Flow and transport modelling in highly heterogeneous geological porous media

Eugenio Pescimoro

June 27, 2023

Acknowledgements

I would like to thank my supervisors Prof. Matteo Icardi, Prof. John Billingham and Dr. Marco Bianchi for the substantial academic support and the several meetings, reviews and feedback that I received during the last three and a half years of research and without whom this work would not have been possible. Dr. Federico Municchi and Prof. Giovanni Porta also deserve a special mention, the first for helping in countless technical issues and sharing culinary experiences, the second for the time invested in draft reviews and improvements that eventually allowed part of this thesis to be published. Among the academic recognitions, I would like to include Ephesia Consult and its employees for an insightful experience into the geostatistic world.

A special thanks to my girlfriend and family for the moral and logistic support that made my PhD life much more enjoyable and to my friends and colleagues with whom I shared the pre-pandemic everyday life on and off campus.

Abstract

Flow and transport processes through porous media are ubiquitous both in natural and industrial environments. Ranging from diffusion in human tissue to oil recovery and CO_2 storage, including the design of porous reactors, geothermal energy production, groundwater remediation, oil recovery and CO_2 storage, the characterisation of fluid flow and solute transport at different scales represents the paradigm to better understand the mechanisms at the base of several processes. Due to the broad spectrum of applications, a vast empirical and numerical research field developed around transport in heterogeneous porous media. While on the numerical side, the mathematical models available for simulating transport at the micro and meso-scales have shown good agreement with the empirical tests, the debate around modelling transport at the macro-scale is still open. One example is the unknown relation between system parameters and their values measured at different scales which is usually addressed as scale effect. Other examples are anomalous or non-Fickian transport phenomena and the validity range of macro-scale transport models. Our study focuses on the impact of the heterogeneous distribution of the subsurface properties on the transport of solute at the macro-scale. Initially we propose an analysis of transport in heterogeneous porous media generated with a random geostatistical algorithm. Subsequently this subject is expanded and applied to a real domain which was surveyed and reconstructed with a high level of resolution. Three-dimensional meso-scale numerical simulations performed with our open-source C++ library, built on top of the finite-volume library OpenFOAM, represent the main source of data to test macro-scale mathematical models.

Contents

List of symbols and abbreviations

- ADE: Advection Dispersion Equations;
- PGS: Pluri-Gaussian Simulations;
- GRF: Gaussian Random Field;
- BTC: Breakthrough Curve;
- PDF: Probability Density Function;
- CDF: Cumulative Density Function;
- σ^2: variance;
- E: Expected value;
- λ: correlation length;
- q: flow rate;
- D: hydrodynamic dispersion;
- D_{mol}: molecular diffusion;
- D_{mec}: mechanical dispersion;
- D_{mac}: macrodispersion;
- k: permeability;
- K: hydraulic conductivity;
- ϕ: porosity;

- v: velocity of the fluid;
- V: Darcy's velocity;
- μ: dynamic viscosity;
- Pe: Péclet number;
- f: PDF of the inverse Gaussian distribution (also known as Wald distribution);
- F: CDF of the inverse Gaussian distribution (also known as Wald distribution);
- ρ: fluid density;

Chapter 1

Introduction

The broad applicability of flow and transport equations is one of the most intriguing and fascinating aspect of an applied mathematic approach to the study of hydrogeology. Indeed, depending on the scale adopted, an extremely broad set of materials can be regarded as porous media, from battery cells to flesh and subsoil. Following a similar parallelism, the equivalence between strongly different types of fluid can be established and the advection-dispersion model reveals its congruity for simulating heat as well as water or oil transport dynamics. Although the advection-dispersion model has been validated in several contexts, the range of represented scales, especially for highly heterogeneous environments like subsurface, is limited. When solute is transported through heterogeneous porous media, it might experience flowing through regions with contrasting geological properties. As a result, depending on sediment geometry and values, the solute concentration in space and time might diverge from the advection-dispersion model description. When the discrepancy between observed transport and model prediction becomes relevant the transport is called anomalous or, in contrast to the Fick law on which the advection-dispersion model is based, non-Fickian.

