0.35. If we set the characteristic times larger than the observation time, the slope decreases. A memory function, with a t_2 slightly larger than peak arrival time, displays a slight decrease in peak concentration and peak arrival time. This means that portion of the local immobile porosity associated to fast characteristic times can be described by the ADE, increas-

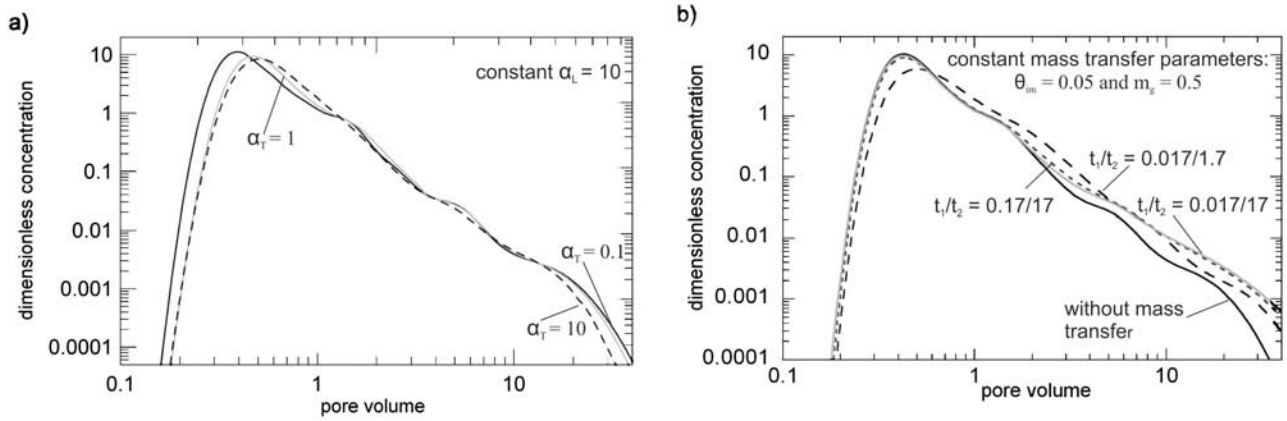

Figure 9. BTCs investigating the local-scale equation. (a) Subgrid heterogeneity is modeled using dispersivity (ADE). When the local transverse dispersivity is varied, the effect on the resulting BTCs is minor. The longitudinal dispersivity was kept constant. (b) A nonlocal small-scale equation is used: a mass transfer term is added with varying memory functions, which causes m_{BTC} to decrease when a large final cutoff time is used.

ing mobile porosity. If t_2 is increased further, peak arrival time is less increased, but and the slope decreases further. If also t_1 is increased by an order of magnitude and the effects are increased slightly. Remember that the assumed slope of the memory function $m_g = 0.5$ leads to a stronger weighting of large characteristic times (small α_n). Still, we find that the addition of a local-scale equation with memory does not make the slope smaller than 2, which would require a much larger immobile porosity and a much larger t_2.

4.4. Comparison With BTCs Using Mass Transfer

[43] Finally we use the memory function derived from the above curves to reproduce the BTC with the homogeneous one-dimensional mass transfer model of equations (1)–(4). The only fitting parameters are dispersivity and mobile porosity. They are fitted against early arrival and peak times. Mobile porosity (0.075–0.16) is in all cases a fraction of the original porosity (0.3). The resulting dispersivity ranges between 10.0 and 34.0, where 10.0 is the local (longitudinal) dispersivity. The slope is fitted very well (Figure 10) using an immobile porosity equal to the difference between total and mobile porosity ($\phi_{im} = 0.3 - \phi_m$). We repeated the model with a lower value of t_1 and we reproduced the BTC identically, but with a smaller dispersivity and a smaller mobile porosity. This means that the early time (high α_n terms) of our memory function cannot be estimated simultaneously with porosity and dispersivity from BTC data. This portion of the memory function equilibrates fast with the mobile region while producing some spreading. Hence, from a fitting point of view, neglecting this portion (i.e., increasing t_1) is virtually identical to increasing mobile porosity and dispersivity. This implies that t_1 is arbitrary, but so are mobile porosity and dispersivity. In fact, our formulation leads to a virtually identical fit setting dispersivity to 10 (local value) and mobile porosity equal to zero. It should be noticed that the early portion of the memory function is likely controlled by diffusion processes at the local (pore) scale, which have not been addressed here. In fact, proper representation of these processes requires pore-scale modeling [e.g., *Bijeljic and Blunt*, 2006; *Tartakovsky and Neuman*, 2008]. The point here is that they cannot be identified from a single, large-scale, BTC.

