

Numerical simulation of non-Fickian transport in geological formations with multiple-scale heterogeneities

Andrea Cortis

Department of Environmental Sciences and Energy Research, Weizmann Institute of Science, Rehovot, Israel

Claudio Gallo

Center for Advanced Studies, Research and Development in Sardinia (CRS4), Pula, Italy

Harvey Scher and Brian Berkowitz

Department of Environmental Sciences and Energy Research, Weizmann Institute of Science, Rehovot, Israel

Received 8 October 2003; revised 9 February 2004; accepted 27 February 2004; published 21 April 2004.

[1] We develop a numerical method to model contaminant transport in heterogeneous geological formations. The method is based on a unified framework that takes into account the different levels of uncertainty often associated with characterizing heterogeneities at different spatial scales. It treats the unresolved, small-scale heterogeneities (residues) probabilistically using a continuous time random walk (CTRW) formalism and the large-scale heterogeneity variations (trends) deterministically. This formulation leads to a Fokker-Planck equation with a memory term (FPME) and a generalized concentration flux term. The former term captures the non-Fickian behavior arising from the residues, and the latter term accounts for the trends, which are included with explicit treatment at the heterogeneity interfaces. The memory term allows a transition between non-Fickian and Fickian transport, while the coupling of these dynamics with the trends quantifies the unique nature of the transport over a broad range of temporal and spatial scales. The advection-dispersion equation (ADE) is shown to be a special case of our unified framework. The solution of the ADE is used as a reference for the FPME solutions for the same formation structure. Numerical treatment of the equations involves solution for the Laplace transformed concentration by means of classical finite element methods, and subsequent inversion in the time domain. We use the numerical method to quantify transport in a two-dimensional domain for different expressions for the memory term. The parameters defining these expressions are measurable quantities. The calculations demonstrate long tailing arising (principally) from the memory term and the effects on arrival times that are controlled largely by the generalized concentration flux term. *INDEX TERMS:* 1832 Hydrology: Groundwater transport; 1829 Hydrology: Groundwater hydrology; 1869 Hydrology: Stochastic processes; *KEYWORDS:* continuous time random walk, non-Fickian transport, nonstationary heterogeneity, numerical simulation

Citation: Cortis, A., C. Gallo, H. Scher, and B. Berkowitz (2004), Numerical simulation of non-Fickian transport in geological formations with multiple-scale heterogeneities, *Water Resour. Res.*, 40, W04209, doi:10.1029/2003WR002750.

1. Introduction

[2] The nature of the paths traveled by a contaminant in an aquifer is strongly influenced by the heterogeneities of the geological formation which determine the underlying flow field. Heterogeneities are present at all scales, from the submillimeter pore scale to the basin scale itself. The ubiquity and high degree of variability in heterogeneities thus rules out the possibility of obtaining complete knowledge of the pore space in which fluid and contaminant are transported.

[3] At the field scale, reasonable definition of the macroscopic characteristics of a geological formation can be

feasible, at a sufficiently coarse resolution, enabling (at least semi) deterministic modeling of flow and transport. However, smaller-scale heterogeneities, which can have a key influence on overall transport behavior, will always remain unresolved, and must be considered in a probabilistic manner. This latter treatment naturally involves a statistical measure, e.g., tracer travel times over a suitable length scale. The basic question, which we address in this paper, is how to integrate quantitative treatment of the small-scale and large-scale heterogeneities, and the effects of their interplay, in modeling overall transport behaviors.

[4] An extensive literature exists which considers a geological medium as a single, stationary domain. In such cases, “uniform” heterogeneities, which can be expressed statistically, are used to account for variability in the domain

properties. The advection-dispersion equation (ADE) with time-independent center of mass velocity, and constant dispersion coefficients, has been the classical tool for predicting contaminant transport. However, solutions of this equation do not correctly quantify many field-scale observations, which are characterized by transport coefficients that have been systematically found to be (length or time) scale-dependent [e.g., *Gelhar et al.* 1992]. This scale-dependent transport, often referred to as “anomalous” or “non-Fickian” transport, is exhibited in breakthrough curves by the appearance of anomalously early and late time arrivals.

[5] More sophisticated stochastic perturbative approaches have also been applied (as reviewed by, e.g., *Dagan* [1989], *Gelhar* [1993], and *Dagan and Neuman* [1997]). These approaches, however, employ an ensemble-averaged ADE, and assume intrinsically that transport is occurring in weakly heterogeneous media (with variance of log-hydraulic conductivity ≤ 1). An ensemble average over the entire domain is only suitable if the size of all the significant heterogeneities is much less than the domain size. With respect to a distribution of heterogeneities with a large variance, characteristic of most naturally occurring geological formations, preferential pathways separate flow and transport into “faster” and “slower” regions; lowest-order perturbation theory is not suitable to describe such behavior. In fact, significant deviations from predicted transport behavior have been observed in field experiments [e.g., *Adams and Gelhar*, 1992; *Boggs et al.*, 1992] as well as in numerical transport simulations [e.g., *Burr and Sudicky*, 1994; *Naff et al.*, 1998; *Salandin and Fiorotto*, 1998; *McLaughlin and Ruan*, 2001; *Pannone and Kitanidis*, 2001; *Dentz et al.*, 2002] and laboratory experiments [*Silliman and Simpson*, 1987; *Levy and Berkowitz*, 2003]. Other related stochastic approaches, deal with solute transport through the ADE and the use of time dependent macrodispersion coefficients [e.g., *Dagan*, 1994]. However, a consistent use of a time dependent dispersion coefficients requires its consideration in the framework of a non local in time transport equation, i.e., the generalized master equation as discussed in section 2.

[6] In the context of heterogeneous media, a different (nonperturbative) stochastic/probabilistic approach, based on the continuous time random walk (CTRW) formalism, has been introduced and successfully applied to hydrogeological systems by *Berkowitz and Scher* [1998, 2001]. These researchers noted that a broad distribution of characteristic timescales, notably power law transport time distributions, can result from a broad distribution of heterogeneity length scales [*Berkowitz and Scher*, 1995]. This phenomenological picture is the basis of CTRW models. Significantly, forms of the ADE, as well as a variety of mobile-immobile and multirate mass transfer models [e.g., *Roth and Jury*, 1993; *Haggerty and Gorelick*, 1995; *Cunningham et al.*, 1997; *Carrera et al.*, 1998; *Haggerty and Gorelick*, 1998] can be derived as special cases within the CTRW framework [e.g., *Berkowitz et al.*, 2002; *Dentz and Berkowitz*, 2003]. The formulation of the fractional differential equations is equivalent to a specific asymptotic case of the CTRW.

[7] Less attention has been devoted to quantitative treatment of transport in nonstationary geological formations,

i.e., in domains that contain both well-defined, large-scale trends and heterogeneities, as well as the inevitable small-scale, unresolved heterogeneities. Most numerical efforts are founded on discretization of the domain of interest into homogeneous zones with prescribed hydrogeological properties; the ADE with constant (or sometimes stochastic) parameters is then applied. In particular, the acquisition of highly detailed hydrogeological measurements is advocated, in order to permit definition of heterogeneities at as high a resolution as possible [e.g., *Koltermann and Gorelick*, 1996; *Labolle et al.*, 1996; *Feehley et al.*, 2000; *Labolle and Fogg*, 2001]. The ADE is then applied to quantify transport through such domains. However, as shown convincingly by, e.g., [*Eggleston and Rojstaczer*, 1998] and [*Sidle et al.*, 1998], unresolved heterogeneities, sometimes called “residues”, will always exist at a scale smaller than that of the measurement resolution. It has been demonstrated [e.g., *Levy and Berkowitz*, 2003] that small-scale heterogeneities existing even in “homogeneous” domains affect significantly the overall transport behavior. Thus the question of how best to integrate resolved and unresolved heterogeneities in transport modeling remains open.

[8] Recently, *Berkowitz et al.* [2002] proposed an alternative way to treat these multiple-scale heterogeneities, based on the CTRW formalism. In this framework, the unresolved small-scale heterogeneities are treated statistically by means of an ensemble averaged master equation (ME). The latter results in a generalized master equation (GME) that contains terms with a nonlocal dependence on time. The non-Fickian behavior arising at this small scale can be traced to a broad distribution of effective local transit times, which gives rise to a memory term [*Dentz et al.*, 2004]. At a larger scale the heterogeneities are treated deterministically and the relevant equation for the mass balance of a nonstationary medium is the classical (ME) [*Risken*, 1989]. Under the assumption of slowly varying concentration over a finite length scale, it is possible to split the solute flux derived from the ME into advective and dispersive parts. The resulting equation is a Fokker-Planck equation which accounts for the transitions in the geological formations in an explicit way. Contrary to the classical ADE case (with spatially dependent dispersivity), the splitting of the ME introduces a new advection term proportional to the divergence of the dispersivity tensor, hereafter referred to as the Fokker-Planck advection term.

[9] Following the ideas presented by *Berkowitz et al.* [2002], we propose to represent the evolution of a contaminant in a two-scale system combining the small scale GME terms, with the large-scale Fokker-Planck advection term into a single equation, hereafter referred to as the Fokker-Planck with memory equation (FPME). This equation is most conveniently solved in Laplace space, and subsequently inverted to the time domain to obtain profiles and breakthrough curves for the concentration. The present paper is concerned with the solution of the FPME for two-dimensional configurations.

[10] In section 2, we discuss derivation of the FPME, and in section 3 we discuss the physical derivation and motivation of the key element determining the residue-scale origin of non-Fickian behavior $\psi(\mathbf{s}, t)$, the probability rate of a local displacement \mathbf{s} at transit time t . In section 4, we discuss numerical solution of the FPME, and provide details

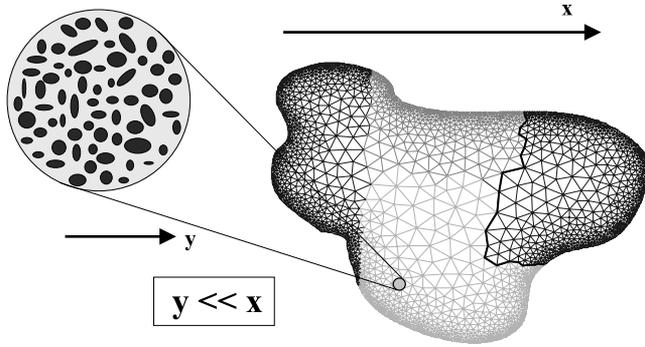


Figure 1. Two-scale system under consideration. The unresolved heterogeneities at the small-scale \mathbf{y} are treated stochastically, whereas the resolved heterogeneities at the large-scale \mathbf{x} are treated deterministically.

of the inverse Laplace transform method we used to obtain the time-dependent concentrations. In section 5 we contrast and discuss the concentration contour plots and breakthrough curves that result from the solutions of the FPME and the ADE, for a two-dimensional field with a range of disorder.

2. Basic Equations

[11] We start from the assumption that our system consists of two separate scales: a small unresolved scale \mathbf{y} and a resolved macroscopic length scale \mathbf{x} . The system under consideration is shown schematically in Figure 1. We also assume that we know exactly the structure of the heterogeneities at the \mathbf{x} scale, but we do not have a perfect knowledge of the heterogeneities at the smaller scale \mathbf{y} . The unresolved scale will be treated in a stochastic framework by means of the CTRW theory, whereas the resolved heterogeneities will be treated deterministically: the two scales will then be matched to obtain the overall governing equation.