1.1 Motivation

Transport of solute in heterogeneous porous media is modelled with good accuracy at the micro (or porous) and meso (or heterogeneity) scale while a long lasting debate still revolves around the models for simulating solute transport at macro (or field) scale 1.1. In this context, the meaning of *heterogeneity* is always associated with discontinuous sediment pattern that happens at the meter scale. While the transport modelling at the micro and meso scales

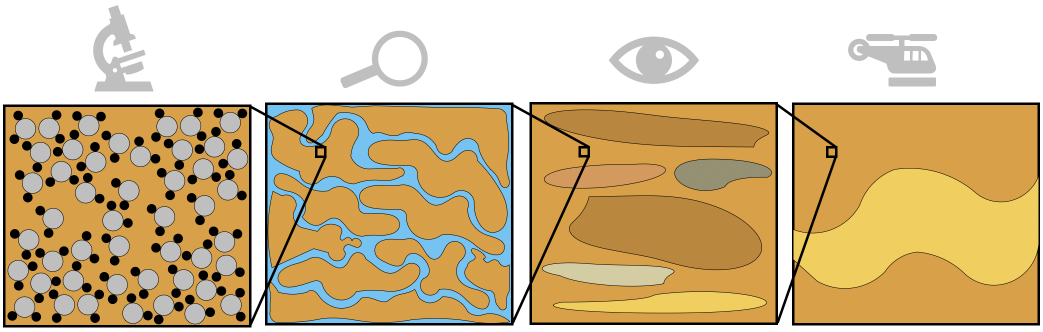

Figure 1.1: The four scales which characterise subsurface formations. From left to right: the nano or continuum scale (10^{-9} m), the micro or porous scale (10^{-4} m), the meso or heterogeneity scale ($10^{-1} - 10^{1}$ m) and the macro or field scale (10^{4} m).

through Navier-Stokes and the Advection-Dispersion equations (ADE) respectively has been widely tested and accepted, the lack of consensus around the macro scale models yielded several approaches that can essentially be subdivided into two main categories: local and non-local transport models. One of the most widely adopted local transport model for the macro scale is the macrodispersion model [1] while some non-local transport models include the multi-rate mass transfer model [2], the continuous time random walk [3] and the moment equation approach [4], just to mention a few. A more comprehensive list can be found in [5]. In this context, the meaning of local and non-local adjectives should not be limited to the spatial domain while it should be extended to the temporal domain as well, thus becoming synonyms of Markovian and non-Markovian models. In other words, local or Markovian models at a given time depend only on the system state at the previous time while the full spatial and temporal history of the system is needed to characterise the state of a non-local or non-Markovian model at any time. Although profoundly different, the applicability of local and non-local models is still quite limited. The unpopularity of local models descends from the strict hypothesis under which they are formulated (e.g. ergodic domain and stationary process) while the non-local models are often difficult to calibrate.

Focusing on the macrodispersion model, in our work we rely on numerical ADE simulations at the meso scale (third panel from left in figure 1.1) to assess the impact of distinct geostatistical parameters on the transport at the macro scale (fourth panel from left in figure 1.1). By quantifying the error that affects the macrodispersion simulations, our study firstly set clear geological guidelines around the validity of the macrodispersion model. Specifically, the

present work aims to

- provide a comprehensive overview of some popular geostatistical, flow and transport models within the geostatistical and hydrogeological context;
- understand the influence of geological and geostatistical parameters on transport and classify their relative importance;
- put the conclusions obtained for synthetic cases to the test by running simulations on a quarry which was excavated, surveyed and whose geological properties were directly tested;
- systematically explore the effect that different types of categorical fields have on transport;
- provide some transport simulation study examples which uncover the potentiality of hydrogeologic studies performed with the openly available `OpenFOAM`®library [6].

The main research question can be formulated as *what is the hierarchy of anomalous transport triggering factors in porous domains characterised by sharply changing geological properties?* Using the present framework as guideline, other open hydrogeological questions can be effectively tackled. Some examples are *how does the lithotype rule affect the transport? what is the role of variable density fluids in triggering non-Fickian transport?*

1.2 Literature review

The literature review in this work focuses on two main areas: geostatistics and flow and transport in porous media.

1.2.1 Geostatistics

Given that the mathematical modelling of subsurface flow and transport heavily depend on the properties assigned to the virtual geological domain and that the spatial pattern followed by the hydraulic properties may vary greatly across a wide range of spatial and temporal scales, the virtual representation of geological formations is of capital importance for solute transport predictions and extremely challenging to be simulated. The widespread response to the complexity and the inaccessibility of the subsurface, is a statistical approach to

geology, namely geostatistics [7]. The complexity of reproducing the subsurface (figure 1.2) has been tackled using different approaches that Koltermann and Gorelick's review [8] classifies into three main groups: structure-imitating, process-imitating and descriptive. Although this classification can be further

Figure 1.2: Sediment outcrops or tunnels provide insights into the otherwise inaccessible subsurface structure. From left to right: the sandstone cliffs on the coast of Slovenia, volcanic rock in the Alcantara valley and the limestone wall of a tunnel through Pasubio mountain.