5. Discussion and Conclusions

[44] BTCs obtained from detailed simulations of transport through intermediate-scale highly heterogeneous media display the type of tailing often observed in field tracer tests. Intermediate-scale heterogeneity means here transport through distances comparable to, or smaller than, the largest heterogeneity scale. That is, transport over media with variability patterns (e.g., high-permeability channels) of size comparable to, or larger than, the size of the transport domain. Given the ubiquity of tailing at all scales, this result suggests that stationary *lnT* fields are rare and lends support to views of heterogeneity evolving over a range of scales [*Neuman*, 1990].

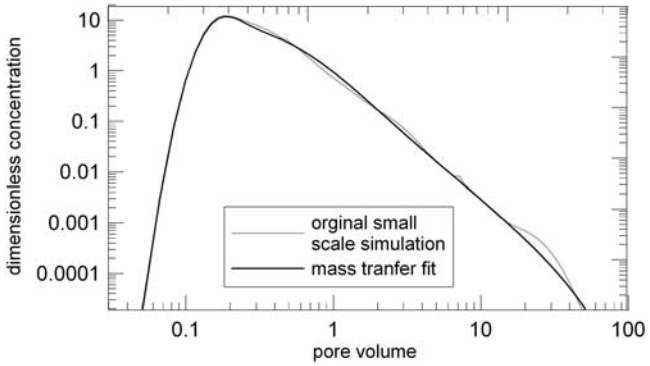

Figure 10. Fitted breakthrough curve using the memory function derived from the heterogeneous small-scale simulations. The only fitting parameters are dispersivity and mobile porosity. As part of the heterogeneity can be modeled either with the mobile zone dispersion or with the memory function, more than one set of parameters fit the original BTC. When t_1 is chosen small, the mobile zone dispersivity is equal to the local dispersivity of the original heterogeneous field simulation.

[45] The above kind of BTCs cannot be accurately modeled with the ADE, because of their non-Fickian nature, which can be well reproduced with nonlocal in time formulations (MRMT, CTRW, fADE). All these formulations require specifying a memory function, whose parameters are linked to those that describe heterogeneity. We have parameterized the memory function in terms of its slope in log-log scale and early and late cutoff times, t_1 and t_2.

[46] The slope of the memory function depends most markedly on connectivity indicators. Our simulations displayed no dependence of the slope on other parameters frequently used in describing heterogeneity, such as variance, Hurst coefficient or correlation distance. Yet, it is clear that they do affect BTC tailing (i.e., tailing disappears if the variance tends to zero and connectivity would have increased if we had used much larger correlation distances).

[47] The slope of the BTC is mildly reduced if nonlocal formulations are adopted for small-scale transport. In fact, if the cutoff time, t_2, for the small-scale transport is much larger than the travel time of the slow flow tubes, small-scale transport will dominate transport over long distances. The issue is nontrivial, nonlocality can be caused by diffusive processes [e.g., *Berkowitz et al.*, 2006], which is predictable and would be naturally described by a memory function at the local scale, and by slow advection which we have extensively discussed here. The problem is that nonlocality caused by diffusion would be scale-independent, while nonlocality caused by slow advection depends on both, mean travel time and distance. Specifically, memory function slope remains unchanged, but immobile porosity and late-time cutoff depend on advection time. Tracer tests are often performed under forced gradient conditions (i.e., velocities much larger than those occurring under natural conditions). The memory function derived from such tracer test should be scaled (i.e., t_2 increased in the same proportion than travel time [*Berkowitz et al.*, 2006]) if caused by slow advection, but not if caused by diffusion. Actually since both effects probably overlay, we would have to split the memory function into diffusion and a slow advection for proper scaling.