2.1. Small Scale

[12] We begin by analyzing the problem at the small-scale \mathbf{y} . As we treat the small scale heterogeneities stochastically, we need to account for the unknown heterogeneities by means of an ensemble average of the classical master equation for the bulk concentration $c_b(\mathbf{y}, t)$

$$\frac{\partial c_b(\mathbf{y}, t)}{\partial t} = - \sum_{\mathbf{y}'} w_{\mathbf{y}}(\mathbf{y}', \mathbf{y}) c_b(\mathbf{y}, t) + \sum_{\mathbf{y}'} w_{\mathbf{y}}(\mathbf{y}, \mathbf{y}') c_b(\mathbf{y}', t), \quad (1)$$

where $w_{\mathbf{y}}(\mathbf{y}, \mathbf{y}')$ is the transition rate (units of reciprocal time) from \mathbf{y}' to \mathbf{y} in a single realization, which varies “randomly” from site to site. It is most expedient in the following to work with the Laplace Transform $\tilde{c}_b(\mathbf{y}, u)$

$$u \tilde{c}_b(\mathbf{y}, u) - c_0(\mathbf{y}) = - \sum_{\mathbf{y}'} w_{\mathbf{y}}(\mathbf{y}', \mathbf{y}) \tilde{c}_b(\mathbf{y}, u) + \sum_{\mathbf{y}'} w_{\mathbf{y}}(\mathbf{y}, \mathbf{y}') \tilde{c}_b(\mathbf{y}', u), \quad (2)$$

where u is the Laplace parameter. It is possible to show [Klafter and Silbey, 1980; Berkowitz et al., 2002] that the

ensemble average of the master equation can be written as a generalized master equation (GME),

$$u \tilde{c}_b(\mathbf{y}, u) - c_0(\mathbf{y}) = - \sum_{\mathbf{y}'} \tilde{\phi}(\mathbf{y}' - \mathbf{y}, u) \tilde{c}_b(\mathbf{y}, u) + \sum_{\mathbf{y}'} \tilde{\phi}(\mathbf{y} - \mathbf{y}', u) \tilde{c}_b(\mathbf{y}', u), \quad (3)$$

where $\tilde{\phi}(\mathbf{s}, u)$ is a memory function which can be expressed as

$$\tilde{\phi}(\mathbf{s}, u) = u \frac{\tilde{\psi}(\mathbf{s}, u)}{1 - \sum_{\mathbf{s}} \tilde{\psi}(\mathbf{s}, u)}, \quad (4)$$

where $\tilde{\psi}(\mathbf{s}, u)$ is the Laplace transform of the probability rate of a displacement \mathbf{s} at time t . As shown by Kenkre et al. [1973] and Shlesinger [1974], the form of equation (4) shows the complete equivalence of the GME and CTRW.

[13] Equation (3) does not separate out the effects of the varying velocity field into advective and dispersive contributions to the tracer transport. This separation can be obtained under the assumption that $c_b(\mathbf{y}, t)$ is slowly varying over a finite length scale for which we can write the following Taylor expansion

$$\tilde{c}_b(\mathbf{y}', u) \approx \tilde{c}_b(\mathbf{y}, u) + (\mathbf{y}' - \mathbf{y}) \cdot \nabla_{\mathbf{y}} \tilde{c}_b(\mathbf{y}, u) + \frac{1}{2} (\mathbf{y}' - \mathbf{y})(\mathbf{y}' - \mathbf{y}) : \nabla_{\mathbf{y}} \nabla_{\mathbf{y}} \tilde{c}_b(\mathbf{y}, u) + \dots \quad (5)$$

This assumption requires only that variations in the spatial particle transitions be relatively small (i.e., that the spatial moments of $\psi(\mathbf{s}, t)$ exist and constitute a length scale over which one can assume the concentration is slowly varying); large fluctuations can still occur in the particle transition times (rates). Under this assumption, following Berkowitz et al. [2002], it is possible to show that the differential equation governing the fluid concentration reads

$$n(u \tilde{c}(\mathbf{y}, u) - c_0(\mathbf{y})) = - \nabla_{\mathbf{y}} \cdot \tilde{\mathbf{j}}(\mathbf{y}, u), \quad (6)$$

where \tilde{c} is the fluid concentration related to the $\tilde{c}_b = n \tilde{c}$, via the porosity n , and the concentration flux $\tilde{\mathbf{j}}(\mathbf{y}, u)$ is defined as

$$\tilde{\mathbf{j}}(\mathbf{y}, u) = \mathbf{q}_{\psi}(u) \tilde{c}(\mathbf{y}, u) - \nabla_{\mathbf{y}} \cdot (\mathbf{D}_{\psi}(u) \tilde{c}(\mathbf{y}, u)). \quad (7)$$

Here $\mathbf{q}_{\psi}(u)$ and $\mathbf{D}_{\psi}(u)$ are the time-dependent fluid flux and generalized dispersion coefficient, respectively:

$$\mathbf{q}_{\psi}(u) = u \frac{\sum_{\mathbf{s}} \tilde{\psi}(\mathbf{s}, u) \mathbf{s}}{1 - \tilde{\psi}(u)}, \quad \mathbf{D}_{\psi}(u) = u \frac{\sum_{\mathbf{s}} \tilde{\psi}(\mathbf{s}, u) \frac{1}{2} \mathbf{s} \mathbf{s}}{1 - \tilde{\psi}(u)}, \quad (8)$$

where $\tilde{\psi}(u) \equiv \sum_{\mathbf{s}} \tilde{\psi}(\mathbf{s}, u)$. We define the transport velocity $\mathbf{v}_{\psi} = \mathbf{q}_{\psi}/n$.

[14] In transport applications the time range of interest is the one in which there has been an accumulation of many particle transitions, i.e., $t/\bar{t} \gg 1$, where \bar{t} is a characteristic time (to be defined in a specific application). In the limit of small $u\bar{t}$ (while retaining the general form of $\tilde{\psi}(\mathbf{s}, u)$), the main u dependence is determined by the denominator in equation (8), i.e., by $\tilde{\psi}(u)$ [Berkowitz et al., 2002].

[15] For convenience and without loss of generality in terms of u [Dentz *et al.*, 2004] we at this point separate the transition rate probability in the form $\tilde{\psi}(\mathbf{s}, u) = \psi(u) p(\mathbf{s})$, $\sum_{\mathbf{s}} p(\mathbf{s}) = 1$. We can now rewrite the generalized fluid flux and dispersion coefficients [Berkowitz *et al.*, 2002] as

$$\mathbf{q}_{\psi}(u) = u \frac{\tilde{\psi}(u)}{1 - \tilde{\psi}(u)} \sum_{\mathbf{s}} p(\mathbf{s}) \mathbf{s}, \quad \mathbf{D}_{\psi}(u) = u \frac{\tilde{\psi}(u)}{1 - \tilde{\psi}(u)} \sum_{\mathbf{s}} p(\mathbf{s}) \frac{1}{2} \mathbf{s} \mathbf{s} \quad (9)$$

and multiplying and dividing by \bar{t} we obtain $\mathbf{q}_{\psi}(u) = \tilde{M}(u) \mathbf{q}$ and $\mathbf{D}_{\psi}(u) = \tilde{M}(u) \mathbf{D}$ where

$$\tilde{M}(u) = \bar{t} u \frac{\tilde{\psi}(u)}{1 - \tilde{\psi}(u)} \quad (10)$$

is a memory function, and

$$\mathbf{q}_{\psi} = \frac{1}{\bar{t}} \sum_{\mathbf{s}} p(\mathbf{s}) \mathbf{s}, \quad \mathbf{D}_{\psi} = \frac{1}{\bar{t}} \sum_{\mathbf{s}} \frac{1}{2} p(\mathbf{s}) \mathbf{s} \mathbf{s}. \quad (11)$$

It will be useful in the following to combine the equations in equation (11) in the form

$$\mathbf{D}_{\psi} = \frac{|\sum_{\mathbf{s}} p(\mathbf{s}) \mathbf{s}| |\sum_{\mathbf{s}} \frac{1}{2} p(\mathbf{s}) \mathbf{s} \mathbf{s}|}{|\sum_{\mathbf{s}} p(\mathbf{s}) \mathbf{s}|} \equiv \alpha_{\psi} |\mathbf{q}_{\psi}| \quad (12)$$

where the tensor

$$\alpha_{\psi} = \frac{\sum_{\mathbf{s}} \frac{1}{2} p(\mathbf{s}) \mathbf{s} \mathbf{s}}{|\sum_{\mathbf{s}} p(\mathbf{s}) \mathbf{s}|}, \quad (13)$$

has dimension of length.

[16] With the above definitions, equation (6) can now be rewritten as

$$n(u\tilde{c}(\mathbf{y}, u) - c_0(\mathbf{y})) = -\tilde{M}(u) \nabla_{\mathbf{y}} \cdot \left[\mathbf{q}_{\psi} \tilde{c}(\mathbf{y}, u) - \mathbf{D}_{\psi} \cdot \nabla_{\mathbf{y}} \tilde{c}(\mathbf{y}, u) \right], \quad (14)$$

which is the equation valid for a non-Fickian transport in an ensemble averaged medium at the \mathbf{y} scale.

[17] Dentz *et al.* [2004] demonstrated that, for the case of an infinite ensemble-averaged medium, the particle tracking simulations of equation (3) match perfectly the solution of equation (14), i.e., the split between the advective and the dispersive part of the total flux derived from the Taylor expansion in equation (5) leads to accurate results for a broad class of $\tilde{\psi}(\mathbf{s}, u)$. Moreover, the solution of equation (14) is equal to the ADE solution divided by $\tilde{M}(u)$ and with u replaced by $u/\tilde{M}(u)$.

[18] Given a unit steady state flux boundary condition $\tilde{j} = u^{-1}$ at the inlet ($x = 0$) and a natural boundary condition $\partial\tilde{c}/\partial x = 0$ at the outlet ($x = 1$), the exact 1-D analytical solution of equation (14) for $\tilde{j}(u)$ at $x = 1$ is

$$\tilde{j}(u) = \frac{1}{u} \frac{2z \exp\left(\frac{1}{2}(\alpha_{\psi}^{-1} + z)\right)}{\left[\exp(z)\left(z + \alpha_{\psi}^{-1} + 2u/(\tilde{M}v_{\psi})\right) + \left(z - \alpha_{\psi}^{-1} - 2u/(\tilde{M}v_{\psi})\right)\right]} \quad (15)$$

where

$$z = \frac{1}{\alpha_{\psi}} \sqrt{1 + \frac{4u\alpha_{\psi}}{\tilde{M}v_{\psi}}}. \quad (16)$$

2.2. Large Scale

[19] In the case of a nonstationary medium, i.e., a medium which has known varying properties at a scale $\mathbf{x} \gg \mathbf{y}$, the appropriate starting point is the master equation (2) at the scale \mathbf{x} , and not the ensemble-averaged master equation (3) [Berkowitz *et al.*, 2002]. This is because the transition rates $w_{\mathbf{x}}(\mathbf{x}', \mathbf{x})$ are assumed to be known, at least in principle. Substituting in equation (2) the same expansion as in equation (5) (for the scale \mathbf{x}), yields a Fokker-Planck equation [Risken, 1989]

$$n(\mathbf{x})(u\tilde{c}(\mathbf{x}, u) - c_0(\mathbf{x})) = -\nabla_{\mathbf{x}} \cdot \left[\mathbf{q}(\mathbf{x}) \tilde{c}(\mathbf{x}, u) - \nabla_{\mathbf{x}} \cdot (\mathbf{D}(\mathbf{x}) \tilde{c}(\mathbf{x}, u)) \right], \quad (17)$$

with time independent transport coefficients and a x -dependent \mathbf{q} and \mathbf{D} . Note that in equation (17), $\mathbf{D}(\mathbf{x})$ is a factor for the concentration \tilde{c} whereas in the classical ADE approach equation (14), $\mathbf{D}(\mathbf{x})$ is a factor for $\nabla\tilde{c}$.

2.3. Integrating the Small and Large Scales

[20] In the present work, we propose to combine the non-Fickian transport induced by the small (and intermediate) scale unresolved heterogeneities, with a Fokker-Planck treatment of the (deterministically known) large scale heterogeneities. The position dependence of $\tilde{\psi}(\mathbf{s}, u; \mathbf{x})$ is carried by $p(\mathbf{s}; \mathbf{x})$, while the $\psi(u; \mathbf{x})$ is considered to be a purely local change in the u dependence and scaling. Introducing the space dependent memory term $\tilde{M}(u; \mathbf{x})$, specific discharge $\mathbf{q}_{\psi}(\mathbf{x})$ and dispersion $\mathbf{D}_{\psi}(\mathbf{x})$ in equation (17), we obtain

$$n(\mathbf{x})(u\tilde{c}(\mathbf{x}, u) - c_0(\mathbf{x})) = -\nabla_{\mathbf{x}} \cdot \left[\tilde{M}(u; \mathbf{x}) \mathbf{q}_{\psi}(\mathbf{x}) \tilde{c}(\mathbf{x}, u) - \nabla_{\mathbf{x}} \cdot (\tilde{M}(u; \mathbf{x}) \mathbf{D}_{\psi}(\mathbf{x}) \tilde{c}(\mathbf{x}, u)) \right] \quad (18)$$

which is a Fokker-Planck equation with a memory term. Developing the second derivative in equation (18) we can write

$$n(\mathbf{x})(u\tilde{c}(\mathbf{x}, u) - c_0(\mathbf{x})) = -\tilde{M}(u; \mathbf{x}) \left[\left(\mathbf{q}_{\psi}(\mathbf{x}) - \nabla_{\mathbf{x}} \cdot \mathbf{D}_{\psi}(\mathbf{x}) \right) \cdot \nabla_{\mathbf{x}} \tilde{c}(\mathbf{x}, u) - \nabla_{\mathbf{x}} \cdot (\mathbf{D}_{\psi}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} \tilde{c}(\mathbf{x}, u)) + \tilde{c}(\mathbf{x}, u) \nabla_{\mathbf{x}}^2 \mathbf{D}_{\psi}(\mathbf{x}) \right]. \quad (19)$$