refined into several sub-groups with multiple algorithms each, the focus of this study is on a very specific class of geostatistical simulations called truncated Pluri-Gaussian simulations (PGS). Amongst the spatial statistical methods, this one is capable of reproducing high contrast values between hydrogeological parameters and thus the one characterised by higher geological realism [9]. Spatial stochastic methods aim to reproduce the range of possible subsurface spatial patterns, by generating several equally probable realisations of geological facies[1]. These multiple images allow us to assess the uncertainty related to flow and transport models. An additional distinction can be made: spatial statistical methods can be Gaussian based (e.g. PGS) or non-Gaussian. The underlying distributions of the random variable in the Gaussian approach is continuous albeit it is suitable to be further divided by thresholds to reproduce categorical variables patterns. Non-Gaussian methods span a wide range of approaches which include (but are not limited to): transition probabilities, indicator variograms, Boolean methods and simulated annealing. Particularly appealing when fulfilling soft data of non-stationary fields is a main requirements, these algorithms go beyond the scope of the present study.

PGS can effectively reproduce heterogeneity fields characterised by sharp interfaces that well fit the purpose of imitating discontinuous subsurface properties. These categorical fields are usually associated with values of perme-

[1]Body of rock or sediments with specified chemical, physical and biological characteristics which distinguish it from adjacent rocks or sediments

ability or porosity. This type of fields, despite their name, is non-Gaussian and it was initially introduced in geostatistics to represent ordered deposits (i.e. depositional or stratified) and later extended to model heterogeneities that are not sequentially ordered [10]. Its most remarkable feature is the subdivision of the multivariate spatial domain into categorical values as a result of a thresholding process that marks the transition from smooth random fields characterised by specific correlation functions and geostatistical parameters to categorical non-Gaussian random fields.

1.2.2 Flow and transport in porous media

Meteorology, hydrology and groundwater hydrology (or hydrogeology) form the triplet of sciences which cover the whole water cycle. The inner difficulty of hydrogeology with respect to the other two branches of water science stems from the impossibility of directly observing the subsurface. Direct (e.g. boreholes) or indirect (e.g. seismic waves) survey are difficult, expensive and, to a certain degree, they still embed uncertainty. Mathematical models which describe subsurface flows and transport phenomena hence become essential to formulate predictions of solute spread through porous media such as soil, sand and rock. However, the modeling approach is not free of challenges. Spatial and temporal scales for flows and transport vary hugely on aquifers, the typical domains of application. Aquifers, geologically defined as subsurface domains able to store and transmit water, may cover scales of 10^4 m (or more) while adsorption and chemical process of interest require the analysis at sub-micro scales on the order of 10^{-7} m. Diffusion and dispersion can be analyzed on the whole set of scales in between, from pore scale (10^{-4} m) to sediment heterogeneities ($10^{-1} - 10^{1}$ m). Consequently, the spatial range of interest spans more than 10 orders of magnitude. The temporal scales which are relevant to transport and chemical modeling purposes show great variations as well: the day scale might be relevant to describe the spread of a contaminant plume through a highly permeable aquifer, nitrification-denitrification processes typically can occur on a seasonal scale [11] while depositional processes may take centuries. Intuitively explained the concept of scales, the idea that a phenomenon, which naturally lives on multiple scales, is conveniently described by different laws at different scales should be intuitive. Of course, valid laws should show consistency between different scales. It then becomes a matter of perspective to focus on a particular range of magnitudes and consequently adopt the appropriate conceptualization for the chosen scale interval. The most popular fluid mechanics examples are probably the Navier-Stokes and Darcy equations: both hold on an interval of several spatial scales, defined by

the limits of the respective conceptual model. For examples, Navier-Stokes is not suitable to describe fluid flows at molecular scale (first panel from left in figure 1.1) because the *continuous medium* hypothesis does not hold. At the same time, Darcy's law cannot be applied at pore scale (second panel from left in figure 1.1) because an extensive property such as permeability has no physical meaning at porous scale.

Introducing the concept of "upscaling" now becomes more natural since it refers to the accurate prediction of flow and transport at large scales based on the interpretation of smaller scale parameters. To "go one scale up" it is essential to develop upscaled mathematical models that is, methods based on averaged small scale properties. As a result, the large scale description becomes computationally convenient with respect to the small scale fully-resolved problem [12]. It is therefore computationally relevant to choose the right model for the right scale: a fully-resolved problem, such as the solution of a molecular model at the pore scale ($10^{-3} - 10^{-5}$ m) or the Navier-Stokes at the heterogeneity scale ($10^{-1} - 10^{1}$ m), provide solutions that are consistent between scales but are not computationally convenient.

