



Mixing in Multidimensional Porous Media: A Numerical Study of the Effects of Source Configuration and Heterogeneity

Alessandra Bonazzi¹ · Marco Dentz² · Felipe P. J. de Barros¹ 

Received: 27 August 2021 / Accepted: 21 June 2022 / Published online: 8 August 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

We investigate transport of an inert solute in multidimensional porous media characterized by spatially variable hydraulic conductivity. Through the use of a GPU-accelerated solute transport simulator based on the Random Walk Particle Tracking technique, we show how different factors such as the degree of heterogeneity, flow dimensionality and source zone configurations impact mixing. Solute mixing is quantified in terms of the temporal evolution of the plume's statistics (mean, variance and probability density function) and the dilution index. Our analysis shows that mixing is strongly affected by the above mentioned factors. We also compare the probability distributions obtained from the numerical simulations with the beta distribution. Despite the discrepancies at very low concentrations, our results show that the fitting with the beta distribution is improved for transport in three-dimensional settings originating from a vertical planar source. To improve the fit at low concentrations, we utilize two one-to-one variable transformation, namely the logarithm and power law transformations. Results demonstrate that the Pareto-type IV and the uniform distributions are capable to capture the lower tail of the cumulative distribution function. Numerical results are validated against existing analytical solution for both homogeneous and heterogeneous media.

Keywords Heterogeneity · Mixing · Random walk particle tracking · Stochastic hydrogeology · Computational modeling · Porous media

1 Introduction

The spatial variability of the hydraulic conductivity in porous formations leads to complex flow patterns which in turn lead to mixing dynamics that differ from the ones observed under uniform flow conditions. Transport in heterogeneous porous media is characterized by early solute breakthrough, tailing behavior of the concentration at late times and irregular spreading and mixing rates (Neuman and Tartakovsky 2009; Dentz et al. 2011; Fiori

✉ Felipe P. J. de Barros
fbarros@usc.edu

¹ Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Kaprielian Hall, 3620 S. Vermont Avenue, Los Angeles, CA 90089-2531, USA

² Spanish National Research Council (IDAEA-CSIC), 08034 Barcelona, Spain

et al 2015). The presence of heterogeneity is also responsible for poorly mixed conditions and spatially dispersed solute plumes (Dentz et al. 2011). The interplay between advective and local-scale dispersive fluxes in heterogeneous porous media results in non-trivial macroscopic transport behavior which is of importance for a broad range of scientific fields and engineering applications (Sahimi 2011). Examples consist of contaminant migration in hydrology (Dagan and Neuman 2005), aquifer remediation (Chapman and Parker 2005), human health-risk assessment (Im et al. 2020), geological storage of CO₂ (Hidalgo and Carrera 2009) and safety assessment of waste repositories (Selroos 1997). Improved understanding of the effects of heterogeneity on mixing is imperative to improve the predictive capabilities of models in the above-mentioned applications.

Many works focused on understanding the role of conductivity heterogeneity on temporal scaling properties of solute mixing. Mixing in heterogeneous porous media has been analyzed in terms of effective dispersion coefficients (Dentz et al. 2000; Fiori and Dagan 2000; Fiori 2001), entropy-based mixing metrics (Kitanidis 1994; de Barros et al. 2015) and the temporal decay of the concentration variability (Kapoor and Kitanidis 1998). Through the use of numerical simulations, Le Borgne et al. (2010) showed how moderate-to-strong levels of heterogeneity induced anomalous temporal scaling for the scalar dissipation rate of a solute plume originating from a line source in a two-dimensional porous medium. Other studies investigated the impact of the sequence of fluid deformation events in two-dimensional flows on mixing metrics (de Barros et al. 2012; Le Borgne et al. 2013, 2015). The effects of flow focusing on the transverse dilution behavior of steady-state plumes have been topic of numerical (de Barros and Nowak 2010; Cirpka et al. 2011) and experimental investigations (Rolle et al. 2009; Gueting and Englert 2013). Kapoor and Kitanidis (1998) analyzed the rate of destruction of the concentration variance by local-scale dispersion through the use of numerical simulations and approximate analytical solutions. Dentz et al. (2018) showed the importance of the initial condition of the solute plume (i.e., its spatial distribution) on the temporal mixing evolution in two-dimensional porous media. Semi-analytical solutions for the statistical description of the concentration field in spatially heterogeneous porous formations are also reported in the literature (Rubin et al. 1994; Fiori and Dagan 2000; Fiori 2001; Tonina and Bellin 2008; Meyer et al. 2010; Dentz and Tartakovsky 2010; Dentz 2012; de Barros and Fiori 2014). Approximate semi-analytical solutions for the dilution index, introduced by Kitanidis (1994), in two- and three-dimensional porous formations are reported in de Barros et al. (2015) and compared with results from numerical simulators (Boso et al. 2013a; de Barros et al. 2015) and field data (de Barros et al. 2015; Soltanian et al. 2020; de Barros and Fiori 2021).

In general, the above-mentioned semi-analytical solutions are based on perturbation theory, and therefore restricted to low-to-moderate levels of hydraulic conductivity heterogeneity. To address mixing in porous formations displaying a stronger degree of heterogeneity (i.e., log-conductivity variance larger than unity), numerical methods are needed. However, traditional Eulerian grid-based numerical approaches are often plagued by oscillations and numerical dispersion which impact the accuracy of the numerical scheme (Zheng et al. 2002; Ferziger et al. 2002; Gotovac et al. 2007). Furthermore, in order to capture the effects of small scale heterogeneity on solute mixing, a fine spatial resolution in the numerical model is required which increases the computational burden associated with flow and transport simulations (Ababou et al. 1989; Bellin et al. 1993). This computational burden is augmented within the context of uncertainty quantification where a Monte Carlo framework is needed (Moslehi et al. 2015). Lagrangian-based methods, such as the Random Walk Particle Tracking (RWPT) technique (Salamon et al. 2006), are an appealing alternative since they are globally mass conservative and no subject to artificial oscillation and numerical dispersion. RWPT was used

to study dispersion in both two- and three-dimensional heterogeneous porous media (Bellin et al. 1992; de Dreuzy et al. 2007; Beaudoin and de Dreuzy 2013). Jankovic et al. (2017) used RWPT to examine the impact of conductivity structure in the mass breakthrough curve in a three-dimensional setting. Sole-Mari et al. (2021), used RWPT simulations to key metrics of transport originating from a source zone occupying the entire inlet area of a bounded three-dimensional porous media displaying non-Gaussian features and a logconductivity variance of unity. Libera et al. (2019) used RWPT to analyze the joint effects of porosity and conductivity variability on both the peak flux-averaged concentration and solute arrival times in three-dimensional porous media for low and high levels of heterogeneity. Using different numerical schemes, including RWPT and smoothed particle hydrodynamics (SPH), Boso et al. (2013b) investigated the impact of heterogeneity in transport metrics such as the second central spatial moments of the solute plume and the dilution index. Results obtained from different numerical schemes were compared and reported for a point-like injection and two-dimensional domains for different logconductivity variances (ranging from 0.2 to 10). Despite the benefits associated with RWPT, there are drawbacks such as the presence of local concentration fluctuations and the need to have a significant number of particles to achieve numerical precision of the concentration field, especially when dealing with heterogeneous porous media (Herrera et al. 2009; Boso et al. 2013b). The need to use a large number of particles has implications on the computational costs associated with transport.

To overcome this challenge and improve the efficiency associated with Lagrangian-based methods, Rizzo et al. (2019) introduced a GPU-accelerated RWPT, denoted as PAR², that 1) enables the use of a large number of particles and 2) is computationally efficient. PAR² has been employed to study hydrogeological connectivity in both Gaussian and non-Gaussian flow fields (Rizzo and de Barros 2019; Morvillo et al. 2021a), aquifer resilience loss and probabilistic-risk analysis (Morvillo et al. 2022), solute transport at the pore-scale (Kamrava et al. 2021) and has been expanded to account for chemical reactions (Morvillo et al. 2021b). The code is open source (see details in Rizzo et al. (2019)) and a step-by-step tutorial on its use and how to connect to existing groundwater flow simulation tools can be found in Morvillo et al. (2022).

The objectives of this work is to provide a systematic numerical investigation of the impact of flow dimensionality, source zone configuration and the degree of heterogeneity on metrics of solute mixing. To achieve our goals, we rely on the RWPT-based simulator PAR² (Rizzo et al. 2019) to show the importance of the aforementioned factors in mixing metrics such as the global spatial mean and variance of the concentration field, the dilution index and the concentration probability distribution function. Results are compared to existing analytical solutions for homogeneous and heterogeneous porous media.

2 Problem Statement

2.1 Flow and Transport Model

We consider a d -dimensional porous medium with constant porosity ϕ and spatially variable (locally isotropic) hydraulic conductivity $K(\mathbf{x})$ with $\mathbf{x} = [x_1, \dots, x_d]^T$ denoting the Cartesian coordinate system and $d = 2$ and 3. The flow field is given by

$$\nabla \cdot \mathbf{q}(\mathbf{x}) = 0 \quad (1)$$

where the specific discharge $\mathbf{q}$ is obtained through Darcy's law

$$\mathbf{q}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x}), \quad (2)$$

with h representing the hydraulic head. Equation (1) is subject to the following boundary conditions: prescribed hydraulic head along the longitudinal direction x_1 and no-flow boundary conditions along directions x_j for $j = 2$ and 3 . These conditions ensure that flow is uniform-in-the-mean along the longitudinal direction. That implies that the mean velocity field is $\langle \mathbf{u} \rangle = (U, 0)$ for $d = 2$ and $\langle \mathbf{u} \rangle = (U, 0, 0)$ for $d = 3$. Here the angled brackets represent the ensemble expected value and U denotes the mean longitudinal velocity given by $U = K_G \mathcal{J} / \phi$ where K_G is the geometric mean of the hydraulic conductivity and $\mathcal{J}$ is the spatially uniform mean hydraulic head gradient.