[21] For the hydrogeological applications considered herein, it can be shown that spatial derivatives of $\tilde{M}(u; \mathbf{x})$ are small and are not included in equation (19) (mass is conserved to an excellent approximation in our results) and we shall assume that the term $\nabla_{\mathbf{x}}^2 \mathbf{D}_{\psi}(\mathbf{x})$ can be neglected, so that equation (19) simplifies to

$$n(\mathbf{x})(u\tilde{c}(\mathbf{x}, u) - c_0(\mathbf{x})) = -\tilde{M}(u; \mathbf{x}) \left[\left(\mathbf{q}_{\psi}(\mathbf{x}) - \nabla_{\mathbf{x}} \cdot \mathbf{D}_{\psi}(\mathbf{x}) \right) \cdot \nabla_{\mathbf{x}} \tilde{c}(\mathbf{x}, u) - \nabla_{\mathbf{x}} \cdot (\mathbf{D}_{\psi}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} \tilde{c}(\mathbf{x}, u)) \right]. \quad (20)$$

Dropping the term $\nabla_{\mathbf{x}}^2 \mathbf{D}_{\psi}$ is reasonable when the \mathbf{D} field can be seen as varying piecewise. Of course, the term should not

be neglected if the dispersion field can be resolved on intermediate scales smaller than the large (facies) scale (but larger than the microscopic heterogeneities). Equation (20) is a general equation which accounts simultaneously for the nonlocal effects of the small-scale, unresolved heterogeneities, and the large-scale, known heterogeneities. Clearly, for a macroscopically homogeneous medium (i.e., no variations in n , \mathbf{q}_ψ , \mathbf{D}_ψ , and $\tilde{M}(u)$ over the scale \mathbf{x}), equation (20) reduces to the form (14) but at the scale \mathbf{x} . Recall that $\tilde{M}(u)$ as given by equation (10) accounts for the small-scale (\mathbf{y}) heterogeneities.

[22] We have thus mapped the effect of the small-scale heterogeneities on the distribution of local transit times into the memory term $\tilde{M}(u; \mathbf{x})$ which is responsible for the anomalous (non-Fickian) dispersion, while the effect of the macroscopic heterogeneities is mapped into the $\mathbf{q}_\psi(\mathbf{x})$ and $\mathbf{D}_\psi(\mathbf{x})$. In the Fokker-Planck equation the effect of the macroscopic heterogeneities are additionally included in the important drift correction term, $\nabla_{\mathbf{x}} \cdot \mathbf{D}_\psi(\mathbf{x})$.

2.4. Special Cases of the FPME

[23] If one discards the drift “correction term”, $\nabla_{\mathbf{x}} \cdot \mathbf{D}_\psi(\mathbf{x})$ and sets $\tilde{M}(u; \mathbf{x}) = 1$, equation (20) is formally identical to the classical ADE equation, and the term $\mathbf{q}_\psi(\mathbf{x})$ can be identified with the classical Darcy velocity, for which we have $\nabla_{\mathbf{x}} \cdot \mathbf{q}_\psi(\mathbf{x}) = 0$. At a local, averaged scale it is still possible to define a dispersivity tensor α_ψ such that $\mathbf{D}_\psi = \alpha_\psi |\mathbf{q}_\psi|$ (see equation (13)). Note, however, that our dispersivity tensor α has a different interpretation altogether with respect to the corresponding dispersivity tensor appearing in the classical ADE [cf., e.g., Bear, 1972; Bear and Bachmat, 1991].

[24] It is also noting that, if and only if $\tilde{M}(u; \mathbf{x}) = u^{1-\beta}$ (which corresponds to a $\psi(t) \sim t^{-(1+\beta)}$ for large t/\bar{t}) with a constant $\beta < 1$, equation (20) reduces to a fractional derivative equation (FDE) [Metzler and Klafter, 2000]

$$n(\mathbf{x}) \frac{\partial c(\mathbf{x}, t)}{\partial t} = - \frac{\partial^{1-\beta}}{\partial t^{1-\beta}} [\mathbf{q}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} c(\mathbf{x}, t) - \nabla_{\mathbf{x}} \cdot (\mathbf{D}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} c(\mathbf{x}, t))]. \quad (21)$$

with the definition of the operator

$$\frac{\partial^{-\beta}}{\partial t^{-\beta}} c(\mathbf{x}, t) \equiv \frac{1}{\Gamma(\beta)} \int_0^t dt' \frac{c(\mathbf{x}, t')}{(t-t')^{1-\beta}} \quad (22)$$

which possesses the important property $\mathcal{L}\{\partial^{-\beta} c(\mathbf{x}, t)/\partial t^{-\beta}\} = u^{-\beta} \tilde{c}(\mathbf{x}, u)$.

[25] We stress that in the theory we develop here the functional form of the memory function, in the FPME, is general and depends only on the particle motion as dictated by the underlying flow patterns and by interactions with the media (as examined in section 3). The FDE is limited to the special case of $\tilde{M}(u; \mathbf{x}) = u^{1-\beta}$ at small $u\bar{t}$ with $\beta < 1$ as indicated above. In this asymptotic limit, the solutions of equation (21) are not necessarily physically meaningful for small t (see Appendix A).

3. Transition Rate Probability $\psi(t)$

3.1. Physical Interpretation

[26] In any particular application of an equation such as equation (20), one must obtain the form of $\psi(\mathbf{y}, t)$ from delineation of the main physical processes. The hydrogeological applications we have considered to date include

transport in fracture networks, “macroscopically homogeneous” sands, and heterogeneous porous media [e.g., Berkowitz and Scher, 1998; Kosakowski et al., 2001; Levy and Berkowitz, 2003].

[27] In order to understand the physical meaning of the transition rate probability $\psi(t)$, it is instructive to review an example of diffusion in a random molecular system [Scher and Lax, 1973]. We then discuss alternative forms of this transition rate probability, as they can be applied to hydrogeological systems.

[28] We start directly from the ME and work with the w transition rates introduced in equation (1). These are the rates at which a particle on one molecular site transfers to another site. The change in the probability $Q(t)$ (normalized concentration) for a particle to remain at a site is determined by the first term on the right-hand side of the ME equation (1)

$$\frac{dQ}{dt} = -Q \sum_j w(r_j), \quad (23)$$

thus

$$Q(t) = \exp\left(-t \sum_j w(r_j)\right). \quad (24)$$

The ensemble average of equation (24) is easily computed and for

$$w(r) = w_M \exp(-r/R_0) \quad (25)$$

one obtains, for large $w_M t$

$$\langle Q(t) \rangle = \exp(-\mathcal{P}(\ln(w_M t))) \quad \text{with} \quad \psi(t) = - \frac{d\langle Q(t) \rangle}{dt} \quad (26)$$

where $\mathcal{P}(x)$ is a polynomial with the leading term $(\eta/3)(\ln(w_M t))^3$ and $\eta \equiv 4\pi\rho R_0^3$ and ρ is equal to the molecular site density. The value of η represents the average number of sites in a sphere of radius R_0 , the effective transfer distance. As η decreases, the fluctuations of neighboring sites within R_0 increases. These fluctuations amplify the spread in the transfer rates $w(r)$ in equation (25), i.e., small variations in intersite separation give rise to large variations in transition times. The key pdf is $\psi(t)$, the distribution of these times while the spatial displacements are slowly varying. Hence η is a measure of the disorder of the system. The “geometry” of the medium is set by ρ and R_0 calibrates the transfer distance.

[29] The example of diffusion in a random molecular system illustrates the important interplay of length scale ℓ and fluctuations. If one considered $\ell \gg (4\pi\rho/3)^{-1/3}$, the average intersite distance, then the transition rates w would involve an average over many sites of the effective transport (local currents) and there would be a suppression of the fluctuations.

[30] Similar considerations hold for chemical transport in a steady flow in disordered porous media or random fracture networks (RFN). In one mechanism the transition rates, w , are a function of the local flow velocity \mathbf{v} and field and laboratory observations of the transport also indicate a sensitivity to fluctuations, i.e., in \mathbf{v} [e.g., Levy and Berkowitz 2003]. The natural length scale ℓ of a porous

medium that is sufficient to capture the fluctuations in \mathbf{v} is one that is large enough to define a local porosity and small enough compared to the coherence lengths of the flow field (the heterogeneity scale). In a RFN one has that ℓ is of the order of the length of the segment between the intersections. These fluctuations in \mathbf{v} determine the spread in transition times. As we assume here smooth variation in the spatial displacements and work in the limit of a spatial continuum, we focus on retaining a good model for $\psi(t; \mathbf{x})$.

[31] As a simple example one assessment of the disorder in a porous medium is the permeability distribution which is often assumed to be lognormally distributed. If we make the assumption, adequate for this discussion, that the velocity distribution resulting from a lognormal permeability distribution follows a similar distribution, and we use $t \propto 1/v$, one has, for large dimensionless times $\tau (\tau \equiv t/\bar{t})$, $\psi(t) = \sqrt{\frac{\alpha}{\pi\tau}} \exp(-\gamma \ln^2 \tau)$ where $\gamma \equiv \frac{1}{2b^2}$ and b is the log standard deviation, which is a measure of the disorder. Again this is a form of $\psi(t) \propto \exp(-\mathcal{P}(\ln(\tau)))$ with γ playing the role of η . Clearly, of course, the actual flow regime influences the relationship between the permeability and velocity distributions.

[32] Generally, the situation is more complicated in that the transition times can be caused by mechanisms other than those strictly of the flow field. The chemicals can diffuse into and out of “stagnant” zones of the medium, and/or adsorb/desorb from the internal surfaces. One option is simply to incorporate these additional effects on overall migration of contaminant in an “effective” form of the transition rate probability $\psi(t)$. Alternatively, especially in terms of a RFN, we can denote such time delays as τ_d and include them in the expression

$$\psi(\mathbf{y}, t) \propto \chi(\mathbf{y}) \int_0^\infty d\xi d\tau_d \Phi_{eff}(\vec{\xi}) \Theta(\tau_d) \delta(t - y\xi - \tau_d) \quad (27)$$

where $\chi(\mathbf{y})$ is a pdf of lengths (e.g., fragments in a RFN), $\Phi_{eff}(\vec{\xi}) = \sum_{AP} \rho(A) \phi(\vec{\xi}; A)$, $\xi \equiv 1/v$, $\vec{\xi} = \hat{\mathbf{y}}$, $\rho(A)$ is a pdf of aperture A , $\phi(\vec{\xi}; A)$ is the pdf of flow velocity corresponding to aperture A (for a RFN); $\Theta(\tau_d)$ is the delay time pdf. If $\Theta(\tau_d) = \delta(\tau_d)$ the expression in equation (27) reduces to our previous one [Scher *et al.*, 2002]. In equation (27) we have a combined effect of transit steps due to flow and multirate transfer.

3.2. Choice of $\psi(t)$

[33] As discussed above, for application of equation (20), we need to choose a transition rate probability density function $\psi(t)$. Clearly, for application to specific, real systems, the functional form of $\psi(t)$ and parameter values within it must be fit or derived from measurable properties of the medium, the flow field and/or the tracer transport itself. Here we consider three possible “generic” forms of $\psi(t)$.

3.2.1. Asymptotic $\psi(t)$

[34] First, a form easily parameterized in Laplace space is

$$\tilde{\psi}(u) = (1 + au + bu^\beta)^{-1}. \quad (28)$$

We shall refer to this model as the “asymptotic” form. This form of $\psi(t)$ is closely related to the general algebraic form $\psi(t) \sim t^{-(1-\beta)}$ which has been a foundation of CTRW theory (as discussed in detail by, e.g., Berkowitz and Scher [2001]).

To illustrate this, note that it is sufficient to add to the Laplace Transform of this algebraic form the first term of the integer power expansion

$$\tilde{\psi}(u) = 1 - bu^\beta - au, \quad (29)$$

which for small u can be approximated by equation (28). These forms have no temporal cutoff and transport remains non-Fickian for all time. The limits of applicability of this function (and in fact of any choice of $\tilde{\psi}(u)$) are discussed in Appendix A.