[48] The minimum slope encountered in this study for all the investigated fields is $m_{BTC} = 2$, while smaller slopes are sometimes observed in field tracer tests. The abundance of slopes close to 2 suggests that we have reached a limit for the type of fields investigated here. Further decrease in slope may be caused by long diffusion times into immobile regions, including heterogeneous diffusivity [*Gouze et al.*, 2008b], by three-dimensionality, especially with variable tortuosity, or by chemical heterogeneity, particularly for solutes that sorb into the least mobile region (e.g., clays).

[49] The scaling behavior of t_2 (late-time cutoff) is quite complex. Besides depending on advection time, t_2 appears to increase linearly with the variance of *lnT*. It is also affected by local-scale transverse dispersion. This implies that theoretical developments are needed before memory functions derived from tracer tests can be safely used for predicting long-term transport.

[50] A final note of caution must be added. The formulations we tested here are nonlocal only in time. They work very well, in the sense that they accurately reproduce BTCs. This is a result of our model setup and the fact that we were only trying to reproduce conservative transport BTCs. Had we tried to simulate reactive transport or spatial distributions of concentrations, nonlocality in space may have been needed. An indicator of this is provided by the small dependence of BTCs on local transverse dispersion, which controls mixing (and thus reactions). Therefore, we conclude that, while spreading can be well modeled with appropriately scaled memory function, mixing may not.

Appendix A: Implementation of a Discrete Mass Transfer Scheme

[51] We describe the implementation of our mass transfer scheme into a standard transport code without solving explicitly for each immobile zone. The exchange term is a sum of exchange terms between the mobile zone and each of the immobile zones f_n:

$$\Gamma = \sum_{n=1}^{N} f_n \tag{A1}$$

The exchange is described by

$$f_n = -\phi_{in}\alpha_n(c_{in} - c) \tag{A2}$$

where c_{in} is the concentration in immobile zone n. Mass balance in the nth immobile zone yields

$$\frac{dc_{in}}{dt} = \alpha_n(c_{in} - c) \tag{A3}$$

[52] The equation may be solved together with equation (1) while treating c_{in} as unknown. However, numerical solution is greatly simplified by eliminating it, much in the spirit of *Carrera et al.* [1998]. Assuming that the time is discretized and that c varies linearly during the time step (i.e., $\partial c/\partial t$ remains constant), we get

$$c_{in}(t) = c^k + (c_{in}^k - c^k)e^{\alpha_n t} + \frac{\partial c}{\partial t}\left[t - \frac{1}{\alpha_n}(1 - e^{-\alpha_n \Delta t})\right] \tag{A4}$$

Substituting $\partial c/\partial t$ by its time discretized approximation, $\Delta c/\Delta t$, and using the resulting value for c_{in} to eliminate it from equation (A2), while evaluating f_n at time $\theta\Delta t$:

$$f_n^{k+\theta} = \phi_{in}\frac{\Delta c}{\Delta t}(1 - e^{-\alpha_n \theta \Delta t}) + \alpha_n \phi_{in}(c_{in}^k - c^k)e^{-\alpha_n \theta \Delta t} \tag{A5}$$

[53] Notice that now f_n (source/sink per unit aquifer volume) is expressed in terms of only one unknown, $\Delta c/\Delta t$. Therefore, adding f_n to any numerical solver requires two operations.

[54] 1. Add the coefficient of $\Delta c/\Delta t$, $\phi_{in}(1 - e^{-\alpha_n \theta \Delta t})$, to the diagonal of the storage matrix for each immobile zone and each node (typically multiplied by the aquifer volume associated to the node, in finite element or finite difference codes)

[55] 2. Add the remaining term, $\alpha_n \phi_{in}(c_{in}^k - c^k)e^{-\alpha_n}\theta\Delta t$, to the right hand side (sinks and sources) vector (also for each immobile zone and node, and multiplied with the associated volume, if needed).