An inert solute is instantaneously released in a domain $\mathcal{V}_o$ located within the porous formation. Depending on the dimensionality of the flow field, the domain $\mathcal{V}_o$ can represent a volume or an area. In this work, we will consider different geometrical configurations for the source zone. The geometrical configurations considered are point, line and planar sources in both two- and three dimensional porous formations. Details pertaining the dimensions of the solute injection zone are provided further below in Sect. 5. The spatiotemporal dynamics of the solute plume's resident concentration c is assumed to be governed by the local advection dispersion equation (ADE):

$$\frac{\partial c(\mathbf{x}, t)}{\partial t} + \mathbf{u}(\mathbf{x}) \cdot \nabla c(\mathbf{x}, t) = \nabla \cdot [\mathbf{D}(\mathbf{x}) \nabla c(\mathbf{x}, t)], \quad (3)$$

where $\mathbf{u} = \mathbf{q} / \phi$ is the velocity vector and $\mathbf{D}$ is the local scale dispersion tensor given by

$$\mathbf{D}(\mathbf{x}) = (\alpha_T |\mathbf{u}(\mathbf{x})| + D_m) \mathbb{1} + \frac{(\alpha_L - \alpha_T)}{|\mathbf{u}(\mathbf{x})|} \mathbf{u}(\mathbf{x}) \mathbf{u}(\mathbf{x})^T \quad (4)$$

with D_m denoting the molecular diffusion coefficient and $\mathbb{1}$ representing the identity matrix. Equation (3) is subject to the following initial condition

$$c(\mathbf{x}, 0) = \begin{cases} c_o & \text{if } \mathbf{x} \in \mathcal{V}_o; \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Due to the spatial randomness of the hydraulic conductivity, the velocity field is spatially variable which impacts solute mixing and spreading rates. In the following subsection, we will describe the details regarding the spatial structure of the K field.

2.2 Stochastic Model

To simulate flow and transport in a spatially random porous medium, we assume that the log-conductivity field, namely $f(\mathbf{x}) = \ln K(\mathbf{x})$, is multivariate Gaussian and statistically stationary. Therefore, f is fully characterized by i) its mean value $\langle f \rangle = \ln K_G$ where K_G is the geometric mean of the conductivity field K_G and ii) its spatial covariance $C_{ff}(\mathbf{r}) = \langle f(\mathbf{x}) f(\mathbf{x}') \rangle$ with $\mathbf{r} = \mathbf{x} - \mathbf{x}'$. The logconductivity variance is given by $\sigma_f^2 \equiv C_{ff}(0)$. In this work, we consider an isotropic correlation length λ . In the following, we adopt an exponential spatial covariance model for C_{ff} (see Ch. 2 of Rubin (2003)) such that

$$C_{ff}(\mathbf{r}) = \sigma_f^2 \exp\left(-\frac{|\mathbf{r}|}{\lambda}\right). \quad (6)$$

2.3 Mixing Metrics

To quantify mixing, we will rely on four descriptors. The first descriptor is the spatially averaged concentration of the plume over volume φ_d of the d -dimensional domain characterized by $c > c^*$ where c^* is a low concentration threshold value. The volume $\varphi_d(t|c^*)$ represents a line in $d = 1$, an area for $d = 2$ and volume for $d = 3$ and is defined as follows

$$\varphi_d(t|c^*) = \int_{\Omega} \mathbb{H}[c(\mathbf{x}, t) - c^*] d\mathbf{x} \quad (7)$$

where $\mathbb{H}[\cdot]$ is the Heaviside function and Ω is the flow domain. This means, $\varphi_d(t|c^*)$ denotes the volume occupied by the solute, or, in other words, the mixing volume.

The spatially averaged concentration is defined by

$$\mu_c(t) = \frac{1}{\varphi_d(t|c^*)} \int_{\varphi_d(t|c^*)} c(\mathbf{x}, t) d\mathbf{x}. \quad (8)$$

Note that for a sufficiently small c^* , the integral on the right side may be approximated by 1 due to mass conservation. Thus, the spatial mean concentration is inversely proportional to the mixing volume $\varphi_d(t|c^*)$.

The second descriptor is the *spatial* variance of the concentration

$$\sigma_c^2(t) = \frac{1}{\varphi_d(t|c^*)} \int_{\varphi_d(t|c^*)} c(\mathbf{x}, t)^2 d\mathbf{x} - [\mu_c(t)]^2, \quad (9)$$

with μ_c given in Eq. (8).

We also examine the concentration probability density function (PDF) $p(c)$. The PDF $p(c)$ is obtained from spatially sampling the concentration point values within the domain $\varphi_d(t|c^*)$, see Eq. (7). It can be formally written as

$$p(c) = \int_{\varphi_d(t|c^*)} \delta[c - c(\mathbf{x}, t)] d\mathbf{x}, \quad (10)$$

where $\delta(c)$ is the Dirac delta distribution.

Finally, the fourth mixing metric investigated in this work is the dilution index E introduced by Kitanidis (1994). The dilution index represents a global measure of dilution. It measures the temporal evolution of the volume occupied by the solute plume and allows to quantify the combined effects of advection and local scale dispersive fluxes on mixing. The dilution index is mathematically expressed as

$$E(t) = \exp[\Lambda(t)], \text{ with} \quad \Lambda(t) = - \int_{\Omega} \frac{c(\mathbf{x}, t)}{M_o} \ln \left[\frac{c(\mathbf{x}, t)}{M_o} \right] d\mathbf{x} \quad (11)$$

where $M_o = c_o \mathcal{V}_o \phi$ is the total mass injected into the porous formation.

We will compute μ_c , σ_c^2 , p and E for different values of σ_f^2 , distinct source dimensions (point, line and planar source zones) and for different flow dimensionality ($d = 2$ and 3).

2.4 Reference Solutions for Homogeneous Media

For reference, we give in the following the expected behaviors of these mixing metrics for homogeneous media in d spatial dimensions. The concentration distribution in response to an instantaneous point source, $c(\mathbf{x}, t = 0) = c_0 \delta(\mathbf{x})$, in an infinite ν -dimensional domain is given by the Gaussian distribution

$$c(\mathbf{x}, t) = \frac{c_0}{(4\pi Dt)^{\nu/2}} \exp \left[-\frac{(\mathbf{x} - \mathbf{ut})^2}{4Dt} \right]. \quad (12)$$

Note that solution for an infinitely extended line source in $d = 2$ spatial dimensions and for an infinitely extended planar source in $d = 3$ is given by expression (12) for $\nu = 1$, the solution for an infinitely extended line source in $d = 3$ spatial dimensions is given by expression (12) for $\nu = 2$. In general, $\nu = d - d_s$, where d is the dimension of space and d_s the dimension of the source distribution. The source dimension is $d_s = 1$ for a line source and $d_s = 2$ for a planar source. In the following we rescale $c \rightarrow c/c_0$.

2.4.1 Concentration PDF

In order to determine the concentration PDF we note that the radius of a concentration isoline measured from the center of mass at $\mathbf{ut}$ is related to the concentration c by

$$r(c) = \sqrt{(4Dt) \ln[c_m(t)/c]}, \quad c_m(t) = \frac{1}{(4\pi Dt)^{\nu/2}}. \quad (13)$$

Thus, we can write Eq. (10) for the concentration PDF in spherical coordinates as

$$p(c) = \frac{d}{r(c^*)} \int_0^{r(c^*)} dr r^{\nu-1} \delta[c - f(r)], \quad f(r) = \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{\nu/2}}. \quad (14)$$

We can evaluate this integral explicitly by using that $\delta[1 - f(r)] = 1/|df/dr| \delta[r - r(c)]$. Thus, we obtain

$$p(c) = \frac{\nu}{2c} \frac{\ln[c_m(t)/c]^{(\nu-2)/2}}{\ln[c_m(t)/c^*]^{\nu/2}} \mathbb{I}(c^* < c \leq c_m(t)). \quad (15)$$

2.4.2 Concentration Moments

The concentration moments are defined in terms of the concentration PDF by

$$\mu^{(k)}(t) = \int_0^\infty dc c^k p(c). \quad (16)$$

Using the explicit expression (15) for the concentration PDF and rescaling the integration variable by $c_m(t)$, we obtain

$$\mu^{(k)}(t) = \frac{c_m(t)^k}{\ln[c_m(t)/c^*]^{\nu/2}} \frac{\nu}{2} \int_{c^*/c_m(t)}^1 dc c^{k-1} \ln(1/c)^{(\nu-2)/2}. \quad (17)$$

We consider situations for which $(c^*/c_m(t) \ll 1)$. In this limit we obtain

$$\mu^{(k)}(t) = A_k c_m(t)^k + \dots, \quad (18)$$

where the dots denote subleading contributions. The constant A_k is defined by

$$A_k = \frac{\nu}{2 \ln(1/c^*)^{\nu/2}} \int_0^1 dc c^{k-1} \ln(1/c)^{(\nu-2)/2}. \quad (19)$$

It depends on the dimension ν of space and the order k of the moment. We denote the mean concentration by $\mu^{(1)} \equiv \mu_c$.

We obtain for the mean concentration in leading order

$$\mu_c(t) = A_1 c_m(t) \sim t^{-\nu/2}, \quad (20)$$

where $\mu_c(t) \equiv \mu^{(1)}(t)$. For the concentration variance, we obtain

$$\sigma_c^2(t) = c_m(t)^2 (A_2 - A_1^2) \sim t^{-\nu}. \quad (21)$$

2.4.3 Dilution Index

The expression for the dilution index is obtained by inserting the Gaussian solution (12) into the definition (11). The explicit analytical solutions are (Kitanidis 1994)

$$E(t) = \frac{\exp(d/2)}{c_m(t)} \sim t^{\nu/2}. \quad (22)$$

Note that it scales as $1/\mu_c(t)$. This is due to the fact that the mean concentration scales as the inverse mixing volume, while the dilution index describes the evolution of the mixing volume.

3 Methods

For all results presented in Sect. 6, the spatially variable logconductivity field f is randomly generated using the SGeMS tool (Remy et al. 2009). SGeMS is based on a sequential Gaussian simulation model (Rubin 2003). The flow Eq. (1) is solved using the finite-difference solver MODFLOW (Harbaugh 2005) with the Python package FloPy (Bakker et al. 2016). The flow domain, $L_1 \times L_2$ (for $d = 2$) or $L_1 \times L_2 \times L_3$ (for $d = 3$), is discretized into a regular grid with numerical grid blocks of dimension Δ .