3.2.2. Truncated Power Law $\psi(t)$

[35] The second form of $\psi(t)$ that we shall consider is the so-called “truncated power law”

$$\psi(t) = \left\{ t_1 \tau_2^{-\beta} \exp(\tau_2^{-1}) \Gamma(-\beta, \tau_2^{-1}) \right\}^{-1} \frac{\exp(-t/t_2)}{(1 + t/t_1)^{1+\beta}} \quad (30)$$

where $\tau_2 \equiv t_2/t_1$ and $\Gamma(a, x)$ is the incomplete Gamma function [Abramowitz and Stegun, 1970]. For $t_1 \ll t \ll t_2$, $\psi(t) \propto (t/t_1)^{-1-\beta}$. In this time regime the transport behavior is anomalous for $0 < \beta < 2$. For transport times $t \gg t_2$ the transport behavior is Fickian. For $\beta > 2$ the transport behavior becomes Fickian already for $t \gg t_1$ and the upper cutoff t_2 plays only a minor role on the transport behavior. Thus the form (30) emphasizes the intermediate range algebraic behavior and the crossover time to normal transport discussed above. This form of $\psi(t)$ is discussed in detail by Dentz *et al.* [2004].

[36] The virtues of these two representative expressions of $\psi(t)$ are their analytical forms for $\tilde{\psi}(u)$, which expedite the extensive use of numerical inverse Laplace transforms (see section 4). In Figures 2a and 2b we compare the behavior of the two different $\psi(t)$ expressions in equations (28) and (30). The parameters of the asymptotic model represented in Figure 2a are $\beta = 0.75$, $b = 10$, $a = b^\beta$. This model is compared to two expressions for the truncated power law $\psi(t)$ for the same value of the β parameter, with $t_1 = 1$ and exponential truncation times τ_2 equal to 10^2 and 10^6 , respectively. In Figure 2b we decrease the value of the power law exponent to $\beta = 0.5$, while keeping all other parameters unchanged. This case corresponds to a more highly disperse system.

[37] The main measurable outputs of our calculations are breakthrough curves. In Figures 2c and 2d we compare the 1-D BTC solutions of equation (15) for the two different $\psi(t)$ expressions in equations (28) and (30), for $\nu_\psi = 1$, and $\alpha_\psi = 0.05$. The long tailing arising from the memory function is clearly evident in Figures 2c and 2d. We recall that the $\beta < 1$ region corresponds to a highly dispersive flow situation. Smaller values of β (Figure 2d) thus correspond to longer tailing. Notice that when the truncation time is much larger than the overall transport time (in this case $t_{max} = 10^3$), the truncated power law model does not yet show the effect of the exponential truncation, and the corresponding breakthrough curve continues to increase in time in the same fashion as the asymptotic model. However, when $\tau_2 \approx t_{max}$ the truncated power law model introduces the transition to a Gaussian-like spreading, i.e., the domain is effectively homogenized by the introduction of the truncation parameter.

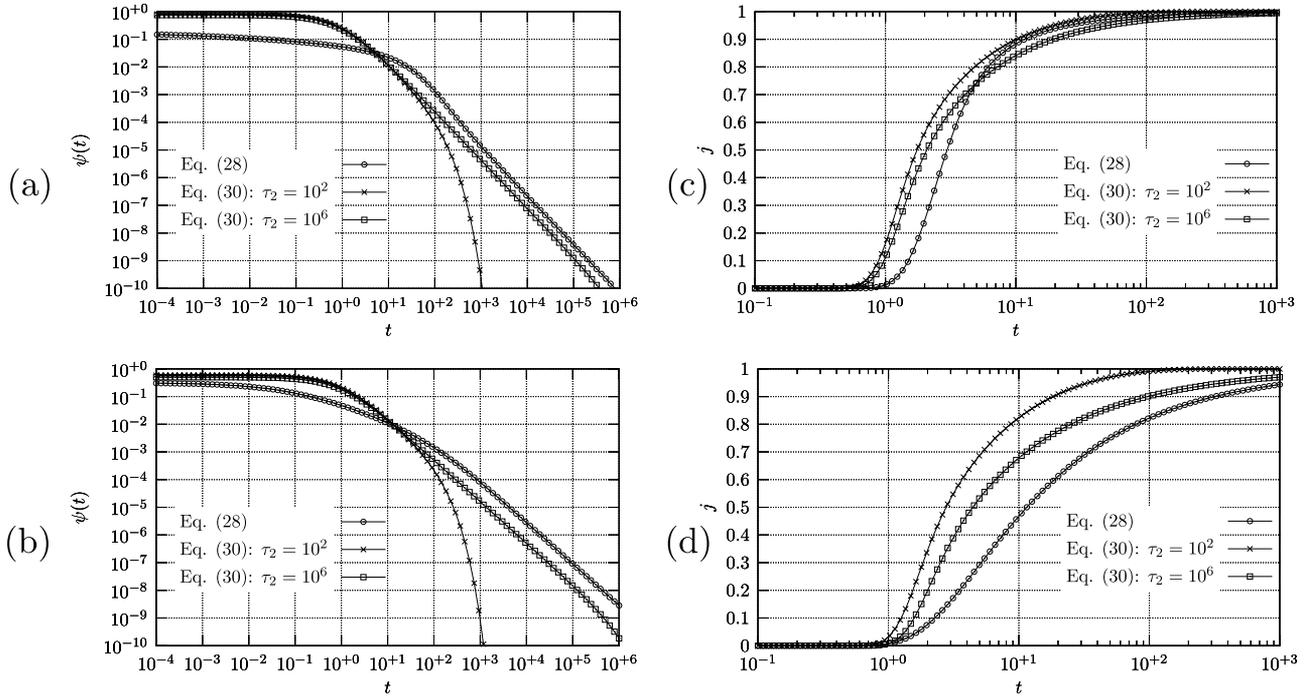


Figure 2. Comparison of the asymptotic equation (28) and truncated power law equation (30) $\psi(t)$ functions. The behavior of both $\psi(t)$ functions is shown for (a) $\beta = 0.75$ and (b) $\beta = 0.5$. The asymptotic $\psi(t)$ is plotted for $b = 10$, $a = b^\beta$; the function (30) is plotted for two different values of the truncation time τ_2 , with $t_1 = 1$. (c and d) Breakthrough curves corresponding to the different $\psi(t)$ in Figures 2a and 2b for a 1-D flow over a unit domain, with a free flow boundary condition at the outlet. The time t is dimensionless.

3.2.3. Modified Exponential $\psi(t)$

[38] The third expression we consider is the one derived [Scher and Lax, 1973] for diffusion in a random molecular system (an asymptotic form of which is discussed in section 3.1)

$$\psi(t) = -\frac{d}{dt} e^{-\eta} \int_0^\infty [1 - e^{-\eta x}] x^2 dx, \quad (31)$$

for which an analytic expression can be obtained

$$\psi(t) = 2 \eta {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix} ; -t \right] e^{-2\eta t} {}_4F_4 \left[\begin{matrix} 1, 1, 1, 1 \\ 2, 2, 2, 2 \end{matrix} ; -t \right], \quad (32)$$

where ${}_pF_q$ is the generalized hypergeometric function defined as

$${}_pF_q \left[\begin{matrix} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{matrix} ; x \right] = \sum_0^\infty \frac{(a_1)_k (a_2)_k \dots (a_p)_k}{(b_1)_k (b_2)_k \dots (b_q)_k} \frac{x^k}{k!}, \quad (33)$$

and $(a)_k = \Gamma(a+k)/\{\Gamma(a)\} = a(a+1) \dots (a+k-1)$ is the Pochhammer symbol, also known as the ‘‘rising factorial’’. We notice that the expression (32) has the structure of an exponential function ($\psi(t) \sim \exp(-t)$) modified by the introduction of the two functions ${}_3F_3$ and ${}_4F_4$. A plot of these two functions is shown in Figure 3. These hypergeometric functions are always positive and tend for small times to the value 1, reducing $\psi(t)$ to a pure exponential.

The limits of these functions for $t \rightarrow \infty$ are also constants equal to $\pi^2/12$ and $\zeta(3)/3$, for ${}_3F_3$ and ${}_4F_4$ respectively, the ζ function being defined as by Abramowitz and Stegun [1970]. We recall that when the $\psi(t)$ is a pure exponential, then the spreading behavior is purely Gaussian. This can be easily seen by substituting the Laplace transform of the pure exponential $\tilde{\psi}(u) = (1+u)^{-1}$ into equation (10) to obtain $\tilde{M} = 1$. Thus equation (32) can actually account for the transition from a purely diffusive regime (at extremely small times when the tracer does not yet ‘‘see’’ the microscopic heterogeneities around the spreading source) to (eventually) another ‘‘homogenized’’ regime in which the length scales involved in the overall transport are much

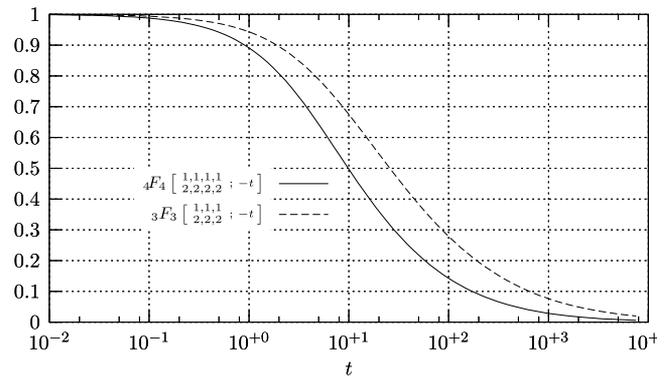


Figure 3. Plot of the ${}_3F_3$ and ${}_4F_4$ functions in equation (32) versus dimensionless time t .

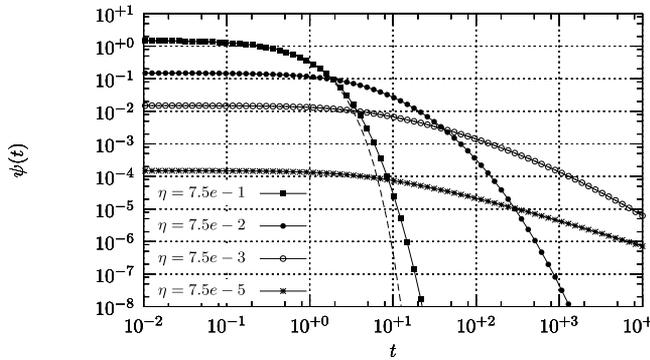


Figure 4. Evolution of the transition probability $\psi(t)$ versus dimensionless time t in equation (32) for different values of the η parameter. The dashed line represents the exponential limit of the function for $\eta = 0.75$.

larger than the correlation length of the disordered porous system. The regime of most practical interest, however, lies in the broad transition between these two extremes in which anomalous transport prevails.

[39] In order to perform the actual numerical computations, we approximate the $\psi(t)$ in equation (32) by joining together two expressions valid for short and long time ranges. Our numerical analysis indicates that the short time behavior is well approximated by the first 28 terms of the Taylor expansion of equation (32) around time $t = 0$ for times $t < t_0$, while the large time ($t > t_0$) asymptotic expression is given by [Scher and Lax, 1973]

$$\frac{1}{t} \frac{\eta}{6} \left(\pi^2 + 6(\gamma + \ln t)^2 \right) e^{-\frac{\eta}{6} (\pi^2 (\gamma + \ln t) + 2(\gamma + \ln t)^3 + 4\zeta(3))}, \quad (34)$$

where γ is the Euler constant. The Laplace transform of this $\psi(t)$ is evaluated numerically for all values of the u Laplace parameter using an adaptive Clenshaw-Curtis quadrature scheme [Trefethen, 2000].

[40] As shown in Figure 4, the $\psi(t)$ in equation (32) exhibits a very wide change in behavior as a function of η . For the largest value of η , $\psi(t)$ is close to a pure exponential decay (dashed line). For lower values of η there are a number of decades of t where $\psi(t)$ can be closely approximated by an algebraic decay, i.e., a straight line with constant slope. For these values of η this slope slowly decreases as t increases. The range of time corresponding to a slope greater than -3 is the range of anomalous transport. For each value of η there is a crossover with increasing t to normal or Gaussian behavior. However the crossover time is a sensitive function of η .

[41] The behavior of $\psi(t)$ in Figure 4 exhibits all the features necessary to understand the subtle aspects of anomalous transport. The key is the role played by disorder. As the disorder increases the spectrum of transition times increases i.e., $\psi(t)$ decays slower over a larger time range. The nearly constant slope in this range is set equal to $-(1 + \beta)$ and β characterizes this time range but is not an intrinsic parameter of the system as is η . Hence anomalous transport ($\beta < 2$) results from an interplay of the disorder of the system and the duration of the observation. At large enough duration (which is often impractically large for a heterogeneous field site) all transport is normal. The often

discussed “preasymptotic” regime usually covers the measurable range.

[42] The breakthrough curves presented in Figure 5 are the solutions of equation (15) for $\nu = 1$ and $\alpha = 0.05$, and show the effect of the disorder parameter η on the spreading. The smaller value of η corresponds to a longer tailing, whereas for larger values the breakthrough curves tend to a more Gaussian-like shape. The η parameter has a clear counterpart in hydrological applications, which will be discussed elsewhere.

