AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2015WR017388

Special Section:
The 50th Anniversary of Water Resources Research

Key Points:
- Recent advances in Lagrangian stochastic methods for groundwater transport
- Systematic presentation of results from theory to applications
- Unified framework in stochastic hydrogeology and health risk analysis

Correspondence to:
A. Fiori,
aldo@uniroma3.it

Citation:
Fiori, A., A. Bellin, V. Cvetkovic, F. P. J. de Barros, and G. Dagan (2015), Stochastic modeling of solute transport in aquifers: From heterogeneity characterization to risk analysis, *Water Resour. Res.*, *51*, 6622–6648, doi:10.1002/2015WR017388.

Received 14 APR 2015
Accepted 26 JUL 2015
Accepted article online 30 JUL 2015
Published online 24 AUG 2015

Stochastic modeling of solute transport in aquifers: From heterogeneity characterization to risk analysis

A. Fiori[1], A. Bellin[2], V. Cvetkovic[3], F. P. J. de Barros[4], and G. Dagan[5]

[1]Dipartimento di Ingegneria, Universita' di Roma Tre, Rome, Italy, [2]Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy, [3]Department of Water Resources Engineering, Royal Institute of Technology, Stockholm, Sweden, [4]Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, California, USA, [5]School of Mechanical Engineering, Tel Aviv University, Ramat Aviv, Israel

Abstract The article presents a few recent developments advanced by the authors in a few key areas of stochastic modeling of solute transport in heterogeneous aquifers. First, a brief review of the Lagrangian approach to modeling plumes longitudinal mass distribution and temporal (the breakthrough curve) mass arrival, is presented. Subsequently, transport in highly heterogeneous aquifers is analyzed by using a recently developed predictive model. It relates the non-Gaussian BTC to the permeability univariate pdf and integral scale, with application to the MADE field observations. Next, the approach is extended to transport of reactive solute, combinnig the effects of the random velocity field and multirate mass transfer on the BTC, with application to mass attenuation. The following topic is modeling of the local concentration field as affected by mixing and dilution due to pore scale dispersion. The results are applied to the analysis of concentration measurements at the Cape Cod field experiment. The last section incorporates the results of the preceding ones in health risk assessment by analyzing the impact of concentration prediction on risk uncertainty. It is illustrated by assessing the effect of identification of macrodispersivity from field characterization and transport modeling, upon the probability of health risk.

1. Introduction

In 1986, WRR has published a special issue (Vol.22, 9S) on "Trends and Directions in Hydrology." Two articles were devoted to the emerging field of stochastic modeling of groundwater flow and transport, covering its first decade [*Dagan*, 1986; *Gelhar*, 1986]. To underscore the pace of development of the subject, a graph with number of articles published by WRR, as function of year was presented [*Dagan*, 1986, Figure 3]: altogether, around 125 papers appeared in the period 1975–1984. In the following 30 years, leading to present days, the field has undergone a tremendous expansion, as witnessed for instance by the publication of 3 monographs [*Dagan*, 1989; *Gelhar*, 1993; *Rubin*, 2003]. It will be difficult to count the number of papers addressing the stochastic approach published by different journals in this period, the leading role of WRR notwithstanding. Nevertheless, to illustrate the point, we shall refer to the citations of the monograph by *Dagan* [1989], which is indicative of the extent of the literature on the topic. Thus, according to the ISI Web of Science, the book was cited around 1300 times in the period 1990–2014, out of which 830 citations were of articles in journals and proceedings categorized as belonging to Water Resources. Among these around 410 papers were published in WRR, and this is of course only a partial count. The expansion of the field was not only quantitative, but also reflected by the diversity of the covered topics.

The present article does not attempt to review this vast opus, but rather concentrates on a few central topics addressed by publications of the authors, focusing on recent developments. Still, these recent works have in common a few principles set forth 30 years ago [e.g., *Dagan*, 1982a, 1984]. Limiting the discussion to 3-D transport, they can be enumerated as follows: spreading of solute plumes is caused by spatial heterogeneity of aquifer properties, primarily of the hydraulic conductivity K; the latter is modeled as space random function characterized by its statistical moments and in particular the univariate PDF and the two point covariance, which have to be identified by field characterization; the Eulerian fluid velocity $\mathbf{V}$ field is also random and has to be derived by solving the equations of flow, in terms of K; transport of a solute plume is solved by the Lagrangian approach (particle tracking in its numerical version) by using $\mathbf{V}$; solute plumes spreading is quantified in terms of spatial or temporal distributions, which are dominated by

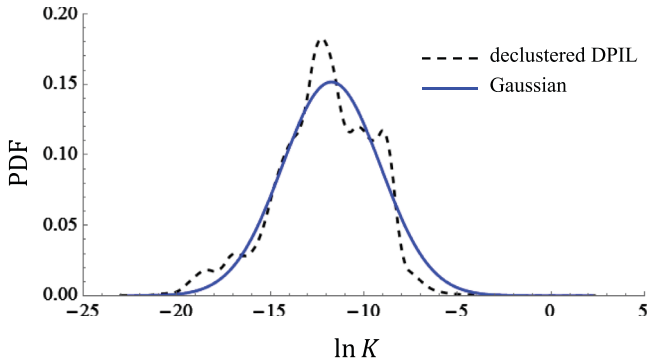

Figure 1. Gaussian distribution (blue) fitted on the declustered DPIL samples of ln K at the MADE site (data from *Bohling et al*. [2012]) and compared to the empirical PDF (adapted from *Fiori et al.*, 2015).

advection; in contrast, the local concentration C is also influenced to a large extent by local pore scale dispersion, which causes dilution; the solution of transport is used for analyzing field experiments as well as for prediction and risk analysis; the solution is based on the underlying characterization of K, as well as on flow and transport initial and boundary conditions.

Along these lines, and focussing on 3-D transport solely, the plan of the article is as follows. Section 2 provides the background by starting with a few field findings, result of K characterization; subsequently it renders the transport equations and the solute concentration distribution, with illustration by a few simple results obtained in the past for spreading in weakly heterogeneous aquifers. Section 3 is concerned with recent developments on transport in highly heterogeneous formations, including the analysis of the *MADE* field experiment. Section 4 deals with reactive transport, concentrating on quantitative characterization of various types of reaction and their incorporation in the pertinent equations. Section 5 addresses the statistical characterization of local concentration and particularly dilution, solutions of transport equations. Section 6 deals with health risk analysis due to aquifer contamination and builds to a large extent on the previous sections. The concluding section 7 summarizes the article.

2. Background

The present section is devoted primarily to the recapitulation of the basic transport equations derived by the Lagrangian approach. The section sets definitions and notations on one hand and serves as starting point for the developments of the following sections, on the other hand. Most of the material can be traced back in different publications, e.g., *Dagan* [1989], and references will be called only in specific cases.

2.1. Aquifer Random Structures

The statistical characterization of the logconductivity spatial distribution $Y(\mathbf{x})$, regarded as a stationary random space function, was the subject of many field investigations. *Freeze* [1975] pointed out that field data indicate that the univariate Y PDF is approximately normal and characterized by $\langle Y\rangle = \ln K_G$ and the variance σ_Y^2, where $\langle\cdot\rangle$ stands for the ensemble mean and K_G is the geometric mean. Tables summarizing values of these parameters for a few aquifers can be found for instance in *Rubin* [2003]. The value of σ_Y^2 is indicative of the level of aquifer heterogeneity and it plays an important role in modeling (see sections 2 and 3).

The Y structure is completely characterized by the joint PDF of its values at any set of points and the associated moments. The simplest is the two point covariance $C_Y = \sigma_Y^2 \rho_Y$, where the autocorrelation ρ_Y depends on the interval (lag) vector $\mathbf{r}$ for stationary random Y. In turn, for an axisymmetric ρ_Y, the horizontal and vertical integral scales I and I_v, respectively, are the important global parameters. Various field identified values, of the order of *meters*, can be found for instance in *Rubin* [2003].

The usual amount of field measurements of K, even for detailed surveys carried out for transport experiments, is not large enough to allow for identification of higher statistical moments. Since numerical simulations and theoretical developments require the complete statistical characterization of Y, it is common to adopt the multi-Gaussian model, i.e. the Y values at any set of points is a multivariate normal vector, such that the knowledge of K_G and C_Y is sufficient to characterize it. Other structural models, like the bimodal one or blocks (section 3) were also considered. Altogether, characterization by field measurements and their analysis is an important and expanding field of research.

To illustrate the univariate field determined Y PDF, we present in Figure 1 the histogram for the *MADE* experimental site (see section 3) which is highly heterogeneous ($\sigma_Y^2 \cong 6.9$). The data were obtained from a large number of measurements obtained with the aid of the recently developed *DPIL* (direct push injection logger)

technology [*Bohling et al.*, 2012]. The fit by a lognormal K distribution seems to be quite accurate, except deviations at the extreme values, which might have been caused by technical problems. A recent study (Fiori et al., submitted manuscript, 2015) determined that their impact on the solute plume prediction is negligible.

In the following we assume Y stationarity and adopt the univariate lognormal model in order to focus the discussion.

2.2. The Eulerian Velocity Field

The first step in modeling transport according to the general approach outlined in the *Introduction* is the derivation of the Eulerian velocity field $\mathbf{V}(\mathbf{x})$ (for simplicity steady flow is assumed). This is achieved by solving the flow equations (Darcy's law and mass conservation)

$$\mathbf{q}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x}) \,,\ \nabla \cdot \mathbf{q}(\mathbf{x}) = 0 \tag{1}$$

where h is the pressure head and $\mathbf{q}$ is the specific discharge vector. The solution of (1) for h is sought in a domain $\mathbf{x} \in \Omega$, for appropriate conditions on the boundary $\partial\Omega$. Subsequently, the Eulerian velocity field is given by $\mathbf{V} = \mathbf{q}/\theta$, where θ is the effective porosity, assumed to be constant.

Due to the randomness of K, equation (1) is stochastic, and the solution $\mathbf{V}(\mathbf{x})$ is random and characterized in terms of various statistical moments. We shall concentrate here on natural gradient flows which can be approximated by a constant mean head gradient $\langle \nabla h \rangle = -\mathbf{J}$ and the corresponding uniform mean velocity $\mathbf{U} = \langle \mathbf{V} \rangle = K_{ef}\mathbf{J}/\theta$. The effective conductivity K_{ef} depends on the K structure and its derivation is the topic of a vast literature. In contrast, a notable example of strongly nonuniform flow, of considerable interest in applications, is well flow. The derivation of the nonstationary $\mathbf{V}(\mathbf{x})$ in this case is a difficult task and the topic is one of ongoing research.

We assume that the random and stationary $\mathbf{V}(\mathbf{x}) = \mathbf{U} + \mathbf{u}(\mathbf{x})$, where $\mathbf{u}$ is the stationary fluctuation, was determined, as a prerequisite to solving transport.

2.3. Formulation of Transport Equations (Lagrangian Approach)

The local concentration $C(\mathbf{x}, t)$ (defined at the Darcy scale) of an inert solute satisfies the Eulerian balance equation

$$\frac{\partial C(\mathbf{x}, t)}{\partial t} + \mathbf{V}(\mathbf{x}) \cdot \nabla C(\mathbf{x}, t) = \nabla.(\mathbf{D}_d C(\mathbf{x}, t)) \tag{2}$$

where $\mathbf{D}_d$ is the tensor of pore scale dispersion coefficients. We refer to (2) as to the *local* Advection-Dispersion Equation (ADE), to be solved under initial condition of injection in a volume Ω_0.

However, we follow the Lagrangian approach (also known as particle tracking in the numerical context) by which the concentration satisfies [see e.g., *Dagan*, 1989]

$$dC(\mathbf{x}, t) = dM_0 \delta[\mathbf{x} - \mathbf{X}_t(t, \mathbf{b})] \tag{3}$$

where dM_0 is the mass of a solute particle injected at $\mathbf{b} \in \Omega_0$ and $\mathbf{x} = \mathbf{X}_t(t, \mathbf{b})$ is the equation of its trajectory. Equation (3) stipulates that the particle moves along the trajectory, while conserving its mass. We consider instantaneous injection at $t = 0$ and initial $C = C_0(\mathbf{b})$, such that $dM_0 = C_0 d\mathbf{b}$, while M_0 is the total mass; it is easy to generalize (3) for continuous injection.

Without loss of generality injection takes place on an area A_0 in the injection plane $x = 0$, normal to the mean velocity $\mathbf{U}(U, 0, 0)$, such that $dM_0 = m_0(\mathbf{b})d\mathbf{b}$, where $\mathbf{b}$ is a 2-D coordinate in A_0 and m_0 is mass/area. Thus, the solution of (3) can be written as follows

$$C(\mathbf{x}, t) = \int_{A_0} m_0(\mathbf{b}) \delta[\mathbf{x} - \mathbf{X}_t(t, \mathbf{b})] d\mathbf{b} \tag{4}$$

The extension to an injection volume Ω_0 can be achieved by considering a plume of infinitesimal thickness db_x such that $m_0 = C_0 db_x$ and subsequent additional integration over b_x in (4).

We consider two modes of injection [e.g., *Kreft and Zuber*, 1978]: in the resident one m_0 is deterministic and given, the simplest case considered here being $m_0 = M_0/A_0 = \text{const}$; in the flux proportional mode m_0 is random and proportional to the local velocity i.e. $m_0 = (M_0/A_0)\ [V_0(\mathbf{b})/\bar{V}_0]$ where $V_0 = V_x(0, \mathbf{b})$. Here and in the sequel we assume that the area A_0 is sufficiently large relative to the integral scale to warrant exchange of

space and ensemble averaging (ergodic plume) such that $\bar{V}_0=(1/A_0)\int_{A_0}V_0(\mathbf{b})d\mathbf{b}\cong U$. The impact of the injection mode is felt up to a distance x for which V_x and $\mathbf{X}_t$ are correlated [see e.g., *Janković and Fiori*, 2010]. The issue of nonergodic plumes and the related uncertainty is a topic of great interest which is not addressed in the present article.

The trajectory equation can be written as follows

$$\frac{d\mathbf{X}_t}{dt}=\mathbf{V}_L(t,\mathbf{b})=\frac{d\mathbf{X}}{dt}+\mathbf{w}_d,\quad \text{with}\quad \frac{d\mathbf{X}}{dt}=\mathbf{V}(\mathbf{X}_t)\quad \text{and}\quad \mathbf{X}_t(0)=\mathbf{b} \tag{5}$$

where $\mathbf{V}_L$ is the Lagrangian velocity, and $\mathbf{w}_d$ is a velocity of a Brownian motion associated with $\mathbf{D}_d$ (2). In the discrete, particle tracking mode, (5) is usually written as $\Delta\mathbf{X}_t=\Delta\mathbf{X}+\Delta\mathbf{X}_d$, with $\Delta\mathbf{X}=\mathbf{V}(\mathbf{X}_t)\Delta t$ whereas $\Delta\mathbf{X}_d$ is a normal vector of zero mean and of variance $\sigma^2_{\Delta X_{d,i}}=2D_{d,i}\Delta t$.

The impact of the pore scale dispersion on the local concentration C is considered in section 5. The discussion here is focused on plume spreading as quantified by the various spatial or temporal moments for which the impact of pore scale dispersion is usually small. Thus $\mathbf{X}_t=\mathbf{X}$ in (5) and the advective trajectory equation becomes

$$\mathbf{X}(t,\mathbf{b})=\mathbf{b}+\mathbf{U}t+\mathbf{X}',\qquad \mathbf{X}'=\int_0^t\mathbf{u}[\mathbf{X}(t')]dt' \tag{6}$$

with $\mathbf{X}'(t,\mathbf{b})$ the random trajectory fluctuation.

Thus, for injection in a finite volume Ω_0 the concentration field is given by

$$C(\mathbf{x},t)=\int_{\Omega_0}C_0(\mathbf{b})\ \delta[\mathbf{x}-\mathbf{X}(t,\mathbf{b})]d\mathbf{b} \tag{7}$$

For $C_0=const$ the integral in (7) becomes an indicator of value C_0 in the volume $\Omega_0(t)$ determined by the fluid trajectories originating from Ω_0, and zero outside. For the assumed incompressible flow and constant θ, the total volume is conserved and so is the mass. Thus, C has a bimodal distribution $C=C_0$ for $\mathbf{x}\in\Omega_0$ and $C=0$ outside; this property will be explored in section 5.