To simulate solute transport (3), we make use of the PAR² code developed Rizzo et al. (2019). PAR² is an open source, GPU-accelerated, simulator based on the Random Walk Particle Tracking (RWPT) (Rizzo et al. 2019). A step-by-step tutorial of PAR², and the details of how to link it to an open source Python package, is reported in Morvillo et al. (2022). PAR² has been employed to study transport at both field and pore scales (Rizzo

et al. 2019; Kamrava et al. 2021). The RWPT is based on the trajectory of the i^{th} solute particle that can be computed using the Itô-Taylor integration scheme (Salamon et al. 2006):

$$\mathbf{X}_i(t + \Delta t) = \mathbf{X}_i(t) + \mathbf{A}(\mathbf{X}_i(t))\Delta t + \mathbf{B}(\mathbf{X}_i(t)) \cdot \xi(t)\sqrt{\Delta t} \quad (23)$$

where ξ is a normally distributed random variable with zero mean and unit standard deviation, Δt is the time step, and the drift vector $\mathbf{A}$ and the displacement matrix $\mathbf{B}$ are defined by:

$$\mathbf{A}(\mathbf{x}) = \mathbf{u}(\mathbf{x}) + \nabla \cdot \mathbf{D}(\mathbf{x}) + \frac{1}{\phi} \mathbf{D}(\mathbf{x}) \cdot \nabla \phi \quad (24)$$

$$2\mathbf{D}(\mathbf{x}) = \mathbf{B}(\mathbf{x}) \cdot \mathbf{B}(\mathbf{x})^T. \quad (25)$$

4 Validation

In this section we validate the results from the numerical simulators with existing analytical solutions for two-dimensional domains. For a homogeneous conductivity field, Kitanidis (1994) derived the following expression for the dilution index for a point source (see also Eq. 22):

$$E(t) = (4\pi Dt)^{d/2} \exp\left(-\frac{d}{2}\right) \quad (26)$$

For a heterogeneous porous medium, de Barros et al. (2015) used perturbation theory to derive a semi-analytical solution for the dilution index. The semi-analytical expression is valid for a point-like instantaneous injection, uniform-in-the-mean flow conditions in the absence of sinks and sources and low-to-mild levels of heterogeneity, i.e., $\sigma_f^2 \lesssim 1$. The semi-analytical solution for the temporal evolution of the dilution index E is given by

$$E(t) = (2\pi)^{d/2} \exp\left(-\frac{d}{2}\right) \prod_{i=1}^d \sqrt{\mathcal{W}_{ii}(t)} \quad (27)$$

where $\mathcal{W}_{ii}$ is the relative particle trajectory covariance along the i^{th} direction given by Fiori (2001)

$$\mathcal{W}_{ii}(t) = \mathcal{X}_{ii}(t) + 2Dt - \mathcal{Z}_{ii}(t;0), \quad (28)$$

where $\mathcal{X}_{ii}$ and $\mathcal{Z}_{ii}$ correspond to the one- and two-particle trajectory covariances respectively (Fiori and Dagan 2000). Note that the “0” present in the two-particle trajectory covariance $\mathcal{Z}_{ii}$ is to emphasize that the function is evaluated for a point-like source, i.e., the separation distance between two particles originally located within the source zone is approximately zero (details provided in de Barros et al. (2015)). The semi-analytical expressions for the particle trajectory covariances $\mathcal{X}_{ii}$ and $\mathcal{Z}_{ii}$ were derived in Fiori and Dagan (2000) and are reproduced in Eqs. (29) and (30). These expressions are valid for low-to-mild levels of heterogeneity and uniform-in-the-mean flow conditions. For an isotropic and constant local dispersion D and point-like injection, the particle trajectories are

$$\mathcal{X}_{ii}(t) = \frac{2^{1+d/2}}{\pi^{d/2}} \int_0^\infty \int_0^t (t-\tau) \cos[k_1 U \tau] e^{-Dk^2 \tau} \hat{u}_{ii}(\mathbf{k}) d\tau d\mathbf{k}, \quad (29)$$

$$\mathcal{Z}_{ii}(t;0) = \left(\frac{2}{\pi}\right)^{d/2} \int_0^\infty \int_0^t \int_0^t \cos[k_1 U(t'-t'')] e^{-Dk^2(t'+t'')} \hat{u}_{ii}(\mathbf{k}) dt' dt'' d\mathbf{k}, \quad (30)$$

where mean longitudinal velocity is denoted by U and $\hat{u}_{ii}(\mathbf{k})$ corresponds to the Eulerian velocity covariance in Fourier space. The wave number vector is given by $\mathbf{k}$ and $k^2 = \sum_{j=1}^\infty k_j^2$ for $j = 1, \dots, d$. Equations (29) and (30) are limited to small to mild levels of heterogeneity. The expression for the velocity covariance in Fourier space, i.e., $\hat{u}_{ij}(\mathbf{k})$, is (Dagan 1984)

$$\hat{u}_{ij}(\mathbf{k}) = U^2 \left(\delta_{ij} - \frac{k_i k_j}{k^2} \right) \left(\delta_{ij} - \frac{k_i k_j}{k^2} \right) \hat{C}_{ff}(\mathbf{k}), \quad (31)$$

where δ_{ij} is Kronecker's delta and $\hat{C}_{ff}(\mathbf{k})$ is the Fourier transform of the logconductivity covariance, Eq. (6). For the purpose of validation, we will restrict ourselves to a two-dimensional setting. Therefore, the Fourier transform of the isotropic logconductivity exponential covariance function for $d = 2$ is (Rubin 2003)

$$\hat{C}_{ff}(\mathbf{k}) = \sigma_f^2 \lambda^2 (1 + k_1^2 \lambda^2 + k_2^2 \lambda^2)^{-3/2}. \quad (32)$$

Figure 1 (left) shows how the flow and transport simulator was able to capture the dilution of a plume in a two-dimensional homogeneous porous medium. The numerical results shows an excellent agreement with the analytical solution, see Eq. (26). The parameters used for the simulations in the homogeneous setting are reported in Table 1.

Next, we compare the numerical results with the semi-analytical solutions derived in de Barros et al. (2015) in a heterogeneous setting, see Eq. (27). The results reported in Fig. 1 (right) are for a two-dimensional heterogeneous aquifer characterized by $\sigma_f^2 = 1$. For this comparison, the spatially random porous medium was generated using HYDRO_GEN

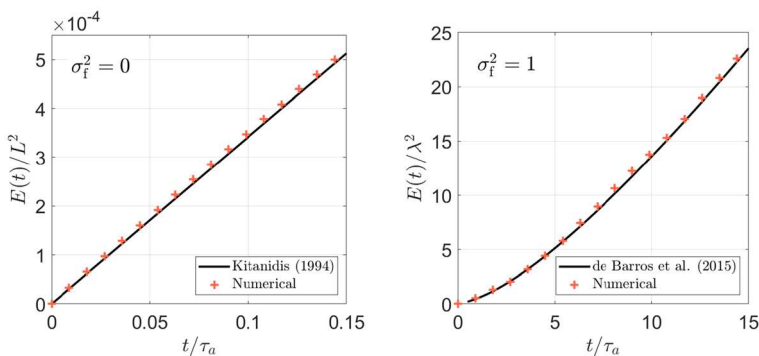


Fig. 1 Comparison between the results from the numerical flow and transport simulators (red crosses) against the analytical results (continuous line) derived by Kitanidis (1994) for a homogeneous medium (left) and the semi-analytical solution derived by de Barros et al. (2015) for a heterogeneous medium characterized by $\sigma_f^2 = 1$ (right). Results depicted for $d = 2$ and a point source. The advective time scale for the homogeneous medium (left) is given by $\tau_a = L_1/U$ where L_1 is the longitudinal dimension of the domain. The advective time scale for the heterogeneous medium (right) is $\tau_a = \lambda/U$ with λ as the correlation scale

Table 1 Input parameters used in the simulations

Parameter	Symbol	Value	Calculated as
Correlation length in x_1 and x_2	λ	10 m	–
Correlation length in x_3 ($d = 3$ only)	λ	10 m	–
Domain length in x_1	L_1	1000 m	100λ
Domain length in x_2	L_2	200 m	20λ
Domain length in x_3 ($d = 3$ only)	L_3	200 m	20λ
Mesh size in x_1 and x_2	Δ	1 m	$\lambda/10$
Mesh size in x_3 ($d = 3$ only)	Δ	1 m	$\lambda/10$
Porosity	ϕ	0.25	–
Mean (geometric) hydraulic conductivity	K_G	1 m/d	–
Mean longitudinal flow velocity	U	0.02 m/d	–
Dispersivity in x_1	α_L	0.1 m^2	0.1Δ
Dispersivity in x_2 (except Sect. 4)	α_T	0.01 m^2	0.01Δ
Dispersivity in x_2 (Sect. 4 only)	α_T	0.1 m^2	0.1Δ
Dispersivity in x_3 ($d = 3$ only)	α_T	0.01 m^2	0.01Δ
Molecular diffusion	D_m	$8.64 \times 10^{-5} \text{ m}^2/\text{d}$	–
Number of particles	N_p	10^7	–

(Bellin and Rubin 1996). Given the stochastic nature of the flow field, we averaged the temporal evolution of the dilution index over five random realizations. The results depicted in Fig. 1 (right) show an excellent agreement with the numerical simulator and the semi-analytical results. This is remarkable given that the perturbation theory-based solution, Eq. (27), is expected to deteriorate for $\sigma_f^2 \gtrsim 1$,

5 Numerical Set-Up

A description of the scenarios investigated is displayed in Fig. 2. As shown in Fig. 2, we will consider both two- and three dimensional flows characterized by two distinct logconductivity variances, namely $\sigma_f^2 = 1$ and 4. These values were selected to represent a mildly heterogeneous case and a highly heterogeneous case. In order to capture the effects of heterogeneity of the flow field on mixing, we consider a ratio of $\lambda/\Delta = 10$. The ratio λ/Δ was selected based on the analysis carried out in the literature (Ababou et al. 1989; Bellin et al. 1993, 1994; de Dreuzy et al. 2007; Moslehi et al. 2015). A detailed numerical error analysis regarding the spatial resolution of the flow field on simulations is reported in Figs. 5, 6 of Leube et al. (2013) and Figs. 2–6 of Moslehi et al. (2015). Figure 2 also provides a graphical representation of the different geometrical configurations for the source zone explored in our study. Details regarding the different scenarios are provide below.

