



Emergence of non-Fickian transport in truncated pluri-Gaussian permeability fields

Eugenio Pescimoro¹ · Matteo Icardi¹ · Giovanni Porta² · Marco Bianchi³

Received: 2 May 2022 / Accepted: 23 September 2022
© Crown 2022

Abstract

We present a numerical simulation study of advective–diffusive scalar transport in three-dimensional high-contrast discontinuous permeability fields, generated with a truncated pluri-Gaussian geostatistical approach. The numerical experiments are run with an Eulerian approach using a novel unified numerical framework based on the finite-volume library OPENFOAM® (Weller et al. in *Comput Phys* 12(6):620–631, 1998), for (1) generating random pluri-Gaussian porous media, (2) solving the steady state Darcy-scale flow, (3) solving the advection diffusion equation, (4) computing post-processing quantities such as first order statistics, spatial probability density functions and breakthrough curves. A range of permeability contrasts, correlation lengths, and Péclet numbers are tested to assess their relative weight on transport control and for the first time, the deviation of a calibrated macrodispersive model from the Fickian transport is quantified. We identify a hierarchy of non-Fickian transport triggering factors. From the tested scenarios, permeability contrast is the main controlling parameter for the anomalous transport behaviour as it enhances the generation of preferential flow paths which are characterised by high advective flow velocities. The Péclet number and the characteristic length at which facies transitions are observed as secondary factors.

✉ Eugenio Pescimoro
eugenio.pescimoro@gmail.com

Matteo Icardi
matteo.icardi@nottingham.ac.uk

Giovanni Porta
giovanni.porta@polimi.it

Marco Bianchi
marcob@bgs.ac.uk

¹ Industrial and Applied Mathematics, University of Nottingham, University Park, Nottingham NG7 2RD, Nottinghamshire, UK

² Civil and Environmental Engineering, Politecnico di Milano, Piazza L. Da Vinci, 32, 20133 Milan, Lombardia, Italy

³ British Geological Survey, Nicker Hill, Keyworth NG12 5GG, Nottinghamshire, UK

Keywords Heterogeneous media · Pluri-Gaussian simulations · Solute transport · OpenFoam · Non-Fickian transport

Mathematics Subject Classification 76

1 Introduction

Subsurface flow and solute transport modelling is used in several engineering and environmental fields (CO₂ storage, groundwater remediation, oil recovery) where mathematical and computational models play a central role in supporting the reliability of analysis and design strategies. The effectiveness of advection–dispersion models in describing solute transport in highly heterogeneous media such as geological formations has been questioned (Adams and Gelhar 1992; Barlebo et al. 2004; Fiori et al. 2016), and the definition of appropriate models and their parameterization remains an open field of research (Zinn and Harvey 2003; Jankovic et al. 2017; Bianchi and Zheng 2016; Yin et al. 2020). An important challenge is how to simulate non-Fickian behaviour, which originates mainly from physical heterogeneities emerging across multiple scales (Dentz et al. 2011; Berkowitz et al. 2006; Gelhar and Axness 1983). Transport is defined as anomalous or non-Fickian when solute plumes and breakthrough curves display a significant departure from the predictions made by an advective–dispersive model where dispersion is expressed with a Fickian analogy, i.e. mechanical dispersion and molecular diffusion are grouped together in a single effective coefficient (Bear 2012).

Approaches to modelling solute transport in heterogeneous porous media largely differ depending on the scale of interest. In this work we start from a mesoscale, which corresponds to a resolution where geological porous media can be described by an equivalent continuum with spatially heterogeneous properties (de Barros et al. 2022; Riva et al. 2008). At this scale, solute transport is governed by two separate mechanisms: advection and local hydraulic dispersion which includes the contributions of molecular diffusion and mechanical dispersion. At the mesoscale, spatial heterogeneity is explicitly represented, most commonly using a statistical model. We then move to macroscale modelling, where the aim is to define an effective model able to describe the dynamics of the system without an explicit description of the underlying heterogeneity. In classical descriptions (Dagan 2012), velocity at these scales may be interpreted as the average Darcy velocity while the hydraulic dispersion coefficient turns into a macrodispersion coefficient, employed to quantify the effect of heterogeneity on solute spreading. This model has been questioned in the literature and alternative non-Fickian effective models have been proposed (Hansen et al. 2018; Zech et al. 2021; Neuman and Tartakovsky 2009). These approaches mainly focused on cases where the underlying (mesoscale) log-conductivity field has a Gaussian distribution. Beyond this specific case, the validity of the Advection Dispersion Equation (ADE) based macrodispersive models are not clearly identifiable a priori, although they are certainly heavily controlled by the degree of heterogeneity of porous media properties (Neuman and Tartakovsky 2009) and their spatial organisation. For example this latter point was recently addressed in Yin et al. (2020), who investigated the

role played by the injection area and the correlation length in activating anomalous transport mechanisms. The persistence of this anomalous transport behaviour at the macroscale can be due, for example, to these regions where the flow paths create preferential fast channels (Edery et al. 2014), a feature that also influences reactive transport settings (Edery et al. 2016, 2021).

In this work we investigate solute transport and the onset of anomalous or non-Fickian transport behaviour in high-contrast heterogeneous permeability fields, generated with the geostatistical pluri-Gaussian truncated (PGS) method (Mariethoz et al. 2009). Solute transport has been widely investigated in continuous Gaussian and non-Gaussian permeability fields (Gotovac et al. 2009; Sole-Mari et al. 2021), and methods have also been proposed to handle non-continuous fields, suitable to reproduce geomaterials where property transition is marked by sharp interfaces (Bianchi and Pedretti 2017). PGS random fields are used in this context to model actual subsurface geological media in a sedimentary setting. In this context this model is used to link an assumed geological architecture or structure, e.g. driven by sedimentological rules, with the spatial distribution of physical properties such as porosity or hydraulic conductivity. This allows to create fields starting from given geological assumptions and explicitly control the connectivity of high- and low-permeability facies. Therefore, PGS can be employed to reproduce and interpret the emergence of non-Fickian transport traits observed in real geological media. Simulation of solute transport in alluvial settings, represented by discontinuous conductivity fields, has been considered by a number of studies in the recent literature. Discontinuous permeability fields with a high connectivity degree and sharp contrast between regions are recognised among the most important factors that regulate the transport of solute (Zhang et al. 2013; Bianchi and Zheng 2016). Facies properties can be qualitatively linked to non-Fickian parameters for alluvial aquifers (Zhang et al. 2014), however such a link remains hard to quantify in a predictive fashion. This is likely due to the fact that several factors can contribute to the emergence of a non-Fickian behaviour of solute travel times. As noted by Zhang et al. (2015), the Péclet number provides useful information on the duration of the anomalous transport while the correlation length controls the connectivity and, therefore, the onset of non-Fickian behaviour.

Starting from these existing results, our aim here is to investigate the connectivity and conductivity contrast thresholds that drive a transition between Fickian and non-Fickian response. Our objective is to then rank the factors triggering transition non-Fickian transport. To this end, we quantify the deviation of the results obtained from numerical simulations in PGS domains from Fickian behaviour by comparing them to the analytical solution of the advection–dispersion equation. We rely on a quantitative approach aimed at capturing the discrepancy between mesoscale simulations and a macrodispersive model, rather than focusing on a detailed characterisation of the processes involved. This allows us to identify the physical and structural thresholds that can lead to non Fickian transport and ultimately contributes to the definition of aquifer typing approaches where the relevance of non Fickian transport features may become identifiable from a knowledge of the field properties.

To achieve these objectives we rely on numerical simulations, by solving the advection–dispersion equation in heterogeneous media using an Eulerian finite volume method. This approach is implemented as a parallel open-source code based on

OPENFOAM® (Weller et al. 1998), as part of the SECUReFoam library (Municchi et al. 2022). The advantages of the Eulerian approach are that it allows the computation of Péclet number and that an accurate simulation of solute low concentration tails does not require a large particle ensemble as with Lagrangian formulations, which have often been used in the recent literature (Edery et al. 2014; Savoy et al. 2017; Dentz et al. 2011). Moreover, the Eulerian description is closer to the experimental conditions where results are often obtained in terms of molar or mass concentration while Lagrangian approaches need to be post-processed to obtain local concentration fields.

From an operational perspective, our approach is based on a single computational framework, including a geostatistical algorithm for permeability field generation, a numerical code for flow and transport simulation, and post-processing tools. This is an interesting feature of our approach as the synthetic generation of realistic geological domains remains one of the main challenges in modelling flow and transport (Heße et al. 2014). Several approaches are available to reproduce complex subsurface structures [sequential Gaussian simulations (Dimitrakopoulos and Luo 2004), Markov chain probability (Carle and Fogg 1997), Multiple-point statistics (Strebelle 2002)] as well as a number of geostatistical open toolboxes [GSLib (Deutsch and Journel 1998), T-PROGS (Carle 1999)]. Nevertheless, few open-source tools exist that provide integrated geostatistical, flow and transport simulation solvers [OpenGeoSys (Kolditz et al. 2012), porousMultiphaseFoam (Horgue et al. 2015), DuMux (Flemisch et al. 2011)].

This work is structured as follows: in Sect. 2 we give the mathematical overview of the problem, in Sect. 3 we describe the testcases and summarise the numerical methodologies. Numerical results and the post-processing are presented in Sect. 4, before we draw conclusions and give some guidelines about the emergence of non-Fickian transport. For the sake of clarity, the terms “facies” (uncountable) and “category” as well as “lithotype” and “truncation” rule will be used interchangeably depending on the context.

2 Methods

We describe here the methods underpinning our numerical simulations. We start by presenting the geostatistical framework and then move to the description of the physical problem, i.e. the flow and transport setting.