Focusing on the resident injection mode and planar injection, the first spatial moment (i.e. the plume centroid equation), is evaluated using (4) and given by

$$\mathbf{R}(t)=\frac{1}{M_0}\int_{A_0}\mathbf{x}C(\mathbf{x},t)d\mathbf{x}=\frac{1}{A_0}\int_{A_0}\mathbf{X}(t,\mathbf{b})d\mathbf{b}\cong\langle\mathbf{X}\rangle=\mathbf{U}t \tag{8}$$

In a similar manner (4) leads to the second central moments as follows

$$S_{ij}(t)=\frac{1}{M_0}\int_{A_0}(x_i-R_i)(x_j-R_j)C(\mathbf{x},t)d\mathbf{x}\cong X_{ij}(t)\equiv\langle X_i'(t,\mathbf{b})X_j'(t,\mathbf{b})\rangle\quad (i,j=1,2,3) \tag{9}$$

An important particular case is of X being characterized by a Gaussian PDF $f(\mathbf{X},t)$, of mean $\mathbf{U}t$ and covariance matrix X_{ij}. Then, by (4) the mean concentration $\langle C(\mathbf{x},t)\rangle=M_0f(\mathbf{x},t)$ satisfies

$$\frac{\partial\langle C(\mathbf{x},t)\rangle}{\partial t}+U\frac{\partial\langle C(\mathbf{x},t)\rangle}{\partial x_1}=\sum_i\sum_j D_{ij}\frac{\partial^2\langle C(\mathbf{x},t)\rangle}{\partial x_i\partial x_j}\quad\text{with}\quad D_{ij}(t)=\frac{1}{2}\frac{dX_{ij}(t)}{dt} \tag{10}$$

Equation (10), though similar to (2), is an *upscaled* ADE and $D_{ij}(t)$, the macrodispersion coefficients, are larger by orders of magnitude than the pore scale $\mathbf{D}_m$. Furthermore, if D_{ij} become constant, transport is coined as Fickian. Determining the conditions under which (10) is obeyed and then relating D_{ij} to the permeability structure, are central topics of stochastic modeling.

A related transport measure is the relative mass arrival at a control plane normal to the mean flow at $x_1=x$

$$M(x,t)=\frac{1}{M_0}\int_x^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty C(\mathbf{x}',t)d\mathbf{x}' \tag{11}$$

which can be rewritten by using (3) and (4) as

$$M(x,t)=\frac{1}{M_0}\int_{A_0} m_0(\mathbf{b})\ H[x-X_1(t,\mathbf{b})]d\mathbf{b} \tag{12}$$

where H is the Heaviside function.

We further narrow the focus by concentrating on the longitudinal spreading since the transverse was found to be of much smaller extent. Defining by $m(x,t)=(1/M_0)\int\int C(x,x_2,x_3,t)\,dx_2dx_3=-\partial M/\partial x$, the longitudinal relative mass distribution, we find from (10) that in the Gaussian case and an ergodic plume for which $m\cong\langle m\rangle$, it satisfies the ADE

$$\frac{\partial m(x,t)}{\partial t}+U\frac{\partial m(x,t)}{\partial x}=D_L\frac{\partial^2 m(x,t)}{\partial x^2}\ \text{ with }\ D_L=\frac{1}{2}\frac{dX_{11}(t)}{dt} \tag{13}$$

where $\alpha_L=D_L/U$ is known as longitudinal macrodispersivity. It is emphasized that pore scale dispersion has generally a small effect upon M or m.

A different and useful related methodology to quantify longitudinal transport is by the breakthrough curve (the BTC) defined as the relative mass discharge at the control plane $\mu(x,t)=\partial M/\partial t$ (dimension $1/T$), such that M is the cumulative BTC.

The BTC can be conveniently related to the Lagrangian kinematics by defining the travel time $\tau(x,\mathbf{b})$ of a fluid particle from the injection plane to the control plane at x by the equation $x-X(\tau,\mathbf{b})=0$. However, the random τ can be derived directly from the Eulerian velocity field by integration along streamlines $y=\eta(x,\mathbf{b})$, $z=\zeta(x,\mathbf{b})$ and using the relationship $d\tau=dx/V_x(x,\eta,\zeta)$ [e.g., *Cvetkovic and Dagan*, 1994].

The mass arrival is related to the travel time by

$$M(x,t)=\frac{1}{M_0}\int_{A_0} m_0(\mathbf{b})H[t-\tau(x,\mathbf{b})]d\mathbf{b} \tag{14}$$

stipulating that solute particles which have crossed the control plane at $t>\tau$ contribute to M. Then, the BTC is given by

$$\mu(x,t)=\frac{\partial M}{\partial t}=\frac{1}{M_0}\int_{A_0} m_0(\mathbf{b})\gamma[t,\tau]d\mathbf{b}\ \text{with}\ \gamma=\delta(t-\tau) \tag{15}$$

The contribution of the elementary mass $dM_0=m_0d\mathbf{b}$ injected at $t=0$ in A_0 to the BTC, through the intersection between the streamtube originating at $d\mathbf{b}$ and the control plane at x, is therefore given by

$$d\mu=\frac{m_0d\mathbf{b}}{M_0}\gamma(t,\tau)\ \ \text{with}\ \gamma=\delta(t-\tau) \tag{16}$$

Equation (16), paralleling (3) for resident concentration, states that for advective transport the mass element $dM_0=m_0d\mathbf{b}$ crosses the control plane at $t=\tau$. It will be generalized in section 4 for reactive transport for which $\gamma(t,\tau)$ has a different expression.

It is seen that γ, as well as μ and M, satisfy the linear equation

$$\frac{\partial\gamma}{\partial t}+\frac{\partial\gamma}{\partial\tau}=0 \tag{17}$$

with boundary condition $\gamma(t,0)=\delta(t)$, which can be easily generalized to continuous injection and will be used as starting point in section 4.

For flux proportional injection and with the definition of the travel time PDF

$$f(\tau,x)=(1/U)\int V_0f(\tau,V_0)dV_0 \tag{18}$$

where $f(\tau,V_0)$ is the joint PDF of first passage time and V_0, we get in (15) by averaging $\langle\mu\rangle=f(\tau,x)$, i.e. the mean BTC is equal to the travel time PDF (for the assumed ergodic plumes $\mu\cong\langle\mu\rangle$). Again, the impact of local-scale dispersion on μ is usually small [see e.g., *Fiori et al.*, 2011].

For Fickian transport $f(\tau,x)$ is Inverse Gaussian of mean $\langle\tau\rangle=x/U$ and variance σ_τ^2, and μ satisfies (13) with $\alpha_L=(U^2/2)d\sigma_\tau^2/dx$.

It is emphasized that in applications in which one is interested in the average concentration in an outlet (e.g., a pumping well) the BTC $\mu = \partial M/\partial t$, or the associated longitudinal relative mass distribution $m = -\partial M/\partial x$, are providing the required information.

2.4. Transport in Weakly Heterogeneous Aquifers

Various simple results were obtained in the past for weakly heterogeneous aquifers for $\sigma_Y^2 < 1$. This has been achieved by a first order expansion of K in the flow equations (1), i.e. $K/K_G = \exp(Y')$ is approximated by $1+Y'$, where $Y' = Y - \langle Y \rangle$. This leads to a great simplification, allowing for instance the derivation of analytical expressions of the macrodispersivity α_L and other results (the topic is covered by an extensive literature summarized in the aforementioned monographs). It is worthwhile to recall a few important results [*Dagan*, 1989]: (i) the trajectory fluctuation X_1' is Gaussian for Y' normal at any time and becomes so for sufficiently large time for any Y' distribution; (ii) the upscaled ADE (13) is obeyed with the longitudinal macrodispersivity growing with time and stabilizing at an asymptotic constant value (Fickian transport) after a "setting time." The shape of $\alpha_L(t)$ is weakly dependent on that of ρ_Y and on anisotropy ratio $f = I_v/I$ [see e.g., *Dagan*, 1989, Figure 4.6.4]; (iii) the asymptotic value is given by $\alpha_L = \sigma_Y^2 I$, a very robust result which is independent of the logconductivity PDF and the anisotropy ratio. Field experiments and numerical simulations validated the first order theoretical results regarding $\alpha_L(t)$ for lognormal K and σ_Y^2 as large as unity.

3. Advective Transport in Highly Heterogeneous Formations

The results obtained for weakly heterogeneous formations, discussed above, were both insightful and useful, being the first available tools for modeling transport in heterogeneous aquifers. They were able to clarify and quantify the dependence of relevant transport quantities (e.g., macrodispersivity and the BTC) on a few key structural features, like e.g., K_G, I and σ_Y^2. However, the first order results have limitations and are generally unable to predict some important transport features that have been observed in numerical simulations and experiments, e.g., the highly skewed and non-Gaussian BTC, apparent anomalous transport and the long pre-Fickian transport regimes, to mention some. Thus, both theoretical and experimental works, and most noteworthy the MADE experiment mentioned in section 2.1, have motivated the need to move beyond the weakly heterogeneous paradigm, into the complex field of flow and transport in highly heterogeneous porous formations (i.e. large σ_Y^2).

Many theoretical models for transport in highly heterogeneous media have been developed in the last two decades [e.g., *Cushman and Ginn*, 1993, 2000; *Berkowitz and Scher*, 1995; *Benson et al.*, 2000; *Harvey and Gorelick*, 2000]. Most models are focused on the BTC prediction, and their common feature is that they use proxy mechanisms and parameters to capture the non-Gaussian behavior of transport, e.g., multirate reactive transport (see section 4). The use of proxy mechanisms circumvents the need to solve numerically the flow and transport equations recapitulated in section 2, which is a formidable task for high σ_Y^2; hence it can be regarded as a "top-down" (or "downward") approach, which is opposite of the "bottom-up" (or "upward") approach in which flow and transport are explicitly solved starting from the K structure [e.g., *Klemeš*, 1983]. In this paper we follow the "bottom-up" approach, as also indicated by the title, and illustrate results by using a theoretical model that we have developed in the last decade, namely MIM-SCA. The model is physically based and explicitly solves both flow and transport for highly heterogeneous systems. As a consequence, the transport predictions are directly related to the underlying flow and K parameters from which the model is built. A brief description of the model follows, while further elaboration and details are provided in previous papers [*Dagan et al.*, 2003; *Fiori et al.*, 2006, 2011; *Cvetkovic et al.*, 2014], and particularly in a concise form in the supporting information of *Cvetkovic et al.* [2014].

The heterogeneous porous formation is modeled as a three-dimensional lattice of adjacent rectangular blocks of longitudinal and transverse side $2I$ and vertical one $2I_v$, of random independent K (see Figure 2). Thus the $K(\mathbf{x})$ field can be represented as a sum of indicators equal to K_i inside the block i, and zero outside, explaining the choice of the denomination MIM (Multi Indicator Model) for such a structure. The model is quite general since it allows to represent a given univariate PDF $f(K)$ (in particular the lognormal distribution) and integral scale I.

Solute is instantaneously injected over a large planar area $A_0 \gg I^2$ transverse to the mean flow U at $x = 0$ and is detected at a control plane, at distance x from the injection plane. The solute BTC $\mu(x,t) \cong \langle \mu(x,t) \rangle$

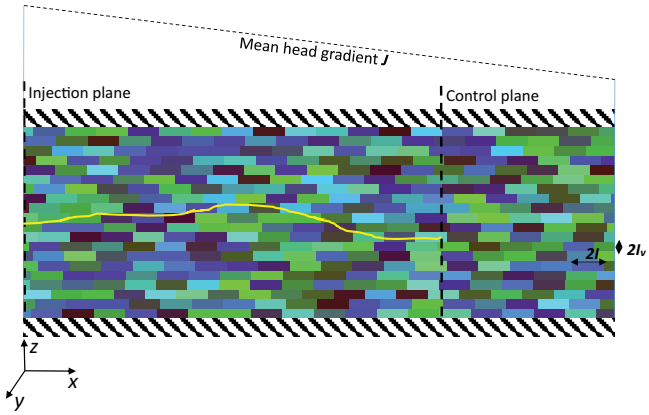

Figure 2. Illustration sketch of the domain and the conceptual Multi Indicator Model (adapted from *Fiori* [2014]).

at the control plane is equivalent to the travel time $\tau(x)$ distribution of solute particles, from the injection to the control plane (see section 2.3). The numerical solution of the flow and transport equations is very demanding in terms of computer resources [see *Janković et al.*, 2003].

A relatively simple solution for flow can be obtained by adopting the self-consistent approach, which is well known in the literature on heterogeneous media, i.e. to superimpose the velocity fields pertaining to isolated blocks submerged into a uniform matrix of conductivity K_0. Then, the self-consistent argument requiring that the mean velocity is equal to U is invoked, rendering the background conductivity K_0 equal to the effective conductivity K_{ef} [e.g., *Dagan*, 1979; *Dykaar and Kitanidis*, 1992; *Jankovic et al.*, 2013]. We employ for K_{ef} the solution by [*Dagan*, 1989, equations (3.4.30) and (3.4.32)] derived for spheroidal inclusions, which is computed by a simple algorithm for any given univariate $f(K)$.

The approach was adapted to transport by assuming that a planar plume is injected far upstream of an isolated block of conductivity K, surrounded by a uniform matrix of conductivity K_0, with uniform velocity U at infinity. After solving for the velocity field, an element of the plume captured by the streamtube crossing the inclusion central plane is followed until its shape stabilizes and travels with velocity U far downstream the inclusion [see e.g., *Dagan and Fiori*, 2003, Figure 4]. With A_w the area of the cross section of the streamtube far downstram from the inclusion (the wake), transport could be characterized by two parameters: the travel time residual τ', which is approximately constant in the wake [*Fiori et al.*, 2007; *Cvetkovic et al.*, 2014, for a cubical inclusion], and the ratio ω between A_w and the inclusion cross sectional area A_c ($A_c=4l^2$ for cubes). Analytical expressions were obtained for $\tau' U/l$ and ω as functions of the conductivity contrast $\kappa=K/K_0$ [*Fiori and Dagan*, 2003; *Cvetkovic et al.*, 2014]. Subsequently, the mass flux for the entire plume is obtained by superimposing the results obtained for isolated blocks and the interaction is accounted for by the self-consistent argument, i.e. by adopting $K_0=K_{ef}$. The final result for the mass flux at a control plane at $x=2Sl$, where S is the number of blocks, is given by [see *Cvetkovic et al.*, 2014, equation (15)]

$$\frac{\langle \mu(x,t)\rangle U}{l}=\left\langle \omega(\kappa_1)\ \ \omega(\kappa_2)\ldots\omega(\kappa_S)\ \ \delta\left\{t-\sum_{s=1}^{S}\left[\frac{2l}{U}+\tau'(\kappa_s)\right]\right\}\right\rangle \tag{19}$$

Thus, (19) states that the total travel time residual is the sum of those pertaining to the S blocks while the contributing mass results from the product of the "capture factors" ω. Subsequently, $\mu\cong\langle\mu\rangle$ is easily determined for a given univariate PDF $f(K)$ by a convolution in the Laplace or the Fourier transforms domains. The same approach can be conveniently cast in a Monte Carlo and random walk procedure [*Cvetkovic et al.*, 2014]. The resulting μ (19) was determined for K lognormal as function of the flow and structural parameters U, σ_Y^2 and l. While (19) is formulated for a control plane at an integer number of blocks, the procedure can be easily extended to any distance by an interpolation procedure. Similarly, the spatial mass distribution $m(x, t)$ can be derived from $\mu(x,t)$ [see e.g., *Fiori et al.*, 2013]. A different model extension, which includes the impact of pore-scale dispersion was presented by *Fiori et al* [2011]. The model was tested against numerical simulations [*Janković et al.*, 2006; *Fiori et al.*, 2006, 2007; *Janković and Fiori*, 2010; *Fiori et al.*, 2011] and laboratory or field experiments [*Fiori et al.*, 2013]. Herein illustrations of a few results.