Two-dimensional case: For $d = 2$, we will consider a point source of dimensions $0.1\lambda \times 0.1\lambda$ centered at $\mathbf{x} = (24, 100)$. We will also consider a line source of length 3λ located at $x_1 = 24$ with transverse dimension $x_2 \in [85, 115]$.

Three-dimensional case: For $d = 3$, we will consider, analogously to the case for $d = 2$, a point source of dimensions $0.1\lambda \times 0.1\lambda$ centered at $\mathbf{x} = (24, 100, 100)$ and also a line source located at $x_1 = 24$ and $x_3 = 100$ with transverse dimension $x_2 \in [85, 115]$.

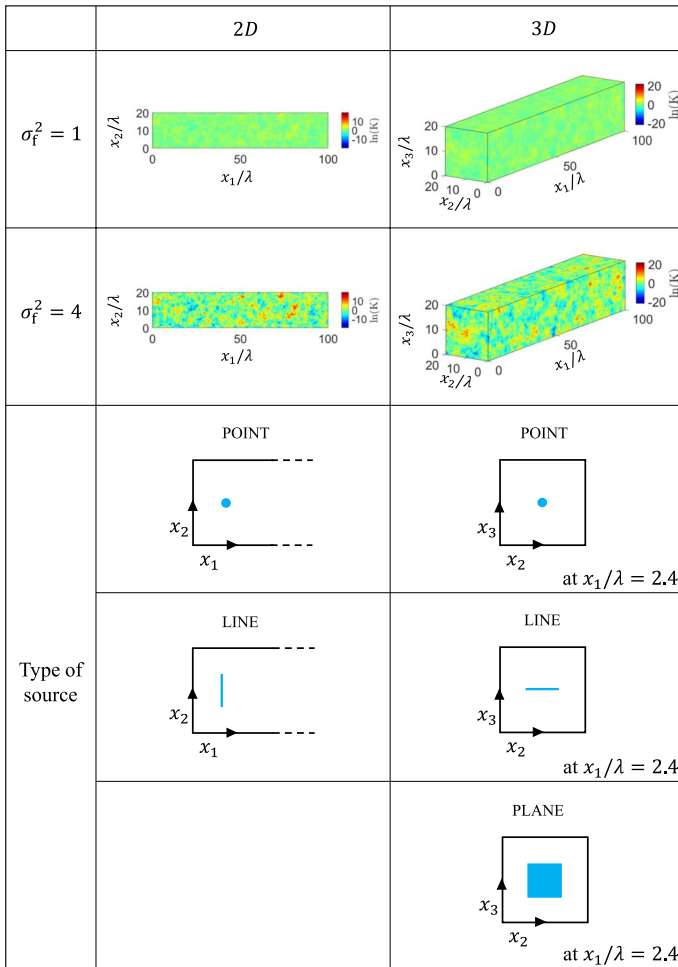


Fig. 2 Summary of the considered scenarios. We analyze transport in domains with $d = 2$ and $d = 3$ characterized by logconductivity fields with variance $\sigma_f^2 = 1$ and $\sigma_f^2 = 4$. In each field we study the impact on solute transport of the source dimension

We will also consider a plane source located at $x_1 = 24$ with transverse dimension $x_2 \in [85, 115]$ and $x_3 \in [85, 115]$ (i.e., of dimensions $3\lambda \times 3\lambda$).

All other parameter values used in the simulations presented in this work are listed in Table 1. All the simulations were performed on a desktop computer equipped with a GPU NVIDIA GeForce GTX 745/PCIe/SSE2 necessary to use PAR² and with 4GB of internal RAM memory. The time required to complete a high-resolution simulation (i.e., $\Delta/\lambda = 0.1$) with 10^7 particles ranged from about three hours for the two-dimensional cases to up to six hours for the three-dimensional cases.

6 Computational Results

Figures 3 and 4 provide examples of the simulated concentration fields in two- and three-dimensional flows for $\sigma_f^2 = 1$ and 4 at the final simulation time. The spatial concentration plume shows more complex features with increasing σ_f^2 that are accompanied with overall lower concentration values, which indicates increased dilution. In the following, we study the mixing and dilution dynamics in a more quantitative fashion in terms of the evolution of the mean concentrations, concentration variance, concentration PDF and dilution index introduced in Sect. 2.3.

6.1 Concentration Mean and Variance

Figures 5 and 6 report the temporal evolution of the spatial concentration mean and variance defined in Eqs. (8) and (9) for $d = 2$ and $d = 3$ spatial dimensions for point, line and planar source distributions. The scalings derived in Sect. 2.4 for homogeneous media indicate the dependence of the decay of the concentration mean and variance on the dimensions of space and the source distribution. Such scalings are reported in Figs. 5 and 6 as dashed lines, and their slope is denoted by the value b . We found that the mean concentration scales as $\mu_c(t) \sim t^{-(d-d_s)/2}$, and the concentration variance as $\sigma_c^2(t) \sim t^{-(d-d_s)}$, where d is the space dimension and d_s the source dimension. This means, that the concentration mean and variance for a homogeneous medium decay faster for higher spatial dimensions and lower source dimension. We observe qualitatively similar dependence on space and source dimension for heterogeneous media, while the decay is in general faster for the heterogeneous than for the homogeneous media. This effect is increasing with increasing σ_f^2 . For $\sigma_f^2 = 1$, we find that the concentration mean and variance for point and line sources in $d = 2$ and 3 decay only slightly faster than expected for homogeneous media. For $\sigma_f^2 = 4$, in contrast we observe a strong acceleration of the decay of concentration mean and variance compared to homogeneous

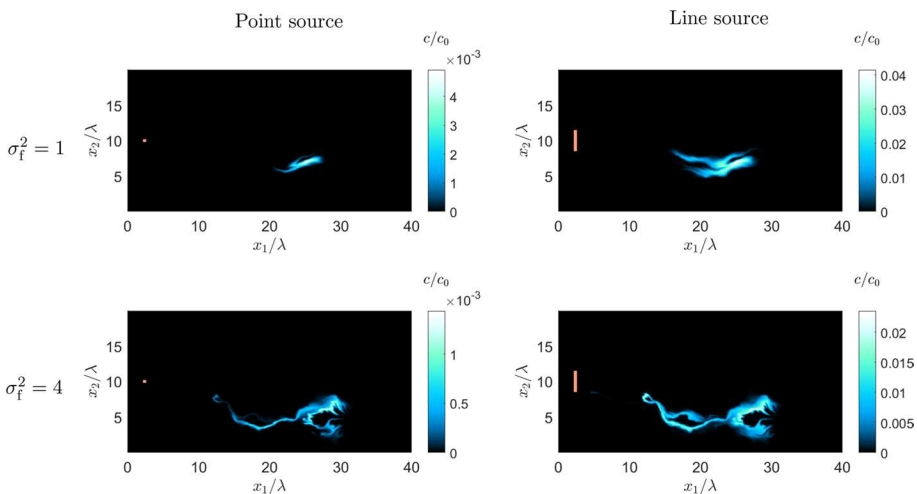


Fig. 3 Concentration field for $d = 2$ at the final simulation time $t/\tau_a = 20$ for point (left column) and line source (right column), in the conductivity field with $\sigma_f^2 = 1$ (top row) and $\sigma_f^2 = 4$ (bottom row). In all four images, the initial position of the solute source is depicted in orange color

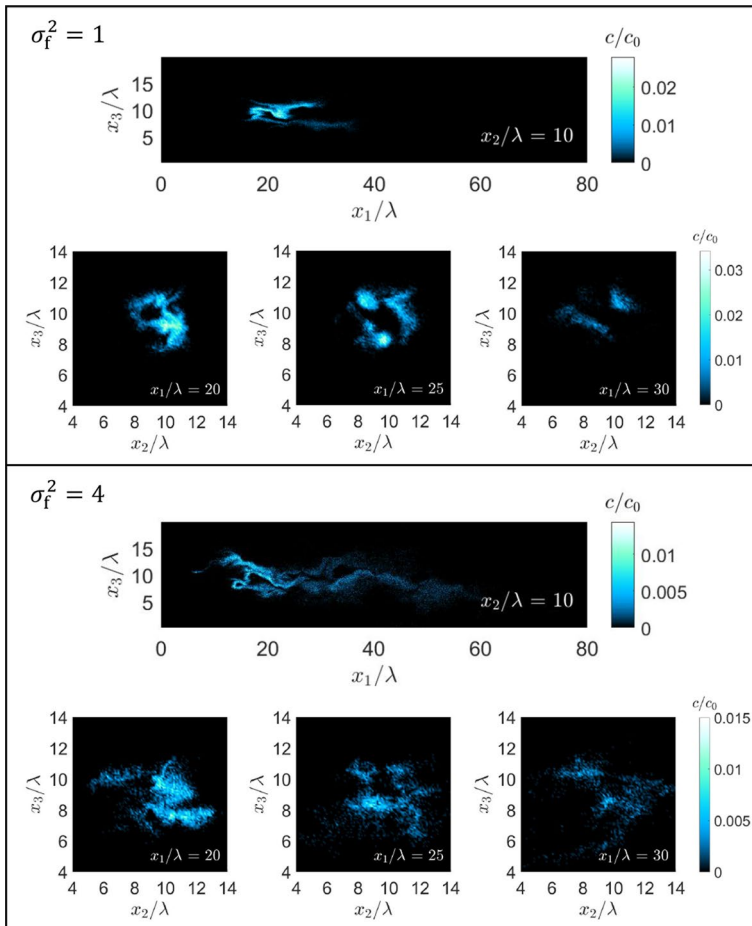


Fig. 4 Concentration fields at the final simulation time $t/\tau_a = 20$ for $d = 3$ with $\sigma_f^2 = 1$ (top box) and $\sigma_f^2 = 4$ (bottom box) for a vertical planar source

media. For the planar source in $d = 3$ dimensions, concentration mean and variance decay significantly faster than for a homogeneous medium both for $\sigma_f^2 = 1$ and 4. Such faster decay can be explained by the fact that a planar source is able to sample more fluctuations of the permeability field. This means that higher concentration gradients have the opportunity to be generated, thus leading to faster mixing. As illustrated in Figs. 3 and 4, the concentration distributions are strongly dispersed for both heterogeneity strengths, and the initial distribution evolves into a lamellar structure. This means, spatial heterogeneity creates strong concentration gradients, which are efficiently degraded through local scale dispersion, thus explaining the faster decay of mean concentration and concentration variance in heterogeneous media compared with homogeneous media.