For the purpose of the present study, nature is observed at three resolution levels: the pore-scale (or micro-scale), the heterogeneity scale (or meso-scale) and the field-scale (or macro-scale). Pore-scale simulations require the virtual representation of the porous matrix where the flow velocity field is provided by the solution of the Navier-Stokes equations [13]. The same physical phenomena at larger scales is well captured by the Darcy equation which is based on macroscopic properties of the porous matrix, e. g. the hydraulic permeability [14]. This is because macroscopic properties are inherently related to microscopic characteristics of the medium, like the case of the hydraulic permeability that arises from the presence of large surface area where no-slip conditions generate linear damping in the momentum equation [15]. However, this in not the only relevant upscaling problem. The complex geometry which characterizes the space heterogeneity of geological media gives rise to recirculating phenomena that limit the applicability of upscaled advection-transport equations in heterogeneous porous media at the field scale [16, 17, 18, 19, 20].

The averaging and homogenisation techniques developed to deal with a multi-scale phenomena like turbulence can be successfully applied to Navier-Stokes equation to obtain the Darcy law [21, 22, 23, 24, 25]. While scale separation in turbulence and transport stems from completely different causes (the former coming directly from the nonlinearity of the Navier-Stokes equation while the latter being a consequence of geometrical properties of the medium), the final goal of the upscaling process in both cases is to achieve a relation between local properties and the upscaled parameters involved in the equations

which describe the macroscopic dynamics [15].

So far we provided an overlook of the upscale meaning in the flow context. When it comes to transport modelling, the advection-dispersion equation structure is essentially preserved across the scales while its coefficients assume different meanings 1.1. At the pore scale, the hydrodynamic dispersion tensor

Table 1.1: Advection-dispersion coefficient meanings at different scales

	micro	*meso*	*macro*
$\boldsymbol{u}$	Navier-Stokes	Darcy's law	Spatial average
$\boldsymbol{D}$	Molecular diffusion	Hydrodynamic dispersion	Macrodispersion

is a diagonal matrix whose coefficients, in case of isotropic molecular diffusion, take the same constant value. At the heterogeneity scale, pore throats are no longer visible but their effect on transport is accounted by summing the mechanical dispersion tensor to the molecular diffusion diagonal matrix. This new tensor is named hydrodynamic dispersion. At the field scale, the macrodispersive model tries to capture the effect of the sediment heterogeneity by redefining the hydrodynamic coefficient as the macrodispersion coefficient, which takes a new statistical meaning.

1.3 Structure of the thesis

The aim of this first chapter is to identify the problem and to provide a description of the classical geostatistical and mathematical approaches to it while introducing the idea of the separation of scales on which following chapters are based. The rest of the study develops around five parts:

- a high level overview of the porous media transport modelling state of the art which can be further subdivided into geostatistical, flow and transport modelling;
- a detailed description of the macrodispersion model, its limitations and the rationale behind the numerical experiments performed in `OpenFOAM`®that is, the establishment of a hierarchy among the non-Fikcian triggering factors. This chapter was published as part of paper and can be read in full at [26];
- description of a real geological site and its reconstruction through the application of geostatistical techniques to field data (Herten analog).

Flow simulations through the permeability field reconstructed by [27] are performed as well and results analysed;

- assessment of geostatistical, flow and transport parameters on transport numerical simulations: the reconstruction of Herten permeability field is adopted as benchmark to evaluate the influence of geostatistical realism, isotropic and anisotropic mechanical dispersion, Péclet number and injection area on transport;

- conclusions.

Most of the experiments and the results presented in this thesis were achieved by collaborating with University of Nottingham and the British Geological Survey staff while using computers and facilities provided by the University of Nottingham. A minor part of this work focused on geostatistical simulations was developed during an external collaboration with Ephesia Consult. More in detail, the study and geostatistical simulations which involve the Herten analog were developed during a two months placement at Ephesia Consult as part of my second year of my PhD. The main areas of the small Swiss-French company are geological consultancy and geostatistical software development. According to the agreement achieved with the company, the objective of the collaboration was a geostatistical simulation suitable for both, their geostatistical demo purposes and our flow and transport tests.

Chapter 2

Geostatistical, flow and transport simulations in heterogeneous porous media

2.1 Geostatistical modeling of heterogeneous reservoirs

The physical and chemical properties of geological domains control the fate of fluid flow and solute transport through them. For this reason, a good understanding of the impact of the geological structure on solute transport requires an in-depth analysis of the relation between geostatistical parameters and macro-transport models. One of the main objective of this study is the assessment of the relation between the geostatistical parameters at the origin of the heterogeneous permeability pattern and the flow and transport.