The first quantity of interest is the longitudinal macrodispersivity α_L, which can be easily obtained from the second temporal moment of μ. Figure 3 displays α_L as function of σ_Y^2 for a few values of $Pe=Ul/D_d$, with D_d the local dispersion coefficient (the procedure is outlined in *Fiori et al.* [2006] and *Janković et al.* [2009]). It is seen for $\sigma_Y^2\lesssim 2$ the weak heterogeneity result $\alpha_L=\sigma_Y^2 l$ (see section 2.4; black dashed line) is an excellent approximation of α_L, well beyond its presumed limits. The reason was explained in *Fiori et al.* [2003] and is

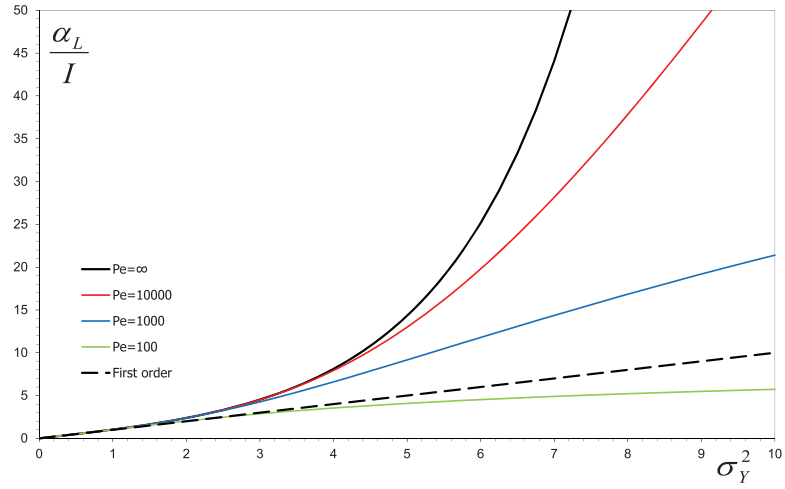

Figure 3. Asymptotic longitudinal macrodispersivity α_L for 3-D heterogeneity as function of the logconductivity variance σ_Y^2, for a few values of the Peclet number $Pe = UI/D_d$. The dashed line depicts the first order result $\alpha_L = \sigma_Y^2 I$ (adapted from *Janković et al.* [2009]).

as follows: the linear expansion around $Y' = 0$ leads to an overestimation of the travel time residual τ' for large Y', while it underestimates τ' for low Y'. When the contributions are symmetrically averaged over high and low Y', like in the normal distribution, there is a compensation of errors, leading to a robust estimate of $D_L = \sigma_\tau^2/(2x) = \alpha_L U$ (13). A similar effect was found numerically by *Bellin et al.* [1992]. Also, consistent with the first order approach [*Berglund and Fiori*, 1997], α_L depends weakly on *Pe* for small to moderate degree of heterogeneity. When heterogeneity is high ($\sigma_Y^2 \gtrsim 2$), α_L becomes large and it strongly depends on *Pe*. The faster than linear growth of α_L with σ_Y^2 is caused mainly by the solute retention in the low conductive zones, where the solute is practically immobilized, leading to a significant growth of τ and therefore of α_L. In such cases, molecular diffusion is quite effective in transferring the solute away from those zones, reducing the growth of τ; this explains the strong impact of *Pe* on macrodispersivity. The solute retention has also a very strong influence on the "setting time" (see section 2.4), which can be extremely large [see e.g., *Fiori et al.*, 2003, Figure 4], such that transport may seem to be anomalous for any reasonable application.

A more complete description of transport is provided by the entire $\mu(x,t)$ (i.e. the BTC). The injection mode (resident concentration or flux proportional) also impacts the BTC [*Demmy et al.*, 1999; *Janković and Fiori*, 2010], and we focus here on the flux proportional injection mode. Figure 4 reproduces the BTC for $\sigma_Y^2 = 1, 2, 4, 8, 16$ and at the control plane distance $x = 80I$ [*Fiori et al.*, 2006]. At such distance, the weak heterogeneity approximation predicts an inverse Gaussian BTC, of mean $t = x/U$. The Gaussian shape is indeed a rather valid approximation for $\sigma_Y^2 \lesssim 1$ (not shown), while it rapidly deteriorates for increasing σ_Y^2, as shown in Figure 4. Thus, the first-order analysis is less robust for the shape of the BTC than for α_L. When $\sigma_Y^2 \gtrsim 1$ the BTC significantly deviates from the symmetric, Gaussian shape, exhibiting a strong asymmetry. The latter manifests in two major features: a strong tailing and a pronounced early peak.

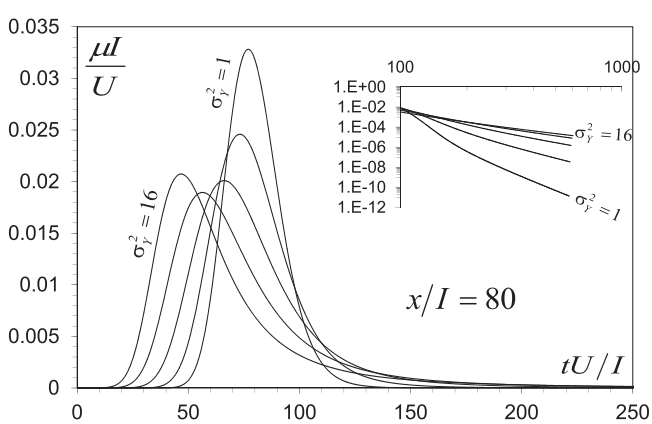

Figure 4. The solute flux $\mu(t;x)$ as function of time for $\sigma_Y^2 = 1, 2, 4, 8, 16$; the control plane is at $x/I = 80$ (adapted from *Fiori et al.* [2006]).

The tailing is due to the late solute arrivals, which in turn are caused by the solute retention by the low conductivity/velocity areas, as previously described. The total mass of the "trapped" solute increases with σ_Y^2, i.e. with the degree of heterogeneity, due to the increase of the relative volume of the low-*K* zones. Such retention mechanisms lead to an apparent anomalous, power-law tailing [e.g.,

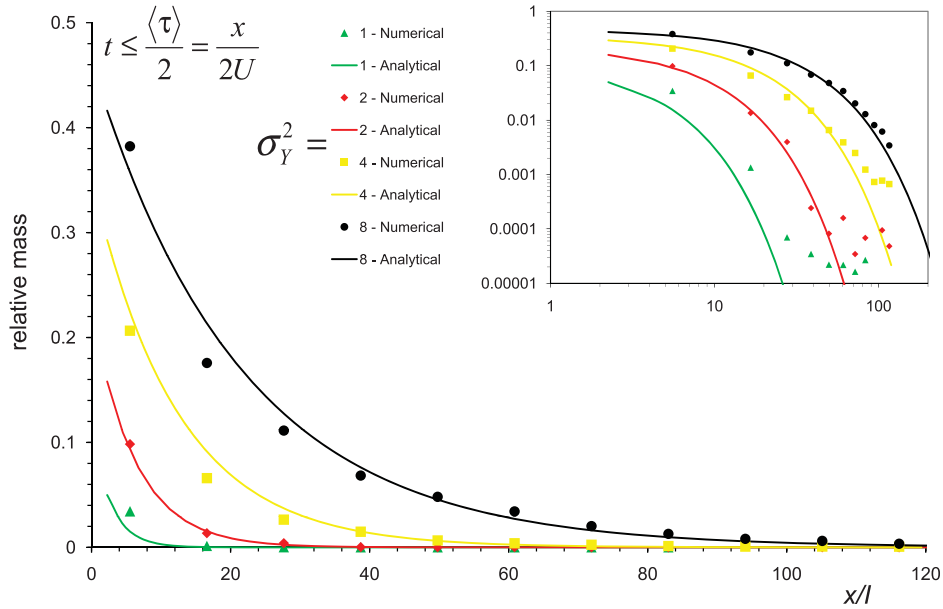

Figure 5. Relative mass arriving at the control plane before $t = \langle\tau\rangle/2 = x/(2U)$ as function of the control plane distance x, for a few values of the log-conductivity variance σ_Y^2; MIM-SCA prediction (solid lines) and numerical results (dots) (adapted from *Fiori and Jankovic* [2012]).

Bouchaud and Georges, 1990], for large σ_Y^2, as visible in the insert of Figure 4. While this circumstance is of some interest for specific applications, e.g., aquifer remediation, the solute mass confined in the tail is anyway rather small.

In turn, the early peak arrival, occurring at $t < x/U$, is caused by the emergence of channeling and preferential flow, particularly in strongly heterogeneous porous media. Connected, high velocity patterns naturally emerge in such media; they are flow-related and their longitudinal extend can be much larger than the conductivity integral scale I [*Fiori and Jankovic*, 2012]. Heterogeneity has indeed a considerable impact on the emergence of channeling patterns, and connectivity is enhanced in 3-D structures as compared to 2-D ones. Thus, a significant fraction of the solute mass travels faster than the mean, reaching the control plane at a time shorter than expected; this circumstance may have an impact on applications, e.g., risk assessment (see section 6). The above feature is further illustrated in Figure 5 which shows the relative mass which arrives at x at times $t \leq \langle\tau\rangle/2$, i.e. with twice the mean velocity U, for a few values of the logconductivity variance σ_Y^2, for both MIM-SCA (solid lines) and numerical simulations (dots). It is seen that more and longer connected channels emerge for increasing σ_Y^2, leading to stronger preferential flow and more mass arriving at the control plane ahead of mean arrival time $\langle\tau\rangle$.

We emphasize, however, that tailing, preferential flow and in general non-Gaussianity are transient features for the setup considered here, and all the BTCs of Figure 4 eventually converge to Gaussian at sufficiently large control plane distances x, with a convergence rate inverse proportional to σ_Y^2. For large σ_Y^2, the convergence to Gaussian can be so slow that transport may look as anomalous for any conceivable application. *Fiori et al.* [2007] have shown that genuine anomalous transport (in which the BTC never converges to Gaussian) for pure advection only occurs for conductivity distributions $f_K \sim K^\delta$ for $K \to 0$, with exponent in the range $-1 < \delta < 0$; in turn, anomalous transport is not possible in presence of local scale dispersion or molecular diffusion, as the latter "destroys" the very long retention times which are responsible for tailing and slow (or lack of) convergence to Gaussian.

Many of the above features occurring in highly heterogeneous formations were observed at the MADE site [*Boggs et al.*, 1992; *Adams and Gelhar*, 1992], a highly heterogeneous formation ($\sigma_Y^2 = 6.6$) which motivated most of the research on transport in highly heterogeneous formation of the last two decades. The first experiment (MADE-1) resulted in a series of 6 snapshots of the longitudinal relative mass distribution $m(x,t)$. MIM-SCA was applied to the MADE experiment by *Fiori et al.* [2013], who employed the K statistics based on DPIL data (Figure 1) [*Bohling et al.*, 2012], while the flow data were inferred from

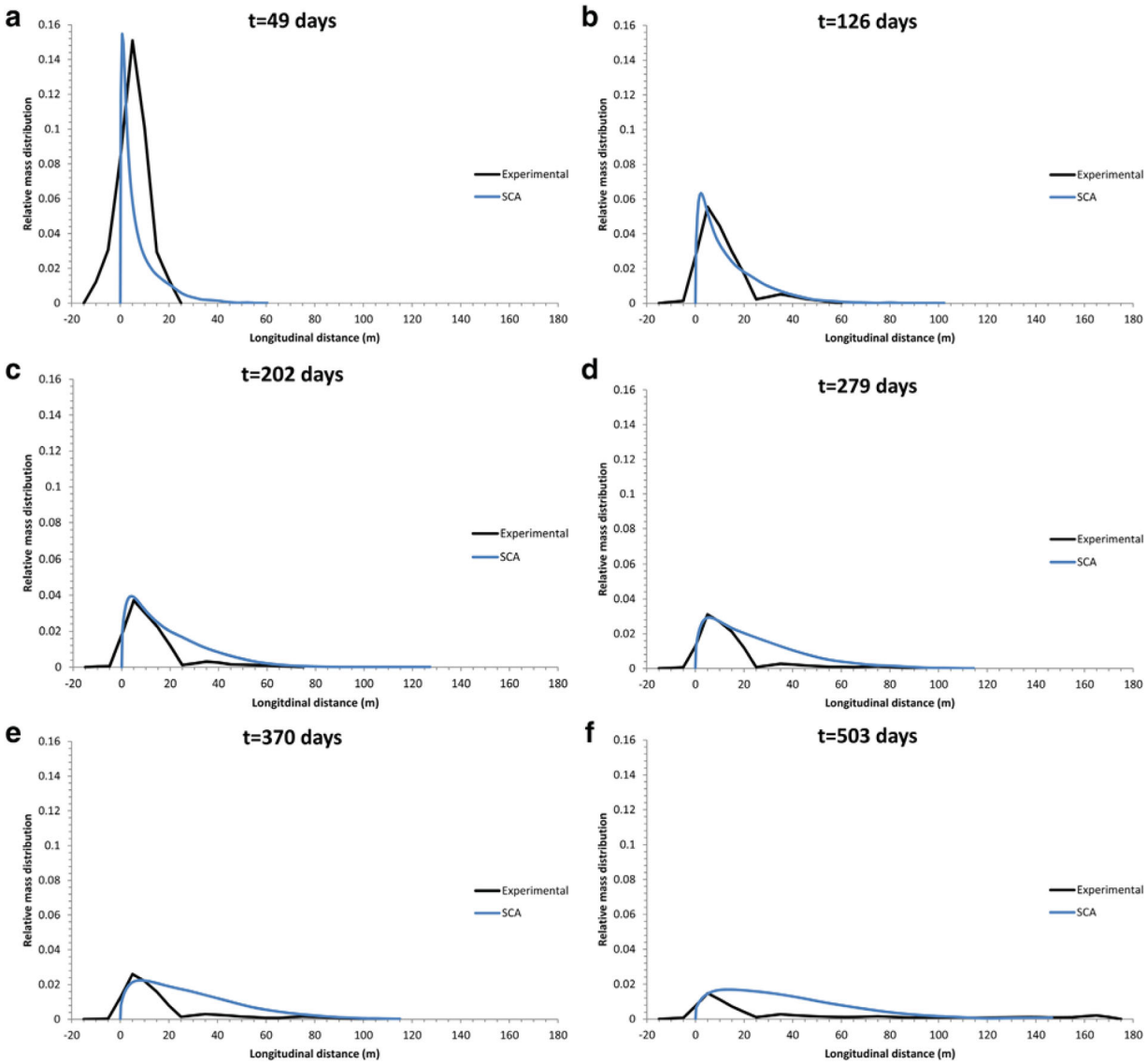

Figure 6. Longitudinal mass distribution $m(x;t)$ at the MADE site: experimental results (black) and theoretical prediction (MIM-SCA, blue line) for the six sampling rounds of the MADE-1 experiment (a)–(f) (adapted from *Fiori et al*. [2013]).

the original papers; the resulting set of parameters was $U = K_{ef}J/\theta = 0.032$ *m/d*, $\sigma_Y^2 = 6.6$, $I = 10.2$ *m*, and no fitting was performed. The theoretical and experimental $m(x,t)$ for the six sampling rounds of the MADE experiment are shown in Figure 6. There was a serious problem with mass recovery, which went down to 43% at $t = 503$ *days*. *Fiori* [2014] argued that a possible explanation for the poor mass recovery, besides those offered by *Adams and Gelhar* [1992], is the difficulty of detection of "fast" channels by the network of monitoring wells; such channels are few and typically carry a larger fraction of mass than their share of their relative volume. Thus, the poor mass recovery is mostly concentrated at the large distances from the injection plane. We note that the spatial behavior of *m* displays an early peak and a long downstream tail features as previously discussed, which are captured by MIM-SCA approximate model. A "perfect" agreement between theory and experiments is not possible for MADE, because of the reasons mentioned in *Fiori et al*. [2013]. Nevertheless, the predictive ability of MIM-SCA is quite remarkable, in view of the complete lack of calibration and use of structural parameters solely. Most important, it may serve as a useful tool to grasp the most significant transport features occurring in heterogenous formations, no matter the degree of heterogeneity, in a relative simple and physically based manner, as briefly summarized in this section.

4. Reactive Transport

Solute transport by groundwater will always involve some degree of mass transfer into essentially immobile zones located on the grains and/or being part of the intra-granular (matrix) porosity. When applying tracer tests for aquifer characterization, the choice of tracers and time scales of the tests are typically such that mass transfer processes are deliberately minimized so that full focus may be set on the physical (hydrodynamic) components of transport (advection and macro-dispersion).

In applications related to groundwater quality with a broad range of contaminants however the relatively large time scales involved will allow for the mass transfer to develop over time. Organic contaminants for instance will in most cases be sorbing to the solid phase and will be subject to some type of loss, or degradation due to microbial activity. Inorganic contaminants such as metals do not degrade however as a rule are subject to relatively strong sorption. An important class of metals are radionuclides that will in most cases strongly sorb as well as be subject to decay. Pathogens on the other hand, will be more or less subject to a mass transfer process in form of attachment-detachment onto the solid phase, but also to a net loss over time.

Applications related to risk assessment that involve organic or inorganic contaminants as well as pathogens, will as a rule involve contaminants/solutes that in some way are transformed (are nonconservative) and/or are subject to mass transfer that combines sorption and diffusion. The sorption process will depend on the micro-scale composition and structure of the solid phase, and may involve complicated geochemical reactions that influence sorption or biodegradation; such complex geochemical reactions however are outside the scope of this work.