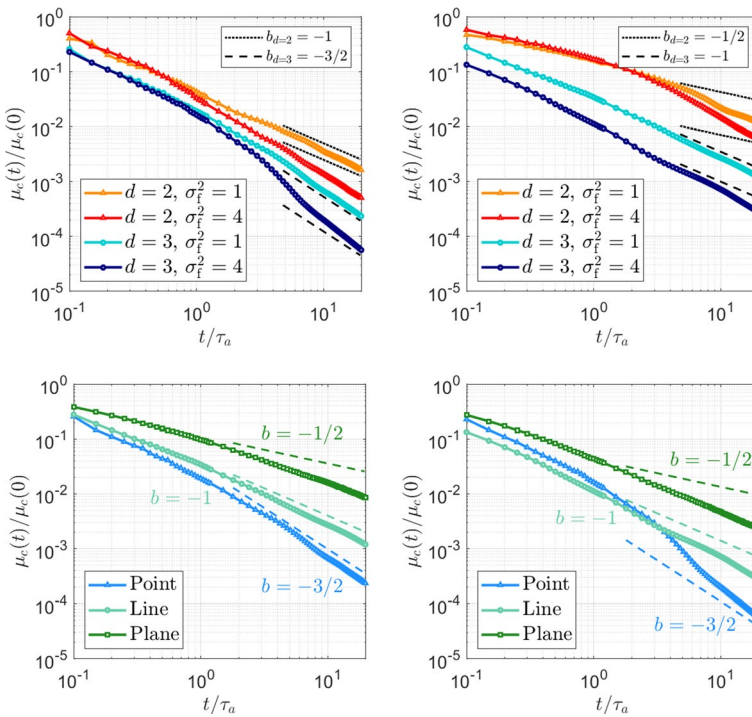


Fig. 5 Mean concentration in time for: point source (top left), line source (top right), 3d field with $\sigma_f^2 = 1$ (bottom left), and 3d field with $\sigma_f^2 = 4$ (bottom right)

6.2 Concentration Distribution

Figure 7 illustrates the concentration PDF, $p(c)$ for point and line sources in both $d = 2$ and 3 spatial dimensions at different times. As discussed in the previous section, at increasing times the concentration mean and variance decrease. The maximum observable concentrations are shifted toward smaller values. At small concentration values, $p(c)$ behaves as predicted for a homogeneous medium as $p(c) \sim c^{-1}$ for all cases under consideration. The solutions for a homogeneous medium are given by Eq. (15). At increasing concentration values, the behaviors expected for a homogeneous medium are different from the observations for the heterogeneous media. For the point source in $d = 2$ and the line source in $d = 3$, we expect $p(c) \sim c^{-1}$ for $c < c_m(t)$. For the point source in $d = 3$, the concentration PDF behaves as $p(c) \sim c(\ln c)^{1/2}$ and for the line source in $d = 2$ as $p(c) \sim c(\ln c)^{-1/2}$. The concentration PDF decreases faster toward smaller maximum values than the corresponding homogeneous solutions.

These behaviors can be understood in the lamellar mixing framework of Villermaux (2012) and Le Borgne et al. (2015) as the result of stretching-enhanced local dispersion. As shown in Figs. 3 and 4, in heterogeneous media, the initially homogeneous solute distribution evolves into a lamellar structure due to the spatial variability of the underlying flow field. Across individual lamellae, the concentration distribution can be approximated by Gaussian distributions characterized by a maximum distribution that depends on the lamella's stretching history. For a lamella that has passed through regions of strong stretching,

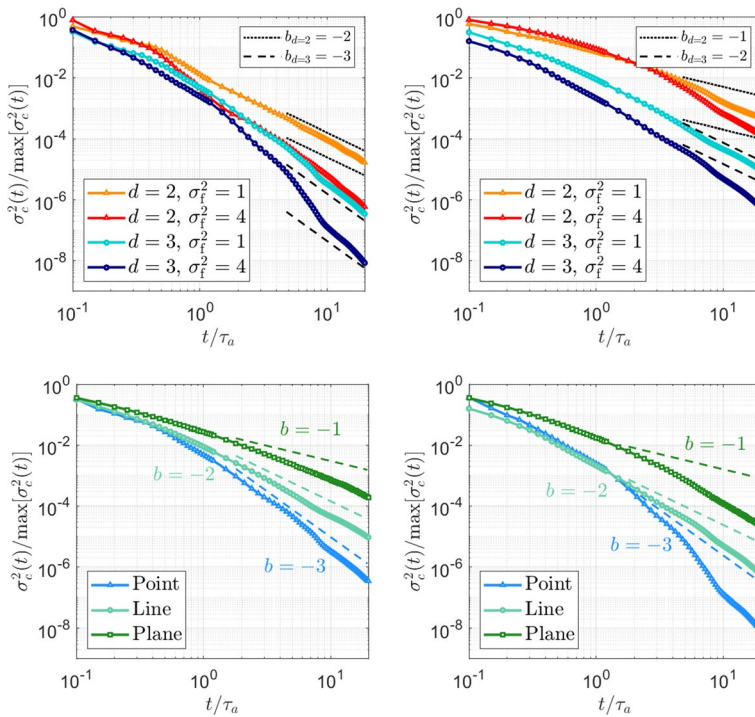


Fig. 6 Concentration variance in time for: point source (top left), line source (top right), three-dimensional field with $\sigma_f^2 = 1$ (bottom left), and three-dimensional field with $\sigma_f^2 = 4$ (bottom right)

the maximum concentration decreases much faster than for the lamella that remains undeformed, as long as the values of local dispersion are large enough to efficiently dissipate concentration gradients. The concentration PDF for an individual lamella is then characterized by (15) for a Gaussian concentration distribution. The full concentration PDF is then constructed by the superposition of lamellae with different stretching histories. This explains the behavior of $p(c) \sim c^{-1}$ at small concentration values, which is characteristic of the Gaussian concentration profile across lamellae, and the decay behavior at large concentration values, which reflects the stretching and deformation history of the lamellae. In this sense, the mixing behavior at the simulation times under consideration here, which are smaller than the characteristic dispersion time over a correlation length, can be understood by the superposition of independent stretched lamellae. At times larger than the characteristic dispersion time scale, lamellae start coalescing, which leads to different behaviors of the concentration PDF as discussed in Le Borgne et al. (2015). However, this regime is beyond the scope of this paper.

Next, we compare the cumulative distribution function (CDF), namely $P(c) = \int_0^c p(\psi) d\psi$, obtained from the numerical simulations with a parametric CDF model, the beta CDF. The beta CDF was originally suggested by Fiori (2001) as a potential distribution to capture the statistics of the concentration field given its flexibility to shift from a bimodal to a unimodal PDF. It has been widely employed to quantify the distribution of the concentration values in heterogeneous domains (Fiorotto and Caroni 2002; Bellin and Tonina 2007) and tested against numerical and analytical solutions (Schwede et al. 2008;

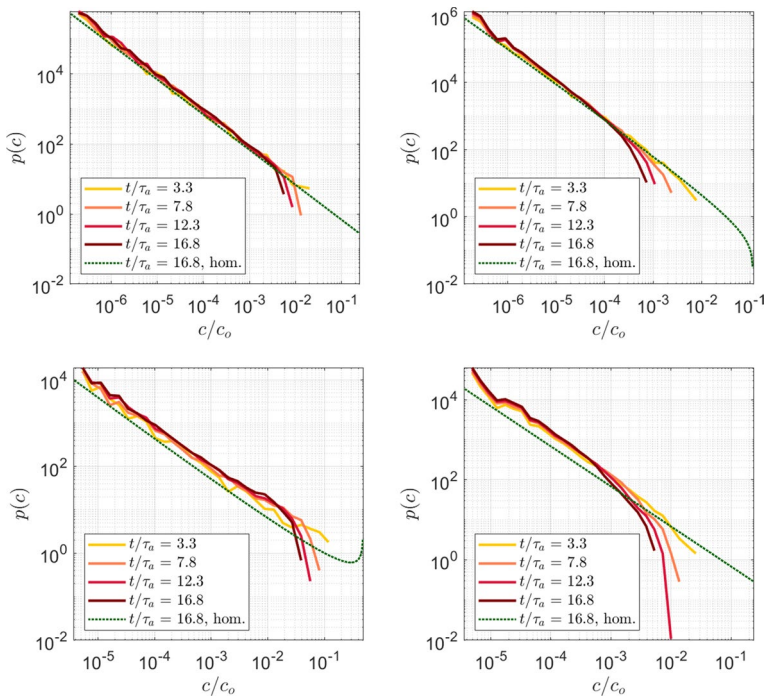


Fig. 7 Concentration PDF for distinct source configurations and flow dimensionality. The results in this Figure are for the fields with $\sigma_f^2 = 1$. Results reported for a point source for $d = 2$ (top left), point source for $d = 3$ (top right), line source for $d = 2$ (bottom left), and line source for $d = 3$ (bottom right). The dotted line represents the concentration PDF in a homogeneous field according to Eq. (15)

Boso et al. 2013a; de Barros and Fiori 2014; Boso and Tartakovsky 2016). The performance of the beta distribution model was also assessed for mixing of fluids with distinct viscosities (Bonazzi et al. 2021) as well as mixing in non-Gaussian random flow fields (de Barros et al. 2022). The beta CDF model is defined by:

$$P(c) = \frac{\Gamma[a_o + b_o]}{\Gamma[a_o]\Gamma[b_o]} \int_0^c dw w^{a_o-1} (1-w)^{b_o-1}, \quad (33)$$

where $\Gamma[z]$ is the Gamma function. The exponents a and b are given in terms of the empirically determined concentration mean μ_c and variance σ_c^2 as

$$a_o = \frac{\mu_c(t)}{\beta}, \quad b_o = \frac{1 - \mu_c(t)}{\beta} \quad \beta = \frac{\sigma_c^2(t)}{\mu_c(t)[1 - \mu_c(t)] - \sigma_c^2(t)}, \quad (34)$$