2.1.1 Random fields

In geostatistical context, a random field $\boldsymbol{Z}$ is defined as a stochastic process indexed by an Euclidean space. Let $(\Omega, \mathcal{F}, \mathbb{P})$[1] be a probability space and $D \subset \mathbb{R}^2$ an Euclidean domain. A stochastic process indexed by an Euclidean space can be thought as a collection of random variables indexed by spatial locations $\boldsymbol{x} \in D$. Each random variable maps a set of possible outcomes Ω to a measurable space, often the real numbers. In this document the bold notation

[1] Ω is the sample space, $\mathcal{F}$ is the set of all the possible Ω subsets and $\mathbb{P}$ is a probability measure which assigns probabilities to all members of $\mathcal{F}$.

$\boldsymbol{Z}$ is used to describe a N-variate Gaussian distribution while Z refers to a single-variate Gaussian distribution. Relevant cases are [28]:

- *stationary* random fields: the mean $\mu(\boldsymbol{x})$ is constant and the covariance depends on $\boldsymbol{r} = \boldsymbol{x}_i - \boldsymbol{x}_j$;
- *isotropic* random fields: the covariance depends on $|\boldsymbol{x}_i - \boldsymbol{x}_j|$.

The covariance function for a random field measures the variation between random variables sampled at two random locations $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ and is defined as

$$C(\boldsymbol{r}) = C(Z(\boldsymbol{x}_i), Z(\boldsymbol{x}_j)) = E\left[(Z(\boldsymbol{x}_i) - E(Z(\boldsymbol{x}_i))) \cdot (Z(\boldsymbol{x}_j) - E(Z(\boldsymbol{x}_j)))\right], \tag{2.1}$$

where E is the expected value of a random variable. In practice, the covariance function computes the covariance value $C(\boldsymbol{r})$ between a pair of variables located at points separated by the vector $\boldsymbol{r}$. The covariance function $C(\boldsymbol{r})$ can also be rewritten as the product between the variance σ^2 and the correlation function $\rho(\boldsymbol{r})$

$$C(\boldsymbol{r}) = \sigma^2 \rho(\boldsymbol{r}). \tag{2.2}$$

It is a well established geostatistical convention to express the spatial relation of a random field in terms of variogram rather than covariance function. Since variogram and covariance function both convey the same information, that is the degree of spatial dependency of a random field between two points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, one of them is sufficient to describe $Z(\boldsymbol{x})$. While the covariance $C(\boldsymbol{r})$ is computed as the expected value of the product between field values at two different locations (2.1), the variogram $2\gamma(\boldsymbol{r})$ shows the variance of the difference between field values at two different locations that is

$$\begin{aligned} 2\gamma(\boldsymbol{r}) = 2\gamma(Z(\boldsymbol{x}_i), Z(\boldsymbol{x}_j)) &= \mathrm{Var}\left[Z(\boldsymbol{x}_i) - Z(\boldsymbol{x}_j)\right] \\ &= E\left[\left(Z(\boldsymbol{x}_i) - Z(\boldsymbol{x}_j) - E\left[Z(\boldsymbol{x}_i) - Z(\boldsymbol{x}_j)\right]\right)^2\right] \\ &= E\left[\left(Z(\boldsymbol{x}_i) - E(Z(\boldsymbol{x}_i)) - Z(\boldsymbol{x}_j) + E(Z(\boldsymbol{x}_j))\right)^2\right]. \end{aligned} \tag{2.3}$$

For a stationary field $E(Z(\boldsymbol{x}_i)) = E(Z(\boldsymbol{x}_j))$ hence the variogram becomes

$$2\gamma(Z(\boldsymbol{x}_i), Z(\boldsymbol{x}_j)) = E\left[\left(Z(\boldsymbol{x}_i) - Z(\boldsymbol{x}_j)\right)^2\right]. \tag{2.4}$$

The variogram differs from the covariance (or correlation) function in that the first measures the *variation* while the second expresses the *similarity*. For example, the variogram computed for two adjacent points that are likely to show similar properties will have a *low* value while the covariance value will be

high. The relation between covariance function and variogram can be deduced applying the covariance definition (2.1) to equation (2.3)

$$\begin{aligned} 2\gamma(Z(\boldsymbol{x}_i), Z(\boldsymbol{x}_j)) &= E\left[(Z(\boldsymbol{x}_i) - E(Z(\boldsymbol{x}_i)))^2\right] + E\left[(Z(\boldsymbol{x}_j) + E(Z(\boldsymbol{x}_j)))^2\right] \\ &\quad - 2E\left[(Z(\boldsymbol{x}_i) - E(Z(\boldsymbol{x}_i))) \cdot (Z(\boldsymbol{x}_j) - E(Z(\boldsymbol{x}_j)))\right] \\ &= C(Z(\boldsymbol{x}_i), Z(\boldsymbol{x}_i)) + C(Z(\boldsymbol{x}_j), Z(\boldsymbol{x}_j)) - 2C(Z(\boldsymbol{x}_i), Z(\boldsymbol{x}_j)) \end{aligned} \tag{2.5}$$

and it is illustrated in figure 2.1.