2.1 Geostatistical model

Permeability fields are generated via the pluri-Gaussian Simulation (PGS) method, i.e. applying a truncation rule to continuous multivariate Gaussian random fields (GRF) (Mariethoz et al. 2009). Fields generated with this approach are characterised by:

- discontinuous permeability fields characterised by a discrete number of zones of uniform permeability whose spatial arrangement is the result of a specific truncation rule [i.e., Lithotype rule (Armstrong et al. 2011)];

- high geological realism since the truncation rule allows simulating observed geometrical relations between geological facies (Koltermann and Gorelick 1996; Linde et al. 2015; Armstrong et al. 2011).

GRFs can be generated in the frequency domain by multiplying independent complex Gaussian random variables by the spectral representation of the covariance function. The spatial field is then reconstructed by applying the inverse Fourier transform to the spectral GRF. To ensure independence of the random field generation from the mesh-discretisation and to allow arbitrary unstructured grids, we apply an explicit discrete inverse Fourier transform discretised with N_f frequencies in each direction Following Mandelbrot and Van Ness (1968); Heße et al. (2014), a discrete-in-frequencies continuous-in-space representation of a complex GRF is therefore given by:

$$Z(\mathbf{x}) = \sum_{j=0}^{N_f} \cos(2\pi \mathbf{a}_j \cdot \mathbf{x}) \sqrt{S(\mathbf{a}_j)} W_j + i \sum_{j=0}^{N_f} \sin(2\pi \mathbf{a}_j \cdot \mathbf{x}) \sqrt{S(\mathbf{a}_j)} W'_j \quad (1)$$

where \mathbf{x} is the position vector, $\mathbf{a}_j = (a_{x,j}, a_{y,j}, a_{z,j})$ is the j^{th} frequency vector, W_j and W'_j are independent complex Gaussian random variable and $S(\mathbf{a}_j)$ is the amplitude of the spectral measure. From Z , we can then extract two independent Gaussian random fields from its real and imaginary parts.

The covariance function of a stationary field quantifies the covariance $\gamma(\mathbf{r})$ between a pair of values of a random variable located at points separated by the distance \mathbf{r} . We denote the correlation function as $\rho(\mathbf{r})$ and the variance as σ^2 (where $\gamma(\mathbf{r}) = \sigma^2 \rho(\mathbf{r})$).

In this work, we assume an exponential correlation function

$$\rho(\mathbf{r}) = 1 - e^{-\sqrt{\frac{r_x^2}{\lambda_x^2} + \frac{r_y^2}{\lambda_y^2} + \frac{r_z^2}{\lambda_z^2}}} \quad (2)$$

with corresponding spectrum

$$S(\mathbf{a}) = \sigma^2 \|\boldsymbol{\lambda}\|^d \frac{\Gamma\left(\frac{d+1}{2}\right)}{\left(\pi \left(1 + a_x^2 \lambda_x^2 + a_y^2 \lambda_y^2 + a_z^2 \lambda_z^2\right)\right)^{\frac{d+1}{2}}}, \quad (3)$$

where $d = 3$ is the number of dimensions, Γ is the Gamma function, $\boldsymbol{\lambda} = (\lambda_x, \lambda_y, \lambda_z)$ are the correlation lengths.

GRFs are continuous fields, but geological media are often characterised by abrupt changes in physical and chemical properties. With the PGS approach discontinuous patterns are reproduced from the truncation of two GRFs according to a *lithotype* or *truncation rule* (Fig. 1), which bins continuous values into a set of categories.

The smooth transition which characterises the GRF is then replaced by $n = (N_r + 1)(N_s + 1)$ categorical values where N_r and N_s are the number of thresholds applied via the truncation rule to the two GRFs. In this sense, the “truncated” adjective refers

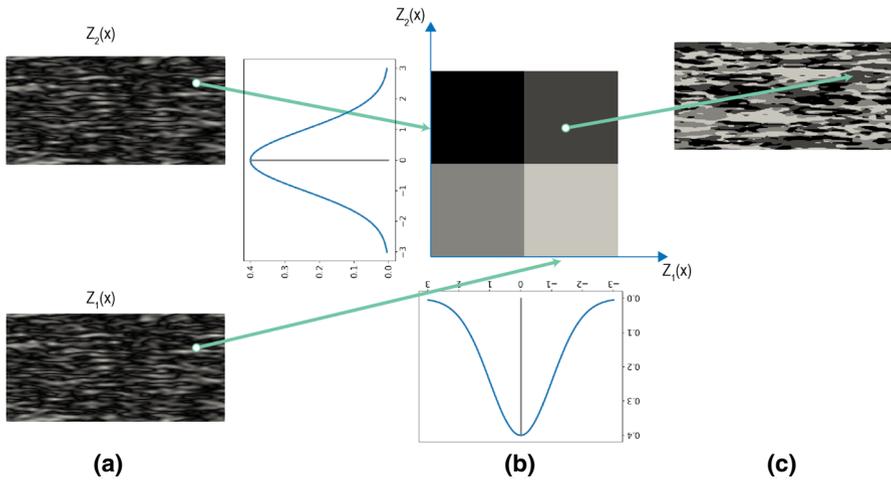


Fig. 1 Truncated pluri-Gaussian simulation. **a** Continuous multivariate Gaussian random fields Z_1 and Z_2 generation; **b** truncation rule for four facies domain and its corresponding thresholds on the Gaussian distribution of the variables; **c** sample of a two-dimensional truncated pluri-Gaussian random field. The arrows indicate the contribution of the two GRFs in assigning a given category at a selected location in space

to a GRF that has been discretised through a binning process. The probability, i.e. the proportion, of the facies φ_i is obtained from

$$p_{\varphi_i}(\mathbf{x}) = [G(r_i) - G(r_{i-1})][G(s_i) - G(s_{i-1})] \quad i = 1 \dots n \tag{4}$$

where n is the number of categories, G is the cumulative Gaussian distribution with the mean and the variance typical of each field. The lithotype rule allows to control the probability of two different categories (or facies) to be in direct contact. This constitutes a fundamental feature as it allows the simulated field to reflect geological transition patterns observed in field data. According to the conceptual steps normally used in PGS geostatistics, transition patterns are captured along the vertical direction by processing field sample information through transition probability matrices (Carle and Fogg 1996; Weissmann et al. 1999) while field observations and/or established conceptual models of geological environments are used as guidance for the estimation of transition patterns in the horizontal directions (Armstrong et al. 2011). In this work, we assume the single truncation diagram, in Fig. 1. In our simulations we vary the correlation lengths λ of the underlying GRFs and the permeability values assigned to different categories. The four categories have equal probability and therefore volumetric fractions $p_{\varphi_i} = 25\%$. The distribution of the multivariate random variables adopted to generate the underlying continuous Gaussian random fields in this study has mean $\mu = 0$ and $\sigma = 1$ and their correlation function is exponential.

2.2 The flow model

We assume fluid flow obeys the standard Darcy's equation which reads

$$\mathbf{V} = -\frac{\mathbf{k}}{\mu}(\nabla p + \rho g \nabla z), \quad (5)$$

where \mathbf{V} is the Darcy velocity vector [LT^{-1}], \mathbf{k} is the permeability tensor [L^2], μ is the dynamic viscosity [$ML^{-1}T^{-1}$], p is the pressure [$MT^{-2}L^{-1}$], ρ is the fluid density [ML^{-3}], g is the gravity constant [LT^{-2}] and $\nabla z = (0, 0, 1)$ [$-$] is an upward unit vector. For this study we set $g = 0$ as any influence of the solute on the liquid density is assumed to be negligible.

The flow solver implemented in OPENFOAM[®] (Weller et al. 1998) is based on Eq. (5) assuming an incompressible fluid. Therefore pressure can be computed according to a Poisson equation

$$\nabla \cdot \mathbf{V} = -\nabla \cdot \frac{\mathbf{k}}{\mu} \nabla p = 0 \quad (6)$$

where we have assumed no sources or sinks are present and the gravity term is zero. The permeability tensor is, from this point, assumed diagonal and isotropic, i.e., $\mathbf{k} = k\mathbb{I}$, \mathbb{I} being the identity matrix. Boundary conditions for the pressure are zero gradient on lateral sides and a fixed gradient of 50 Pa/m in the longitudinal direction.

2.3 Local transport model

The advective flux per unit area \mathbf{J}_{adv} [LT^{-1}] is the product of the advective Darcy velocity \mathbf{V} [LT^{-1}] and solute concentration c [$-$]

$$\mathbf{J}_{adv} = \mathbf{V}c. \quad (7)$$

In line with previous work (Edery et al. 2014), we neglect mechanical dispersion and model the diffusive fluxes \mathbf{J}_{mol} [LT^{-1}] as

$$\mathbf{J}_{mol} = -\phi \mathbf{D}_{mol} \nabla c \quad (8)$$

where \mathbf{D}_{mol} [L^2T^{-1}] is the molecular diffusion tensor and ϕ is the porosity of the medium. Summing up the advective and diffusive fluxes, the conservation of mass yields the advection–diffusion equation, which, for the case of isotropic diffusion and porosity and no source/sink terms is

$$\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{v}c) - \mathbf{D}_{mol} \nabla^2 c = 0 \quad (9)$$

where $\mathbf{v} \equiv \mathbf{V}/\phi$ is the fluid velocity, i.e. the velocity that would be measured by a flow meter in the porous domain and $\mathbf{D}_{mol} = D\mathbb{I}$. We impose a constant concentration on the whole inlet face of the domain (9) and zero gradient on all the other sides.