4. Finite Element Discretization of the Equations

[43] Analytical solutions for equation (17) exist only for very particular choices of the memory function and for specific boundary conditions. A numerical treatment for general forms of this equation is therefore needed.

[44] Broadly speaking, any already existing numerical scheme for the solution of the time-independent ADE can be easily modified to find a solution for the Laplace-transformed concentration $\tilde{c}(\mathbf{x}, u)$. The treatment of the divergence of the dispersion tensor in the advective part of equation (20), however, requires particular care. In fact, if the field-scale domain is represented (as is often the case in geological descriptions) as an ensemble of different regions with piecewise constant properties, then the derivative of the dispersion is different from zero only at the discontinuity interfaces and can only be defined in the distributions sense. We write this generalized derivative as

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{D}_\psi(\mathbf{x}) = \Delta \mathbf{D}_\psi(\mathbf{x}_0) \delta(\mathbf{x}_0), \quad (35)$$

where $\Delta \mathbf{D}_\psi(\mathbf{x}_0)$ is the dispersion jump, and $\delta(\mathbf{x}_0)$ is the Dirac delta function at the discontinuity point \mathbf{x}_0 . If we seek a solution of equation (18) in the strong sense, the most obvious way to deal with the infinity involved with the Dirac delta function is to approximate the jump with a linear variation of the dispersion in a small region about the interface.

[45] It is also possible to account for the discontinuities explicitly by seeking a solution to the PDE equation (18) in a weak sense, using a Finite Element Method (FEM) formulation. For a careful discretization of the numerical

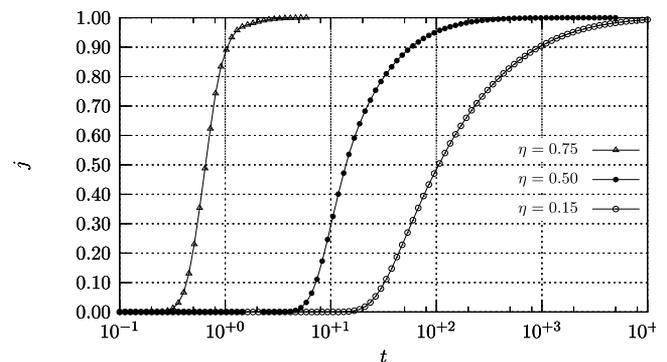


Figure 5. Breakthrough curves relative to the transition probability function $\psi(t)$ in equation (32) versus dimensionless time t for different values of the disorder parameter η , with $\nu_\psi = 1$ and $\alpha_\psi = 0.05$.

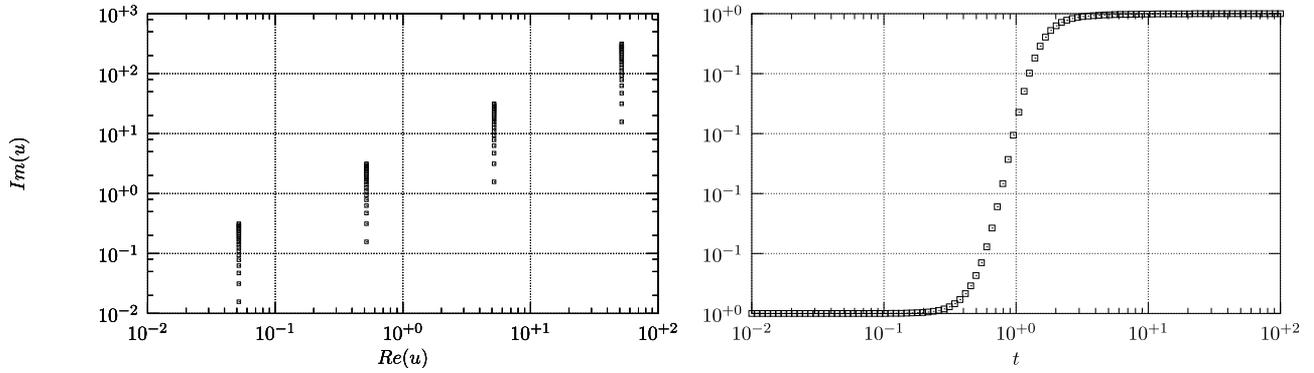


Figure 6. (left) Set of Laplace u parameters needed to obtain (right) the smooth breakthrough curve.

mesh in the interface region, the two methods turn out to give numerically equivalent results. In the present work, we use the linear approximation approach. The concentration $\tilde{c}(\mathbf{x}, u)$ solution of equation (18) is subsequently inverted in the time domain by means of an inverse Laplace transform algorithm.

[46] The inversion for the Laplace transform involves finding the solution $f(t)$ of an integral equation of the first kind [Krylov and Skoblya, 1977]:

$$\int_0^{\infty} f(t)e^{-ut}dt = F(u), \quad (36)$$

where $F(u)$ is a given function of the complex parameter u . In our case, the function to be inverted $\tilde{c}(\mathbf{x}, u)$, is not specified in analytical form, but only through the numerical values obtained from the solution of equation (18). We chose the *de Hoog et al.* [1982] Laplace inversion algorithm. This algorithm makes use of complex valued Laplace parameters. This algorithm works as follows. The time vector for which we want to obtain the concentrations is split into sections of the same order of magnitude, and individual sections are inverted at a given time. Simultaneous inversion for times covering several orders of magnitudes gives inaccurate results for the small times. The Laplace parameter u is expressed as

$$u = \frac{1}{2T}(-\log_{10} \epsilon + i2\pi\mathbf{r}) \quad (37)$$

where T is the maximum of the considered section of the time vector, and $\mathbf{r} = [1, 2, \dots, 2M + 1]$ so that there are $2M + 1$ terms in the Fourier series expansion. The ϵ parameter is typically of the order of 10^{-9} . Thus, to obtain the smooth breakthrough curve (solving equation (15) for

$v_{\psi} = 1$ and $\alpha_{\psi} = 0.05$) in Figure 6 (right), for example, we need to evaluate the FPME at the complex u parameters shown in Figure 6 (left).

5. Results and Discussion

[47] We present a model nonstationary, heterogeneous system as an example of the procedure one can perform with actual field data. The hydrological application here is represented by the distribution in values of porosity, permeability and dispersivity (see Table 1). In this example mode we choose a two-dimensional (2-D) domain Ω , depicted in Figure 7 for a series of numerical computations for transport to illustrate all the new features of our approach with the FPME (i.e., the interplay of local disorder, generalized concentration flux term and macroscopic heterogeneity).

[48] The unit (1×1) domain is subdivided into macroscopically homogeneous regions of three different materials $\Omega_{1,2,3}$ whose porosity, permeability and dispersivity are given in Table 1. We impose constant pressure boundary conditions equal to 1 and 0 on the inlet and outlet boundaries, respectively, and a no-flow over the remaining por-

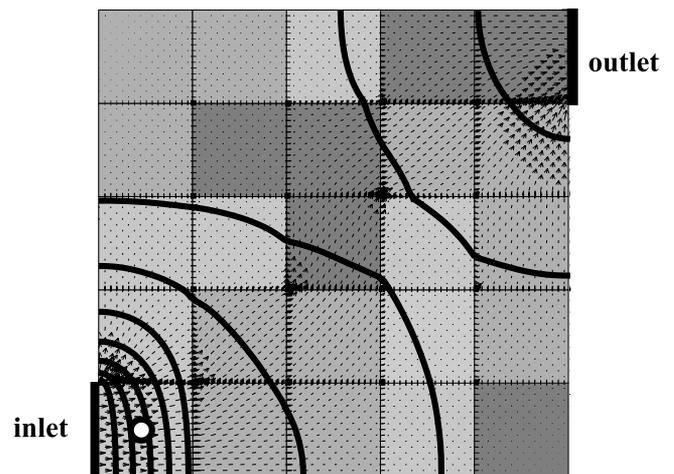


Figure 7. Nonstationary domain used for the computations. The three regions $\Omega_{1,2,3}$ are represented in the light, medium, and dark shading, respectively. The overall region spans 1×1 nondimensional units. The point source is located at the coordinates $(0.1, 0.1)$. The inlet and outlet regions are also indicated.

Table 1. Values of the Permeability K , Longitudinal and Transverse Dispersivity α_{ψ}^l and α_{ψ}^t , and Porosity n for the Three Homogeneous Facies $\Omega_{1,2,3}$

	Ω_1	Ω_2	Ω_3
K_x	0.10	0.50	0.40
K_y	0.10	0.50	0.40
α_{ψ}^l	0.05	0.025	0.10
α_{ψ}^t	0.005	0.0025	0.01
n	0.30	0.35	0.40

Table 2. Values of Coefficients for the Transition Probability Functions $\psi(t)$ in Equation (30) and Equation (28) Coefficients for the Three Homogeneous Facies $\Omega_{1,2,3}$ for the Two Cases (\mathcal{L}) ($\beta > 1$) and (\mathcal{S}) ($\beta < 1$)

	Ω_1	Ω_2	Ω_3
Equation (28), (\mathcal{L})			
a	1.00	1.10	1.20
b	-0.50	-0.55	-0.60
β	1.25	1.35	1.45
Equation (28), (\mathcal{S})			
a	0.10	0.15	0.20
b	1.00	1.50	2.00
β	0.60	0.70	0.80
Equation (30), (\mathcal{L})			
t_1	0.10	0.15	0.20
t_2	1×10^6	5×10^6	10×10^6
β	1.25	1.35	1.45
Equation (30), (\mathcal{S})			
t_1	0.10	0.15	0.20
t_2	1×10^6	5×10^6	10×10^6
β	0.60	0.70	0.80

tions of the boundary. We assume an initial resident concentration $c_0(\mathbf{x}) = 0$, a constant concentration at the injection point $(0.1, 0.1)$, a free boundary condition at the outlet boundary, and a no-flux boundary condition on the remainder of the domain. The approximation of the interface zone between the different macroscopically homogeneous regions is obtained by linearizing the spatial derivative of the dispersivity over a small region about the interface itself.

[49] The first step in the actual computations is the solution of the flow field $\mathbf{q}(\mathbf{x})$ by means of any classical solver. We make the approximation $\mathbf{q}(\mathbf{x}) \sim \mathbf{q}_\psi(\mathbf{x})$. The second step is the solution of equation (20) for all values of the u variable indicated by the de Hoog algorithm: the $\alpha_\psi(\mathbf{x})$ field is assumed to be known from field measurements, as are all the spatially dependent parameters in $\bar{M}(u, \mathbf{x})$. The final step is a numerical Laplace inverse transform of the Laplace transformed solutions for every point of interest in space.

[50] We solve the FPME in equation (20) for two specific forms of $\psi(t)$, namely the asymptotic form (given by equation (28)) and the truncated power law form (given by equation (30)). These functions were chosen because their properties have been investigated extensively [see, e.g., Berkowitz and Scher, 2001; Dentz et al., 2004]. Within the context of these $\psi(t)$ functions, we employ two sets (\mathcal{L} and \mathcal{S}) of β values in the regions $\Omega_{1,2,3}$ (see Table 2). In set \mathcal{L} the β coefficients are larger than one, whereas in set \mathcal{S} they are smaller than one. We recall that $\beta < 1$ is typical of highly dispersive (non-Fickian) transport, while $1 < \beta < 2$ is associated with moderately dispersive systems; $\beta > 2$ leads to Fickian transport.

[51] For each set \mathcal{L} and \mathcal{S} of β coefficients, tracer transport was calculated for the FPME, and compared to tracer behavior as determined from application of the standard ADE model (i.e., by using equation (20) and setting the memory term to unity and the drift correction term $\nabla_{\mathbf{x}} \cdot \mathbf{D}_\psi(\mathbf{x})$ to zero). Note that, as discussed in section 2.4, the definitions of the dispersivity in Table 1 are different with respect to the FPME and the ADE. More specifically, α_ψ , as it is used in equation (20) for the FPME is given by equation (13). In contrast, α as applied

to the classical ADE is the conventional ‘‘dispersivity’’ coefficient.

[52] In the following discussion, we present concentration contour plots of tracer in the 2-D domain, as a function of time, as well as breakthrough curves associated with the tracer arrival at the outlet boundary of the domain. We begin by showing time frames of the isoconcentration lines, in Figures 8 and 9, as the tracer plume migrates through the system. Three time frames (for early, intermediate, and late times) are reported for the sake of comparison among the different cases.