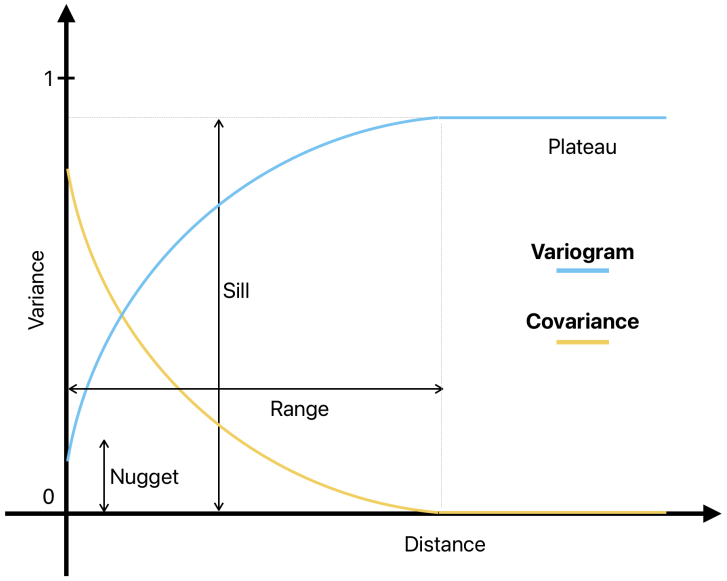

Figure 2.1: Geostatistical terminology to describe the variogram: *range* refers to the separation distance between two points beyond which no correlation is expected to be found; *nugget* is the value of the variance computed for a pair of adjacent points; *sill* is the value taken from the variance when it is computed for any pair of two points that are separated by a distance longer than the range. The variogram conveys the same information as the covariance function and it is the standard in the geological practice.

2.1.2 Gaussian random fields

Gaussian Random Fields (GRFs) are the most widely used random fields because of their good mathematical properties, such as being composed by a collection of random variables whose Probability Density Function (PDF) is

fully determined by the mean and the variance. In order for a random field to be Gaussian, its variables follow a Gaussian distribution *i.e.* their probability density function (PDF) should be

$$f(\boldsymbol{x}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\boldsymbol{x}-\mu}{\sigma}\right)^2}. \tag{2.6}$$

This entails that the collection of random variables

$$\boldsymbol{Z} = [Z(\boldsymbol{x}_1), Z(\boldsymbol{x}_2), .., Z(\boldsymbol{x}_N)]$$

behaves like an N-variate Gaussian distribution for any $\boldsymbol{x}_1, .., \boldsymbol{x}_N \in D$. That is, $\boldsymbol{Z} \sim N(\boldsymbol{\mu}, C)$ where the mean vector $\boldsymbol{\mu}$ and the covariance matrix C have entries $\mu_i = \mu(\boldsymbol{x}_i)$ and $c_{ij} = C(\boldsymbol{x}_i, \boldsymbol{x}_j) \quad \forall i, j = 1, .., N$. As an example, in figure 2.2 we show some two-dimensional realisations of GRFs using Gaussian (2.17), Exponential (2.18) and Matérn (2.19) covariances. Notwithstanding

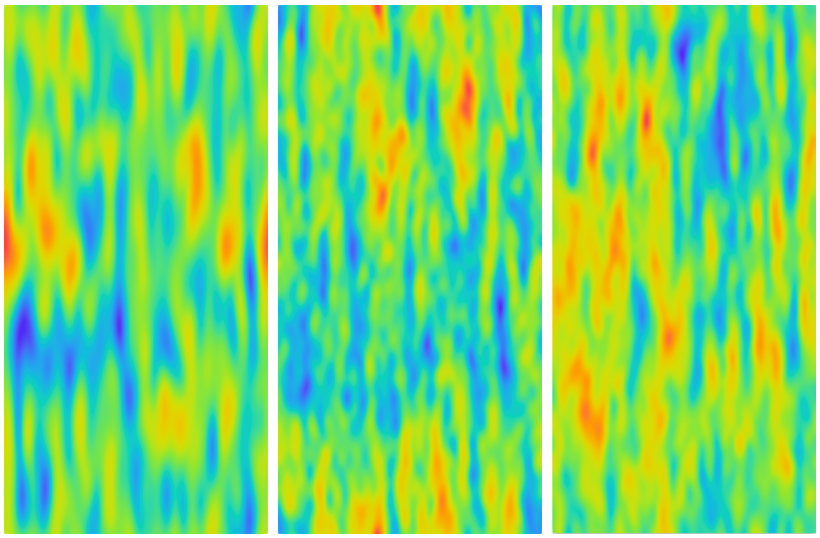

Figure 2.2: Continuous GRFs characterised by three different covariances. From left to right: Gaussian, exponential and Matérn covariance.