In this study, to focus on the effects of the heterogeneity, the geostatistical parameters and the Péclet number, we have made strong assumptions on the permeability (isotropic and diagonal), porosity (constant) and neglected mesoscopic dispersion. Whilst preliminary tests suggested these do not impact the main findings of this work, the investigation of these processes may be tackled in future contributions.

2.4 Macrodispersion model

Transport mechanisms described so far characterise the transport behaviour at the mesoscale, i.e. where geological and flow resolution allows for heterogeneity to be modelled explicitly. However, macroscale models aim to provide an overall description while using an effective/upscaled advection–dispersion equation neglecting heterogeneity. Here, we only focus on transport along the main velocity direction and the longitudinal dispersion processes, therefore we will compare our results with a one-dimensional advection–dispersion equation:

$$\frac{\partial C}{\partial t} + \overline{v_x} \frac{\partial C}{\partial x} - D_{xx}^{mac} \frac{\partial^2 C}{\partial x^2} = 0, \quad (10)$$

where C is the section-averaged concentration, D_{xx}^{mac} is the longitudinal component of the macrodispersion tensor and $\overline{v_x}$ indicates the spatial average of the longitudinal component of the velocity. Macrodispersion in Fickian transport models can be predicted or inferred. Predictive macrodispersion estimates are often evaluated computing the product between a typical length scale and an average velocity (13) while inferred macrodispersion assessments can be performed using the moments' method (20) or applying the least square method to the breakthrough curve, as illustrated in Sect. 2.5.1.

2.5 Quantities of interest

The record of the section-averaged concentration in time at a control section (e.g. outlet boundary or an arbitrary point) constitutes the breakthrough curve (BTC). Under a continuous injection, the BTC is equivalent to the cumulative density function (CDF) of the arrival times of the solute mass ($F(t)$) while its time derivative, which is a concentration rate, is the probability density function (PDF) of the arrival times ($f(t)$). These functions are typically obtained by injecting a pulse in time or a constant concentration at the inlet (or an injection point).

To enable the comparison between simulations considering different parameters and different duration, we consider a dimensionless time T obtained by dividing t by the average travel time, calculated as the ratio between the longitudinal domain dimension and the average fluid velocity. This quantity is equivalent to the injected pore volume. The section averaged concentration at the outlet is non-dimensionalised by dividing it by the single inlet concentration and is represented by \bar{c} .

In the post processing phase of the simulation results, the following quantities were estimated:

Péclet number

$$Pe_x [-] = \frac{\overline{v_x} \lambda_x}{D_{mol}}; \tag{11}$$

effective permeability

$$k_x^{eff} [m^2] = -\frac{\overline{v_x} \mu}{\frac{\partial p}{\partial x} - \rho g}; \tag{12}$$

nominal macrodispersion

$$D_{mac}^{ij} [m^2/s] = \phi \lambda^T V \tag{13}$$

where λ and V are typical lengths and velocity vectors. Equation (13) allows the macrodispersion matrix to be approximated a priori starting from geostatistical (correlation length λ) and flow (velocity V) data, independent of the BTC data. Concentration data coming from the BTC constitutes the basis for the methods adopted to estimate the macrodispersion from the mesoscale simulations, as illustrated in Sect. 2.5.1.

2.5.1 Breakthrough curve and inverse Gaussian approximation

The mass arrival time distribution simulated with the one-dimensional advection–dispersion equation is the inverse Gaussian distribution. This corresponds to the analytical solution of Eq. (10) in a semi-infinite one-dimensional domain with a Dirac-delta initial condition. For practically relevant parameters, this is almost indistinguishable from the solution on a finite domain with a Dirac-delta (in time) concentration injection at the inlet. For our problem with a continuous injection at the inlet, due to the linearity of the problem, the BTC is well approximated by the integral in time of the Inverse Gaussian distribution, computed for a fixed section in space (the outlet in our case). When transport behaviour is Fickian, we can approximate the experimental BTCs with the cumulative density function of the Inverse Gaussian distribution as

$$F(T; \mu_1, \nu) = \bar{c} = \Phi \left(\sqrt{\frac{\nu}{T}} \left(\frac{T}{\mu_1} - 1 \right) \right) + e^{\frac{2\nu}{\mu_1}} \Phi \left(-\sqrt{\frac{\nu}{T}} \left(\frac{T}{\mu_1} + 1 \right) \right) \tag{14}$$

where Φ is the standard normal cumulative distribution function, μ_1 is the first order statistical moment of the concentration rate distribution and ν is a shape parameter. The PDF of the solute arrival times can be obtained through a time derivative of (14) and corresponds to the PDF of the solute arrival times. This PDF is expressed as (Tartakovsky and Dentz 2019)

$$f(T; \mu_1, \nu) = \frac{\partial \bar{c}}{\partial T} = \sqrt{\frac{\nu}{2\pi T^3}} \exp \left[-\frac{\nu(T - \mu_1)^2}{2\mu_1^2 T} \right]. \tag{15}$$

Other analytical solutions are available for different boundary conditions on finite domains (Van Genuchten 1982). For the purposes of this paper, we will only consider the Inverse Gaussian model as a reference for Fickian transport due to its simpler analytical formula more suitable to fitting and moment matching. The macrodispersive solution is generally a good approximation (Berkowitz et al. 2006) if

- domain is large;
- experiment time is long;
- domain’s properties are ergodic.

IN the assumption of a Fickian model such as (15) arrival times display a sharp and exponential tail as $t \rightarrow \infty$. Non-Fickian transport processes have a clear impact on the shape of the PDF of the arrival times: early arrival concentrations raise the PDF peak and power low scaling emerges prior to exponential decay (Berkowitz et al. 2006; Ederly et al. 2014).

The moments’ method

Following Yu et al. (1999); Kreft and Zuber (1978), the estimation of the statistical moments of the cumulative Inverse Gaussian is performed by approximating its parameters μ_1 and ν as

$$E[\bar{c}] = \mu_1 \tag{16}$$

$$Var[\bar{c}] = \mu_2 - \mu_1^2 = \frac{\mu_1^3}{\nu}. \tag{17}$$

To compute the first and second order moments we used

$$\begin{aligned} \mu_1 &= \int_0^{+\infty} fT dT = \int_0^{+\infty} F'T dT = - \int_0^{+\infty} F dT + [FT]_0^{+\infty} \\ &= - \sum_{i=0}^{+\infty} F_i \Delta T + F_{+\infty} T_{+\infty}, \end{aligned} \tag{18}$$

$$\begin{aligned} \mu_2 &= \int_0^{+\infty} fT^2 dT = \int_0^{+\infty} F'T^2 dT = -2 \int_0^{+\infty} FT dT + [FT^2]_0^{+\infty} \\ &= -2 \sum_{i=0}^{+\infty} F_i T_i \Delta T + F_{+\infty} T_{+\infty}^2. \end{aligned} \tag{19}$$

The estimated effective velocity and macrodispersion coefficient can be estimated from the statistical moments as

$$V_x = \frac{L_x}{\mu_1} \tag{20}$$

$$D_{xx}^{mac} = \frac{\mu_2 V^3}{2L_x} \tag{21}$$

where L_x is the distance between the inlet and outlet sections (in our case the domain length). To quantify the distance between the numerical outputs and the Inverse Gaus-

sian approximation, a normalised error e was defined as

$$e = \frac{||\bar{c}(T) - F(T)||}{||\bar{c}(T)||} \cdot 100. \quad (22)$$

Least squares estimation

Parameter estimation is performed by minimising the least squared error between numerical data and the models (14) and (15). Under the assumption of identically distributed and uncorrelated errors this procedure corresponds to a maximum likelihood estimation. This procedure is applied to three types of data

- probability density function of the solute arrival times, obtained by numerical differentiation of the BTC values at the outlet;
- cumulative density function of the concentration arrival times (i.e. the BTC itself);
- PDF of the arrival times obtained in an interval of 0.5 dimensionless time unit, centred around the peak of the probability density function of the arrival times.

In the first and third case, the analytical function used as reference to perform the least squares fitting is the probability density function of the inverse Gaussian distribution given by Eq. (15) while for the second case the analytical function is Eq. (14). For all cases the analysis was performed using Python library `lmfit` constraining the estimation so that μ_1 and ν were always non-negative. Values of the estimated parameters uncertainty are also obtained from the diagonal entries of the parameters covariance matrix computed by `lmfit` and were used to assess the reliability of the estimate. The initial values for the least square estimation were set equals to values computed for μ_1 and ν with the moments' method.

3 Numerical experiments

Geostatistical, flow and transport numerical simulations were conducted over hexahedral domains which represent a portion of the subsurface with dimensions $(L_x/l, L_y/l, L_z/l) = (2, 1, 1)$ where L_i are the dimensions of the domain and we took $l = L_y = L_z$ as the reference length. The mesh is unstructured and characterised by cubic cells of dimension $d/l = 0.01$, so that the total number of cells is 2×10^6 .

The permeability distribution within the domain corresponds to the field generated with a PGS simulation while porosity is assumed homogeneous over the domain. All the simulated fields considered in this study share the correlation function reported in Eq. (2), the number of permeability zones, as well as the volumetric proportion for each of the facies (see Table 1). We investigate the variability of the observed output and of the estimated parameters as a function of three inputs: geostatistical parameters (e.g., correlation length used to generate the conductivity fields), hydraulic properties (i.e., permeability) and transport regime, defined in terms of Pe .