[53] Figure 8 shows the migration of tracer plumes through the domain shown in Figure 7, as given by solution of the FPME for the asymptotic (Figure 8a) and truncated power law (Figure 8b) $\psi(t)$ functions, with $\beta > 1$ (set \mathcal{L} , see Table 2), along with that given by solution of the ADE (Figure 8c). The choice of $\beta > 1$ was made in order to yield timescales for the rate of tracer advance that are of the same order of magnitude as those obtained from solution of the ADE. As discussed above, this range of β values is applicable to systems with moderate dispersivity properties. Deviation of tracer transport behavior from that described by the ADE is apparent at all times. In particular, differences are clearly evident in comparing the early and late time tracer patterns for the asymptotic $\psi(t)$ function (Figure 8a) to that for the ADE solution (Figure 8c). With respect to the contour plots for the truncated power law $\psi(t)$ function (Figure 8b), a retardation of the tracer at all times is evident; this is due to the fact that the exponential cutoff time t_2 is much larger than t_{max} , the maximum time spread from first tracer arrival to last arrival for the considered parameters. Note also that, when compared to the ADE solution in Figure 8c, the isoconcentration contours in Figures 8a and 8b are ‘‘compressed’’ along the interfaces of maximum dispersivity contrast. This is due to the effect of the ‘‘drift correction’’ term. Finally, comparison of the contour patterns for the ADE (Figure 8c) to those for the FPME solutions (Figures 8a and 8b) indicates that, particularly at the intermediate time, the tracer follows a relatively straight path toward the outlet region, which is determined essentially by the underlying flow streamlines.

[54] We consider now the FPME solutions for the asymptotic (Figure 9a) and truncated power law (Figure 9b) $\psi(t)$ functions, with $\beta < 1$ (set \mathcal{S}). Note that relative to Figure 8, the timescales for the overall rate of tracer advance increase by orders of magnitude (and thus the contours for the ADE solution are totally different; see Figure 8c). In fact, although the actual contour patterns are totally different, the early time plots shown in Figures 9a and 9b correspond to those of the late time plots shown in Figures 8a and 8b. This dramatic effect is due to the highly anomalous (non-Fickian) dispersion characteristics of media with values of $\beta < 1$. Comparing Figures 9a and 9b, the simulated migration patterns are more similar than those for the case $\beta > 1$ (Figures 8a and 8b), especially at small and intermediate times. This can be attributed to the highly dispersive and relatively slow overall rate of tracer advance; in this range of β values, and for the cutoff time t_2 used for the truncated power law $\psi(t)$ function, the two $\psi(t)$ functions are similar (recall, e.g., Figure 2). However, for larger times (third time frame), the truncated power law $\psi(t)$ model advances the tracer faster, with respect to the asymptotic

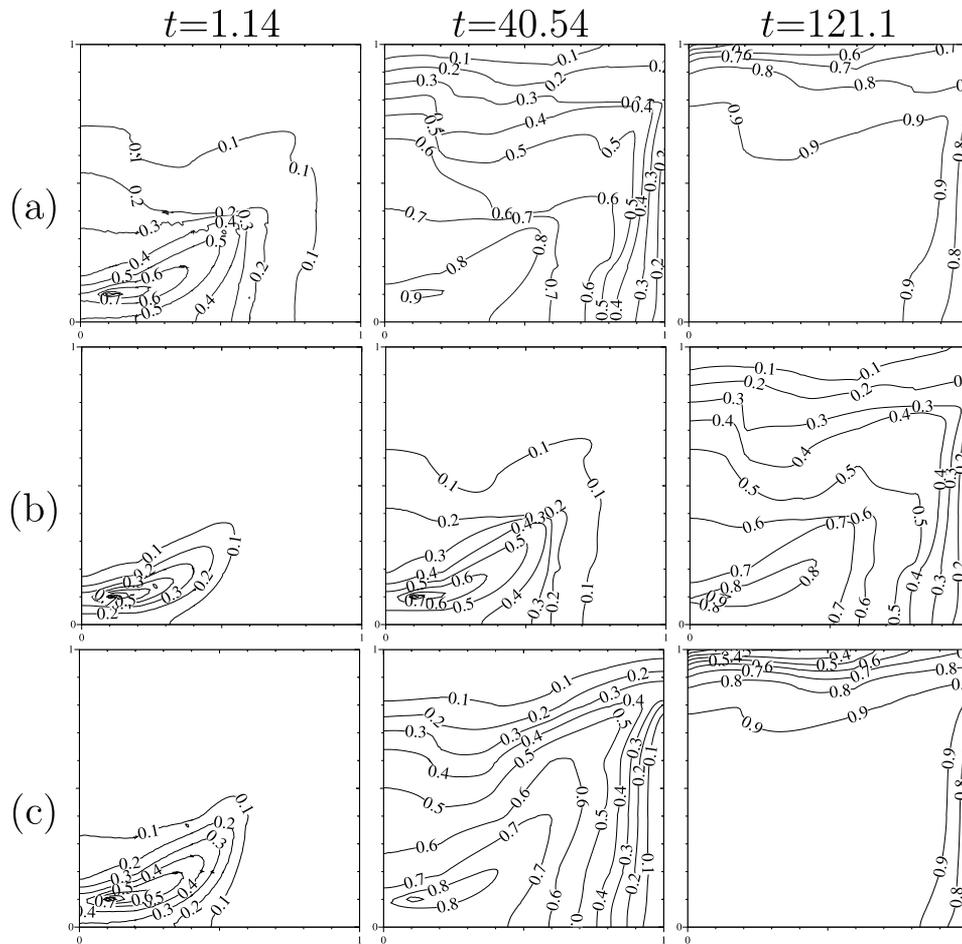


Figure 8. Contour plots of tracer concentration for plume migration through the heterogeneous domain shown in Figure 7. Shown here are solutions of (a) FPME with the asymptotic $\psi(t)$ equation (28) (set \mathcal{L} parameters), (b) FPME with the truncated power law $\psi(t)$ equation (30) (set \mathcal{L} parameters), and (c) ADE.

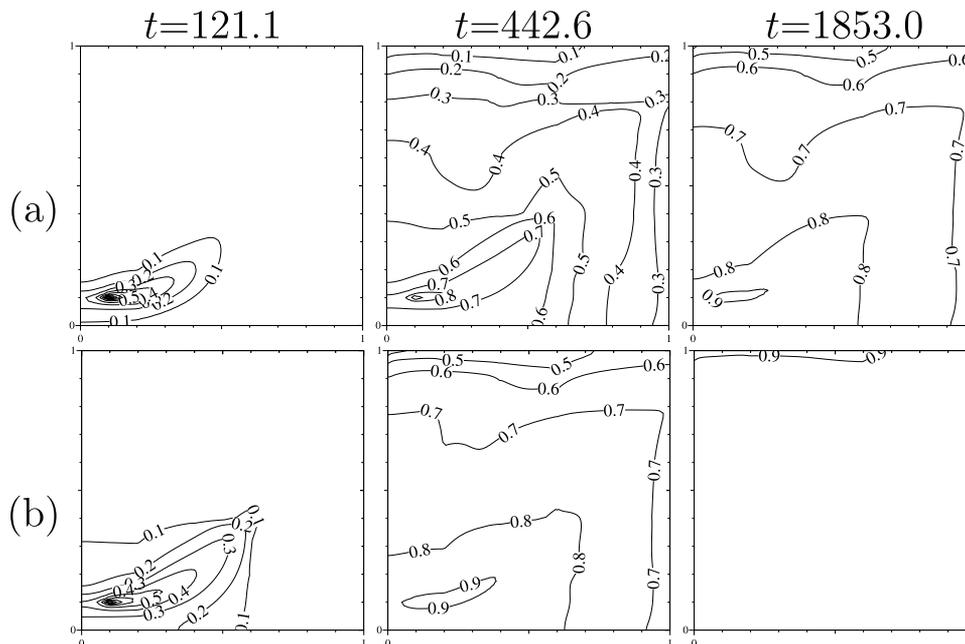


Figure 9. Contour plots of tracer concentration for plume migration through the heterogeneous domain shown in Figure 7. Shown here are solutions of (a) FPME with the asymptotic $\psi(t)$ equation (28) (set \mathcal{S} parameters) and (b) FPME with the truncated power law $\psi(t)$ equation (30) (set \mathcal{S} parameters).

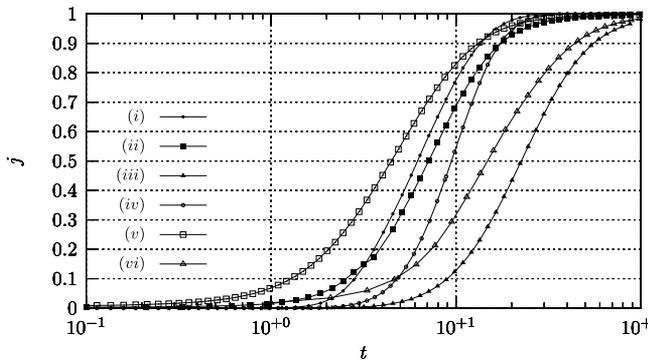


Figure 10. Breakthrough curves for the heterogeneous domain in Figure 7 (set \mathcal{L}). Curves are as follows: i, ADE (Figure 8c); ii, FPME with asymptotic $\psi(t)$ equation (28) (Figure 8a); iii, FPME with truncated power law $\psi(t)$ equation (30) (Figure 8b); iv, ADE with drift correction (Figure 12a); v, FPME with asymptotic $\psi(t)$ without drift correction (Figure 12b); vi, FPME with truncated power law $\psi(t)$ without drift correction (Figure 12c).

model. This is because the truncation time t_2 is of the order of the overall transport and therefore the behavior of the tracer in the truncated power law is already Fickian at these times. In contrast, longer tailing arises when transport is non-Fickian. Finally, as before, it is possible to discern the effects of the dispersivity interfaces on the contour patterns.

[55] We now analyze the different simulated transport behaviors by examining the breakthrough curves (BTCs) associated with the simulations shown in Figures 8 and 9. By integrating the tracer concentration flux over the outlet boundary, the corresponding BTCs are given in Figures 10 and 11, respectively. Other BTCs shown in Figure 10 are derived from simulations shown in Figure 12; these simulations and BTCs will be discussed below.

[56] Considering first Figure 10, curves i and ii, it is clear that the FPME solution with the asymptotic $\psi(t)$ function differs in a subtle yet significant manner from the ADE: the BTC that results from solution of the FPME for the asymptotic $\psi(t)$ function possesses an earlier initial arrival time and a longer, late time tailing than the BTC deriving from solution of the ADE. We have examined this issue in detail elsewhere on the basis of experimental measurements [e.g., Berkowitz *et al.*, 2000; Levy and Berkowitz, 2003]. In contrast, the BTC for the truncated power law $\psi(t)$ function is retarded in comparison to both the asymptotic function and the ADE. Detailed analyses of these behaviors, emphasizing how the choice of β changes the character of the BTCs, are given by Margolin and Berkowitz [2000] and Dentz *et al.* [2004] for the asymptotic and truncated power law $\psi(t)$, respectively, for stationary domains. The BTCs shown in Figure 11 are similar in shape to each other, as noted for the patterns in Figure 9. Note again the difference in timescale between Figures 10 and 11 and the fact that in highly heterogeneous systems (for which it can be expected that $\beta < 1$) the BTC predicted by the ADE differs by about one and one half orders of magnitude.

[57] Previous analysis of laboratory and field data have demonstrated the ability of CTRW-based solutions to accurately capture the early and late arrival times as well as the mean behavior of a plume [e.g., Berkowitz and Scher, 1998;

Kosakowski *et al.*, 2001; Levy and Berkowitz, 2003]. We thus stress again that the ADE solutions provided herein should be considered only as a reference, because of their familiarity. However, we advocate the use of the FPME solution: the FPME as derived in section 2 includes critical terms that affect, both subtly and significantly, tracer migration behavior.

[58] We now examine further the effect of the choice of t_2 in the truncated power law $\psi(t)$ function. From Figures 10 and 11, it is seen that the times for the relative concentration to reach unity are, respectively, of the order of 10^2 and 10^5 . In the simulations, the value of t_2 is of the order of 10^5 – 10^6 (Table 2). Thus, for the simulation based on $\beta < 1$ (Figures 9b and 11, curve ii), the cutoff time t_2 is roughly of the same order of magnitude as the overall breakthrough time. In such cases, transport evolves from non-Fickian to Fickian, as the domain becomes “homogenized” relative to the migrating tracer. In contrast, comparing Figures 8b and 10, curve iii, the time for the relative concentration to reach unity is of the order of 10^2 , while the value of t_2 is of the order of 10^6 . Here the tracer reacts strongly to the heterogeneities in domain (relative to the overall transport length scale). In other words, when the truncation time is much larger than the overall transport time, transport remains highly non-Fickian, and the domain is not homogenized. A detailed analysis of the effects of t_2 , relative to the overall transport time, is given by Dentz *et al.* [2004].