[56] We do not solve here explicitly for immobile concentrations. If needed, as for reactive transport simulations, this can be easily done as an extra step using equation (A4) after the system has been solved and $\Delta c/\Delta t$ is known.

[57] **Acknowledgments.** This work was partly funded by ENRESA and the European Commission through the project FUNMIG and by the Spanish agency for scientific research (CICYT) through project CTM2007-66724. The first author acknowledges additional financial support by AGAUR (Agència de Gestió d'Ajuts Universitaris i de Recerca) of the Catalan regional government. The review process was particularly lively, with all three reviewers (Andreas Cortis, Roy Haggerty, and an anonymous one), associate editor (Olaf Cirpka), and executive editor (Brian Berkowitz) providing a large number of substantial positive criticisms. We gratefully acknowledge their input.

References

Alcolea, A., J. Carrera, and A. Medina (2008), Regularized pilot points method for reproducing the effect of small scale variability: Application to simulations of contaminant transport, *J. Hydrol.*, *355*(1–4), 76–90, doi:10.1016/j.jhydrol.2008.03.004.

Becker, M. W., and A. Shapiro (2000), Tracer transport in fractured crystalline rock: Evidence of nondiffusive breakthrough tailing, *Water Resour. Res.*, *36*(7), 1677–1686.

Becker, M. W., and A. M. Shapiro (2003), Interpreting tracer breakthrough tailing from different forced-gradient tracer experiment configurations in fractured bedrock, *Water Resour. Res.*, *39*(1), 1024, doi:10.1029/2001WR001190.

Benson, D. A., S. W. Wheatcraft, and M. M. Meerschaerdt (2000), The fractional-order governing equation of Lévy motion, *Water Resour. Res.*, *36*(6), 1413–1423, doi:10.1029/2000WR900032.

Berkowitz, B., and H. Scher (1997), Anomalous transport in random fracture networks, *Phys. Rev. Lett.*, *79*(20), 4038–4041, doi:10.1103/PhysRevLett.79.4038.

Berkowitz, B., and H. Scher (1998), Theory of anomalous chemical transport in fracture networks, *Phys. Rev. E*, *57*(5), 5858–5869.

Berkowitz, B., and H. Scher (2009), Exploring the nature of non-Fickian transport in laboratory experiments, *Adv. Water Resour.*, in press.

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

Berkowitz, B., S. Emmanuel, and H. Scher (2008), Non-Fickian transport and multiple-rate mass transfer in heterogeneous media, *Water Resour. Res.*, *44*, W03402, doi:10.1029/2007WR005906.

Bijeljic, B., and M. J. Blunt (2006), Pore-scale modeling and continuous time random walk analysis of dispersion in porous media, *Water Resour. Res.*, *42*, W01202, doi:10.1029/2005WR004578.

Carrera, J. (1993), An overview of uncertainties in modelling groundwater solute transport, *J. Contam. Hydrol.*, *13*(1–4), 23–48, doi:10.1016/0169-7722(93)90049-X.

Carrera, J., X. Sánchez-Vila, I. Benet, A. Medina, G. Galarza, and J. Guimera (1998), On matrix diffusion: Formulations, solution methods and qualitative effects, *Hydrogeol. J.*, *6*(1), 178–190, doi:10.1007/s100400050143.

Cirpka, O. A., and A. J. Valocchi (2007), Two-dimensional concentration distribution for mixing-controlled bioreactive transport in steady state, *Adv. Water Resour.*, *30*, 1668–1679, doi:10.1016/j.advwatres.2006.05.022.

Cortis, A., and B. Berkowitz (2004), Anomalous transport in "classical" soil and sand columns, *Soil Sci. Soc. Am. J.*, *68*, 1539–1548.