Based on the assigned permeability values, we distinguish two cases: low and high contrast. For both cases the permeability values k_i assigned to the four geomaterials considered are evenly spaced on a logarithmic scale. However, for the low permeability contrast case the four permeability values range between 10^{-10} and 10^{-13} m^2 with

Table 1 Parameters kept constant throughout the simulations

<i>Geostatistical parameters</i>	
Correlation function	Exponential
Number of facies	4
Volumetric proportion	25%
<i>Flow parameters</i>	
Pressure gradient [Pa/m]	50
<i>Transport parameters</i>	
Fixed inlet concentration [-]	1

a relative ratio $\log_{10}(k_i/k_{i+1}) = 1$ while for the high permeability contrast case permeability values range between 10^{-9} and 10^{-15} m² and $\log(k_i/k_{i+1}) = 2$. Boundary conditions for the pressure are set as zero gradient along the lateral boundaries and a one dimensional pressure gradient of 50 Pa/m aligned with the longitudinal direction. A constant concentration is imposed on the inlet face, the remaining boundaries are considered impermeable.

The simulation workflow is divided into three steps

- geostatistics: the permeability domain is generated using the truncated pluri-Gaussian algorithm;
- flow: Eq. (5) is solved with the prescribed boundary conditions and provides the steady state flow field;
- transport: advection–dispersion transient model is solved with continuous injection for each simulation time step.

The simulations are run within the open-source OPENFOAM[®]-based library SECUREFoam (Municchi et al. 2022) which includes the `setRandomField` utility for truncated pluri-Gaussian simulations, `simpleDarcyFoam` and `adaptiveScalarTransportFoam` solvers for flow and transport simulations. Most of the simulations were run in parallel on 96 cores split between 8 HPC nodes. An adaptive time step tied to the Courant number was implemented together with an automatic check on the section-average outlet concentration value which stopped the transport simulation when a value of 0.99 on the outlet boundary was reached. In this setting, the overall simulation time ranges between 1 and 7 hours depending on the permeability contrast adopted, with high contrast cases being characterised by larger CPU costs. Transport simulation are the most expensive of the three simulation steps, accounting for between the 70 and 95 % of the total computational time for the low or high permeability contrast setting, respectively. Steady state flow has been solved by discretising the Darcy equation combined with mass balance, in a primal (non-mixed) form, i.e., with the pressure as the only variable. This exactly satisfies the mass balance at the faces, as in the finite volume framework the velocity is discretised as fluxes over the faces and so is the divergence term in the pressure equation. In terms of computational time, this means that the solution for the flow field is typically achieved in a few minutes, while the geostatistical simulation and post-processing could take up to 1h approximately.

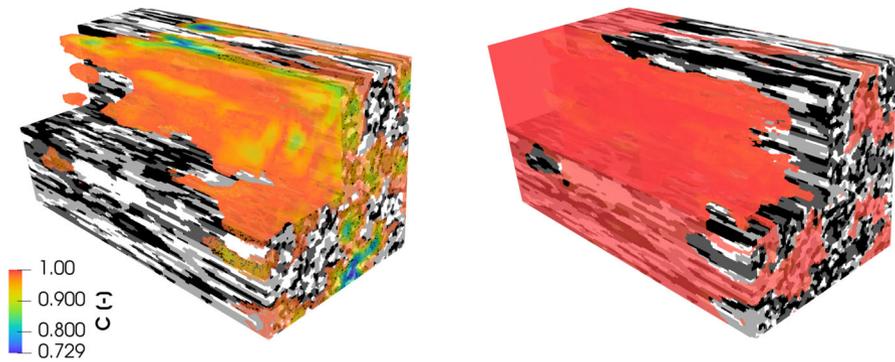


Fig. 2 Solute plume distribution in high contrast permeability domain at late time. Domain sizes are $2 \times 1 \times 1$ and correlation lengths along the three directions are set to $(0.8, 0.1, 0.1)$. On the left panel it is possible to observe how low solute concentration values (0 blue–0.99 red) are confined to low permeability regions (10^{-12} [m²] dark grey - 10^{-13} [m²] black) while on the right panel high concentration values (0.99 blue–1.00 red) are highlighted and their spatial distribution clearly show that saturated zones are concentrated in highly permeable regions (10^{-10} [m²] dark grey– 10^{-11} [m²] black)

4 Results

The results presented in this section aim to assess the impact of the parameters related to the PGS fields on solute transport processes. To this end, first, we compare the PDFs of velocity point values obtained in the considered fields. Then, we move to the analysis of the transport simulations and we provide a qualitative assessment of the variability exhibited by results obtained from realisations of the conductivity fields generated with identical geostatistical parameters (Sect. 4.2). Finally, we analyse the impact of three physical parameters on arrival time PDFs, namely permeability contrast (Sect. 4.3), the longitudinal correlation length used to generate the fields (Sect. 4.4) and Péclet number (Sect. 4.5).

Before moving to these detailed analyses, Fig. 2 illustrates the simulation of the solute plume at late times through a PGS field with high contrast permeability and characterised by longitudinal correlation length of 0.8 m. Figure 2 on the left highlights the regions where concentration values falls beneath the 0.99 threshold while the right panel visualises the regions where concentration values fall between 0.99 and 1. It is possible to observe that the transport of the solute is facilitated in the high permeable regions of the domain (white and light grey) while low permeability ones (dark grey and black) form a flow barrier that impede advective solute transport.

4.1 Velocity PDFs

The velocity PDF has a direct influence on non-Fickian transport features (Comolli et al. 2019) as large differences between permeability values may induce bi- or multi-modal velocity distributions (Yin et al. 2020). The PDFs of point velocity values with increasing longitudinal correlation lengths λ_x are shown in Fig. 3. The longitudinal velocity V_x normalised with the average longitudinal velocity is reported on the hor-

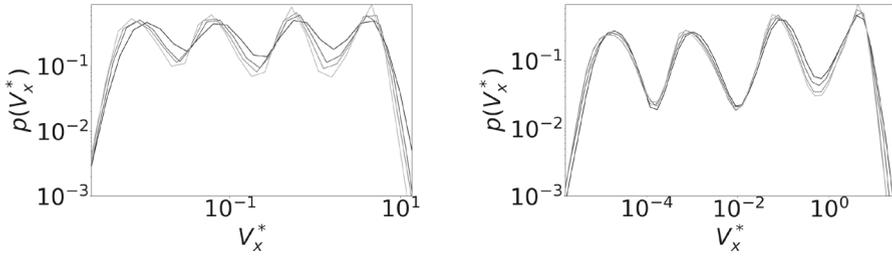


Fig. 3 PDFs of the longitudinal velocity component for low and high contrast. Velocity PDFs are shown for low (left) and high (right) permeability contrast. The correlation lengths λ_x span between 0.4 (darker lines) and 1.0 (lighter lines)

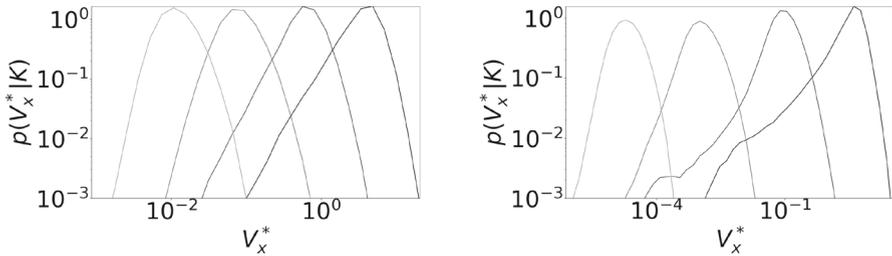


Fig. 4 Conditional PDFs of the longitudinal velocity values in low (left) and high (right) permeability contrast. Results are shown for correlation lengths $\lambda_x = 0.4$. The curves are shown with different colours depending on the facies permeability, i.e., lighter colours correspond to low permeability and darker colours to high permeability media

horizontal axis of Figs. 3 and 4 as V_x^* . The vertical axis of Fig. 3 reports the probability density distribution as a function of the longitudinal velocity $p(V_x^*)$. As expected, the velocity distributions in Fig. 3 are comparable for different correlation lengths. All distributions show four peaks of similar height corresponding to the four facies that populate the domain. However, as the correlation length increases, the four peaks become sharper reflecting the formation of preferential flow paths where velocities are lumped around the mean velocity of a given facies, each corresponding to a mode of the distributions (Fig. 4), similar to what is shown for a bimodal permeability distribution by Yin et al. (2020). This result also indicates that with a decrease of λ_x the distribution of velocity values progressively converges towards a uniform distribution across the whole range. Comparing the amplitude of the peaks in the two panels of Fig. 3 we observe that as the permeability contrast increases the four modes of the distribution appear more distinct for high contrast than for low contrast. Note also that the high contrast distribution spans a much wider interval of velocity values as compared to the low contrast one.

Figure 4 reports the conditional PDFs $p(V_x^* | k = k_i)$, where each distribution considers only longitudinal velocities values computed in cells associated with a given facies ($i = 1 \dots 4$). Velocities in highly permeable regions show an asymmetric distribution, characterised by a pronounced peak and a leftward tail. Conversely, velocity values observed in the low permeability regions tend to assume a symmetric and compact distribution. This distinct behaviour is particularly evident in high contrast media.

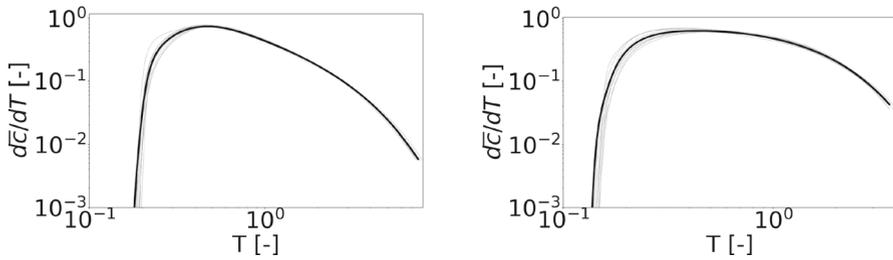


Fig. 5 PDFs of solute arrival times associated with 10 realisations with the same correlation length ($\lambda_x = 0.8$ m). Low and high permeability contrast on the left and right side of the panel respectively

This means that high-permeability regions may feature a broad distribution of velocity values because of the overall connectivity of the field. Highly connected regions give rise to fast channels in formations featuring large values of k but poorly connected regions may also involve high-permeability cells.