Our present discussion builds on three-dimensional fluid trajectories (or stream tubes) as a basic hydrodynamic feature of groundwater flow in heterogeneous aquifers. The main focus is quantifying expected BTC for single species, denoted by $\langle\mu\rangle$ that averages μ (15)-(16) (with $\mu \cong \langle\mu\rangle$ for the assumed ergodic plumes), generalized however for mass transfer as well as linear decay or degradation. Emphasis will be on linear models for low (trace) concentrations, and on (semi)analytical expressions that are general yet simple for implementation.

4.1. Transport With Mass Transfer Along Trajectories

Along the lines of the preceding sections, we consider the transport of a dissolved tracer in a spatially variable three-dimensional, steady-state velocity field $\mathbf{V}(\mathbf{x})$ with a mean drift set parallel to the x-axis, i.e., $U \equiv \langle V_1\rangle$, and $\langle V_2\rangle = \langle V_3\rangle = 0$; the initial tracer concentration in the domain is null. In a heterogeneous aquifer the variability of $\mathbf{V}$ is primarily due to variability in the hydraulic conductivity K, whereas in a fractured rock formation, $\mathbf{V}$ is usually considered nonzero only in fractures along a three-dimensional network with random hydraulic properties.

It has been shown in section 2 that the BTC $\mu(x,t)$ can be expressed with the aid of the function γ for unit mass pulse injection with $\gamma(t) = \delta(t)$ at $\tau = 0$ (or $x = 0$), into an initially solute-free domain (see equation (16)). In this section, γ is generalized for reactive transport by rewriting (17) as follows

$$\frac{\partial\gamma}{\partial t} + \frac{\partial\gamma}{\partial\tau} = -\frac{\partial\gamma^\star}{\partial t} \quad \text{with boundary condition } \gamma(t,0) = \delta(t) \tag{20}$$

where $\gamma^\star$ is the contribution to $d\mu$ at the control plane at x, resulting from the solute undergoing reaction. The simplest kinetic model is of first-order mass transfer with a single (or uni-) rate [*Cvetkovic and Shapiro*, 1990; *Cvetkovic and Dagan*, 1994]

$$\frac{\partial\gamma^*}{\partial t} = k_0(A\gamma - \gamma^*) \tag{21}$$

where k_0 [1/T] is the backward exchange rate and Ak_0 is the forward exchange rate. Thus, one of the illustrations considered by *Cvetkovic and Dagan* [1994] is of spheres of equal radius occupying the fraction A of immobile volume, while k_0 was related to the diffusion between the mobile and immobile phases. Toward solving (20), equation (21) is rewritten in the Laplace Transform (LT) domain as

$$\hat{\gamma}^* = \frac{k_0 A}{s + k_0}\hat{\gamma} \tag{22}$$

where s is the LT variable and "hat" denotes LT.

Table 1. Memory Function Derived From a Multirate Model With the Pareto Type I Rate Distribution[a]

	Multirate Mass Transfer With Pareto Type I Distribution
$\phi(k)$	$\nu k_0^{\nu} k^{-\nu-1}$ $(k \geq k_0)$
Memory function (Laplace Transform domain)	
$\hat{g}(s)$	${}_2F_1(1, \nu, \nu+1, -s)$
Memory function (real domain)	
$g(t)$	$\nu E_{\nu}(t)$
Limits of t	
$t \to 0$ $(0 < \nu < 1)$	$g \to t^{\nu-1}$
$t \to 0$ $(\nu > 1)$	$g \to$const.
$t \to \infty$	$g \to 0$
Limits of ν	
$\nu = 1/2$	Corresponds to Fickian diffusion into a finite matrix
$\nu \neq 1/2$	Corresponds to non-Fickian diffusion into a finite matrix
$\nu \gg 1$	Corresponds to first-order kinetics

[a]The memory function g, times t and t_0, the Laplace Transform variable s are all normalized with the characteristic rate k_0.

Equation (22) is a particular case of the generalized linear equilibrium sorption model proposed in chromatography [*Villermaux*, 1974]

$$\hat{\gamma}^* = \hat{g}\,\hat{\gamma} \tag{23}$$

where $g(t)$ is the "memory function" [e.g., *Cvetkovic*, 2012]. In the equilibrium case there is no memory, such that $\hat{g} = A$ and $g(t) = A\delta(t)$, whereas for (22) $\hat{g} = k_0 A/(s+k_0)$, $g(t) = k_0 A \exp(-k_0 t)$. Thus, the solution of (20)−(21) in the LT domain with (23) is given by

$$\hat{\gamma}(s,\tau) = \exp\{-\tau s[1+\hat{g}(s)]\} \tag{24}$$

Cvetkovic and Dagan [1994, Figure 2] have explored the properties of $\gamma(t,\tau)$, displaying the transition from advection $\gamma = \delta(t-\tau)$ to equilibrium (retardation) $\gamma = \delta[t-(1+A)\tau]$.

The solution $\gamma(x,t)$ based on the unirate linear model (22) cannot reproduce the behavior of diffusive mass transfer especially into immobile zones of complex structure. Exchange with multiple rates was proposed as a heuristic model for capturing trapping in solids [*Noolandi*, 1977; *Schmidlin*, 1977] and applied more recently for groundwater transport [*Haggerty and Gorelick*, 1995]; *Carrera et al.*, 1998; *Haggerty and Meigs*, 2000; *Cvetkovic and Haggerty*, 2002]; it is modeled by the memory function as

$$\hat{g}(s) = A\int \frac{k}{s+k}\phi(k)\ dk \ \text{ i.e. } \ g(t) = A\int k\ \phi(k)\ e^{-kt}\,dk \tag{25}$$

with multiple sites of different relative volume and rate k. The most general model considered here is of a continuous k and corresponding $A\phi(k)$, where $\phi(k)$ [T] stands for the PDF of exchange rates k. When normalized, the memory function is a PDF of return times for tracer particles to the mobile zone, once these have been immobilized. It can also be interpreted as the kinetic partitioning in the LT domain as noted in (22) [*Villermaux*, 1974]. Though different distributions $\phi(k)$ may be considered [*Haggerty and Meigs*, 2000; *Cvetkovic and Haggerty*, 2002; *Cvetkovic*, 2012], for convenience we adopt in this work a truncated power-law in form of a Pareto type I distribution, i.e.,

$$\phi(k) = \frac{\nu\, k_0^{\nu}}{k^{1+\nu}};\ \ k \geq k_0 \tag{26}$$

where $\nu > 0$ is the exponent and k_0 [1/T] is the smallest exchange rate (the properties are provided in Table 1). Clearly, the kinetics of mass transfer in (24) is controlled by $\phi(k)$. The simplest case $\phi(k) = \delta(k-k_0)$ yields (22).

The solution γ (24) is conditional on a given τ, pertaining to a streamline, which itself is a random variable due to spatial variability of K, as discussed in section 2. With the dependence of the discharge μ upon γ given in (15), its mean is obtained by averaging γ over the τ ensemble. The PDF $f(\tau,x)$ is defined by (18) for flux proportional injection. Hence, ensemble averaging in (24) yields

$$\langle\hat{\gamma}(s,x)\rangle = \langle\exp\{-\tau s[1+\hat{g}(s)]\}\rangle = \hat{f}[s(1+\hat{g}(s)),x] \tag{27}$$

with $m_0 = (M_0/A_0)V_0/U$ in (15) while $\hat{f}$ stands for the LT with respect to τ. For a large injection area A_0 we obtain from (27) for the BTC

$$\hat{\mu}(x,s) \cong \langle\hat{\mu}(x,s)\rangle = \frac{1}{A_0}\int_{A_0}\langle\hat{\gamma}(s,x)\rangle d\mathbf{b} = \hat{f}[s(1+\hat{g}(s)),x] \tag{28}$$

The solution (28) can be easily generalized for continuous injection at $x = 0$ or for nonuniform mass distribution over A_0. Hence, the general equation (28) averages the mass flux (discharge) over three-dimensional

random trajectories, accounting for mass transfer along them; this is referred to as the Lagrangian Stochastic Advection-Reaction (LaSAR) methodology, or approach.

4.2. Illustration Examples

In the following we wish to illustrate the applicability of (28), by selecting specific forms of f and g. Let the hydrodynamic transport to be governed by the special case of Fickian macrodispersion with $f(\tau,x)$ being the inverse Gaussian PDF for advective travel time τ

$$f(\tau,x)=\frac{x}{2(\pi\alpha_L U)^{1/2}\tau^{3/2}}\exp\left[-\frac{(x-U\tau)^2}{4\alpha_L U\tau}\right] \text{ i.e. } \hat{f}(s,x)=\exp\left[\frac{x}{2\alpha_L}\left(1-\frac{4\alpha_L s}{U}\right)\right] \tag{29}$$

that is obtained as the solution of the advection-dispersion equation (ADE) with injection and detection in the flux [*Kreft and Zuber*, 1978]. Note that the coefficient of variation for τ with PDF defined in (29), CV_τ, is related to the macro-dispersivity as $CV_\tau=\sqrt{2\,\alpha_L/x}$.

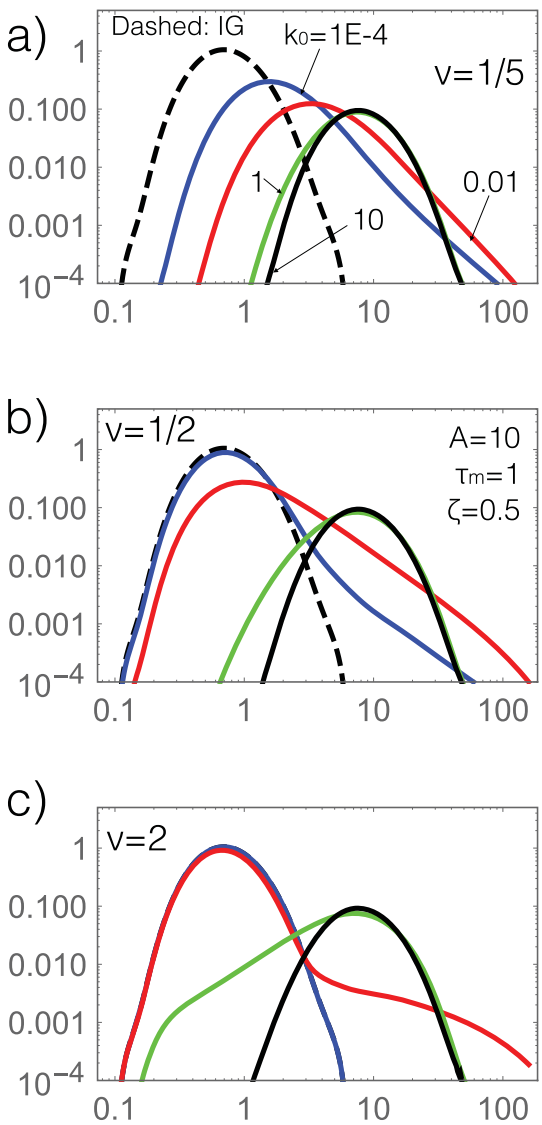

Figure 7. Expected breakthrough curves (28) with f (29) and g given in Table 1, for different values of the dimensionless cut-off rate k_0, and for different multirate exponents: (a) $\nu=1/5$; (b) $\nu=1/2$; (c) $\nu=2$. The case of pure advection-dispersion governed by the ADE (i.e., inverse-gaussian distribution, (29)) is also included as the dashed curve.

As for the memory function, we adopt $\phi(k)$ as the Pareto type I PDF (26), which yields a close-form expression for the memory function in real and LT domains (Table 1).

4.2.1. Expected Breakthrough

Expected breakthrough (or mass discharge) was obtained by numerical inversion of the LT (27), with (25)-(26) and g is given in Table 1. Consider the case of relatively large macro-dispersion with $\alpha_L/x=0.125$ where $\langle\tau\rangle\equiv\tau_m=x/U$ is to be used for normalizing time. The black dashed curve in Figure 7 shows the dimensionless expected breakthrough curve for pure advection-dispersion, i.e., without mass transfer.

In Figures 7, we show the effect of the two key mass transfer parameters k_0 (lower bound cut-off rate) and the exponent ν ($\nu=1/2$ corresponds to mass transfer by Fickian diffusion), for a wide range of k_0 and ν. The exponent ν has a strong effect only for small k_0; when k_0 is large, equilibrium mass transfer is attained and ν has no impact (black curve in Figure 7). For small k_0, kinetic effects are strong and the BTC is very sensitive to ν. For $\nu=2$, the mass transfer corresponds to the first-order kinetic model. Fickian mass transfer with $\nu=1/2$ and small k_0 corresponds to diffusion in a large matrix (blue and red curves in Figure 7b). As the volume of the immobile zone diminishes by increasing k_0, the BTC converges to the first-order exchange and eventually to the equilibrium case (black curves in Figure 7).

4.2.2. Moment Analysis

An interesting perspective on the expected BTC and the effect of coupled processes is obtained by analyzing first few moments. Moments are readily obtained from (28) by derivation. The first three noncentral moments of μ (28) are:

$$m_1=(1+A);\quad m_2=\frac{2A\nu}{k_0(1+\nu)}+(1+A)^2\,(1+\mathrm{CV}_\tau^2) \tag{30}$$

$$m_3=\frac{6A\nu}{k_0^2(2+\nu)}+\frac{6A(1+A)\nu}{k_0(1+\nu)}(1+\mathrm{CV}_\tau^2)+(1+A)^3\,(1+3\,\mathrm{CV}_\tau^2+3\,\mathrm{CV}_\tau^4) \tag{31}$$

where from the coefficient of variation and the skewness of the BTC are readily computed. Note that all times are normalized by $\langle\tau\rangle\equiv\tau_m$.

Relative effects of hydrodynamic transport and mass transfer are illustrated in Figure 8, for two levels of macro-dispersion with $\alpha_L/x=0.005$ and $\alpha_L/x=0.125$, respectively. Both the CV and skewness of the BTC clearly indicate that roughly $k_0=1$ (dimensionless k_0 is in effect the *Damkohler* number) defines the limit for the significance of mass transfer kinetics. For low k_0 the CV and skewness are dominated by mass transfer with macro-dispersion having a small effect, whereas for large k_0 the shape of the BTC is entirely controlled by macro-dispersion, irrespective of the nature of the mass transfer as controlled by exponent ν. The CV is mildly sensitive to the exponent ν, whereas skewness is controlled entirely by k_0.

4.2.3. Solute Attenuation

Hydrodynamic transport (advection and dispersion), mass transfer and transformation (loss) jointly define an attenuation (or self-purification) process where the total mass (relative to M_0) discharged at the control plane x over all time $M(x)=\int_0^\infty \mu(t,x)dt$ is less than the total injected mass M_0, i.e., $M<1$. Full self-purification is achieved over a given domain if $M\approx 0$ and no self-purification is present if $M\approx 1$. A convenient measure of the coupled processes is the attenuation index defined at x as $a_I\equiv -\ln(M)$ [*Cvetkovic*, 2011a].

Let the solute be subject to linear decay with λ [1/T] as the decay (loss) rate for transport quantified by (28); we shall assume for simplicity that λ is the same in the mobile and immobile phases. The expected BTC with decay is obtained from (27) by replacing s with $s+\lambda$. In the limit $s\to 0$, we obtain the solute mass fraction discharged across the control plane at x, i.e., $M/M_0=\hat{\mu}(\lambda)$, and

$$a_I(x)\equiv -\ln\left[\mu(\lambda,x)\right]=\frac{x}{2\alpha_L}\left\{\left[1+\frac{4\alpha_L\lambda(1+\hat{g}(\lambda))}{U}\right]^{1/2}-1\right\} \tag{32}$$

where $\hat{g}(\lambda)$ is the LT of the memory function as a function of the decay rate λ since $s\to 0$. Equation (32) is a particular case of a more general formulation which can be used for relatively simple computations to study for instance the sensitivity of contaminant attenuation (or aquifer self-purification capacity) to a wide range of models and parameters, such as Fickian or non-Fickian macroscopic dispersion [*Cvetkovic*, 2011b], kinetically controlled mass transfer and linear decay. For simplicity, we shall assume for illustration that transport is Fickian, and consider mass transfer rates distributed by the Pareto type I PDF with Fickian diffusive exchange ($\nu=1/2$) (Table 1) and for fixed $A=10$.

Figure 9 shows the sensitivity of the attenuation index a_I (32) on the macro-dispersivity and the cutoff rate k_0. Clearly, the effect of macro-dispersivity depends on the rate k_0: for a large $k_0=100$, equilibrium prevails and the attenuation is strongest. With onset of kinetics and decreasing k_0, attenuation diminishes as does the impact of macro-dispersivity. Thus the largest effect of macro-dispersion is under conditions of small or nonexistent kinetic effects; the kinetic effect enhances overall dispersion of transport and reduces the relative impact of macro-dispersivity.