Note that the symbol b_o here has a different meaning than the slope b in Figs. 5 and 6. The results displayed in Fig. 8 show the comparison between the CDF obtained from the numerical simulations and the beta CDF model (34) fitted with Eqs. (8) and (9), i.e., we calculated the spatial average and spatial variance of the concentration in the mixing volume obtained from the numerical simulations and used those values to estimate the parameters a_o and b_o in Eq. (34). Figure 8 illustrates the comparison in both two-dimensional (Fig. 8, left) and three-dimensional (Fig. 8, right) domains for distinct source

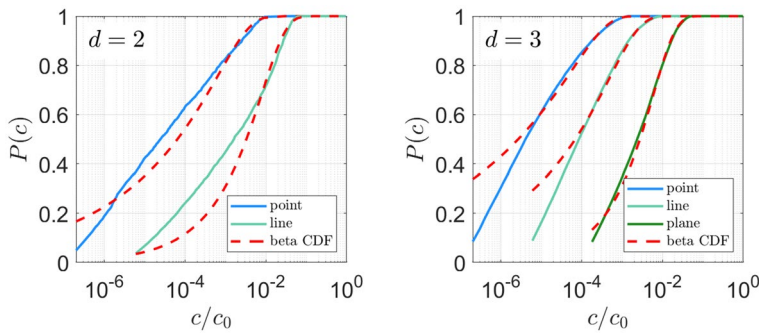


Fig. 8 Comparison between the concentration cumulative distribution function (CDF) and beta distribution model for $d = 2$ (left) and $d = 3$ (right). The results in this Figure are for the fields with $\sigma_f^2 = 1$ at time $t/\tau_a = 12.3$. Results computed for different source zone configurations (point, line and vertical planar sources)

configurations at $t/\tau_a = 12.3$. In agreement with previous findings reported in the literature (Fiorotto and Caroni 2002; Boso et al. 2013a; de Barros and Fiori 2014), the beta distribution fails to capture the probabilities at low concentrations. This is particularly evident in the case of two-dimensional domains (see Fig. 8, left). The fitting of the beta distribution improves with an increase in the dimensionality of the porous medium (compare Fig. 8 left and right). As depicted in Fig. 8 (right), the agreement between the beta CDF model and the empirical CDF is augmented when the solute source zone is a vertical plane.

The discrepancy between the beta model and the data at small concentrations can be traced back to the fact that the mixing behavior at the small concentration fringes of the solute plume is dominated by diffusion. We further explore this discrepancy at low concentrations by analyzing the third (skewness) and fourth (kurtosis) statistical moments of the concentration data at $t/\tau_a = 12.3$ for a log-conductivity variance of unity. Table 2 provides a comparison between the higher order moments obtained from the fitted beta CDF and the empirical CDF obtained from the high resolution numerical simulations. The higher moments of the beta distributions were calculated from the distribution coefficients in Eq. (34). As expected, the higher order moments estimated from the fitted beta model are different than the ones characterizing the empirical distribution.

Table 2 Skewness and kurtosis of the concentration empirical distribution (originated from the numerical simulations) and of the fitted beta distribution at $t/\tau_a = 12.3$. The ratios of the higher order moments between the empirical and the beta distributions are also reported. The skewness and kurtosis of the empirical distribution obtained from the numerical simulations are represented by $\mu^{(3)}$ and $\mu^{(4)}$ whereas the ones for the beta distribution are $\mu_\beta^{(3)}$ and $\mu_\beta^{(4)}$

Simulation	Skewness			Kurtosis		
	$\mu^{(3)}$	$\mu_\beta^{(3)}$	$\mu^{(3)}/\mu_\beta^{(3)}$	$\mu^{(4)}$	$\mu_\beta^{(4)}$	$\mu^{(4)}/\mu_\beta^{(4)}$
$d = 2$, point	3.05	4.55	0.67	12.54	30.94	0.41
$d = 2$, line	1.82	2.92	0.62	5.82	12.58	0.46
$d = 3$, point	4.15	5.14	0.81	22.94	39.71	0.58
$d = 3$, line	3.79	4.23	0.90	20.17	26.71	0.76
$d = 3$, plane	2.28	2.72	0.84	8.85	10.98	0.81

To overcome the limitations of the beta distribution at capturing the probabilities of low concentration occurrences, we explore alternative distributions. To do so, we propose to employ two one-to-one variable transforms, namely the power law and logarithm transforms. We transform the concentration data originated from the numerical simulations as follows: $\sqrt{c}$ and $\ln[c]$. The results depicted in Fig. 9 show a good performance, within the low concentration range, of the Pareto-type IV distribution for the $\sqrt{c}$ and a uniform distribution for $\ln[c]$. Table 3 reports the relative error between the estimates for the probability that $c < c^*$ obtained from the fitted concentration CDFs and the empirical CDF. The relative error is given by $\epsilon = 100 \times |P(c^*) - \hat{P}(c^*)|/P(c^*)$, where $P(c)$ is the empirical concentration CDF obtained from the numerical simulations and $\hat{P}(c)$ is the fitted CDF (i.e., beta, uniform or Pareto-type IV). The concentration threshold c^* is defined as follows:

$$c^*(t) = c_{\min}(t) + \eta[c_{\max}(t) - c_{\min}(t)], \quad (35)$$

where $c_{\max}$ and $c_{\min}$ are the maximum and minimum concentrations within in the plume at time t . The quantities are defined as:

$$c_{\min}(t) = \min_{\mathbf{x} \in \varphi_d} c(\mathbf{x}, t); \quad (36)$$

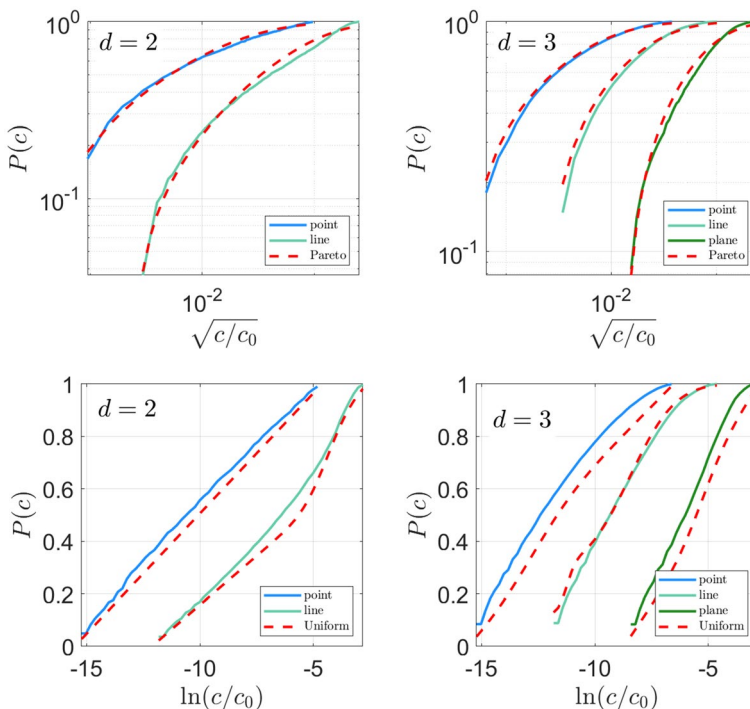


Fig. 9 Comparison between the concentration cumulative distribution function (CDF) and the Pareto distribution model for $d = 2$ (top left) and $d = 3$ (top right), and between the concentration CDF and the uniform distribution model for $d = 2$ (bottom left) and $d = 3$ (bottom right). Results are for the fields with $\sigma_f^2 = 1$ at time $t/\tau_a = 12.3$

Table 3 Relative errors ϵ between the fitted concentration CDFs and the empirical CDF obtained from the numerical simulations. Results computed for different values of η , see Eq. (35)

Simulation	c^*	ϵ for beta	ϵ for Pareto-type IV	ϵ for uniform
$\eta = 0.01$				
$d = 2$, point	9.5×10^{-5}	13.97	2.20	7.60
$d = 2$, line	6.4×10^{-4}	36.55	6.99	13.27
$d = 3$, point	1.5×10^{-5}	2.02	2.12	1.59
$d = 3$, line	1.1×10^{-4}	2.58	6.38	3.32
$d = 3$, plane	7.8×10^{-4}	9.47	3.87	6.25
$\eta = 0.05$				
$d = 2$, point	4.7×10^{-4}	5.80	4.21	5.12
$d = 2$, line	3.1×10^{-3}	13.67	12.01	15.63
$d = 3$, point	7.2×10^{-5}	2.29	1.19	0.95
$d = 3$, line	5.2×10^{-4}	1.75	2.07	1.00
$d = 3$, plane	3.2×10^{-3}	4.59	11.87	4.21
$\eta = 0.1$				
$d = 2$, point	9.5×10^{-4}	2.11	3.69	5.27
$d = 2$, line	6.3×10^{-3}	3.20	11.55	9.51
$d = 3$, point	1.4×10^{-4}	0.77	0.90	0.21
$d = 3$, line	1.0×10^{-3}	1.38	0.42	0.81
$d = 3$, plane	6.2×10^{-3}	1.83	6.58	1.91

$$c_{\max}(t) = \max_{\mathbf{x} \in \varphi_d} c(\mathbf{x}, t), \quad (37)$$

where φ_d is defined in Eq. (7). For the results present in Table 3, we compute c^* for $\eta = 0.01, 0.05$ and 0.1 . The results reported in Table 3 inform that at low concentrations, both the Pareto-type IV and uniform distributions perform well. As expected, the fit of the beta distribution improves as c^* increases.

6.3 Dilution Index

The temporal evolution of the dilution index defined by Eq. (11) is shown in Fig. 10. The evolution of the dilution index mirrors the evolution of the mean concentration. Both quantities are related to the mixing volume. The mean concentration is inversely proportional to the mixing volume, while the dilution index is directly a measure for the mixing volume (Kitanidis 1994). For a homogeneous medium, the dilution index scales as $E(t) \sim t^{(d-d_s)/2}$. Thus, it increases faster with increasing dimension d of space and decreasing source dimension d_s .