4.2 Variability of transport behaviour across multiple realizations

Figure 5 displays the overlap of the PDF of the solute arrival times, obtained taking the time derivative of the BTC ($d\bar{c}/dT$) obtained from 10 realisations of permeability fields generated with the same geostatistical parameters. The observed variability tends to be greater for early times while at later times the different realizations attain similar values. This behaviour is the result of a continuous injection in the whole inlet face, where the solute broadly explores the facies' heterogeneities as the solute fill the whole domain. Local injections in high/low conductivity have been considered in previous works (Zhang et al. 2015) and may display a larger variability within the sample. A comparative study on the averaging property of Eulerian simulations in local injection setting will be considered for future works. The outlined behaviour does not show qualitatively relevant differences between low (left panel) and high (right panel) permeability contrast. However, we observe a slightly larger spread in computed early solute arrivals for the high contrast as compared to the low contrast case (see the rising limb of the curves in Fig. 5). Because in the following we focus on the assessment of the macroscopic response of the system and the departure from a Fickian macrodispersive model, we deem a single realisation to be representative of the response of the system to various combinations of the investigated parameters.

4.3 Effect of permeability contrast

Figure 6 illustrates the effect of the permeability contrast between facies on transport, by comparing the results of transport simulations performed on two geological domains with identical arrangement but assuming low and high permeability contrast. Geostatistical, flow and transport parameters associated with the results in Fig. 6 are shown in Table 2.

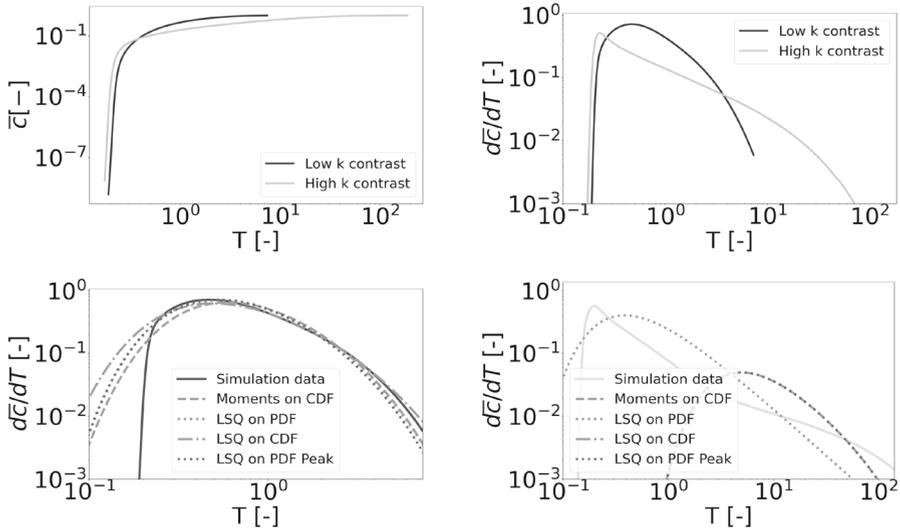


Fig. 6 Top panels: breakthrough curves (left) and their time derivatives (right) simulated on identical geological structure with low and high permeability contrast between the facies. Bottom panels: the curves are overlapped with the corresponding Inverse Gaussian approximations via least square (LSQ) or moments' method estimation in low (left) and high (right) permeability contrast case

The simulated BTCs (i.e., CDFs of the solute arrival times) are shown in the top left panel of Fig. 6 while on top right panel of Fig. 6 the corresponding time derivative are shown, these latter corresponding to the PDFs of arrival times. The dotted and dashed lines in the bottom panels of Fig. 6 are computed applying the moment's method and the least square method illustrated in Sect. 2.5.1 to the results of the numerical experiments. A Fickian model based on the Inverse Gaussian distribution yields a reasonable fitting of the numerical data for the low contrast simulation, where the permeability contrast remains within one order of magnitude (see Fig. 6, bottom left and the first error column in Table 3). In this case the match between the numerical simulation and the Inverse Gaussian distribution is satisfactory especially for the peak and the right tail of the distribution, representing late arrivals. While Eq. (14) represents the analytical solution for the ADE in an infinite domain, in our case the optimal analytical solution should consider the semi-infinite boundary condition adopted for the concentration. The analytical expression can be found in Van Genuchten (1982). In the low contrast case the results obtained with different estimation methods are self-consistent, i.e. least squares and moments methods yield similar outcomes. For high permeability contrast, the Inverse Gaussian distribution cannot match the simulated data (Fig. 6, bottom right), regardless of the method used to estimate its parameters (Least Square or Moments method).

In summary, Fig. 6 suggests that as the permeability contrast increases, the evolution of the solute concentration shows significant departure from the Fickian model.

A detailed analysis of these results is shown in Table 3. Our results suggest that the level of accuracy of macrodispersion models in capturing transport behaviour decreases with the permeability contrast. Table 3 reports the values of macrodispersion

Table 2 Geostatistical and flow parameters for the low and high permeability contrast fields used to generate results reported in Figure 6. The flow parameters were defined by equations (11) and (12)

	Low contrast	High contrast
<i>Geostatistical parameters</i>		
λ_x [m]	0.8	0.8
$\log(k_{i+1}/k_i)$ [-]	1	2
$Var(\log(k))$ [$\log m^4$]	6.5	26.1
<i>Flow parameters</i>		
$k_{eff,x}$ [m^2]	2.11×10^{-11}	1.49×10^{-10}
\bar{V}_x [m/s]	1.07×10^{-6}	9.64×10^{-6}
Pe_x [-]	21113	59633

The flow parameters were defined by Eqs. (11) and (12)

Table 3 Simulated (*Reference solution*) and estimated (*Methods 1–4*) values for average Darcy velocity \bar{V}_x , macrodispersion D_{mac} and relative breakthrough error e values

	Low contrast			High contrast		
	\bar{V}_x [m/s]	D_{mac} [m^2/s]	e [%]	\bar{V}_x [m/s]	D_{mac} [m^2/s]	e [%]
<i>Reference solution</i>	1.07×10^{-6}	8.57×10^{-7}	–	9.64×10^{-6}	7.71×10^{-6}	–
<i>Method 1</i>	6.94×10^{-7}	4.97×10^{-7}	16.20	7.03×10^{-7}	1.59×10^{-6}	79.5
<i>Method 2</i>	2.08×10^{-7}	1.81×10^{-7}	10.71	3.79×10^{-5}	8.79×10^{-7}	48.1
<i>Method 3</i>	6.50×10^{-7}	6.09×10^{-7}	11.94	6.36×10^{-7}	2.33×10^{-6}	126.7
<i>Method 4</i>	2.42×10^{-7}	1.13×10^{-7}	11.93	8.95×10^{-8}	3.74×10^{-5}	49.2

The *reference solution* values represent the average longitudinal velocity and the nominal macrodispersion as from Eq. (13). *Method 1* is the moments method, *Method 2*, 3 and 4 correspond to least squares method applied to the simulated PDF, CDF and the PDF peak data. As a result of the unsuitability of the Inverse Gaussian model to the describe the PDF in high permeability contrast scenarios, some values (*italic*) are characterised by extremely large standard deviations ($\sigma > 10^6$)

parameters computed through approximation (13) (considered as a reference value) and compare them with the the estimated ones. Transport parameters estimations are closer to the reference values for the low contrast if compared to the high contrast cases. Moreover the estimates obtained through least squares in the high contrast case are generally affected by large confidence bounds (i.e., they are indicated in italic in Tabel 7) thus the estimated values cannot be considered as reliable.

4.4 Effect of spatial correlation

Transport simulations are performed on PGS domains sharing comparable geostatistical parameters (Table 4) while increasing longitudinal correlation lengths (Fig. 7). These provide interesting insights into the transition from Fickian to anomalous transport in relation to the connectivity degree of the sediment structure (Fig. 8). We emphasise here that the correlation length mentioned here is the one employed to generate the continuous Gaussian random fields which are then employed to generate

Table 4 Geostatistical and flow parameters for low permeability contrast simulations. The flow parameters are defined by equations (11) and (12)

	$\lambda_x = 0.4$	$\lambda_x = 0.6$	$\lambda_x = 0.8$	$\lambda_x = 1.0$
<i>Geostatistical parameters</i>				
$\log(k_i/k_{i+1})$	1	1	1	1
$Var(\log(k))$ [$\log m^4$]	6.7	6.6	6.6	6.6
<i>Flow parameters</i>				
$k_{eff,x}$ [m^2]	1.75×10^{-11}	1.99×10^{-11}	2.14×10^{-11}	2.27×10^{-11}
Pe_x [-]	3508	5972	8571	11355

The flow parameters are defined by Eqs. (11) and (12)