4.3. Generalization

The preceding derivations can be generalized in a few directions. Thus, nonlinear reaction is a possible one, e.g., the case of nonlinear equilibrium considered by *Dagan and Cvetkovic* [1996]. A different extension is for nonuniform reaction properties in the domain $0-x$, unlike the uniform one assumed in (21). This can be achieved by segmentation of the interval between $x=0$ and x, assuming different reaction parameters along each segment and then deriving $\hat{\mu}$ as a convolution [e.g., *Molin and Cvetkovic*, 2010; *Molin et al.*, 2010]. Finally, the Lagrangian formulation of transport along trajectories using advective travel time and LaSAR methodology can also be used for general reactions [*Cvetkovic*, 1997; *Malmstrom and Martinet*, 2004] These extensions are not considered in the present article.

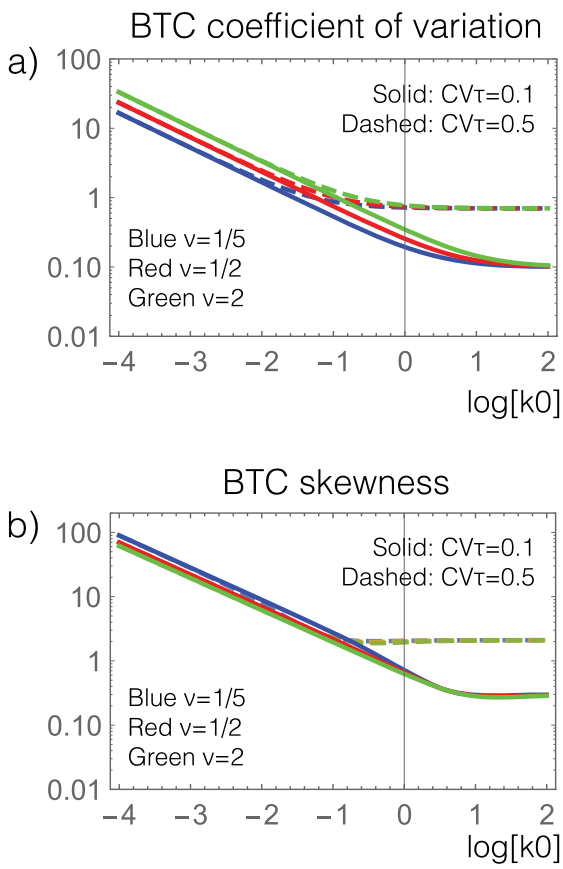

Figure 8. Coefficient of variation and skewness of the expected breakthrough curve computed from (30)-(31), for different multirate exponent ν and different advective travel time coefficient of variation CV_τ, as a function of the dimensionless multirate cut-off rate k_0: (a) coefficient of variation; (b) skewness.

5. Local Concentration and Dilution

In the preceding sections we considered the impact of formation heterogeneity, pore scale dispersion and reactions on the the global transport measures like the longitudinal mass distribution $m(x,t)$ or the BTC $\mu(x,t)$. These are robust measures which are sufficient to determine, for instance, the concentration in the water extracted by a pumping well which captures the contaminant plume. Furthermore, as shown in section 3, pore scale dispersion has a negligible impact on μ for weakly heterogeneous aquifers while it affects primarily the long tail (of very low volume) in highly heterogeneous one. In contrast, in the present section we consider the spatial distribution of the local concentration C within the plume, whose maximal admissible value may serve for regulatory purposes. This topic, still under active research, is also related to the determination of effective transverse pore scale dispersion coefficients at field scale.

5.1. Concentration Probability Density Function

We consider first the PDF of local concentration C (of support volume of Darcy's scale) which is a random spatial variable, whose PDF is given as follows, in absence of pore-scale dispersion [*Dagan*, 1982b]

$$f(C)=\left[1-\frac{\langle C\rangle}{C_0}\right]\delta(C)+\frac{\langle C\rangle}{C_0}\delta(C-C_0) \qquad (33)$$

where C_0 is the initial constant concentration and $\langle C\rangle$ is the ensemble mean concentration. For Fickian transport the latter is Gaussian, solution of (10), and characterized entirely by the macrodispersion tensor. For highly heterogeneous formations $\langle C\rangle$ can be determined approximately from $\mu(x,t)$ by adopting some assumptions on the plume transverse shape, task awaiting further investigations. The mean concentration can be also regarded as representative of a space average over a sufficiently large volume, as discussed in the sequel.

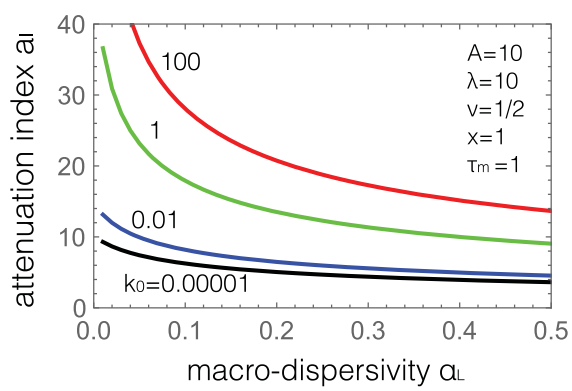

Figure 9. Illustration of the attenuation index (32) as a function of dimensionless macro-dispersivity α_L/x, for different dimensionless cut-off rates k_0, and selected values of A, λ, ν, x and τ_m.

As shown in the discussion in section 2 following equation (7), (33) is exact irrespective of the type of heterogeneity characterizing the spatial distribution of K. Hence, the variance and the coefficient of variation resulting from (33) are given by

$$\sigma_C^2=\langle C\rangle(C_0-\langle C\rangle)\ ,\ CV_C=\sqrt{\frac{C_0}{\langle C\rangle}-1}. \qquad (34)$$

Equation (34) shows that uncertainty is larger at the fringe of the plume and it increases with time. The larger CV_C, less representative is $\langle C\rangle$ of the actual local

concentration. In particular, CV_C computed at the plume center $\mathbf{x}=\overline{\mathbf{b}}+\mathbf{U}\,t$, where $\overline{\mathbf{b}}$ is the center of mass of the injection volume, increases with time without bounds [*Fiori and Dagan*, 2000]. Notwithstanding the theoretical value of this solution as an upper bound, the effect of pore-scale dispersion should be included due to its effect in controlling, through dilution, the local concentration. Indeed, the solution for pure advective transport predicts the spreading of the plume in the form of "thin fingers" [see e.g., *Jankovic et al.*, 2003, Figure 2] and at their boundary the concentration jumps from C_0 to zero. Transverse pore scale dispersion is effective in smearing these discontinuities. However, the actual value of the coefficients at field scale, though small, of order $10^{-3}m$, is still much larger than the laboratory determined values, and difficult to identify.

The inclusion of pore scale dispersion changes dramatically the behavior of CV_C at large times. *Fiori and Dagan* [2000] obtained the following approximate expression of CV_C at large times

$$CV_C \rightarrow \frac{f\,Pe}{2\sqrt{2}\,\Phi_1^{(3D)}}\frac{lnt'}{t'} \quad \text{with } t'=tU/I \tag{35}$$

for three-dimensional heterogeneous formations. In (35) $f=I_v/I$ is the anisotropy ratio and and the function $\Phi_1^{(3D)}$ assumes values between 0 and 1, depending on f and $Pe=U\,I/D_d$, where D_d is the pore scale dispersion coefficient assumed isotropic and equal to the transverse component, since the longitudinal one has a small to negligible impact [*Fiori*, 1996].

Pore scale dispersion causes the coefficient of variation to decline, at large times, as lnt'/t'. *Fiori and Dagan* [2000] noticed that this asymptotic limit is attained at exceedingly large times (i.e. for $t' > 1000$). In addition, CV_C peaks at intermediate times ($t' \simeq 10$) and its value increases with Pe. The main conclusion of this analysis is that uncertainty of point concentration may be large at any time relevant to applications. The effect of conditioning on solute concentration, which may introduce some benefits, reducing for example, uncertainty in proximity of the measurement points [see e.g., *Graham and McLaughlin*, 1989; *Deng et al.*, 1993; *Dagan et al.*, 1996], is not considered here.

Obtaining the concentration PDF at a given location in the presence of pore scale dispersion proved to be difficult. It can be argued that in case of a uniform injection it should be bounded by the binary distribution (33) at the initial time and a Gaussian distribution for $t \rightarrow \infty$, when the plume is very diluted and the dominant dispersion mechanism is pore scale dispersion. In view of these constraints, and other evidences emerging from the analysis of the evolution in time of the maximum concentration in the field scale tracer tests at Borden and Cape Cod, *Fiori* [2001] argued that the Beta distribution can be a good model for the concentration PDF. Numerical validation of this hypothesis has been provided in a few subsequent works [e.g., *Fiorotto and Caroni*, 2002; *Tonina and Bellin*, 2008]. A theoretical demonstration supporting the choice of the Beta distribution, based on a phenomenological equation resembling the conditions in which transport takes place in heterogeneous formations, has been proposed by *Bellin and Tonina* [2007]. This is an active field of research and although recent results show that the Beta model approximates very well the concentration PDF, its generality cannot be confirmed [see e.g., *Dentz and Tartakovsky*, 2010; *Boso et al.*, 2014]. In a recent study, *de Barros and Fiori* [2014] developed a general expression of the concentration CDF whose theoretical moments recovers the expressions of the moments obtained from the Beta PDF.

Along these lines the PDF of the reduced variable $z=C/C_0$ assumes the following form [*Bellin and Tonina*, 2007]

$$f(z)=\frac{\Gamma[p+q]}{\Gamma[p]\,\Gamma[q]}z^{-1+p}(1-z)^{-1+q} \tag{36}$$

where $\Gamma[r]=\int_0^\infty t'^{-1}e^{-t}dt$ is the Euler Gamma function and the constants p and q are given by $p=\langle Z\rangle/\beta_o$ and $q=(1-\langle Z\rangle)/\beta_o$, where β_o depends on the variance of Z, $\sigma_Z^2=\sigma_C^2/C_0^2$, through the following expression

$$\beta_o=\frac{\sigma_Z^2}{\langle Z\rangle[1-\langle Z\rangle]-\sigma_Z^2} \tag{37}$$

The coefficient β_o given by (37) is the main factor controlling the shape of the concentration PDF (36). In the absence of pore scale dispersion $\sigma_Z^2 \rightarrow \langle Z\rangle[1-\langle Z\rangle]$, such that $\beta_o \rightarrow \infty$ and consequently $f(z) \rightarrow \langle Z\rangle\delta(1-z)+[1-\langle Z\rangle]\delta(z)$, which is the dimensionless version of equation (33), as required. On the other hand for $\sigma_Z^2 \rightarrow 0$, $\beta_o \rightarrow 0$ such that the PDF of z tends to a normal distribution with mean $\langle Z\rangle$ and

variance σ_Z^2, which leads to: $f(z)=\delta[z-\langle Z\rangle]$, since for $\sigma_Z^2 \to 0$ the normal distribution converges to the Dirac delta distribution.

Pore scale dispersion controls the transition from the binary distribution, characteristic of the initial condition, to the unimodal normal resembling shape occurring at larger times and further to the ergodic deterministic solution $C=\langle C\rangle$ observed at exceedingly large times when $\sigma_Z^2 \to 0$. The setting times at which these conditions are reached reduces with *Pe*. Section 6 will illustrate how to employ the concentration PDF to evaluate human health risk. Similarly, *Rubin et al*. [1994] and *Bellin et al*. [1994] showed that the sampling volume has an impact similar to pore scale dispersion on the concentration PDF, with the important difference that it manifests since the injection and not only after a rather long setting time.

The support volume of local concentration has a characteristic size equal to the Darcy's scale. When larger sampling devices are used, the concentration variance reduces as an effect of averaging over a larger support volume. The effect of the sampling volume on the concentration moments depends chiefly on two factors: the source dimensions and the size of the sampling device. The effect of the sampling support volume shall be briefly illustrated in section 6.

5.2. Frequency Distribution of Solute Concentration

When dealing with risk analysis and remediation, a quantity of interest is the fraction of the plume volume where a given threshold concentration is exceeded. The larger such a volume, the larger is the risk that this threshold is actually exceeded at a given location. The volume in which a given concentration C_{tr} is exceeded is given by definition (see details in *Boso et al*. [2013]) $V(C_{tr})=\int_\Omega H[C(\mathbf{x},t)-C_{tr}]d\,\mathbf{x}$, with Ω the flow domain and H is the Heaviside function. Accordingly, the probability of not exceeding a given concentration C_{crit}, is defined as

$$\mathcal{F}_C(C \le C_{crit})=1-\frac{V(C_{crit})}{V(C^*)} \tag{38}$$

where $V(C^*)$ is the volume of the plume, defined as the one in which the concentration is above the detection limit C^*. An approximation of $\mathcal{F}_C$ is provided by the frequency distribution of a sample of concentration measurements under the hypothesis that their supporting volume is the same and that they are evenly distributed through the plume, such that the entire plume is sampled.

By using numerical simulations *Bellin and Tonina* [2007] showed that the frequency (38) is well approximated by the Beta model (36) with the parameters p and q obtained by replacing the mean $\langle Z\rangle$ and variance σ_Z^2 into the expressions for p and q and (37) by the corresponding spatial moments

$$\bar{Z}(t)=\frac{1}{V(C^*)}\int_{V(C^*)} C(\mathbf{x},t)\,d\mathbf{x};\quad S_Z^2(t)=\frac{1}{V(C^*)}\int_{V(C^*)} C(\mathbf{x},t)^2\,d\mathbf{x}-\bar{Z}(t)^2 \tag{39}$$

The moments (39) can be obtained either by theoretical models of concentration distribution (as will be shown later) by numerical simulations or by using experimental data. *Bellin and Tonina* [2007] showed that the Beta model (36), with the parameters obtained as described above, represents accurately the frequency distribution obtained by solving numerically the transport equation in a heterogeneous formation. Similarly, the Beta model (36) obtained by substituting the sample moments, i.e. the moments computed by using the concentration measurements included in the data set, into the expressions (37) of p and q provides a good match with the experimental frequency distribution at Cape Cod [*Bellin and Tonina*, 2007], as shown in Figure 10.

These results show that the key parameters for approximating the frequency distribution of the solute concentration are the spatial mean and variance of the reduced variable $Z(\mathbf{x},t)=C(\mathbf{x},t)/C_0$, both changing in time. While the spatial mean reduces continuously, the spatial variance first increases, reaches a peak, and than decreases progressively to zero at exceedingly large times, when the plume behaves as a Gaussian plume.

An effective theoretical model of the spatial moments can be obtained by replacing the actual concentration C in equation (39) with the following approximation of the ensemble mean of the Lagrangian concentration [*Fiori*, 2001; *Boso et al*., 2013] based on first order approximation

$$C(\mathbf{x},t) \cong \langle C_{\mathcal{L}}(\xi,t;\mathbf{c}=0)\rangle = C_0 \prod_{i=1}^{m} \eta_i(\xi_i,t) \tag{40}$$

where

$$\eta_i(\xi_i,t) = \frac{1}{2}\left\{ erf\left[\frac{\xi_i + L_i/2}{\sqrt{2W_{t,ii}(t;0)}}\right] - erf\left[\frac{\xi_i - L_i/2}{\sqrt{2W_{t,ii}(t;0)}}\right]\right\} \tag{41}$$

and C_0 is the constant initial concentration in a rectangular volume of sides L_i $(i=1,2,3)$. In equation (41) $W_{t,ii}(t;0)=X_{ii}(t)+2D_{d,i}t-Z_{ii}(t;0)$, such that $W_{t,ii}=0$ in the absence of pore scale dispersion. The variable $\xi=\mathbf{x}-\mathbf{P}(\mathbf{c},t)$ is the spatial coordinate relative to the centroid $\mathbf{P}$ of the particle that at time $t=0$ was released at $\mathbf{x}=\mathbf{c}$ within the source volume, the latter being defined as follows [*Fiori*, 2001]: $\mathbf{P}(\mathbf{c},t)=\lim_{\Delta\to 0}\int_{\Delta} n\, C_0(\mathbf{c}+\mathbf{y})\mathbf{X}_t(t;\mathbf{c}+\mathbf{y})\, d\,\mathbf{y}$, where Δ is a small volume of solute, which in the definition of the Lagrangian concentration is assumed to be composed by a large number of particles (for additional details, see *Fiori* [2001]). In words, dilution is modeled by following each Darcy-scale solute particle belonging to the initial plume and focusing on the mass exchanges between the examined particle and the surrounding ones.