While the dependence on space and source dimensions is qualitatively similar for heterogeneous media, dilution is significantly accelerated compared to homogeneous media (note that in Fig. 10, analogously to Figs. 5 and 6, b denotes the slope of the homogeneous scaling). In analogy to the behaviors observed for the mean concentration, dilution increases with increasing variance σ_f^2 , this means with increasing level of heterogeneity (see top row, Fig. 10). This is because the macro-scale spreading of the solute plume increases with an increase in σ_f^2 . Larger values of σ_f^2 implies more erratically shaped plumes and an increase

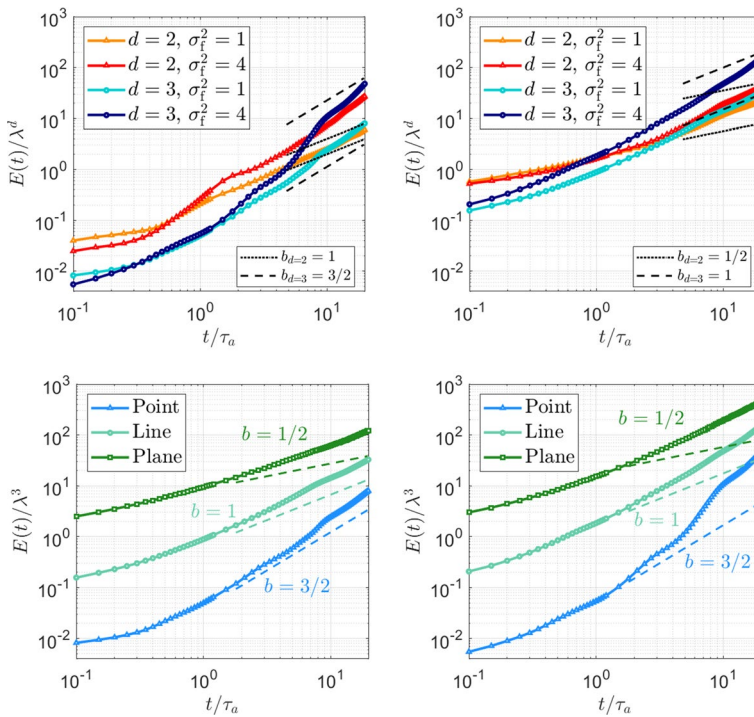


Fig. 10 Dilution index in time for: point source (top left), line source (top right), 3d field with $\sigma_f^2 = 1$ (bottom left), and 3d field with $\sigma_f^2 = 4$ (bottom right)

in the plume's surface area with the surrounding (i.e., ambient) fluid. As a consequence, diffusive and local-scale dispersive solute mass fluxes are augmented. Comparing, for instance, the plots in the bottom row of Fig. 10, one can observe that, for all source dimensions, the rate of increase in the dilution index is higher when compared to the homogeneous scaling depicted by the dashed lines. The strongest increase, compared to the behavior for homogeneous media, is observed for the planar source case (Fig. 10, bottom). This is in agreement with what was observed for the mean concentration results reported in Sect. 6.1.

Close inspection of Fig. 10 (bottom row) also reveals that the source zone mainly impacts the rate of dilution. The rate of dilution is highest for a point-like injection. For this case, the particles that constitute the point-like source, in a three-dimensional setting, will all occupy the same position at $t = t_0$. However, at the next considered time, $t_1 > t_0$, each particle will have moved in a different direction due to the combination of advection and dispersion. This means that the mixing volume at t_1 would have significantly increased (due to spreading in all three directions) when compared to the initial volume at t_0 , when it was just a point.

7 Summary

Modeling solute mixing in spatially heterogeneous porous media flows is a challenging task. The multiscale variability of the hydraulic conductivity impacts both spreading and mixing rates of the solute body. The presence of preferential flow paths (i.e., hydrogeological connectivity), low-conductivity zones and abrupt facies transitions require numerical models to have a high spatial resolution. The need for such fine spatial resolution increases the computational costs associated with flow and transport simulations, especially when dealing with three-dimensional stochastic systems.

In this work, we performed high resolution numerical simulations to study the temporal evolution of the mixing behavior of a solute plume in heterogeneous porous media. Mixing is quantified in terms of the solute plume's concentration mean and variance, probability distribution and the dilution index. Our analysis showed how mixing is affected by different factors such as the dimensionality of the flow domain, the degree of heterogeneity in the conductivity field and the geometrical configuration of the source zone.

With few exceptions, most studies up-to-date have focused their attention solely to studying mixing in spatially heterogeneous two-dimensional domains for point and line sources. In this work, we provide a systematic investigation of mixing in both two- and three-dimensional spatially heterogeneous porous media flows using a computationally efficient transport simulator denoted PAR² developed in Rizzo et al. (2019). PAR² code is an open source, GPU accelerated transport simulator based on the random walk particle tracking technique. Morvillo et al. (2022) provides additional the details as well as a step-by-step tutorial on how to connect the GPU accelerated particle tracking code employed in this work with other tools (such as the flow simulator and the random conductivity field generator). All codes can be also obtained from the information provided in Morvillo et al. (2022). We tested the accuracy of the numerical results with analytical and semi-analytical results reported in the literature for both homogeneous and heterogeneous porous media.

Our results highlight the importance of the source configuration and flow dimensionality on the temporal evolution of different mixing patterns. In particular, hydraulic conductivity fields with $d = 3$ exhibit faster decays in time of both μ_c and σ_c^2 respect to $d = 2$ fields with the same level of heterogeneity. Moreover, such decays of μ_c and σ_c^2 happen faster for a point source and slower for a planar source, analogously to the analytical results in a homogeneous field. The source configuration also affects dilution, since the dilution index is higher for a planar source and lower for a point source, although such difference seems to decreases when heterogeneity is increased. These results are in agreement with the theoretical analysis of Dentz and de Barros (2013) where it was analytically shown that both source zone configuration and flow spatial dimension impact the self-averaging transport behavior of the solute plume. Here we expand this theoretical analysis for different mixing metrics and higher levels of heterogeneity in the conductivity field.

Finally, we compared the statistics of the concentration values with the beta PDF model. Using a parametric model to approximate the concentration PDF is appealing given that one can estimate the concentration mean and variance from a limited data set (i.e., due to high costs of data acquisition). Previous works showed the suitability of the beta PDF model to capture the main features of the concentration field statistics for both point-like and line sources (Bellin and Tonina 2007; Boso et al. 2013a). Our results show that the

performance of the beta PDF improves when the flow field is three-dimensional and transport originates from a vertical planar source. Given that the performance of the beta distribution deteriorates for low concentration values, we propose two alternative distributions to better capture the probabilities of observing $c \ll c_0$. Our results show that the Pareto-type IV distribution performs well after employing a power-law transform to the concentration data. Similarly, the uniform distribution also performs well at low concentrations when a logarithm transform is applied to the concentration data. The analysis carried out in this study shows a promising path in employing a mixture distribution approach to model the statistics of solute concentration field.

Acknowledgements The authors want to thank Prof. Alberto Bellin from the University of Trento (Italy) for a fruitful discussion regarding geostatistics and for providing the *Hydro_Gen* code. The authors would like to thank all the Reviewers and the Editors for the constructive comments.

Funding FPJdB acknowledges the financial support provided by the NSF grant number 1654009.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

References

- Ababou, R., McLaughlin, D., Gelhar, L.W., Tompson, A.F.B.: Numerical simulation of three-dimensional saturated flow in randomly heterogeneous porous media. *Trans. Porous Media* **4**(6), 549–565 (1989)
- Bakker, M., Post, V., Langevin, C.D., Hughes, J.D., White, J.T., Starn, J.J., Fienen, M.N.: Scripting modflow model development using python and flopy. *Groundwater* **54**(5), 733–739 (2016)
- Beaudoin, A., de Dreuz, J.R.: Numerical assessment of 3-d macrodispersion in heterogeneous porous media. *Water Resour. Res.* **49**(5), 2489–2496 (2013)
- Bellin, A., Rinaldo, A., Bosma, W.J.P., van der Zee SEATM, Rubin, Y.: Linear equilibrium adsorbing solute transport in physically and chemically heterogeneous porous formations: 1. analytical solutions. *Water Resour. Res.* **29**(12): 4019–4030 (1993)
- Bellin, A., Rubin, Y.: *Hydro_gen*: a spatially distributed random field generator for correlated properties. *Stoch. Hydrol Hydraul.* **10**(4), 253–278 (1996)
- Bellin, A., Tonina, D.: Probability density function of non-reactive solute concentration in heterogeneous porous formations. *J. Contam. Hydrol.* **94**(1–2), 109–125 (2007)
- Bellin, A., Salandin, P., Rinaldo, A.: Simulation of dispersion in heterogeneous porous formations: Statistics, first-order theories, convergence of computations. *Water Resour. Res.* **28**(9), 2211–2227 (1992)
- Bellin, A., Rubin, Y., Rinaldo, A.: Eulerian-lagrangian approach for modeling of flow and transport in heterogeneous geological formations. *Water Resour. Res.* **30**(11), 2913–2924 (1994)
- Bonazzi, A., Morvillo, M., Im, J., Jha, B., de Barros, F.P.J.: Relative impacts of permeability heterogeneity and viscosity contrast on solute mixing. *Phys. Rev. Fluids* (2021)
- Boso, F., Tartakovsky, D.M.: The method of distributions for dispersive transport in porous media with uncertain hydraulic properties. *Water Resour. Res.* **52**(6), 4700–4712 (2016)
- Boso, F., de Barros, F.P.J., Fiori, A., Bellin, A.: Performance analysis of statistical spatial measures for contaminant plume characterization toward risk-based decision making. *Water Resour. Res.* **49**(6), 3119–3132 (2013)
- Boso, F., Bellin, A., Dumbser, M.: Numerical simulations of solute transport in highly heterogeneous formations: a comparison of alternative numerical schemes. *Adv. Water Resour.* **52**, 178–189 (2013)
- Chapman, S.W., Parker, B.L.: Plume persistence due to aquitard back diffusion following dense nonaqueous phase liquid source removal or isolation. *Water Resour. Res.* **41**(12) (2005)
- Cirpka, O.A., de Barros, F.P.J., Chiogna, G., Rolle, M., Nowak, W.: Stochastic flux-related analysis of transverse mixing in two-dimensional heterogeneous porous media. *Water Resour. Res.* **47**(6) (2011)
- Dagan, G., Neuman, S.P.: *Subsurface flow and transport: a stochastic approach*. Cambridge University Press (2005)
- Dagan, G.: Solute transport in heterogeneous porous formations. *J. Fluid Mech.* **145**, 151–177 (1984)