[59] Having seen that the FPME (20) can provide a rich spectrum of transport behaviors, for which the classical ADE is a particular case, we focus our interest now on examining the relative influence, on overall tracer transport, of the memory and drift correction terms within the FPME. Recall that neither of these terms appears in the usual ADE formulations. We therefore solve for transport through the domain considered previously (Figure 7), but with the memory term set to unity in the FPME (which $\Omega_{1,2,3}$ domains are forced to act as purely homogeneous units with Fickian tracer migration). We then solve the FPME with the memory term defined for both (asymptotic and truncated power law) choices of $\psi(t)$, using the parameter set \mathcal{L} , and with the drift correction term set to zero. As before, we include here the simulated tracer transport using the ADE as a familiar reference. Figure 12 presents the resulting sets of contour plots. Note that the timescales are

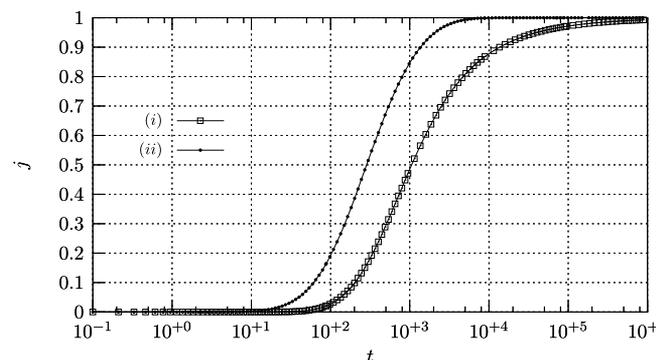


Figure 11. Breakthrough curves for the heterogeneous domain in Figure 7 (set \mathcal{S}). Curves are as follows: i, FPME with asymptotic $\psi(t)$ equation (28) (Figure 9a); ii, FPME with truncated power law $\psi(t)$ equation (30) (Figure 9b).

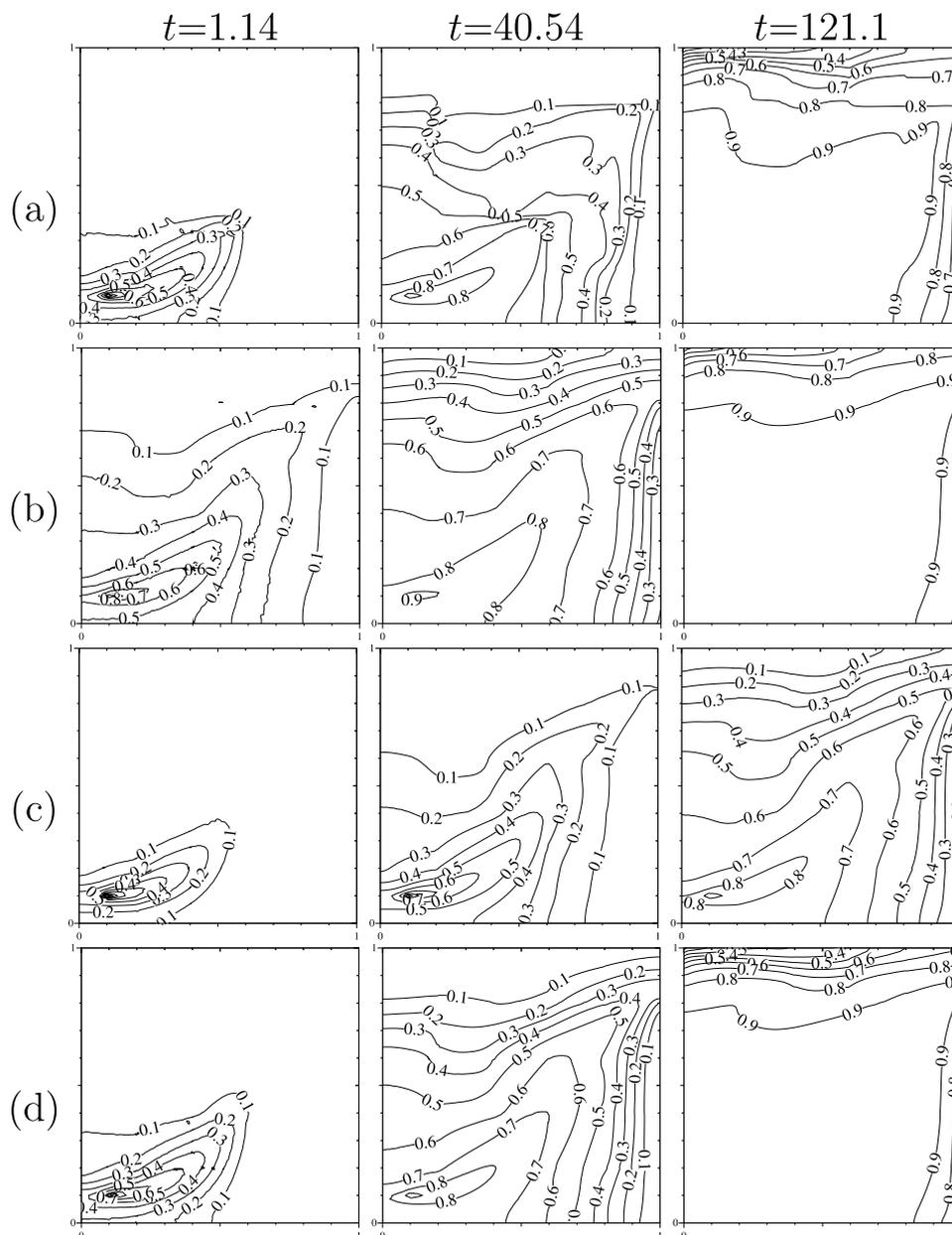


Figure 12. Contour plots of tracer concentration for plume migration through the heterogeneous domain shown in Figure 7. Shown here are solutions of (a) ADE with drift correction (i.e., FPME with $\bar{M} = 1$), (b) FPME with the asymptotic $\psi(t)$ equation (28) (set \mathcal{L} parameters) without drift correction, (c) FPME with the truncated power law $\psi(t)$ equation (30) (set \mathcal{L} parameters) without drift correction, and (d) ADE.

identical (as they should be) to those shown in Figure 8. The corresponding BTCs for these cases are shown in Figure 10, curves iv, v, and vi.

[60] The contours shown in Figure 12a represent the solution to the FPME model, with the memory term set to unity. The effect of the interface contrast in the dispersivities (compare to Figure 7) can be clearly seen, particularly in the intermediate time frame: in the central region of the domain, where the α value increases, the contour lines from 0.4 to 0.7 are compressed and the tracer tends to find an alternative path along directions of minimum dispersivity contrast. We note that similar compression of the intermediate time contour lines, in the central region of the domain, was also

observed in Figure 8a. From Figure 10, curve iv, the overall effect of the drift correction term on the integrated breakthrough curve, relative to the BTC solution of the ADE, is to cause a retardation in the mean arrival time. Long-time tailing effects are not observed here because the local memory terms do not play any role.

[61] In the next two sets of contour plots in Figure 12, the drift correction term is set to zero, and for values of $\beta > 1$, the FPME is solved for the asymptotic (Figure 12b) and truncated power law (Figure 12c) $\psi(t)$ functions. The contour plots in Figures 12b, 12c, and 12d indicate that when the drift correction is disregarded, the tracer chooses to follow the most direct path to the outlet region. In

Figure 12b we notice that a breakthrough of the contour line 0.1 occurs already for early times ($t = 1.14$) (compare to 8a), which indicates that ignoring the drift correction leads to an underestimate of the mean arrival time. The corresponding BTCs appear in Figure 10, curves v and vi. In both cases, comparing to Figure 10, curves ii and iii, it is evident that removal of the drift correction term leads principally to notably shorter mean arrival times.

6. Summary and Conclusions

[62] Using a continuous time random walk framework, we have developed a multiscale Fokker Planck with memory equation (FPME) which accounts for tracer transport subject to both unresolved, small-scale heterogeneities and known, large-scale heterogeneities. The resulting equation has unique features, memory and drift terms, not considered in other formulations such as those mentioned in the Introduction. In particular, the approach described here allows description of the complete spatial and temporal evolution of the tracer plume, and is not limited to analysis of evolution of the moments. The memory term accounts for the unresolved heterogeneities at the small scale whereas the drift correction properly describes the tracer movement across a macroscopic discontinuity in the dispersivity. The classical ADE is seen to be a special case of the FPME valid only for purely homogeneous domains. The memory term is derived from specification of the transition rate probability density function, which stems from a statistical description of the small-scale heterogeneities and their effect on tracer transport. We have considered three forms of this function which are relevant to hydrogeological applications. In the case of stationary disordered domains these forms have accounted for results of both laboratory and field observations.

[63] We have also introduced a computationally efficient numerical solution method to solve the FPME. A key feature is that this solution method is compatible with existing frameworks and software packages; only the “module” that solves the actual tracer transport in the medium need be replaced with the solution methods provided here. We illustrate the influence of the different measurable parameters introduced in the memory function, notably the β parameter controlling the power law like tailing, by solving the FPME over a macroscopically heterogeneous domain and comparing it to the ADE solution for reference. The calculations demonstrate that long tailing arises (principally) from the memory term, while the effects on arrival times are controlled largely by the drift term correction. The effect of a β coefficient smaller than one has particularly dramatic effects on the retardation of a migrating contaminant plume. The effect of the divergence of the dispersivity tensor (drift correction) is to enhance this overall retardation.

Appendix A: Limitations on the Choice of $\tilde{\psi}(u)$

[64] In this appendix we discuss some limitations to be considered for a physically meaningful definition of the probability rate $\psi(t) \equiv \mathcal{L}^{-1}[\tilde{\psi}(u)]$. Not all possible choices for $\tilde{\psi}(u)$ are valid. The probability rate $\psi(t)$ must be positive, normalized and bounded for all times t . When $\psi(t) < 0$ the memory function $\tilde{M}(u)$ (see definition in equation (10)) can become negative for values of u restricted to the real axis and also vanish at some point, which can introduce a

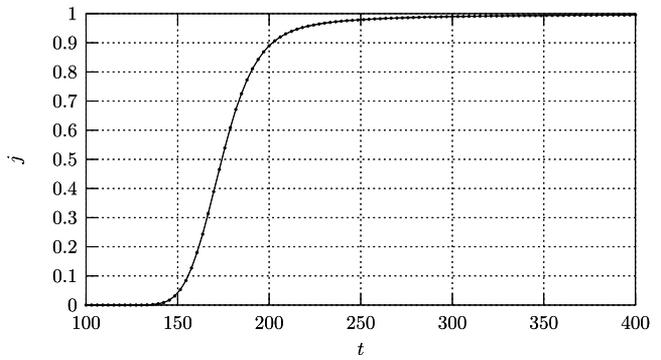


Figure 13. Breakthrough curves for the $\psi(t)$ in equations (A1) (solid line) and (A5) (dots).

singularity in a solution of the FPME [Dentz *et al.*, 2004]. Interestingly, $\tilde{M}(u) = 0$ or $1/\tilde{M}(u) = 0$ do not yield singularities of $\tilde{j}(u)$ in equation (15).

[65] For example, we note that the LT of the probability rate $\tilde{\psi}(u)$ described in equation (28) (repeated here for convenience)

$$\tilde{\psi}(u) = [1 + au + bu^\beta]^{-1}. \quad (\text{A1})$$

does not yield a valid, physically meaningful $\psi(t)$ for some combinations of the parameters a , b , and β . For instance, the combination $\beta = 1.5$ and $a = b = 1$ leads to negative values of $\psi(t)$, for times larger than $t \sim 5.3$; this probability has of course to be rejected. On the other hand, while not strictly well-defined in the usual mathematical sense, a $\psi(t)$ which is negative for times much smaller than the time span of interest does not necessarily represent a problem from the computational point of view. When $1 < \beta < 2$, $a > 0$, and $b < 0$, it is possible to find a characteristic time τ_0 such that $\psi(\tau) < 0$, for $\tau < \tau_0$. We can estimate τ_0 as $1/u_0$, where u_0 is the value that sets $au + bu^\beta = 0$ in equation (A1) to obtain $\tau_0 \equiv (-b/a)^{1/\beta-1}$. For $u \gtrsim u_0$ the denominator of equation (A1) vanishes. Formally in this situation it is not possible to define the inverse Laplace transform of equation (A1) because the singularities of equation (A1) appear in the right half ($\text{Re}(u) > 0$) of the u complex plane. In this context, we consider a situation in which the combination of parameters is $\beta = 1.32$, $a = 4.19 \times 10^{-1}$, $b = -5.39 \times 10^{-2}$, the transport velocity $v_\psi = 5.4991 \times 10^{-3}$ and the dispersivity $\alpha_\psi = 2.1934 \times 10^{-3}$. The BTC obtained from equation (15) corresponding to these values is plotted in Figure 13(solid line). In this case the characteristic time $\tau_0 \sim 10^{-3}$, whereas the time span of interest is $\tau \sim 100-400$. The mathematical issue here can be stated in a simple, nonrigorous way (a rigorous account is given by Zemanian [1968]): in a \mathcal{L}^{-1}

$$f(t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} F(u)e^{ut} du \quad (\text{A2})$$

the asymptotic small u behavior of $F(u)$ can determine the large t behavior of $f(t)$. The well-known Tauberian theorem is a good example of this relationship: if

$$F(u) \sim u^{-\beta} A(1/u) \quad u \rightarrow 0 \quad \beta > 0 \quad (\text{A3})$$

then

$$f(t) \sim t^{\beta-1}A(t)/\Gamma(\beta) \quad t \rightarrow \infty \quad (\text{A4})$$

where $A(t)$ is a slowly varying function of t , e.g., $A \propto \ln t$.