Equation (40) resembles the solution of the ADE in a homogeneous formation with the second spatial plume moments replaced by the moments W_{ii}. Similarly to the concentration variance discussed in section 5.1, the hypothesis of vanishing distance between the two particles is employed here in the expression for Z_{ii} in order to allow for the analytical treatments leading to equation (40). Notice that this expression cannot provide the estimate of the concentration at a given fixed position, but it is very useful to provide global estimates, such at the frequency distribution of the concentration and global dilution metrics, such as the dilution index, as will be discussed in the ensuing section.

5.3. Dilution Index

Dilution is an important parameter representative of the decay of local concentration and therefore its quantification has attracted a great deal of attention, following the introduction of the dilution index by *Kitanidis* [1994] and the following application to reactive solutes by *Kapoor et al*. [1997]. According to *Kitanidis* [1994] the dilution index is defined as

$$E(t) = \exp\left\{-\int_{\Omega} \frac{C_{\mathcal{L}}(\xi,t)}{M_0} \ln\left[\left(\frac{C_{\mathcal{L}}(\xi,t)}{M_0}\right)\right] d\xi\right\} \tag{42}$$

where $M_0=C_0\,V_0$ is the mass of solute released instantaneously within the volume V_0 and Ω is the flow domain. The dilution index, given by equation (42) is the exponential of the plume's entropy.

The theoretical framework developed by *Fiori* [2001] and already used in section 5.2 can be applied also to obtain a first-order semi-analytical solution of the dilution index in a heterogeneous formation for instantaneous solute injection within a small source V_0 in an unbounded domain [*de Barros et al*., 2015]:

$$E(t) = \exp\left\{-\sum_{i=1}^{N}\int_{-\infty}^{\infty} \eta_i(\xi_i,t)\ln \eta_i(\xi_i,t) d\xi_i\right\} \tag{43}$$

This expression has been obtained under the same simplifying assumptions employed in the expression for the concentration variance and frequency distribution discussed above and has been successfully verified against accurate numerical solutions in two-dimensional heterogeneous aquifers [*de Barros et al*., 2015] and experimental data from the first Cape Cod experiment, as discussed in the sequel.

5.4. Application to Cape Cod Data

The first Cape Cod experiment [*LeBlanc et al*., 1991; *Hess et al*., 1992] provided a valuable data set to validate the previous theoretical developments and the Beta model introduced in section 5.1. The aquifer is weakly heterogeneous ($\sigma_Y^2=0.24$) with a pronounced vertical anisotropy, i.e. $I=3.5\,m$ and $I_v=0.19\,m$ [*Hess et al*., 1992], $f=I_v/I=0.054$. The source volume dimensions are $L_3=6.3\,I_v$ in the vertical direction and $L_1=L_2=1.1\,I$ in the horizontal direction. The relatively large vertical dimension poses a limitation to the use of the above models, which have been developed under the assumption of a small source volume. Details on the

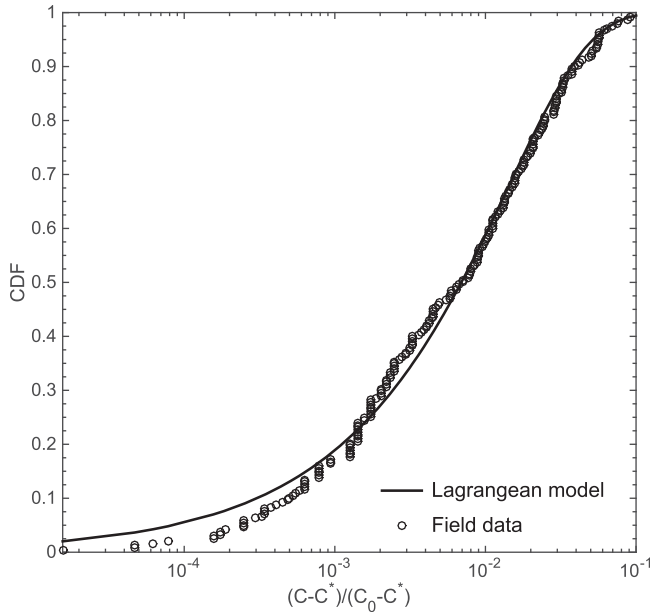

Figure 10. Empirical (experimental) and theoretical concentration CDFs of the solute plume of the first Cape Cod experiment at 237 days from injection. $C^*=0.5\,mg/l$ is the threshold below which concentration measurements have been considered not reliable by *LeBlanc et al*. [1991].

computation of the one and two-particle moments, based on the experimental analysis of the tracer test presented by *Garabedian et al*. [1991] are provided in the work by *Boso et al*. [2013].

Figure 10 compares the CDF obtained with the Beta model and the experimental concentration CDF 237 days after the injection, obtained by neglecting concentrations smaller than $0.5\,mg/l$, identified as not reliable by *Garabedian et al*. [1991]. The Beta model has been applied by substituting into (40) the approximated expressions for one- and two-particle moments obtained as shown in *Boso et al*. [2013]. The resulting estimate of the concentration has been used into equation (39) to obtain the spatial concentration mean and variance, which are then used into equation (36) of the concentration PDF, which integrated provides the theoretical CDF shown in Figure 10 with a continuous line. The match between the experimental and the theoretical CDF, estimated by the Beta model is remarkable, providing a solid experimental validation of the theoretical developments outlined in section 5.1.

The evolution of the dilution index at the Cape Cod provided by equation (43), with the one- and two-particle moments obtained as described above for the computation of the concentration CDF, is compared in Figure 11 with the estimate of the dilution index obtained by *Thierrin and Kitanidis* [1994] from the concentration measurements and recently updated by *de Barros et al*. [2015].

Also in this case, the Lagrangian concentration provides a satisfying estimate of the overall spatial distribution of the solute concentration, such as to represent satisfacorily the dilution index. Though the computations are based on the Lagrangian coordinates ξ_i, their uncertainty, however, becomes immaterial for global quantities such as the frequency distribution and the dilution index.

6. Probabilistic Human Health Risk Assessment

One of the primary concerns related to subsurface contamination is its impact on human health. Improper waste disposal of toxic chemicals or accidental spills pose a risk to groundwater resources and consequently human health. For such reasons, it is important to properly predict solute migrational patterns and the pollutant concentration at an environmentally sensitive location (e.g., control plane or observation well). Once the solute plume reaches the environmentally sensitive target, humans can be exposed to contamination through different pathways such as water ingestion, dermal contact and inhalation. Previous sections presented methods to predict flow and transport in heterogeneous aquifers. In this section, we focus in applying these concepts to assess human health risk in a probabilistic manner. These stochastic tools may serve risk managers to assess the likelihood that a contamination event will lead to adverse health effects in humans.

Let C_e denote the concentration that humans will be exposed to at a environmentally sensitive target during a certain exposure time window. In applications, C_e is taken as the peak of the flux-averaged concentration breakthrough curve at a given location while in some studies, C_e is the maximum of the exposure duration averaged concentration breakthrough curve [see *Maxwell and Kastenberg*, 1999; *Hassan et al*., 2001; *Maxwell et al*., 2008]. The definition of C_e will depend on the application at hand and the regulatory agency risk

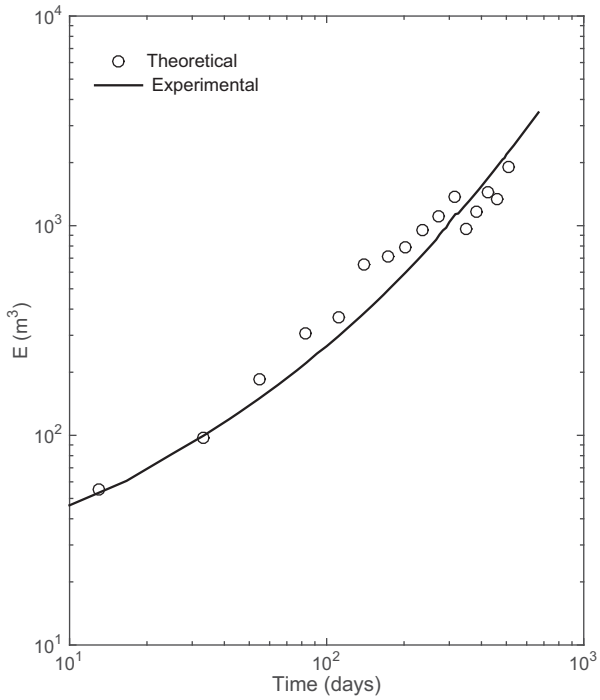

Figure 11. Temporal evolution of Dilution index E in the first Cape Cod experiment. The theoretical expression (43) with Pe=380 (solid line) is compared with the recent estimate based on experimental proposed by [*de Barros et al.*, 2015] (open symbols), which revisits the original estimate by *Thierrin and Kitanidis* [1994].

guidelines. For the purpose of this work, we will consider the peak of the concentration breakthrough curve at the environmentally sensitive target.

As described in sections 1 and 2, the lack of a full characterization in $Y(\mathbf{x})$ leads to uncertainty in $\mathbf{V}(\mathbf{x})$ and consequently in $C(\mathbf{x},t)$ [e.g., *Dagan*, 1989; *Rubin*, 2003]. Therefore C_e is also a random quantity since it is directly linked to $C(\mathbf{x},t)$. Human physiological and behavioral parameters are other sources of uncertainty in the overall health risk metric [e.g., *Maxwell and Kastenberg*, 1999; *Maxwell et al.*, 1999; *Tartakovsky*, 2007; *de Barros and Rubin*, 2008; *Tartakovsky*, 2013]. This implies that individuals consuming contaminated groundwater also contribute to the overall uncertainty in risk assessment since human physiology and toxicology are not fully understood and each individual may have different exposure and response to a certain chemical. As a consequence, human health risk (which depends on C_e and health-related parameters) is subject to uncertainty and quantified stochastically through its low order statistical moments or through its PDF (or equivalently, its cumulative distribution function, the CDF).

For the purpose of illustration, let us consider a scenario where an individual is exposed to a hypothetical carcinogenic substance through the water ingestion pathway. In this case, the risk metric of interest is the increased lifetime cancer risk (ILCR). The ILCR quantifies the probability that an exposed individual will contract cancer over a lifetime [*USEPA*, 1989]. In general, the USEPA recommends 10^{-6} or 10^{-4} as threshold values for ILCR [*USEPA*, 1989, 2001]. Normally, if ILCR $>10^{-6}$, remediation actions are required. Note that other measures of health risk could be employed *in lieu of* the ILCR. In order to quantify ILCR, we adopt the linear ILCR model recommended by the USEPA [*USEPA*, 1989]

$$\mathrm{ILCR} = \beta C_e \text{ with } \beta > 0, \tag{44}$$

where the parameter β contains all the health-related variables relevant for the water ingestion pathway. It is defined as

$$\beta = f_m \times CPF \times \left(\frac{IR}{BW}\right) \times \frac{EF \times ED}{AT} \tag{45}$$

with f_m denoting the metabolized fraction [*McKone and Bogen*, 1991], CPF is the cancer potency factor for the specific contaminant, IR is the water ingestion rate, BW is the body weight, EF corresponds to the exposure frequency, ED is the exposure duration and the average lifetime is given by AT. Details pertaining these parameters can be found in the literature [e.g., *USEPA*, 1989, 2001]. Given the stochasticity of (44), ILCR will be characterized through its CDF, namely $\mathcal{F}_r$:

$$\mathcal{F}_r(r_o) = \mathrm{Prob}[\mathrm{ILCR} \le r_o] \tag{46}$$

where r_o is a given regulatory threshold value (with typical values ranging from 10^{-6} to 10^{-4}). Equation (46) is a useful tool since it allows contaminant site managers to evaluate the probability that ILCR is below a given

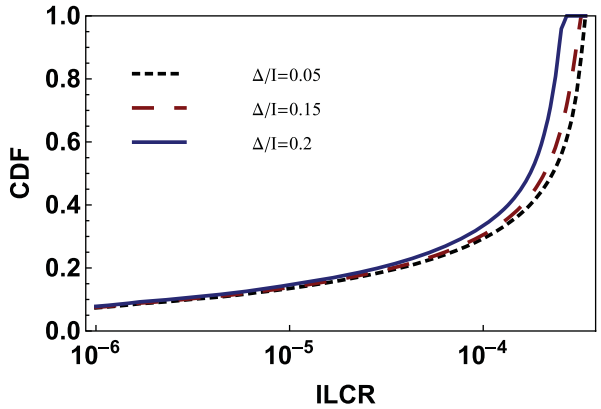

Figure 12. Cumulative distribution function (CDF) for the increased lifetime cancer risk (ILCR) as a function of the vertical scale of sampling device Δ, see section 5.1. Results obtained for a 3-D heterogeneous flow field with $\sigma_Y^2=0.2$, isotropic Y covariance (exponential), point source injection and $Pe=10^3$. The risk CDF was determined using the methodology developed in *de Barros and Fiori* [2014]. The dimensionless vertical sampling device length is given by Δ/I.

regulatory threshold value r_o. The risk CDF (46) can be determined by establishing a probabilistic map between C_e and ILCR, see equation (44):

$$\mathcal{F}_r(r_o) = \text{Prob}\left[C_e \le \frac{r_o}{\beta}\right] = \mathcal{F}_C\left(\frac{r_o}{\beta}\right) \quad (47)$$

where the concentration CDF is given by $\mathcal{F}_C$, see section 5.1. The ILCR CDF (47) can be evaluated through Monte Carlo simulations [e.g., *Maxwell and Kastenberg*, 1999; *Maxwell et al.*, 1999; *de Barros et al.*, 2009] or analytically by mapping random variables [*de Barros and Fiori*, 2014] or assuming a CDF model for $\mathcal{F}_r$ [e.g., *Andričević and Cvetković*, 1996; *de Barros and Rubin*, 2008]. Analytical or semi-analytical expressions for the first two statistical moments of risk (e.g., mean and variance) are also available in the literature [*Andričević and Cvetković*, 1996; *de Barros and Rubin*, 2008; *de Barros and Fiori*, 2014]. Incorporating the effects of human physiological and behaviorial uncertainties can be achieved by marginalization of the risk CDF (47) as done in *de Barros and Rubin* [2008] and by using a nested Monte Carlo approach [*Maxwell and Kastenberg*, 1999].

Determining the CDF in (47) requires knowledge of the concentration CDF $\mathcal{F}_C$. Section 5.1 provides details on how to statistically characterize the concentration at any point in space and time. One approach is to assume the CDF model of $\mathcal{F}_C$ (e.g., Beta, lognormal, etc) and use the concentration mean and variance as fitting parameters (see section 5.1). The second approach is to use a fully derived $\mathcal{F}_C$ by mapping random variables [e.g., *Dentz and Tartakovsky*, 2010; *Boso et al.*, 2014; *de Barros and Fiori*, 2014].

If the sampling support volume is small and $\sigma_Y^2 < 1$, then $\mathcal{F}_C$ can be approximated by the Beta CDF described in section 5.1 [*Fiori*, 2001; *Bellin and Tonina*, 2007]. As the sampling volume increases, the concentration CDF departs from the bimodal distribution, with large predominance of the extreme values $C=0$ and $C=C_0$, to a unimodal, Gaussian-like distribution. The departure from bimodality reflects the larger range of concentration values and a larger entropy. In particular, *Rubin et al.* [1994] and *Bellin et al.* [1994] showed that if the sampling volume increases, the uncertainty in the concentration can be approximated by a Gaussian PDF (for $\sigma_Y^2 < 1$) [see also *Schwede et al.*, 2008]. As discussed in *de Barros and Fiori* [2014], the support volume of the sampling device will have a strong effect on the tails of the concentration CDF which are related to extreme events and of relevance to human health risk assessment.

Using the Lagrangian concentration framework, see section 5, *de Barros and Fiori* [2014] fully derived $\mathcal{F}_C$ as a function of sampling device volume which can be substituted in equation (47). Figure 12 illustrates the ILCR CDF based on the $\mathcal{F}_C$ developed in *de Barros and Fiori* [2014]. Results are depicted for a nonreactive substance released from point source in a steady-state flow in a 3-D aquifer with $\sigma_Y^2 = 0.2$, isotropic exponential covariance and $Pe = 10^3$. The health risk CDF is shown in Figure 12 for different vertical sizes of the sampling device Δ. For this specific illustration, the concentration CDF is evaluated at a fixed time ($tU/I = 0.5$) at an environmentally sensitive target centered with the contaminant source. The value of β is consistent with the data in the literature [*Maxwell and Kastenberg*, 1999; *de Barros et al.*, 2009] and is approximately equal to 0.0012 in all upcoming illustration examples.