Cortis, A., and C. Knudby (2006), A continuous time random walk approach to transient flow in heterogeneous porous media, *Water Resour. Res.*, *42*, W10201, doi:10.1029/2006WR005227.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, 465 pp., Springer, Berlin.

De Simoni, M., J. Carrera, X. Sánchez-Vila, and A. Guadagnini (2005), A procedure for the solution of multicomponent reactive transport problems, *Water Resour. Res.*, *41*, W11410, doi:10.1029/2005WR004056.

De Simoni, M., X. Sanchez-Vila, J. Carrera, and M. W. Saaltink (2007), A mixing ratios–based formulation for multicomponent reactive transport, *Water Resour. Res.*, *43*, W07419, doi:10.1029/2006WR005256.

Dentz, M., and B. Berkowitz (2003), Transport behavior of a passive solute in continuous time random walks and multirate mass transfer, *Water Resour. Res.*, *39*(5), 1111, doi:10.1029/2001WR001163.

Dentz, M., H. Kinzelbach, S. Attinger, and W. Kinzelbach (2000), Temporal behavior of a solute cloud in a heterogeneous porous medium: 1. Point-like injection, *Water Resour. Res.*, *36*(12), 3591–3604.

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*, 155–173.

Di Donato, G., E. Obi, and M. J. Blunt (2003), Anomalous transport in heterogeneous media demonstrated by streamline-based simulation, *Geophys. Res. Lett.*, *30*(12), 1608, doi:10.1029/2003GL017196.

Farrell, J., and M. Reinhard (1994), Desorption of halogenated organics from model solids, sediments, and soil under unsaturated conditions: 2. Kinetics, *Environ. Sci. Technol.*, *28*(1), 63–72.

Freyberg, D. L. (1986), A natural gradient experiment on solute transport in a sand aquifer: 2. sSpatial moments and the advection and dispersion of nonreactive tracers, *Water Resour. Res.*, *22*(13), 2031–2046.

Gaillard, B., D. Getto, D. Marion, S. Cottez, and J. P. Sauty (1990), Opérations de multiraçages en écoulement influencé par pompage pour la détermination des paramètres hydrodynamiques de l'aquifère superficiel du site d'El Cabril (Espagne), technical report, Enresa, Madrid.

Gelhar, L.-W. (1993), *Stochstic Subsurface Hydrology*, Prentice-Hall, Old Tappan, N. Y.

Gómez-Hernández, J., and A. Journel (1993), Joint sequential simulation of multigaussian fields, in *Geostat Troia 1992*, vol. 1, edited by Soares, pp. 85–94, Kluwer, Acad., Dordrecht, Netherlands.

Gouze, P., T. Le Borgne, R. Leprovost, G. Lods, T. Poidras, and P. Pezard (2008a), Non-Fickian dispersion in porous media: 1. Multiscale measurements using single-well injection withdrawal tracer tests, *Water Resour. Res.*, *44*, W06426, doi:10.1029/2007WR006278.

Gouze, P., Y. Melean, T. Le Borgne, M. Dentz, and J. Carrera (2008b), Non-Fickian dispersion in porous media explained by heterogeneous microscale matrix diffusion, *Water Resour. Res.*, *44*, W11416, doi:10.1029/2007WR006690.

Guimerà, J., and J. Carrera (2000), A comparison of hydraulic and transport parameters measured in low-permeability fractured media, *J. Contam. Hydrol.*, *41*(3–4), 261–281, doi:10.1016/S0169-7722(99)00080-7.

Hadermann, J., and W. Heer (1996), The Grimsel (Switzerland) migration experiment: Integrating field experiments, laboratory investigations and modelling, *J. Contam. Hydrol.*, *21*(1–4), 87–100, doi:10.1016/0169-7722(95)00035-6.

Haggerty, R., and S. M. Gorelick (1995), Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity, *Water Resour. Res.*, *31*(10), 2383–2400.