[66] To further demonstrate that $\psi(t) < 0$ for small times (well below the time range of interest) does not cause essential difficulties, we point out that it is possible to make the probability rate $\psi(t)$ positive for all time by adding a second-order correction in the denominator:

$$\tilde{\psi}'(u) = [1 + au + bu^{\beta} + \tau_0 u^2]^{-1}. \quad (\text{A5})$$

To within $\mathcal{O}(u^2)$, equations (A1) and (A5) have the same leading asymptotic small u ($u \leq 10^{-2}$) dependence. In Figure 13 we present the BTC corresponding to the $\psi(u)$ in equation (A5) (dots). The two solutions are essentially identical. Thus, for all practical purposes, the original expression in (38) can be used regardless of its negative values for large u (corresponding to small times). The formal approach [Zemanian, 1968] avoids the large u behavior by using a relation between $f(t)$ and $F(u)$ where $F(u)$ in equation (A2) is analytic in an open strip $\sigma_1 < \text{Re}(u) < \sigma_2$.

[67] We point out that the probability rates in equations (30) and (32) are positive for all values of their respective parameters and are thus not subject to the limitations of equation (A1). It is obviously better to work with a well-posed $\psi(t)$ at the outset. In general one can develop an asymptotic series for $F(u)$ in the $u \rightarrow \infty$ limit by integrating a Taylor expansion of $f(t)$ in equation (36),

$$F(u) \sim \int_0^{\infty} dt e^{-ut} (f(0) + tf_1 + \frac{1}{2}t^2 f_2 + \dots) \sim f(0)/u + f_1/u^2 \dots \quad (\text{A6})$$

where $f_1 \equiv df/dt|_0$ and $f_2 \equiv d^2f/dt^2|_0$. Hence $\tilde{\psi}(u) u, \infty \rightarrow \psi(0)/u$ and therefore $\tilde{M}(u) \xrightarrow{u, \infty} \psi(0)$. One can avoid difficulties if $\psi(0) > 0$ and $\psi(t)$ is well behaved for small t . It should be pointed that this is not the case for the FDE equation (21) with $\beta < 1$. The equivalent to the asymptotic limit of small u defining the operators in the FDE corresponds to a $\tilde{\psi}(u) = 1/(1 + u^{\beta})$, which does not have a finite $\psi(0)$, i.e., $\psi(t)$ is unbounded as $t \rightarrow 0$. The solution of the FDE is thus physically meaningful only for large t .

[68] **Acknowledgments.** The financial support of BP International, the Israel Science Foundation (grant 147/01), the European Commission-Marie Curie Individual Fellowship programme, and the Sardinian Regional Authorities is gratefully acknowledged.

References

- Abramowitz, M., and I. Stegun (1970), *Handbook of Mathematical Functions*, Dover, Mineola, N. Y.
- Adams, E. E., and L. W. Gelhar (1992), Field study of dispersion in a heterogeneous aquifer: 2. Spatial moment analysis, *Water Resour. Res.*, **28**, 3293–3308.
- Bear, J. (1972), *Dynamics of Fluids in Porous Media*, Elsevier Sci., New York.
- Bear, J., and Y. Bachmat (1991), *Introduction to Modeling of Transport Phenomena in Porous Media*, Kluwer Acad., Norwell, Mass.
- Berkowitz, B., and H. Scher (1995), On characterization of anomalous dispersion in porous and fractured media, *Water Resour. Res.*, **31**, 1461–1466.
- Berkowitz, B., and H. Scher (1998), Theory of anomalous chemical transport in random fracture networks, *Phys. Rev. E*, **57**, 5858–5869.
- Berkowitz, B., and H. Scher (2001), The role of probabilistic approaches to transport theory in heterogeneous media, *Transp. Porous Media*, **42**, 241–263.
- Berkowitz, B., H. Scher, and S. E. Silliman (2000), Anomalous transport in laboratory-scale, heterogeneous porous media, *Water Resour. Res.*, **36**, 149–158.
- Berkowitz, B., J. Klafter, R. Metzler, and H. Scher (2002), Physical pictures of transport in heterogeneous media: Advection-dispersion, random walk and fractional derivative formulations, *Water Resour. Res.*, **38**, 1191, doi:10.1029/2001WR001030.
- Boggs, J. M., S. C. Young, and L. C. Beard (1992), Field study of dispersion in a heterogeneous aquifer: 1. Overview and site description, *Water Resour. Res.*, **28**, 3281–3291.
- Burr, D. T., and E. A. Sudicky (1994), Nonreactive and reactive solute transport in three-dimensional heterogeneous porous media: Mean displacement, plume spreading, and uncertainty, *Water Resour. Res.*, **30**, 791–815.
- Carrera, J., X. Sánchez-Vila, I. Benet, A. Medina, G. Galarza, and J. Guimerá (1998), On matrix diffusion: Formulations, solution methods, and qualitative effects, *Hydrogeol. J.*, **6**, 178–190.
- Cunningham, J. A., C. J. Werth, M. Reinhard, and P. V. Roberts (1997), Effects of grain-scale mass transfer on the transport of volatile organics through sediments: 1. Model development, *Water Resour. Res.*, **33**, 2713–2726.
- Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer-Verlag, New York.
- Dagan, G. (1994), The significance of heterogeneity of evolving scales to transport in porous formations, *Water Resour. Res.*, **30**, 3327–3336.
- Dagan, G., and S. P. Neuman (Eds.) (1997), *Subsurface Flow and Transport: A Stochastic Approach*, Cambridge Univ. Press, New York.
- de Hoog, F. R., J. H. Knight, and A. N. Stokes (1982), An improved method for numerical inversion of Laplace transforms, *SIAM J. Sci. Stat. Comput.*, **3**, 357–366.
- 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 (2002), Temporal behavior of a solute cloud in a heterogeneous porous medium: 3. Numerical simulations, *Water Resour. Res.*, **38**(7), 1118, doi:10.1029/2001WR000436.
- Dentz, M., A. Cortis, H. Scher, and B. Berkowitz (2004), Time behavior of solute transport in heterogeneous media: Transition from anomalous to normal transport Adv, *Water Resour.*, **27**(2), 155–173, doi:10.1016/j.advwatres.2003.11.002.
- Eggleston, J., and S. Rojstaczer (1998), Identification of large-scale hydraulic conductivity trends and the influence of trends on contaminant transport, *Water Resour. Res.*, **34**, 2155–2168.
- Feehley, C. E., C. Zheng, and F. J. Molz (2000), A dual-domain mass transfer approach for modeling solute transport in heterogeneous aquifers: Application to the Macrodispersion Experiment (MADE) site, *Water Resour. Res.*, **36**, 2501–2515.
- Gelhar, L. W. (1993), *Stochastic Subsurface Hydrology*, Prentice-Hall, Old Tappan, N. J.
- Gelhar, L. W., C. Welty, and K. R. Rehfeldt (1992), A critical review of data on field-scale dispersion in aquifers, *Water Resour. Res.*, **28**, 1955–1974.
- 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**, 2383–2400.
- Haggerty, R., and S. M. Gorelick (1998), Modeling mass transfer processes in soil columns with pore-scale heterogeneity, *Soil Sci. Soc. Am. Proc.*, **62**, 62–74.
- Kenkre, V. M., E. W. Montroll, and M. F. Shlesinger (1973), Generalized master equations for continuous-time random walks, *J. Stat. Phys.*, **9**, 45–50.
- Klafter, J., and R. Silbey (1980), Derivation of continuous-time random walk equations, *Phys. Rev. Lett.*, **44**, 55–58.
- Koltermann, C. E., and S. M. Gorelick (1996), Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating and descriptive approaches, *Water Resour. Res.*, **32**, 2617–2658.
- Kosakowski, G., B. Berkowitz, and H. Scher (2001), Analysis of field observations of tracer transport in a fractured till, *J. Contam. Hydrol.*, **47**, 29–51.
- Krylov, V. I., and N. S. Skoblya (1977), *A Handbook of Methods of Approximate Fourier Transformation and Inversion of the Laplace Transform*, Mir, Moscow.

- Labolle, E. M., and G. E. Fogg (2001), Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, *Transp. Porous Media*, 42, 155–179.
- Labolle, E. M., G. E. Fogg, and F. B. Tompson (1996), Random-walk simulation of transport in heterogeneous porous media: Local mass-conservation problem and implementation methods, *Water Resour. Res.*, 32, 583–593.
- Levy, M., and B. Berkowitz (2003), Measurement and analysis of non-Fickian dispersion in heterogeneous porous media, *J. Contam. Hydrol.*, 64, 203–226.
- Margolin, G., and B. Berkowitz (2000), Application of continuous time random walks to transport in porous media, *J. Phys. Chem. B*, 104, 3942–3947. (Correction, *J. Phys. Chem. B*, 104, 8762, 2000).
- McLaughlin, D., and F. Ruan (2001), Macrodispersivity and large-scale hydrogeologic variability, *Transp. Porous Media*, 42, 133–154.
- Metzler, R., and J. Klafter (2000), The random walk's guide to anomalous diffusion: A fractional dynamics approach, *Phys. Rep.*, 339, 1–77.
- Naff, R., D. F. Haley, and E. A. Sudicky (1998), High-resolution Monte Carlo simulation of flow and conservative transport in heterogeneous porous media: 2. Transport results, *Water Resour. Res.*, 34, 679–697.
- Pannone, M., and P. K. Kitanidis (2001), Large-time spatial covariance of concentration of conservative solute and application to the Cape Cod tracer test, *Transp. Porous Media*, 42, 109–132.
- Risken, H. (1989), *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer Ser. Synergetics, vol. 18, 2nd ed., Springer-Verlag, New York.
- Roth, K., and W. A. Jury (1993), Linear transport models for adsorbing solutes, *Water Resour. Res.*, 29, 1195–1203.
- Salandin, P., and V. Fiorotto (1998), Solute transport in highly heterogeneous aquifers, *Water Resour. Res.*, 34, 949–961.
- Scher, H., and M. Lax (1973), Stochastic transport in a disordered solid. II. Impurity conduction, *Phys. Rev. B*, 7, 4502–4519.
- Scher, H., G. Margolin, and B. Berkowitz (2002), Towards a unified framework for anomalous transport in heterogeneous media, *Chem. Phys.*, 284, 349–359.
- Shlesinger, M. F. (1974), Asymptotic solutions of continuous-time random walks, *J. Stat. Phys.*, 10, 421–434.
- Sidle, C., B. Nilson, M. Hansen, and J. Frederica (1998), Spatially varying hydraulic and solute transport characteristics of a fractured till determined by field tracer tests, Funen, Denmark, *Water Resour. Res.*, 34, 2515–2527.
- Silliman, S. E., and E. S. Simpson (1987), Laboratory evidence of the scale effect in dispersion of solutes in porous media, *Water Resour. Res.*, 23, 1667–1673.
- Trefethen, L. N. (2000), *Spectral Methods in Matlab*, Soc. for Ind. and Appl. Math., Philadelphia, Pa.
- Zemanian, A. H. (1968), *Generalized Integral Transformations*, Wiley-Interscience, Hoboken, N. J.
-
- B. Berkowitz, A. Cortis, and H. Scher, Department of Environmental Sciences and Energy Research, Weizmann Institute of Science, Rehovot 76100, Israel. (brian.berkowitz@weizmann.ac.il; andrea.cortis@weizmann.ac.il; harvey.scher@weizmann.ac.il)
- C. Gallo, CRS4, Parco Scientifico e Tecnologico, Località Pixina Manna Edificio 1, C.P. 25, I-09010 Pula (CA), Italy. (fender@crs4.it)