Figure 12 shows that the effect of the sampling device becomes noticeable at high risk values (ILCR $> 10^{-4}$). Higher values of ILCR are associated with events with high concentration. For larger sampling volumes, the probability that the lower concentration values are observed increases. This is mainly due to two mechanisms. First, when the scale of the sampling device increases, the variability of the concentration is reduced. Second, for this specific case of a point source release, the sampling device helps to diminish the concentration by sampling more streamlines. For ILCR $< 10^{-4}$, the solute concentration values are small and the risk CDFs

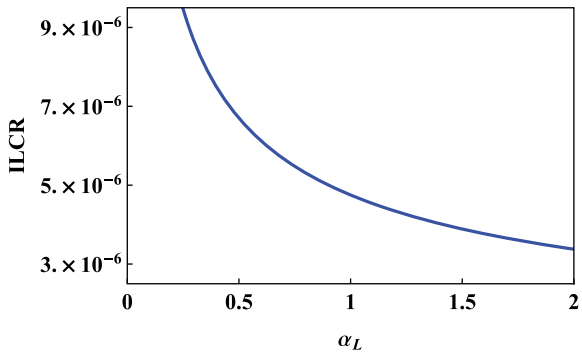

Figure 13. Expected increased lifetime cancer risk as a function of the longitudinal macrodispersivity α_L. Results obtained for an ergodic plume released in a 3-D heterogeneous flow field. Transport is purely advective ($Pe \to \infty$).

display similar values for the sampling device volume values investigated. Additional discussion on the impact of the sampler's volume on concentration statistics can be found in the literature [e.g., *Dagan*, 1989; *Rubin et al.*, 1994; *Rubin*, 2003; *Tonina and Bellin*, 2008].

Alternatively, the ILCR can be stochastically characterized in terms of the statistics of the peak flux averaged concentration observed at the control plane. The flux averaged concentration [*Kreft and Zuber*, 1978] can be expressed as $C_f(t;x)=M_0\,\mu(t;x)/Q_w$ where M_0 is the initial total solute mass and μ is the total relative solute mass discharge crossing the control plane situated at the longitudinal position x and Q_w is the steady state volumetric water discharge. The statistical moments of μ have been subject of intense study [e.g., *Dagan et al.*, 1992; *Cvetkovic et al.*, 1992; *Andricevic and Cvetkovic*, 1998] as discussed in sections 2–4 and applied in human health risk assessment [*Andričević and Cvetković*, 1996; *de Barros and Rubin*, 2008; *Molin and Cvetkovic*, 2010; *Molin et al.*, 2010]. As described in section 2 (see equation (18) and the ensuing paragraph), $\langle\mu\rangle = f(t;x)$, the travel time distribution, and closed form expressions were given for Inverse Gaussian (appropriate to weak heterogeneity by equation (29)) and by the convolution (19) for high heterogeneity. We can obtain the mean and variance of ILCR by replacing C_e, in equation (44), with the peak of the flux averaged concentration at the control plane

$$\langle \mathrm{ILCR}\rangle = \frac{\beta M_0}{Q_w}\left\langle \max_{t>0}\ \mu(t;x)\right\rangle \quad \text{and} \quad \sigma^2_{\mathrm{ILCR}} = \left(\frac{\beta M_o}{Q_w}\right)^2 \mathrm{Var}\left[\max_{t>0}\ \mu(t;x)\right] \tag{48}$$

with $\langle \max_{t>0}\mu(t;x)\rangle$ and $\mathrm{Var}[\max_{t>0}\mu(t;x)]$ corresponding to the mean and variance of the maximum value (over time) of total relative solute discharge crossing the control plane. It is emphasized that in the present article we focused the discussion on ergodic plumes for which $\mu \cong \langle\mu\rangle$ such that σ^2_{ILCR} is negligible. Analysis of the impact of nonergodic plumes, not discussed here, can be found, e.g., in *Cvetkovic et al.* [1992], *Andricevic and Cvetkovic* [1998], *Cvetkovic and Dagan* [1994] and *Fiori et al.* [2002].

We illustrate the use of (48) for an ergodic contaminant plume (e.g., with characteristic dimensions much larger than I) for which $\mu(t;x) \approx \langle\mu(t;x)\rangle = f(t;x)$ such the expected cancer risk is

$$\mathrm{ILCR} = \frac{\beta M_o}{Q_w}\max_{t>0} f(t;x). \tag{49}$$

Figure 13 shows the expected ILCR based on equation (49). The computation of the ILCR is obtained by using the inverse Gaussian PDF (29), which is fully characterized by the uniform-in-the-mean longitudinal velocity U and the macrodispersivity $\alpha_L = \sigma_Y^2 I$ for weakly heterogeneous aquifers (see section 2). The control plane is situated at $x = 20I$ m, the mean longitudinal velocity is $U = 4.5$ m/d and the injected mass is $M_o = 2500$ g. The result depicted in Figure 13 illustrates how the ILCR is diminished with increasing α_L thus emphasizing the importance of characterizing the structure of Y and appropriate models which relate it to transport, in order to quantify risk.

As an additional remark, we point out that an appealing feature of the stochastic theory presented in sections 2–5 is the flexibility to incorporate site characterization data which is important to reduce the uncertainty in the health risk metric and other environmental performance metrics [e.g., *Rubin*, 1991; *Rubin and Dagan*, 1992; *de Barros et al.*, 2012]. Relevant questions associated with site characterization and risk analysis are: *How much site characterization is needed in order to provide a reliable risk estimation? Given that financial resources are limited, how should risk managers and decision makers allocate resources toward uncertainty reduction in the final risk estimation?* Addressing the latter is particularly challenging given the contribution of additional sources of uncertainty (e.g., health parameters) in the overall risk end-point. Hence, the stochastic framework is

particularly important to help decision makers best allocate resources to reduce both the uncertainty of the prediction of interest (i.e. human health risk) and the likelihood of wrong decision.

The effect of hydrological data acquisition in human health risk error reduction was investigated in the literature [*Maxwell et al*., 1999; *de Barros and Rubin*, 2008; *de Barros et al*., 2009, 2012]. For instance, the Lagrangian framework presented in sections 2–4 was employed in *de Barros and Rubin* [2008] to analyze the impact of conditioning the ILCR CDF (47) on hydraulic conductivity measurements. Several works [*Maxwell et al*., 1999; *de Barros and Rubin*, 2008; *de Barros et al*., 2009, 2012] showed how key parameters controlled the worth of hydrogeological data in reducing the uncertainty of human health risk predictions. These works showed that the value of hydrogeological information in human health risk was particularly sensitive to the scale of the solute plume, the sampling device support volume, the well pumping rate, local scale dispersion and finally, the location of the environmentally sensitive target. *de Barros et al*. [2009] discussed the importance of the ratio between the scale of the capture zone (induced by the pumping rate of the well) and the scale of the source zone on determining the value of Y measurements in reducing the overall risk uncertainty. The role of plume ergodicity in site characterization was also analyzed within the context of human health risk [*Maxwell et al*., 1999; *de Barros et al*., 2009]. It was shown that under nonergodic transport conditions, uncertainty reduction in human health risk benefited more from hydrogeological data acquisition as opposed to additional characterization from the human health-related parameters [*de Barros et al*., 2009].

The need for additional site characterization data will also depend on the type of adverse health effect caused by the contaminant and the manner in which the concentration is being modeled (resident vs flux-averaged) [*de Barros and Rubin*, 2008]. If the prediction of interest is chronic health effects due to long exposure periods, then global features of the breakthrough curve may suffice. On the other hand, if one is interested in acute effects in short exposure periods (e.g., such as a neurotoxin), data acquisition campaigns aimed at characterizing local fluctuations of the conductivity field and solute transport pathways may improve the prediction of peak concentrations at the environmentally sensitive target. The key point is that site characterization efforts will depend on the prediction goal upon which decisions will be made [see discussions in *de Barros and Rubin*, 2008; *de Barros et al*., 2009; *Nowak et al*., 2010].

7. Summary

The present article can be regarded as a follow up of the review of the first decade of stochastic modeling of flow and transport [*Dagan*, 1986], published in "Trends and Directions in Hydrology" special issue of WRR. It follows the same basic tenets: aquifer logconductivity Y is modeled as a stationary random space function; its univariate pdf and two point covariance are determined by field characterization; the velocity field is determined by solving the stochastic flow equations, for a mean uniform head gradient; transport is modeled in the Lagrangian framework (particle tracking); the solution for the solute concentration serves for probabilistic risk assesment.

The present article is not a comprehensive review, but it rather reports recent advances forwarded by the authors in a few key areas, as detailed by sections in the following.

Section 2 recapitulates the basic approach and serves as starting point for the rest. It emphasizes quantification of advective transport of plumes by the BTC (breakthrough curve) at control planes, in terms of fluid particles travel time, besides the usual spatial mass distribution measures. It makes a clear distinction between two notions, which sometimes lack a clear definition in the literature: on one hand, the *local* stochastic ADE at Darcy's scale, with advection driven by the random Eulerian velocity field and mixing modeled by pore scale dispersion, and on the other hand the deterministic *upscaled* ADE with advection by the constant mean velocity and spreading by macrodispersion. The solution of the latter leads to Gaussian plumes and is valid under limiting conditions, the main one being weak heterogeneity ($\sigma_Y^2 \lesssim 1$). Another important well known property revealed by the first order analysis is that longitudinal macrodispersivity depends, besides σ_Y^2, on the longitudinal integral scale of Y, and to a lesser extent on the shape ot the two point covariance or the anisotropy ratio (this reminds the dependence of spreading in turbulence upon the large eddies).

Section 3 deals with transport modeling in highly heterogeneous aquifers ($\sigma_Y^2 > 1$), topic of intensive recent research. It is focussed on the results obtained by the authors for the aquifer structure coined as MIM (Multi Indicator Model): an ensemble of rectangular blocks of independent random Y and longitudinal size $2I$. It can

model heterogeneous aquifers of arbitrary Y univariate pdf and integral scale I. An approximate analytical solution of flow and transport is obtained by the SCA (self-consistent approximation), as superposition of solutions for isolated blocks surrounded by a homogeneous matrix of effective properties. The MIM-SCA model is applied to the MADE experiment in a predictive mode, by using the field identified Y parameters and mean head gradient. The predicted plumes longitudinal mass distribution agree satisfactorily with the measured ones, with no parameter fitting. The MIM-SCA model provides insight into a few major features like the non-Gaussian, highly skewed BTC, with a long tail caused by the large residence time in low conductivity blocks. It reveals the presence of flow induced high velocity channels which carry more mass than their relative volume, explaining fast arrival at the control plane. Pore scale dispersion, which has negligible effect on the BTC for weak heterogeneity, impacts the tail by transferring solute from the low conductivity blocks.

Section 4 extends the model to transport of reactive solutes. The main approach, coined as LaSAR (Lagrangian Stochastic Advection-Reaction), incorporates in the BTC both advection (in terms of the random fluid particles travel time) and reaction by a time dependent source term. The most general considered model for reaction is linear kinetic of first-order, multirate mass transfer. The latter is charachterized by a few parameters: the ratio between the total immobile to mobile volumes and a pdf continuous distribution of transfer rates between sites, modeled as a truncated power-law in form of a Pareto type I distribution. An analytical solution is obtained in the Laplace domain which is illustrated by a few applications: the impact of different parameters (spreading by macrodispersivity and the reaction parameters) on the BTC, similarly for BTC moments and the attenuation of plume mass by decay added to the other mechanisms.

Section 5 deals with modeling of local concentration C, as affected by dilution caused by pore scale dispersion (primarily the transverse one). While the pdf of C is bimodal and of maximal variance at the source, it tends asymptotically to to a Gaussian one, governed by pore scale dispersion, at very large travel times. A Beta distribution is proposed to describe the transition, with parameters depending on the C mean and variance, which in turn depend on advective spreading and pore scale dispersion. Additional measures to characterize dilution are the relative volume exceeding a given threshold and the dilution index. The latter are computed approximately and the outcome is compared favorably with the detailed measurements at the Cape Cod site. These are advances in a topic which is undergoing active research.

The last section 6, integrates the results of the former ones, by examining the impact of contaminant concentration upon risk to human health. The central topic is the dependence of the probability of increased lifetime cancer risk (ILCR) upon concentration. The uncertainty stems from both the one affecting concentration, solution of transport, and the parameters quantifying health risk. By adopting a linear model between ILCR and threshold concentration, the expression of the ILCR cumulative probability distribution (CDF) is considered in a few applications. Thus, one of major interest, is the dependence of the CDF on σ_Y^2 via the rate of spreading as characterized by macrodispersivity, justifying the effort invested in site characterization and development of transport models. Other examined applications are the impact of the sampled volume of contaminated water upon risk uncertainty as well as the reduction of the latter by conditioning on site data collection.

The partial list of topics examined in the paper, with a selective presentation reflecting the authors own works, underscores the tremendous progress achieved by stochastic modeling in the last 30 years. While all the sections bring up recent advances, it is clear that none can be viewed as concluding the topic.

It seems to us that future work has to be concentrated in two main directions. The first one is to integrate the accumulated considerable knowledge, as illustrated by the present study, and available in many others, toward application by practitioners. The other one is further advancement by research on the many issues which are still open, a few of them indicated by the developments presented by this article.

Acknowledgments
The contribution by VC was supported in part by the Swedish Nuclear Fuel and Waste Management Co. (SKB), Stockholm, Sweden. All data for this paper are properly cited and referred to in the reference list.

References

Adams, E. E., and L. W. Gelhar (1992), Field study of dispersion in a heterogeneous aquifer: 2. Spatial moments analysis, *Water Resour. Res.*, *28*(12), 3293–3307.
Andričević, R., and V. Cvetković (1996), Evaluation of risk from contaminants migrating by groundwater, *Water Resour. Res.*, *32*(3), 611–621.
Andricevic, R., and V. Cvetkovic (1998), Relative dispersion for solute flux in aquifers, *J. Fluid Mech.*, *361*, 145–174.
Bellin, A., and D. Tonina (2007), Probability density function of non-reactive solute concentration in heterogeneous porous formations, *J. Contam. Hydrol.*, *94*(1), 109–125.