- de Barros, F.P.J., Fiori, A.: 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**(5), 4018–4037 (2014)
- de Barros, F.P.J., Nowak, W.: On the link between contaminant source release conditions and plume prediction uncertainty. *J. Contam Hydrol.* **116**(1–4), 24–34 (2010)
- de Barros, F.P.J., Dentz, M., Koch, J., Nowak, W.: Flow topology and scalar mixing in spatially heterogeneous flow fields. *Geophys. Res. Lett.* (2012). <https://doi.org/10.1029/2012GL051302>
- de Barros, F.P.J., Fiori, A., Boso, F., Bellin, A.: A theoretical framework for modeling dilution enhancement of non-reactive solutes in heterogeneous porous media. *J. Contam. Hydrol.* **175**, 72–83 (2015)
- de Barros, F.P.J., Guadagnini, A., Riva, M.: Features of transport in non-gaussian random porous systems. *Int. J. Heat Mass Trans.* **184**(122), 244 (2022)
- de Dreuzy, J.R., Beaudoin, A., Erhel, J.: Asymptotic dispersion in 2d heterogeneous porous media determined by parallel numerical simulations. *Water Resour. Res.* **43**(10), (2007)
- de Barros F.P.J, Fiori, A.: On the maximum concentration of contaminants in natural aquifers. *Trans Porous Media* pp 1–18 (2021)
- Dentz, M., de Barros, F.P.J., Le Borgne, T., Lester, DR.: Evolution of solute blobs in heterogeneous porous media. *J. Fluid Mech.* 853 (2018)
- Dentz, M., Tartakovsky, DM.: Probability density functions for passive scalars dispersed in random velocity fields. *Geophys. Res. Lett.* 37(24)(2010)
- Dentz, M.: Concentration statistics for transport in heterogeneous media due to stochastic fluctuations of the center of mass velocity. *Adv. Water Resour.* **36**, 11–22 (2012)
- Dentz, M., de Barros, F.P.J.: Dispersion variance for transport in heterogeneous porous media. *Water Resour. Res.* **49**(6), 3443–3461 (2013)
- Dentz, M., Kinzelbach, H., Attinger, S., Kinzelbach, W.: Temporal behavior of a solute cloud in a heterogeneous porous medium: 1. point-like injection. *Water Resour. Res.* **36**(12), 3591–3604 (2000)
- Dentz, M., Le Borgne, T., Englert, A., Bijeljic, B.: Mixing, spreading and reaction in heterogeneous media: a brief review. *J. Contam. Hydrol.* **120**, 1–17 (2011)
- Ferziger, J.H., Perić, M., Street, R.L.: Computational methods for fluid dynamics, vol. 3. Springer (2002)
- Fiori, A.: The Lagrangian concentration approach for determining dilution in aquifer transport: theoretical analysis and comparison with field experiments. *Water Resour. Res.* **37**(12), 3105–3114 (2001)
- Fiori, A., Dagan, G.: Concentration fluctuations in aquifer transport: a rigorous first-order solution and applications. *J. Contam. Hydrol.* **45**(1–2), 139–163 (2000)
- Fiori, A., Bellin, A., Cvetkovic, V., de Barros, F.P.J., Dagan, G.: Stochastic modeling of solute transport in aquifers: from heterogeneity characterization to risk analysis. *Water Resour. Res.* **51**(8), 6622–6648 (2015)
- Fiorotto, V., Caroni, E.: Solute concentration statistics in heterogeneous aquifers for finite pecllet values. *Transp. Porous Media* **48**(3), 331–351 (2002)
- Gotovac, H., Andricevic, R., Gotovac, B.: Multi-resolution adaptive modeling of groundwater flow and transport problems. *Adv. Water Resour.* **30**(5), 1105–1126 (2007)
- Guetting, N., Englert, A.: Hydraulic conditions at the source zone and their impact on plume behavior. *Hydrogeol. J.* **21**(4), 829–844 (2013)
- Harbaugh, A.W.: MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process. US Department of the Interior, US Geological Survey Reston, VA (2005)
- Herrera, P.A., Massabó, M., Beckie, R.D.: A meshless method to simulate solute transport in heterogeneous porous media. *Adv. Water Resour.* **32**(3), 413–429 (2009)
- Hidalgo, J.J., Carrera, J.: Effect of dispersion on the onset of convection during co2 sequestration. *J. Fluid Mech.* (2009)
- Im, J., Rizzo, C.B., de Barros, F.P.J.: Resilience of groundwater systems in the presence of bisphenol a under uncertainty. *Sci. Total Environ.* **727**(138), 363 (2020)
- Jankovic, I., Maghrebi, M., Fiori, A., Dagan, G.: When good statistical models of aquifer heterogeneity go right: the impact of aquifer permeability structures on 3D flow and transport. *Adv. Water Resour.* **100**, 199–211 (2017)
- Kamrava, S., Im, J., de Barros, F.P.J., Sahimi, M.: Estimating dispersion coefficient in flow through heterogeneous porous media by a deep convolutional neural network. *Geophys. Res. Lett.* **48**(18), e2021GL094443 (2021)
- Kapoor, V., Kitanidis, P.K.: Concentration fluctuations and dilution in aquifers. *Water Resour. Res.* **34**(5), 1181–1193 (1998)
- Kitanidis, P.K.: The concept of the dilution index. *Water Resour. Res.* **30**(7), 2011–2026 (1994)

- Le Borgne, T., Dentz, M., Villiermaux, E.: Stretching, coalescence, and mixing in porous media. *Phys. Rev. Lett.* **110**(20): 204,501(2013)
- Le Borgne, T., Dentz, D., Bolster, D., Carrera, J., de Dreuzy, J.R., Davy, P.: Non-fickian mixing: temporal evolution of the scalar dissipation rate in heterogeneous porous media. *Adv. Water Resour.* **33**, 1468–1475 (2010). <https://doi.org/10.1016/j.advwatres.2010.08.006>
- Le Borgne, T., Dentz, M., Villiermaux, E.: The lamellar description of mixing in porous media. *J. Fluid Mech.* **770**, 458–498 (2015)
- Leube, P.C., de Barros, F.P.J., Nowak, W., Rajagopal, R.: Towards optimal allocation of computer resources: trade-offs between uncertainty quantification, discretization and model reduction. *Environ. Model. Softw.* **50**, 97–107 (2013)
- Libera, A., Henri, C.V., de Barros, F.P.J.: Hydraulic conductivity and porosity heterogeneity controls on environmental performance metrics: implications in probabilistic risk analysis. *Adv. Water Resour.* **127**, 1–12 (2019)
- Meyer, D.W., Jenny, P., Tchelepi, H.A.: A joint velocity-concentration pdf method for tracer flow in heterogeneous porous media. *Water Resour. Res.* **46**(12), (2010)
- Morvillo, M., Bonazzi, A., Rizzo, C.B., de Barros, F.P.J.: Improving the computational efficiency of first arrival time uncertainty estimation using a connectivity-based ranking monte carlo method. *Stoch. Environ. Res. Risk Assess.* **35**(5), 1039–1049 (2021)
- Morvillo, M., Rizzo, C.B., de Barros, F.P.J.: A scalable parallel algorithm for reactive particle tracking. *J. Comput. Phys.* **446**(110), 664 (2021)
- Morvillo, M., Im, J., de Barros, F.P.J.: VisU-Hydra: a computational toolbox for groundwater contaminant transport to support risk-based decision making. *Front. Earth Sci.* **10**, 916198 (2022). <https://doi.org/10.3389/feart.2022.916198>
- Moslehi, M., Rajagopal, R., de Barros, F.P.J.: Optimal allocation of computational resources in hydrogeological models under uncertainty. *Adv. Water Resour.* **83**, 299–309 (2015)
- Neuman, S.P., Tartakovsky, D.M.: Perspective on theories of non-fickian transport in heterogeneous media. *Adv. Water Resour.* **32**(5), 670–680 (2009)
- Remy, N., Boucher, A., Wu, J.: *Applied geostatistics with SGeMS: A user's guide*. Cambridge University Press(2009)
- Rizzo, C.B., de Barros, F.P.J.: Minimum hydraulic resistance uncertainty and the development of a connectivity-based iterative sampling strategy. *Water Resour. Res.* **55**(7), 5593–5611 (2019)
- Rizzo, C.B., Nakano, A., de Barros, F.P.J.: Par2: parallel random walk particle tracking method for solute transport in porous media. *Comput. Phys. Commun.* **239**, 265–271 (2019)
- Rolle, M., Eberhardt, C., Chiogna, G., Cirpka, O.A., Grathwohl, P.: Enhancement of dilution and transverse reactive mixing in porous media: experiments and model-based interpretation. *J. Contam. Hydrol.* **110**(3–4), 130–142 (2009)
- Rubin, Y.: *Applied stochastic hydrogeology*. Oxford University Press (2003)
- Rubin, Y., Cushey, M., Bellin, A.: Modeling of transport in groundwater for environmental risk assessment. *Stoch. Hydrol. Hydraul.* **8**(1), 57–77 (1994)
- Sahimi, M.: *Flow and transport in porous media and fractured rock: from classical methods to modern approaches*. John Wiley & Sons (2011)
- Salamon, P., Fernández-García, D., Gómez-Hernández, J.J.: A review and numerical assessment of the random walk particle tracking method. *J. Contam. Hydrol.* **87**(3–4), 277–305 (2006)
- Schwede, R.L., Cirpka, O.A., Nowak, W., Neuweiler, I.: Impact of sampling volume on the probability density function of steady state concentration. *Water Resour. Res.* **44**(12), (2008)
- Selroos, J.O.: A stochastic-analytical framework for safety assessment of waste repositories: 1. theory. *Groundwater* **35**(3), 468–477 (1997)
- Sole-Mari, G., Riva, M., Fernández-García, D., Sanchez-Vila, X., Guadagnini, A.: Solute transport in bounded porous media characterized by generalized sub-gaussian log-conductivity distributions. *Adv. Water Resour.* **147**(103), 812 (2021)
- Soltanian, M.R., Behzadi, F., de Barros, F.P.J.: Dilution enhancement in hierarchical and multiscale heterogeneous sediments. *J. Hydrol.* **587**(125), 025 (2020)
- Tonina, D., Bellin, A.: 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 (2008)
- Villiermaux, E.: Mixing by porous media. *C R Mécanique* **340**, 933–943 (2012)
- Zheng, C., Bennett, G. D., *Applied contaminant transport modeling*. Wiley-Interscience, New York (2002)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.