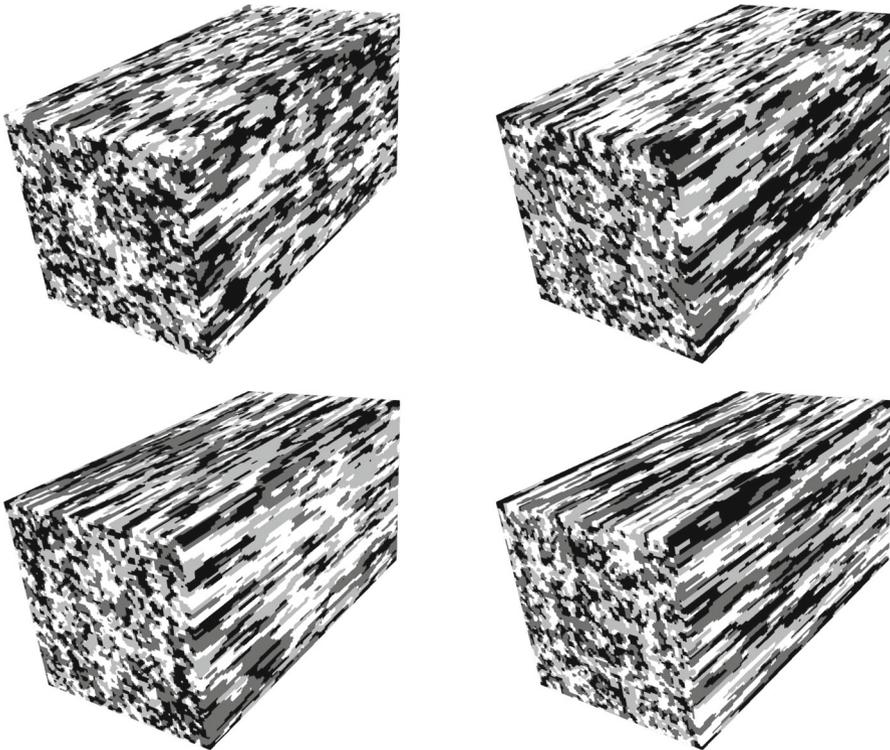


Fig. 7 Truncated pluri-Gaussian permeability fields with increasing longitudinal correlation length λ_x . Clockwise order from top left panel: $\lambda_x = 0.4$, $\lambda_x = 0.6$, $\lambda_x = 0.8$, $\lambda_x = 1.0$

the conductivity fields (see Fig. 1). This length can be interpreted as the characteristic length over which facies’ transitions are observed.

Figure 8 displays the PDFs of the solute arrival times obtained for a range of values assigned to λ_x . As the longitudinal correlation length increases, the magnitude of the peak value increases and the peak shifts towards earlier arrival times. This can be explained observing that, with increasing correlation in the longitudinal direction, the

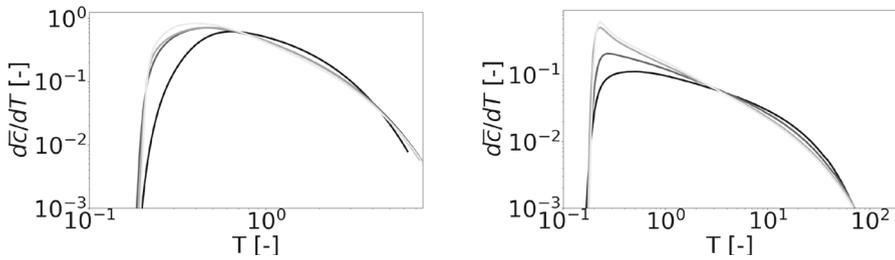


Fig. 8 Arrival time PDFs computed as $d\bar{c}/dT$ for low (left panel) and high (right panel) permeability contrast for as a function of the longitudinal correlation length λ_x . From the darkest to the lightest curve the longitudinal correlation length $\lambda_x = 1$ increases with evenly spaced interval from 0.4 to 1

Table 5 Geostatistical and flow parameters^a for high permeability contrast simulations

	$\lambda_x = 0.4$	$\lambda_x = 0.6$	$\lambda_x = 0.8$	$\lambda_x = 1.0$
<i>Geostatistical parameters</i>				
$\log(k_i/k_{i+1})$	2	2	2	2
$Var(\log(k))$ [$\log m^4$]	26.1	26.2	25.9	26.2
<i>Flow parameters</i>				
$k_{eff,x}$ [m^2]	1.50×10^{-10}	1.76×10^{-10}	1.93×10^{-10}	2.05×10^{-10}
Pe_x [-]	29953	52919	77127	102719

The flow parameters were defined by Eqs. (11) and (12) and are computed in the longitudinal direction of the flow. Longitudinal correlation length increases from 0.4 to 1.0

connectivity between highly permeable facies favours the formation of fast channels where advection prevails over diffusion thus leading to early arrivals. This effect is more evident for the high contrast scenario (right side of Fig. 8) and is reflected by the solute arrivals PDFs trends: as the domain connectivity increases, the initial concentration peak rises while the central segment of the curve highlights a power law response.

Tables 4 and 5 report relevant geostatistical and flow simulation parameters which are required to interpret the results of the average Darcy velocity and macrodispersion estimation process provided in Tables 6 and 7. As a result of the emergence of preferential flow-paths, the effective permeability shows a positive trend for increasing correlation lengths.

Comparing the relative error computed for low and high permeability contrast cases, it is clear that the reliability of the Fickian model at the macroscale decreases with the permeability contrast (see Fig. 9). We also observe a mild increasing trend of the relative error with increasing values of the correlation length. This trend is justified by the role of the preferential flow-paths which facilitate fast advective flow that make overall solute behaviour anomalous.

Summarising the results of our numerical experiments, we observe that both correlation length and permeability contrast are triggering factors for non-Fickian transport behaviour. Increasing the correlation length by a factor 2 induces a 15-20 % increase

Table 6 Simulated (*Reference solution*) and estimated (*Methods 1–4*) values for average Darcy velocity \bar{V}_x , macrodispersion D_{mac} and relative breakthrough error e values in low permeability contrast simulations

Low contrast												
$\lambda_x = 0.4$												
	\bar{V}_x [m/s]	D [m ² /s]	e [%]	$\lambda_x = 0.6$			$\lambda_x = 0.8$			$\lambda_x = 1.0$		
				\bar{V}_x [m/s]	D [m ² /s]	e [%]	\bar{V}_x [m/s]	D [m ² /s]	e [%]	\bar{V}_x [m/s]	D [m ² /s]	e [%]
<i>Ref. s.</i>	8.77×10^{-7}	3.51×10^{-7}	–	9.95×10^{-7}	5.97×10^{-7}	–	1.07×10^{-6}	8.56×10^{-7}	–	1.14×10^{-6}	1.14×10^{-6}	–
<i>M. 1</i>	5.39×10^{-7}	2.68×10^{-7}	12.9	6.26×10^{-7}	4.54×10^{-7}	16.3	6.94×10^{-7}	4.97×10^{-7}	16.2	7.60×10^{-7}	6.15×10^{-7}	19.1
<i>M. 2</i>	1.30×10^{-6}	6.18×10^{-8}	7.3	1.91×10^{-6}	1.68×10^{-7}	11.0	2.08×10^{-6}	1.81×10^{-7}	10.7	2.48×10^{-6}	2.32×10^{-7}	16.7
<i>M. 3</i>	5.12×10^{-7}	3.17×10^{-7}	28.2	5.87×10^{-7}	5.55×10^{-7}	11.8	6.50×10^{-7}	6.09×10^{-7}	11.9	7.10×10^{-7}	7.60×10^{-7}	15.8
<i>M. 4</i>	2.03×10^{-6}	1.31×10^{-7}	23.7	2.32×10^{-6}	1.27×10^{-7}	13.0	2.41×10^{-6}	1.13×10^{-7}	11.9	2.68×10^{-6}	1.17×10^{-7}	16.4

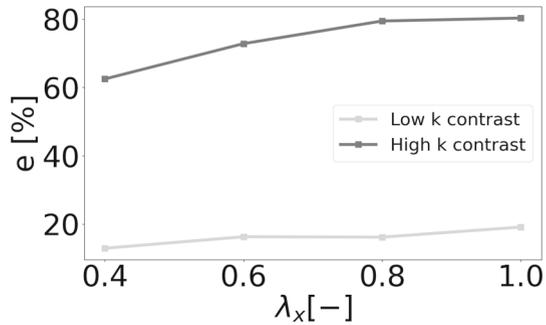
Method 1 is the moments method, *Method 2*, *3* and *4* correspond to least squares method applied to the simulated PDF, CDF and the PDF peak data

Table 7 Simulated (*Reference solution*) and estimated (*Methods 1–4*) values for average Darcy velocity \bar{V}_x , macrodispersion D_{mac} and relative error e values obtained for high permeability contrast simulations

High contrast												
$\lambda_x = 0.4$												
$\lambda_x = 0.6$				$\lambda_x = 0.8$				$\lambda_x = 1.0$				
	\bar{V}_x [m/s]	D [m ² /s]	e [%]	\bar{V}_x [m/s]	D [m ² /s]	e [%]	\bar{V}_x [m/s]	D [m ² /s]	e [%]	\bar{V}_x [m/s]	D [m ² /s]	e [%]
<i>Ref. s.</i>	7.49×10^{-6}	3.00×10^{-6}	–	8.82×10^{-6}	5.29×10^{-6}	–	9.65×10^{-6}	7.72×10^{-6}	–	1.03×10^{-6}	1.03×10^{-6}	–
<i>M. 1</i>	4.89×10^{-7}	6.27×10^{-7}	62.5	6.02×10^{-7}	1.05×10^{-6}	72.8	7.03×10^{-7}	1.60×10^{-6}	79.5	8.08×10^{-7}	1.91×10^{-6}	80.3
<i>M. 2</i>	1.52×10^{-5}	9.08×10^{-7}	44.2	2.56×10^{-5}	1.58×10^{-6}	54.01	3.79×10^{-5}	8.79×10^{-7}	48.1	4.02×10^{-5}	8.32×10^{-7}	42.9
<i>M. 3</i>	4.51×10^{-7}	8.57×10^{-7}	87.6	5.52×10^{-7}	1.46×10^{-6}	108.8	6.36×10^{-7}	2.34×10^{-6}	126.7	7.24×10^{-7}	2.94×10^{-6}	130.1
<i>M. 4</i>	4.89×10^{-7}	6.27×10^{-7}	70.4	6.02×10^{-7}	1.05×10^{-6}	59.2	8.95×10^{-8}	3.74×10^{-5}	49.2	4.22×10^{-5}	3.53×10^{-7}	43.1

Method 1 is the moments method, *Methods 2, 3 and 4* correspond to least squares method applied to the simulated PDF, CDF and the PDF peak data

Fig. 9 Longitudinal correlation length vs relative error computed as the departure from Fickian model (Eq. 22) with parameters estimated with the moments method (Method 1)



in the observed error. Thus, the correlation length appears to have a milder effect if compared with the permeability contrast.