Bellin, A., P. Salandin, and A. Rinaldo (1992), Simulation of dispersion in heterogeneous porous formations: Statistics, first-order theories, convergence of computations, *Water Resour. Res.*, *28*(9), 2211–2227.
Bellin, A., Y. Rubin and A. Rinaldo (1994), Eulerian-Lagrangian approach for modeling of flow and transport in heterogeneous geological formations, *Water Resour. Res.*, *30*(11), 2913–2924.
Benson, D., S. Wheatcraft, and M. Meerschaert (2000), Application of a fractional advection-dispersion equation, *Water Resour. Res.*, *36*(6), 1403–1412.
Berglund, S. and A. Fiori (1997), Influence of transverse mixing on the breakthrough of sorbing solute in a heterogeneous aquifer, *Water Resour. Res.*, *33*(3), 399–405.
Berkowitz, B., and H. Scher (1995), On characterization of anomalous dispersion in porous and fractured media, *Water Resour. Res.*, *31*(6), 1461–1466.
Boggs, J. M., S. C. Young, L. M. Beard, L. W. Gelhar, K. R. Rehfeldt and E. E. Adams (1992), Field study of dispersion in a heterogeneous aquifer: 1. Overview and site description, *Water Resour. Res.*, *28*(12), 3281–3291.
Bohling, G. C., G. Liu, S. J. Knobbe, E. C. Reboulet, D. W. Hyndman, P. Dietrich, and J. J. Butler (2012), Geostatistical analysis of centimeter-scale hydraulic conductivity variations at the made site, *Water Resour. Res.*, *48*, W02525, doi:10.1029/2011WR010791.
Boso, F., F. P. J. de Barros, A. Fiori and A. Bellin (2013), Performance analysis of statistical spatial measures for contaminant plume characterization toward risk-based decision making, *Water Resour. Res.*, *49*, 3119–3132, doi:10.1002/wrcr.20270.
Boso, F., S. V. Broyda, and D. M. Tartakovsky (2014), Cumulative distribution function solutions of advection-reaction equations with uncertain parameters, *Proc. R. Soc. A*, *470*(2166), doi:10.1098/rspa.2014.0189.
Bouchaud, J.-P., and A. Georges (1990), Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications, *Phys. Rep.*, *195*(4), 127–293.
Carrera, J., X. Sanchez-Vila, I. Benet, A. Medina, G. Galarza, and G. Guimera (1998), On matrix diffusion: Formulations, solutions and qualitative effects, *Hydrogeol. J.*, *6*, 178–190.
Cushman, J., and T. Ginn (1993), On dispersion in fractal porous media, *Water Resour. Res.*, *29*(10), 3513–3515.
Cushman, J., and T. Ginn (2000), Fractional advection-dispersion equation: A classical mass balance with convolution-Fickian flux, *Water Resour. Res.*, *36*(12), 3763–3766.
Cvetkovic, V. (1997), Transport of reactive solutes, in *Subsurface Flow and Transport: The Stochastic Approach*, edited by G. Dagan and S. P. Neuman, pp. 133–145, Cambridge Univ. Press, Mass.
Cvetkovic, V. (2011a), Tracer attenuation in groundwater, *Water Resour. Res.*, *47*, W12541, 2057–2067, doi:10.1029/2011WR010999.
Cvetkovic, V. (2011b), The tempered one-sided stable density: A universal model for hydrological transport?, *Environ. Res. Lett.*, *6*, 034008, doi:10.1088/1748-9326/6/3/034008.
Cvetkovic, V. (2012), A general memory function for modeling mass transfer in groundwater transport, *Water Resour. Res.*, *48*, W04528, doi:10.1029/2011WR011657.
Cvetkovic, V., and G. Dagan (1994), Transport of kinetically sorbing solute by steady random velocity in heterogeneous porous formations, *J Fluid Mech.*, *265*, 189–215.
Cvetkovic, V., and R. Haggerty (2002), Transport with exchange in disordered media., *Phys. Rev. E*, *65*, 1–9.
Cvetkovic, V., and A. Shapiro (1990), Mass arrival of sorptive solute in heterogeneous porous media, *Water Resour. Res.*, *26*(9), 2057–2067.
Cvetkovic, V., A. Shapiro, and G. Dagan (1992), A solute flux approach to transport in heterogeneous formations: 2. Uncertainty analysis, *Water Resour. Res.*, *28*(5), 1377–1388.
Cvetkovic, V., A. Fiori and G. Dagan (2014), Solute transport in aquifers of arbitrary variability: A time-domain random walk formulation, *Water Resour. Res.*, *50*, 5759–5773, doi:10.1002/2014WR015449.
Dagan, G. (1979), Models of groundwater flow in statistically homogeneous porous formations, *Water Resour. Res.*, *15*(1), 47–63.
Dagan, G. (1982a), Stochastic modeling of groundwater flow by unconditional and conditional probabilities: 1. Conditional simulation and the direct problem, *Water Resour. Res.*, *18*(4), 813–833.
Dagan, G. (1982b), Stochastic modeling of groundwater flow by unconditional and conditional probabilities: 2. The solute transport, *Water Resour. Res.*, *18*(4), 835–848.
Dagan, G. (1984), Solute transport in heterogeneous porous formations, *J Fluid Mech.*, *145*, 151–177.
Dagan, G. (1986), Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale, *Water Resour. Res.*, *22*(9S), 120S–134S.
Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer, Berlin.
Dagan, G. and V. Cvetkovic (1996), Reactive transport and immiscible flow in geological media. i. general theory, *Proc. R. Soc. London, Ser. A*, *452*, 1471–1480.
Dagan, G. and A. Fiori (2003), Time-dependent transport in heterogeneous formations of bimodal structures: 1. The model, *Water Resour. Res.*, *39*(5), 1112, doi:10.1029/2002WR001396.
Dagan, G., V. Cvetkovic and A. Shapiro (1992), A solute flux approach to transport in heterogeneous formations 1. The general framework, *Water Resour. Res.*, *28*(5), 1369–1376.
Dagan, G., I. Butera and E. Grella (1996), Impact of concentration measurements upon estimation of flow and transport parameters: The Lagrangian approach, *Water Resour. Res.*, *32*(2), 297–306, doi:10.1029/95WR02716.
Dagan, G., A. Fiori and I. Janković (2003), Flow and transport in highly heterogeneous formations: 1. Conceptual framework and validity of first-order approximations, *Water Resour. Res.*, *39*(9), 1268, doi:10.1029/2002WR001717.
de Barros, F. P. J., and A. Fiori (2014), First-order based cumulative distribution function for solute concentration in heterogeneous aquifers: Theoretical analysis and implications for human health risk assessment, *Water Resour. Res.*, *50*, 4018–4037, doi:10.1002/2013WR015024.
de Barros, F. P. J., and Y. Rubin (2008), A risk-driven approach for subsurface site characterization, *Water Resour. Res.*, *44*, W01414, doi:10.1029/2007WR006081.
de Barros, F. P. J., Y. Rubin and R. Maxwell (2009), The concept of comparative information yield curves and its application to risk-based site characterization, *Water Resour. Res.*, *45*, W06401, doi:10.1029/2008WR007324.
de Barros, F. P. J., S. Ezzedine, and Y. Rubin (2012), Impact of hydrogeological data on measures of uncertainty, site characterization and environmental performance metrics, *Adv. Water Resour.*, *36*, 51–63.
de Barros, F. P. J., A. Fiori, F. Boso and A. Bellin (2015), A theoretical framework for modeling dilution enhancement of non-reactive solutes in heterogeneous porous media, *J. Contam. Hydrol.*, *175–176*, 72–83.
Demmy, G., S. Berglund, and W. Graham (1999), Injection mode implications for solute transport in porous media: Analysis in a stochastic Lagrangian framework, *Water Resour. Res.*, *35*(7), 1965–1973.

Deng, F. W., J. H. Cushman, and J. W. Delleur (1993), Adaptive estimation of the log fluctuating conductivity from tracer data at the Cape Cod site, *Water Resour. Res.*, *29*(12), 4011–4018, doi:10.1029/93WR02480.

Dentz, M., and D. Tartakovsky (2010), Probability density functions for passive scalars dispersed in random velocity fields, *Geophys. Res. Lett.*, *37*, L24406, doi:10.1029/2010GL045748.

Dykaar, B. B., and P. K. Kitanidis (1992), Determination of the effective hydraulic conductivity for heterogeneous porous media using a numerical spectral approach: 2. Results, *Water Resour. Res.*, *28*(4), 1155–1166.

Fiori, A. (1996), Finite-Peclet extensions of *Dagan*'s solutions to transport in anisotropic heterogeneous formations, *Water Resour. Res.*, *32*(1), 193–198.

Fiori, A. (2001), The Lagrangian concentration approach for determining dilution in aquifer transport: Theoretical analysis and comparison with field experiments, *Water Resour. Res.*, *37*(12), 3105–3114.

Fiori, A. (2014), Channeling, channel density and mass recovery in aquifer transport, with application to the made experiment, *Water Resour. Res.*, *50*, 9148–9161, doi:10.1002/2014WR015950.

Fiori, A. and G. Dagan (2000), Concentration fluctuations in aquifer transport: A rigorous first-order solution and applications, *J. Contam. Hydrol.*, *45*(1), 139–163.

Fiori, A. and G. Dagan (2003), Time-dependent transport in heterogeneous formations of bimodal structures: 2. Results, *Water Resour. Res.*, *39*(5), 1125, doi:10.1029/2002WR001398.

Fiori, A. and I. Jankovic (2012), On preferential flow, channeling and connectivity in heterogeneous porous formations, *Math. Geosci.*, *44*(2), 133–145.

Fiori, A., S. Berglund, V. Cvetkovic and G. Dagan (2002), A first-order analysis of solute flux statistics in aquifers: The combined effect of pore-scale dispersion, sampling, and linear sorption kinetics, *Water Resour. Res.*, *38*(8), 1137, doi:10.1029/2001WR000678.

Fiori, A., I. Janković and G. Dagan (2003), Flow and transport in highly heterogeneous formations: 2. Semianalytical results for isotropic media, *Water Resour. Res.*, *39*(9), 1269, doi:10.1029/2002WR001719.

Fiori, A., I. Janković and G. Dagan (2006), Modeling flow and transport in highly heterogeneous three-dimensional aquifers: Ergodicity, Gaussianity, and anomalous behavior 2. Approximate semianalytical solution, *Water Resour. Res.*, *42*, W06D13, doi:10.1029/2005WR004752.

Fiori, A., I. Janković, G. Dagan and V. Cvetković (2007), Ergodic transport through aquifers of non-Gaussian log conductivity distribution and occurrence of anomalous behavior, *Water Resour. Res.*, *43*, W09407, doi:10.1029/2007WR005976.

Fiori, A., I. Jankovic and G. Dagan (2011), The impact of local diffusion upon mass arrival of a passive solute in transport through three-dimensional highly heterogeneous aquifers, *Adv. Water Resour.*, *34*(12), 1563–1573.

Fiori, A., G. Dagan, I. Jankovic and A. Zarlenga (2013), The plume spreading in the made transport experiment: Could it be predicted by stochastic models?, *Water Resour. Res.*, *49*, 2497–2507, doi:10.1002/wrcr.20128.

Fiori, A., E. Volpi, A. Zarlenga, and G. C. Bohling (2015), Gaussian or non-Gaussian log conductivity distribution at the MADE site: What is the impact on the breakthrough curve?, *J. Contam. Hydrol.*, *179*, 23–34, doi:10.1016/j.jconhyd.2015.05.004.

Fiorotto, V., and E. Caroni (2002), Solute concentration statistics in heterogeneous aquifers for finite peclet values, *Transp. Porous Media*, *48*(3), 331–351, doi:10.1023/A:1015744421033.

Freeze, R. (1975), A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res.*, *11*(5), 725–741.

Garabedian, S. P., D. R. LeBlanc, L. W. Gelhar and M. A. Celia (1991), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 2. Analysis of spatial moments for a nonreactive tracer, *Water Resour. Res.*, *27*(5), 911–924, doi:10.1029/91WR00242.

Gelhar, L. W. (1986), Stochastic subsurface hydrology from theory to applications, *Water Resour. Res.*, *22*(9S), 135S–145S.

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

Graham, W., and D. McLaughlin (1989), Stochastic analysis of nonstationary subsurface solute transport: 2. Conditional moments, *Water Resour. Res.*, *25*(11), 2331–2355, doi:10.1029/WR025i011p02331.

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

Harvey, C., and S. M. Gorelick (2000), Rate-limited mass transfer or macrodispersion: Which dominates plume evolution at the macrodispersion experiment (made) site?, *Water Resour. Res.*, *36*(3), 637–650.

Hassan, A., R. Andricevic, and V. Cvetkovic (2001), Computational issues in the determination of solute discharge moments and implications for comparison to analytical solutions, *Adv. Water Resour.*, *24*(6), 607–619.

Hess, K. M., S. H. Wolf, and M. A. Celia (1992), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 3. Hydraulic conductivity variability and calculated macrodispersivities, *Water Resour. Res.*, *28*(8), 2011–2027, doi:10.1029/92WR00668.

Janković, I. and A. Fiori (2010), Analysis of the impact of injection mode in transport through strongly heterogeneous aquifers, *Adv. Water Resour.*, *33*(10), 1199–1205.

Janković, I., A. Fiori and G. Dagan (2003), Flow and transport in highly heterogeneous formations: 3. Numerical simulations and comparison with theoretical results, *Water Resour. Res.*, *39*(9), 1270, doi:10.1029/2002WR001721.

Janković, I., A. Fiori and G. Dagan (2006), Modeling flow and transport in highly heterogeneous three-dimensional aquifers: Ergodicity, Gaussianity, and anomalous behavior 1. Conceptual issues and numerical simulations, *Water Resour. Res.*, *42*, W06D12, doi:10.1029/2005WR004734.

Janković, I., A. Fiori and G. Dagan (2009), The impact of local diffusion on longitudinal macrodispersivity and its major effect upon anomalous transport in highly heterogeneous aquifers, *Adv. Water Resour.*, *32*(5), 659–669.

Jankovic, I., A. Fiori and G. Dagan (2013), Effective conductivity of isotropic highly heterogeneous formations: Numerical and theoretical issues, *Water Resour. Res.*, *49*, 1178–1183, doi:10.1029/2012WR012441.

Kapoor, V., L. W. Gelhar, and F. Miralles-Wilhelm (1997), Bimolecular second-order reactions in spatially varying flows: Segregation induced scale-dependent transformation rates, *Water Resour. Res.*, *33*(4), 527–536, doi:10.1029/96WR03687.

Kitanidis, P. K. (1994), The concept of the dilution index, *Water Resour. Res.*, *30*(7), 2011–2026, doi:10.1029/94WR00762.

Klemeš, V. (1983), Conceptualization and scale in hydrology, *J. Hydrol.*, *65*(1), 1–23.

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. Sci.*, *33*(11), 1471–1480.

LeBlanc, D. R., S. P. Garabedian, K. M. Hess, L. W. Gelhar, R. D. Quadri, K. G. Stollenwerk, and W. W. Wood (1991), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 1. Experimental design and observed tracer movement, *Water Resour. Res.*, *27*(5), 895–910, doi:10.1029/91WR00241.

Malmstrom, G. D., M., and P. Martinet (2004), Modeling expected solute concentration in randomly heterogeneous flow systems with multi-component reactions, *Environ. Sci. Technol.*, *38*.

Maxwell, R., S. Carle, and A. Tompson (2008), Contamination, risk, and heterogeneity: On the effectiveness of aquifer remediation, *Environ. Geol.*, *54*(8), 1771–1786.

Maxwell, R. M., and W. E. Kastenberg (1999), Stochastic environmental risk analysis: An integrated methodology for predicting cancer risk from contaminated groundwater, *Stochastic Environ. Res. Risk Assess.*, *13*(1), 27–47.

Maxwell, R. M., W. E. Kastenberg, and Y. Rubin (1999), A methodology to integrate site characterization information into groundwater-driven health risk assessment, *Water Resour. Res.*, *35*(9), 2841–2885.

McKone, T., and K. Bogen (1991), Predicting the uncertainties in risk assessment, *Environ. Sci. Technol.*, *25*(10), 1674–1681.

Molin, S., and V. Cvetkovic (2010), Microbial risk assessment in heterogeneous aquifers: 1. Pathogen transport, *Water Resour. Res.*, *46*, W05518, doi:10.1029/2009WR008036.

Molin, S., V. Cvetkovic, and T. Stenström (2010), Microbial risk assessment in heterogeneous aquifers: 2. Infection risk sensitivity, *Water Resour. Res.*, *46*, W05519, doi:10.1029/2009WR008039.

Noolandi, J. (1977), Multiple-trapping model of anomalous transit-time dispersion in a-se, *Phys. Rev. B*, *16*, 4466–4473.

Nowak, W., F. de Barros, and Y. Rubin (2010), Bayesian geostatistical design: Task-driven optimal site investigation when the geostatistical model is uncertain, *Water Resour. Res.*, *46*, W03535, doi:10.1029/2009WR008312.

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

Rubin, Y. (2003), *Applied Stochastic Hydrogeology*, Oxford Univ. Press, Oxford, U. K.

Rubin, Y. and G. Dagan (1992), Conditional estimation of solute travel time in heterogeneous formations: Impact of transmissivity measurements, *Water Resour. Res.*, *28*(4), 1033–1040.

Rubin, Y., M. Cushey and A. Bellin (1994), Modeling of transport in groundwater for environmental risk assessment, *Stochastic Hydrol. Hydraul.*, *8*(1), 57–77.

Schmidlin, F. W. (1977), Theory of trap-controlled transient photoconduction, *Phys. Rev. B*, *16*, 2362–2385.

Schwede, R. L., O. A. Cirpka, W. Nowak, and I. Neuweiler (2008), Impact of sampling volume on the probability density function of steady state concentration, *Water Resour. Res.*, *44*, W12433, doi:10.1029/2007WR006668.

Tartakovsky, D. (2007), Probabilistic risk analysis in subsurface hydrology, *Geophys. Res. Lett.*, *34*, L05404, doi:10.1029/2007GL029245.

Tartakovsky, D. (2013), Assessment and management of risk in subsurface hydrology: A review and perspective, *Adv. Water Resour.*, *51*, 247–260.

Thierrin, J. and P. K. Kitanidis (1994), Solute dilution at the borden and cape cod groundwater tracer tests, *Water Resour. Res.*, *30*(11), 2883–2890, doi:10.1029/94WR01983.

Tonina, D., and A. Bellin (2008), Effects of pore-scale dispersion, degree of heterogeneity, sampling size, and source volume on the concentration moments of conservative solutes in heterogeneous formations, *Adv. Water Resour.*, *31*(2), 339–354.

USEPA (1989), Risk assessment guidance for Superfund Volume I: Part A, human health manual, *Tech. Rep. Rep.EPA/540/1–89/002*.

USEPA (2001), Risk assessment guidance for Superfund: Volume III—Part A, process for conducting probabilistic risk assessment, *Tech. Rep. Rep.EPA 540/R-02/002*.

Villermaux, J. (1974), Deformation of chromatographic peaks under the influence of mass transfer phenomenon, *J. Chromatogr. Sci.*, *12*, 822–831.