the visible differences between the fields in figure 2.2, they are all Gaussian as their random variables share the same Gaussian PDF. The differences between the spatial patterns stem from the different covariance functions: the smooth field on the left panel in 2.2 is generated using a Gaussian covariance 2.17 function while the covariance of the field in the central panel is exponential 2.18 which makes the spatial pattern more convoluted. The field on the right panel is characterised by Matérn covariance 2.19 which in terms of smoothness falls between the Gaussian and the exponential.

2.1.3 Truncated Gaussian random fields

GRFs such as the one in figure 2.2 are continuous functions while geological media are often characterised by abrupt changes in physical and chemical properties (figure 2.4). For this reason, Truncated Gaussian Simulations (TGSs) were deemed capable of reproducing the complex pattern of the sediments' properties with higher fidelity than continuous GRFs. The idea behind the algorithm is to simulate one (or more than one, as we will see in the PGS section 2.1.4) Gaussian variable $Z(\boldsymbol{x}_i)$ at every point in the study area and then use a *truncation* or *lithotype rule* to convert these values into facies. In other words, a *truncation* rule maps a set of continuous numbers Gaussian distributed to a discrete number of sediment categories. For example, as a result of the binning process which applies three thresholds to the continuous GRFs of figure 2.2 we obtain a contour map characterised by four intervals. If constant values are assigned to the four regions between the contour lines the result graphically appears as the truncated GRFs in figure 2.3. These colours

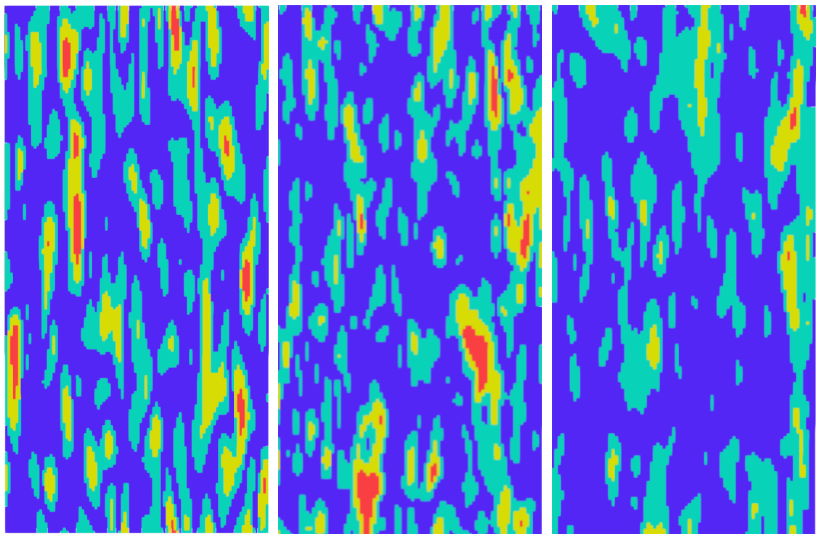

Figure 2.3: Truncated GRFs characterised by three different covariances. From left to right: Gaussian, exponential and Matérn covariance.

can be associated to different properties of the geological media such as permeability or porosity and the interval definition is called *truncation rule*. In this sense, the "truncated" adjective refers to a GRF that has been "binned". As it will be explained more in the details in section 4.2, the TGS algorithm relies on information provided by field samples and processed using transition probability matrices [29, 30] to capture the vertical transition pattern between

facies (i.e. facies that can or cannot be adjacent) while it uses conceptual models like the one in figure 2.4 for the same information on the horizontal direction. Categorical random fields can honour two important observed

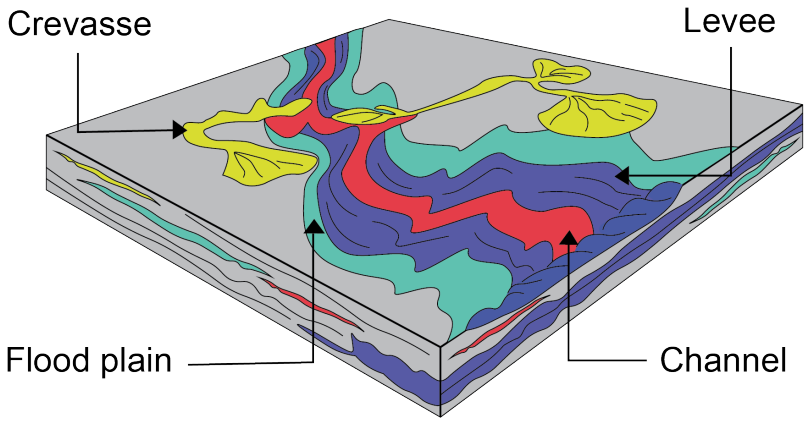

Figure 2.4: Conceptual geological model of alluvial sediments. Since boreholes can only capture the vertical alternation of facies, a conceptual model scheme is needed to determine the horizontal contact between facies.