4.5 Effect of Péclet

We analyse here the effect of the Péclet number on our results. This means that not only the effect of different permeability fields (as in Sect. 4.3) or correlation lengths (Sect. 4.4) were tested, but also the effect of the compound-specific diffusion coefficient. Our tests were conducted by decreasing the molecular diffusion coefficient of one magnitude order at each simulation that correspond to an increase of the Péclet (Pe) of one magnitude order at each simulation. The average Darcy velocity and the correlation length λ_x are kept constant. The variations in Pe have a marked influence on the right tail of the arrival times distributions, as shown in Fig. 10. This result is in agreement with Zhang et al. (2015) and can be explained observing that diffusion effects become apparent in late arrivals, while early arrivals are conversely driven by advection-dominated processes. Left and right panels in Fig. 10 show the combined effect of increasing Péclet (darker to lighter curves) in a low (left) and high (right) permeability contrast domain. We observe that both trends display a linear trend in Fig. 10, indicating a proportionality $e \sim a \log(Pe)$, where the constant a depends on the assumed permeability contrast. Thus, the analysis yields comparable conclusions with the ones obtained in Sect. 4.4: while the interplay between increasing Péclet numbers and the departure from Fickian behaviour is clear, the increase of permeability contrast still appears to play a predominant role.

Figure 11 shows that the increase in permeability contrast by one order of magnitude exhibits a stronger control on transport behaviour than the increase in Péclet number by one magnitude order as for comparable Pe , the error associated with low permeability contrast simulations (light curve) is always lower than the error associated with the high permeability contrast (dark curve). It is interesting to note that the permeability contrast appears to control the rate at which the Fickian model error decreases for decreasing Pe . The error is here represented by taking the approximation resulting from the moments' method (M.1 in Tables 6 and 7).

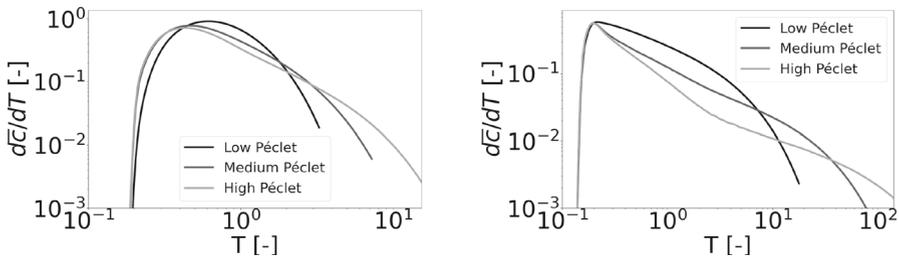
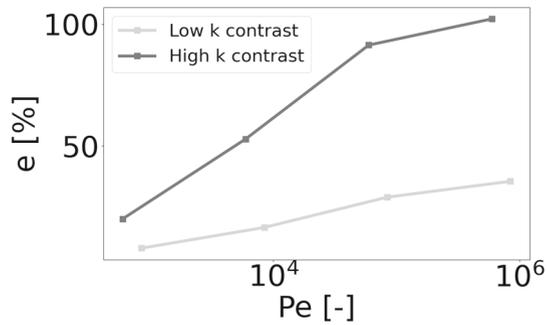


Fig. 10 Arrival times PDFs in low (left) and high (right) permeability contrast domains sharing the same geological structure ($\lambda_x = 0.8$ m) while characterised by a different Pécellet. In case of low permeability contrast, the Pécellet number ranges between 8×10^2 and 8×10^4 while for high permeability contrast the Pécellet ranges between 6×10^3 and 6×10^5 . Significant Pécellet variation for these simulations was obtained by changing the molecular diffusion coefficient

Fig. 11 Pécellet number vs relative error computed as in Eq. (22). The interplay between increasing Pécellet numbers, permeability contrast and relative error is qualitatively similar to the one exhibited by the increasing longitudinal correlation length in Fig. 8



5 Conclusions

In this work we have explored the impact of permeability contrast and, similarly to Zhang et al. (2014, 2015) and Yin et al. (2020), correlation length and Pécellet number on the emergence of non-Fickian transport in random discontinuous permeability fields.

The following conclusions can be drawn from the interpretation of our results:

- our results combine error and uncertainty quantification metrics to assess departure from a Fickian transport regime in PGS fields. Emergence of non Fickian transport is quantified upon relying on relative error with respect to the prediction of a macrodispersive solution, where this latter can be obtained with diverse estimation strategies. Large relative errors and large confidence intervals for estimated parameters are indicative of the unsuitability of the Inverse Gaussian distribution in interpreting the outcomes of high-resolution numerical simulations, thus indicating non-Fickian response;
- a Fickian macrodispersive model can match with reasonable accuracy solute arrival times in ergodic domains featuring conductivity values distributed over up to four orders of magnitudes. In such conditions the Fickian model underestimates early arrival times, but can capture with good accuracy the peak and late arrivals. Overall observed errors are in the order of 10-20 %. Lowest errors are obtained when the characteristic size associated with the medium heterogeneity is much smaller than the distance travelled by the solute;

- for ergodic domains, a hierarchy of non-Fickian triggering factors can be established: permeability contrast plays a primary role in determining the fate of the solute, while correlation length and Péclet number can be both considered secondary non-Fickian transport triggering factors;
- fluid velocity PDFs support the prevalence of permeability contrast over correlation length in triggering non-Fickian transport. The velocity distribution is strongly modified by permeability contrast, displaying a much larger spread in high contrast media if compared with low contrast ones. Conversely, increasing the correlation length only slightly affects the shape of the flow velocity PDFs. Interesting insights are also gained upon considering velocity distributions separately by facies. In both high and low contrast media flow velocity values in low permeability regions are homogeneously distributed around the corresponding peak values. Conversely, flow velocities in permeable facies display increasing skewness with increasing permeability contrast indicating the occurrence of low velocity regions in highly permeable media. This element could be further exploited in the context of macroscopic non-Fickian parameterisation, which will be considered in future works;
- the BTC variability observed between multiple realisations of the same geological setting is more evident at early times while it tends to disappear at late times;
- the relative importance of diffusion and advective processes, captured by Pe , plays an important role in the solute transport response. Yet, the influence of Pe on the accuracy of a macrodispersive model is markedly influenced by the assumed permeability contrast. Our results suggest a logarithmic trend $e \sim a \log(Pe)$ where the constant a is proportional to the assumed permeability contrast;
- while for Fickian or moderately non-Fickian transport the different parameter estimation methods (method of moments or least-squares-based methods) are equivalent, when a macrodispersion approximation is sought for significantly non-Fickian curves, the choice of the fitting method is crucial as it can lead to very different effective parameters and fitted curves. Although this is expected, due to the lack of validity of the underlying model, it has important consequences for practitioners that are nevertheless forced to use and fit macrodispersion effective parameters. Here, the method of moments is built to preserve accurately the statistics but it could predict poorly the early arrival peak as well as the long tails.

Future works will include the extension to more realistic injection scenarios, variable density and hydrodynamic dispersion models, investigating the effect of different lithotype rules, as well as interpreting the non-Fickian transport results with more complex anomalous transport models including spatial Markov processes (Sherman et al. 2021) and Generalised Multi-Rate Transfer equations (Municchi and Icardi 2020).

Acknowledgements We would like to especially thank Prof. John Billingham for his insightful suggestions and for his detailed proofreading of this paper.

Funding This work was supported by the Royal Society through the grant No. IES\R3\170302.