Haggerty, R., S. A. McKenna, and L. C. Meigs (2000), On the late-time behavior of tracer test breakthrough curves, *Water Resour. Res.*, *36*(12), 3467–3479.

Haggerty, R., C. F. Harvey, C. Freiherr von Schwerin, and L. C. Meigs (2004), What controls the apparent timescale of solute mass transfer in aquifers and soils? A comparison of experimental results, *Water Resour. Res.*, *40*, W01510, doi:10.1029/2002WR001716.

Kitanidis, P. K. (1988), Prediction by the method of moments of transport in a heterogeneous formation, *J. Hydrol.*, *102*, 453–473.

Knudby, C., and J. Carrera (2005), On the relationship between indicators of geostatistical, flow and transport connectivity, *Adv. Water Resour.*, *28*, 405–421, doi:10.1016/j.advwatres.2004.09.001.

Kosakowski, G., B. Berkowitz, and H. Scher (2001), Analysis of field observations of tracer transport in a fractured till, *J. Contam. Hydrol.*, *47*(1), 29–51.

Kreft, A., and A. Zuber (1978), On the physical meaning of the dispersion equation and its solutions for different initial and boundary conditions, *Chem. Eng.*, *33*, 1471–1480.

Lallemand-Barres, P., and P. Peaudecerf (1978), Recherche des relations entre a valeur de la dispersivite macroscopique d'un milieu aquifere, ses autres caracteristiques et les conditions de mesure, *Bur. Rech. Geol. Min. Fr., Sect.*, *3*(4), 277–284.

Le Borgne, T., and P. Gouze (2008), Non-Fickian dispersion in porous media: 2. Model validation from measurements at different scales, *Water Resour. Res.*, *44*, W06427, doi:10.1029/2007WR006279.

Le Borgne, T., M. Dentz, and J. Carrera (2008a), Spatial Markov processes for modeling Lagrangian particle dynamics in heterogeneous porous media, *Phys. Rev. E*, *78*, 026308.

Le Borgne, T., M. Dentz, and J. Carrera (2008b), Lagrangian statistical model for transport in highly heterogeneous velocity fields, *Phys. Rev. Lett.*, *101*, 090601.

Levy, M., M. B. BerkowitzLevy, and B. Berkowitz (2003), Measurement and analysis of non-Fickian dispersion in heterogeneous porous media, *J. Contam. Hydrol.*, *64*(3–4), 203–226.

Luo, J., M. Dentz, J. Carrera, and P. Kitanidis (2008), Effective reaction parameters for mixing controlled reactions in heterogeneous media, *Water Resour. Res.*, *44*, W02416, doi:10.1029/2006WR005658.

McKenna, S., L. Meigs, and R. Haggerty (2001), Tracer tests in a fractured dolomite: 3. Double-porosity, multiple-rate mass transfer processes in convergent flow tests, *Water Resour. Res.*, *35*(7), 1143–1154.

Medina, A., and J. Carrera (1996), Coupled estimation of flow and solute transport parameters, *Water Resour. Res.*, *32*(10), 3063–3076.

Meigs, L. C., and R. L. Beauheim (2001), Tracer tests in a fractured dolomite: 1. Experimental design and observed tracer recoveries, *Water Resour. Res.*, *37*(5), 1113–1128.

Neretnieks, I., and A. Rasmuson (1984), An approach to modelling radionuclide migration in a medium with strongly varying velocity and block sizes along the flow path, *Water Resour. Res.*, *20*(12), 1823–1836.

Neretnieks, I., T. Eriksen, and P. Tahtinen (1982), Tracer movement in a single fissure in granitic rock: Some experimental results and their interpretation, *Water Resour. Res.*, *18*(4), 849–858.

Neuman, S. P. (1990), Universal scaling of hydraulic conductivities and dispersivities in geologic media, *Water Resour. Res.*, *26*(8), 1749–1758.

Neuman, S. P., and V. Di Federico (2003), Multifaceted nature of hydrogeologic scaling and its interpretation, *Rev. Geophys.*, *41*(3), 1014, doi:10.1029/2003RG000130.