Declaration

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adams, E.E., Gelhar, L.W.: Field study of dispersion in a heterogeneous aquifer: 2. Spatial moments analysis. *Water Resour. Res.* **28**(12), 3293–3307 (1992)
- Armstrong, M., Galli, A., Beucher, H., Loc'h, G., Renard, D., Doligez, B., Eschard, R., Geffroy, F.: *Pluri-gaussian Simulations in Geosciences*. Springer, Berlin (2011)
- Barlebo, H.C., Hill, M.C., Rosbjerg, D.: Investigating the macrodispersion experiment (made) site in columbus, mississippi, using a three-dimensional inverse flow and transport model. *Water Resour. Res.* **40**(4) (2004). <https://doi.org/10.1029/2002WR001935>
- Bear, J.: *Hydraulics of Groundwater*. Courier Corporation, Chelmsford (2012)
- Berkowitz, B., Cortis, A., Dentz, M., Scher, H.: Modeling non-Fickian transport in geological formations as a continuous time random walk. *Rev. Geophys.* **44**(2) (2006). <https://doi.org/10.1029/2005RG000178>
- Bianchi, M., Zheng, C.: A lithofacies approach for modeling non-Fickian solute transport in a heterogeneous alluvial aquifer. *Water Resour. Res.* **52**(1), 552–565 (2016)
- Bianchi, M., Pedretti, D.: Geological entropy and solute transport in heterogeneous porous media. *Water Resour. Res.* **53**(6), 4691–4708 (2017)
- Carle, S.F.: *T-Progs: Transition Probability Geostatistical Software*, vol. 84. University of California, Davis (1999)
- Carle, S.F., Fogg, G.E.: Transition probability-based indicator geostatistics. *Math. Geol.* **28**(4), 453–476 (1996)
- Carle, S.F., Fogg, G.E.: Modeling spatial variability with one and multidimensional continuous-lag markov chains. *Math. Geol.* **29**(7), 891–918 (1997)
- Comolli, A., Hakoun, V., Dentz, M.: Mechanisms, upscaling, and prediction of anomalous dispersion in heterogeneous porous media. *Water Resour. Res.* **55**(10), 8197–8222 (2019). <https://doi.org/10.1029/2019WR024919>
- Dagan, G.: *Flow and Transport in Porous Formations*. Springer, Berlin (2012)
- de Barros, F.P.J., Guadagnini, A., Riva, M.: Features of transport in non-Gaussian random porous systems. *Int. J. Heat Mass Transf.* **184**, 122244 (2022). <https://doi.org/10.1016/j.ijheatmasstransfer.2021.122244>
- 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)
- Deutsch, C.V., Journel, A.G.: *Gslib. Geostatistical Software Library and User's Guide*, vol. 369. Oxford University Press, Oxford (1998)
- Dimitrakopoulos, R., Luo, X.: Generalized sequential Gaussian simulation on group size ν and screen-effect approximations for large field simulations. *Math. Geol.* **36**(5), 567–591 (2004)
- Ederly, Y., Guadagnini, A., Scher, H., Berkowitz, B.: Origins of anomalous transport in heterogeneous media: structural and dynamic controls. *Water Resour. Res.* **50**(2), 1490–1505 (2014)
- Ederly, Y., Porta, G.M., Guadagnini, A., Scher, H., Berkowitz, B.: Characterization of bimolecular reactive transport in heterogeneous porous media. *Transp. Porous Media* **115**(2), 291–310 (2016). <https://doi.org/10.1007/s11242-016-0684-0>
- Ederly, Y., Stolar, M., Porta, G., Guadagnini, A.: Feedback mechanisms between precipitation and dissolution reactions across randomly heterogeneous conductivity fields. *Hydrol. Earth Syst. Sci.* **25**(11), 5905–5915 (2021). <https://doi.org/10.5194/hess-25-5905-2021>
- Fiori, A., Cvetkovic, V., Dagan, G., Attinger, S., Bellin, A., Dietrich, P., Zech, A., Teutsch, G.: Debates–stochastic subsurface hydrology from theory to practice: the relevance of stochastic subsurface

- hydrology to practical problems of contaminant transport and remediation. What is characterization and stochastic theory good for? *Water Resour. Res.* **52**(12), 9228–9234 (2016)
- Flemisch, B., Darcis, M., Erbertseder, K., Faigle, B., Lauser, A., Mosthaf, K., Müthing, S., Nuske, P., Tatomir, A., Wolff, M.: Dumux: Dune for multi-phase, component, scale, physics,... flow and transport in porous media. *Adv. Water Resour.* **34**(9), 1102–1112 (2011)
- Gelhar, L.W., Axness, C.L.: Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resour. Res.* **19**(1), 161–180 (1983)
- Gotovac, H., Cvetkovic, V., Andricevic, R.: Flow and travel time statistics in highly heterogeneous porous media. *Water Resour. Res.* **45**(7) (2009) <https://doi.org/10.1029/2008WR007168>
- Hansen, S.K., Haslauer, C.P., Cirpka, O.A., Vesselinov, V.V.: Direct breakthrough curve prediction from statistics of heterogeneous conductivity fields. *Water Resour. Res.* **54**(1), 271–285 (2018). <https://doi.org/10.1002/2017WR020450>
- Heße, F., Prykhodko, V., Schlüter, S., Attinger, S.: Generating random fields with a truncated power-law variogram: a comparison of several numerical methods. *Environ. Model. Softw.* **55**, 32–48 (2014)
- Horgue, P., Soullaine, C., Franc, J., Guibert, R., Debenest, G.: An open-source toolbox for multiphase flow in porous media. *Comput. Phys. Commun.* **187**, 217–226 (2015)
- 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)
- Kolditz, O., Bauer, S., Bilke, L., Böttcher, N., Delfs, J.-O., Fischer, T., Görke, U.J., Kalbacher, T., Kosakowski, G., McDermott, C.: Opengeosys: an open-source initiative for numerical simulation of thermo-hydro-mechanical/chemical (thm/c) processes in porous media. *Environ. Earth Sci.* **67**(2), 589–599 (2012)
- Koltermann, C.E., Gorelick, S.M.: Heterogeneity in sedimentary deposits: a review of structure-imitating, process-imitating, and descriptive approaches. *Water Resour. Res.* **32**(9), 2617–2658 (1996)
- Kreft, A., Zuber, A.: On the physical meaning of the dispersion equation and its solutions for different initial and boundary conditions. *Chem. Eng. Sci.* **33**(11), 1471–1480 (1978)
- Linde, N., Renard, P., Mukerji, T., Caers, J.: Geological realism in hydrogeological and geophysical inverse modeling: a review. *Adv. Water Resour.* **86**, 86–101 (2015)
- Mandelbrot, B.B., Van Ness, J.W.: Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10**(4), 422–437 (1968)
- Mariethoz, G., Renard, P., Cornaton, F., Jaquet, O.: Truncated plurigaussian simulations to characterize aquifer heterogeneity. *Groundwater* **47**(1), 13–24 (2009)
- Municchi, F., Icardi, M.: Generalized multirate models for conjugate transfer in heterogeneous materials. *Phys. Rev. Res.* **2**(1), 013041 (2020)
- Municchi, F., Pescimoro, E., Hidalgo, J., Icardi, M.: SECUREFOAM V0.1 (2022). <https://doi.org/10.5281/zenodo.6958098>
- Neuman, S.P., Tartakovsky, D.M.: Perspective on theories of non-Fickian transport in heterogeneous media. *Adv. Water Resour.* **32**(5), 670–680 (2009)
- Riva, M., Guadagnini, A., Fernandez-Garcia, D., Sanchez-Vila, X., Ptak, T.: Relative importance of geostatistical and transport models in describing heavily tailed breakthrough curves at the lauswiesen site. *J. Contam. Hydrol.* **101**(1–4), 1–13 (2008). <https://doi.org/10.1016/j.jconhyd.2008.07.004>
- Savoy, H., Kalbacher, T., Dietrich, P., Rubin, Y.: Geological heterogeneity: goal-oriented simplification of structure and characterization needs. *Adv. Water Resour.* **109**, 1–13 (2017)
- Sherman, T., Engdahl, N.B., Porta, G., Bolster, D.: A review of spatial markov models for predicting pre-asymptotic and anomalous transport in porous and fractured media. *J. Contam. Hydrol.* **236**, 103734 (2021)
- 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**, 103812 (2021). <https://doi.org/10.1016/j.advwatres.2020.103812>
- Strebelle, S.: Conditional simulation of complex geological structures using multiple-point statistics. *Math. Geol.* **34**(1), 1–21 (2002)
- Tartakovsky, D.M., Dentz, M.: Diffusion in porous media: phenomena and mechanisms. *Transp. Porous Media* **130**(1), 105–127 (2019)
- Van Genuchten, M.T.: Analytical Solutions of the One-Dimensional Convective-Dispersive Solute Transport Equation, vol. 1661. US Department of Agriculture, Agricultural Research Service, Beltsville (1982)

- Weissmann, G.S., Carle, S.F., Fogg, G.E.: Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics. *Water Resour. Res.* **35**(6), 1761–1770 (1999)
- Weller, H.G., Tabor, G., Jasak, H., Fureby, C.: A tensorial approach to computational continuum mechanics using object-oriented techniques. *Comput. Phys.* **12**(6), 620–631 (1998)
- Yin, M., Zhang, Y., Ma, R., Tick, G.R., Bianchi, M., Zheng, C., Wei, W., Wei, S., Liu, X.: Super-diffusion affected by hydrofacies mean length and source geometry in alluvial settings. *J. Hydrol.* **582**, 124515 (2020)
- Yu, C., Warrick, A., Conklin, M.: A moment method for analyzing breakthrough curves of step inputs. *Water Resour. Res.* **35**(11), 3567–3572 (1999)
- Zech, A., Attinger, S., Bellin, A., Cvetkovic, V., Dagan, G., Dentz, M., Dietrich, P., Fiori, A., Teutsch, G.: A comparison of six transport models of the made-1 experiment implemented with different types of hydraulic data. *Water Resour. Res.* **57**(5), 2020–028672 (2021). <https://doi.org/10.1029/2020WR028672>
- Zhang, Y., Green, C.T., Fogg, G.E.: The impact of medium architecture of alluvial settings on non-Fickian transport. *Adv. Water Resour.* **54**, 78–99 (2013)
- Zhang, Y., Green, C.T., Baeumer, B.: Linking aquifer spatial properties and non-Fickian transport in mobile-immobile like alluvial settings. *J. Hydrol.* **512**, 315–331 (2014)
- Zhang, Y., Green, C.T., Tick, G.R.: Peclet number as affected by molecular diffusion controls transient anomalous transport in alluvial aquifer-aquitard complexes. *J. Contam. Hydrol.* **177**, 220–238 (2015)
- Zinn, B., Harvey, C.F.: When good statistical models of aquifer heterogeneity go bad: a comparison of flow, dispersion, and mass transfer in connected and multivariate gaussian hydraulic conductivity fields. *Water Resour. Res.* **39**(3) (2003). <https://doi.org/10.1029/2001WR001146>

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