Neuman, S. P., and D. Tartakovsky (2009), Perspective on theories of anomaloustransport in heterogeneous media, *Adv. Water Resour.*, in press.

Rasmuson, A., and I. Neretnieks (1986), Radionuclide migration in strongly fissured zones: The sensitivity to some assumptions and parameters, *Water Resour. Res.*, *22*(4), 559–569.

Sánchez-Vila, X., and J. Carrera (1997), Directional effects on convergent flow tracer tests, *Math. Geol.*, *29*(4), 551–569.

Sánchez-Vila, X., and J. Carrera (2004), On the striking similarity between the moments of breakthrough curves for a heterogeneous medium and a homogeneous medium with a matrix diffusion term, *J. Hydrol.*, *294*(1–3), 164–175, doi:10.1016/j.jhydrol.2003.12.046.

Sánchez-Vila, X., I. Colominas, and J. Carrera (1993), FAITH: User's guide, Tech. Univ. of Catalonia, Barcelona, Spain.

Shapiro, A. M. (2001), Effective matrix diffusion in kilometer-scale transport in fractured crystalline rock, *Water Resour. Res.*, *37*(3), 507–522.

Shapiro, A. M., R. A. Renken, R. W. Harvey, M. R. Zygnerski, and D. W. Metge (2008), Pathogen and chemical transport in the karst limestone of the Biscayne aquifer: 2. Chemical retention from diffusion and slow advection, *Water Resour. Res.*, *44*, W08430, doi:10.1029/2007WR006059.

Tartakovsky, A. M., and S. P. Neuman (2008), Effects of Peclet number on pore-scale mixing and channeling of a tracer and on directional advective porosity, *Geophys. Res. Lett.*, *35*, L21401, doi:10.1029/2008GL035895.

Taylor, G. I. (1953), Dispersion of soluble matter in solvent flowing slowly through a tube, *Proc. R. Soc. London, Ser. A.*, *219*, 186–203.

Valocchi, A. J. (1985), Validity of the local equilibrium assumption for modeling sorbing solute transport through homogeneous soils, *Water Resour. Res.*, *21*(6), 808–820.

Wang, P. P., C. Zheng, and S. M. Gorelick (2005), A general approach to advective-dispersive transport with multirate mass transfer, *Adv. Water Resour.*, *28*, 33–42, doi:10.1016/j.advwatres.2004.10.003.

Werth, C. J., J. A. Cunningham, P. V. Roberts, and M. Reinhard (1997), Effects of grain-scale mass transfer on the transport of volatile organics through sediments: 2. Column results, *Water Resour. Res.*, *33*(12), 2727–2740.

Zhang, Y., B. Baeumer, and D. A. Benson (2006), Relationship between flux and resident concentrations for anomalous dispersion, *Geophys. Res. Lett.*, *33*, L18407, doi:10.1029/2006GL027251.

Zhang, Y., D. A. Benson, and B. Baeumer (2007), Predicting the tails of breakthrough curves in regional-scale alluvial systems, *Ground Water*, *45*(4), 473–484, doi:10.1111/j.1745-6584.2007.00320.x.

Zinn, B., and C. F. Harvey (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.

Zinn, B., L. C. Meigs, C. F. Harvey, R. Haggerty, W. J. Peplinski, and C. F. von Schwerin (2004), Experimental visualization of solute transport and mass transfer processes in two-dimensional conductivity fields with connected regions of high conductivity, *Environ. Sci. Technol.*, *38*(14), 3916–3926, doi:10.1021/es034958g.

J. Carrera, Institute of Environmental Analysis and Water Studies, CSIC, c/ Lluís Solà i Sabaras s/n, E-08028 Barcelona, Spain.

X. Sánchez-Vila, Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia, Gran Capita S/N, E-08034 Barcelona, Spain.

M. Willmann, Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia, Jordi Girona 1–3, D2, E-08034 Barcelona, Spain. (matthias.willmann@upc.edu)