features: the contact sequence between the facies and their volumetric proportions. The theory which follows refers to the case where only one GRF is truncated multiple times, which provides self-embedded or orderly facies (figure 2.3) i.e. structure where one facies always share its boundary at most with two and only two facies which remain the same over the domain. This is useful to reproduce geological domains where the sequence stratigraphy is clear, for instance when sandstone is followed by shaly sandstone then shale so that shale will never touch sandstone. However, repeating the truncation process for two GRFs at time allow the simulation of more complex structures (figure 2.5).

We now define a truncated Gaussian random field. The result of a mapping process which associate the values of a continuous N-variate Gaussian random field $\boldsymbol{Z}$ to a partition $B(\boldsymbol{x})$ of the n-dimensional real space $\mathbb{R}^n$ is called truncated Gaussian random field. In this study the n-dimensional real space $\mathbb{R}^n$ is either the real coordinate plane $\mathbb{R}^2$ or three dimensional space $\mathbb{R}^3$. Following [31], F_i is the i^{th}-sediment category which seats at a point $\boldsymbol{x} \in \mathbb{R}^n$ and that is associated to the partition $B_i(\boldsymbol{x})$ through the truncation rule. In other words, the value of $\boldsymbol{Z}$ at location $\boldsymbol{x}$ is mapped into one of the sediment categories F_i

if the value falls into the interval defined by the partition $B_i(\boldsymbol{x})$ that is

$$F_i(\boldsymbol{x}) \triangleq Z(\boldsymbol{x}) \in B_i(\boldsymbol{x}). \tag{2.7}$$

The probability of having the category F_i at the point $\boldsymbol{x}$ corresponds to its volumetric proportion that is, the ratio between volume occupied by F_i and the whole domain volume F. Adopting the notation by [31], we can formulate it as

$$p_{F_i} = E[1_{F_i}(\boldsymbol{x})] = p(t_{i-1} < \boldsymbol{x} \leqslant t_i) \tag{2.8}$$

where p and E are referring to the probability and the expected value of a random variable while 1_{F_i} is an indicator function that is one when F_i is found at the position $\boldsymbol{x}$ and zero everywhere else. For fixed threshold values t_i, all the values taken by $Z(\boldsymbol{x})$ in between two thresholds are set constant. The smooth transition which characterizes the GRF is then replaced by $i+1$ categorical values. Calling 1_{F_i} the facies indicator function we can define i facies as

$$1_{F_i}(\boldsymbol{x}) = 1 \Leftrightarrow t_{i-1} \leqslant Z(\boldsymbol{x}) \leqslant t_i \tag{2.9}$$

whose probability p_{F_i} are

$$p_{F_i}(\boldsymbol{x}) = p(t_{i-1} \leqslant Z(\boldsymbol{x}) \leqslant t_i) = \tag{2.10}$$

$$= p(-\infty < Z(\boldsymbol{x}) \leqslant t_i) - p(-\infty < Z(\boldsymbol{x}) \leqslant t_{i-1}). \tag{2.11}$$

However, what happens most of the time in practice is that we want to retrieve the thresholds from the experimentally known facies proportion. Once the sequence between facies has been established, the one-to-one relation between facies proportions and thresholds is relatively easy to be inverted. Being $G(t)$ the cumulative PDF of the Gaussian distribution, the probability of the facies i is computed as

$$p_{F_i}(\boldsymbol{x}) = G(t_i) - G(t_{i-1}). \tag{2.12}$$

By inverting this relationship we can deduce the thresholds

$$\begin{aligned} t_1 &= G^{-1}[p_{F_1}(\boldsymbol{x})] \\ t_2 &= G^{-1}[p_{F_1}(\boldsymbol{x}) + p_{F_2}(\boldsymbol{x})] \\ t_i &= G^{-1}[p_{F_1}(\boldsymbol{x}) + p_{F_2}(\boldsymbol{x}) + .. + p_{F_i}(\boldsymbol{x})]. \end{aligned} \tag{2.13}$$

For example, to retrieve the thresholds associated to a geological domain characterised by three sediment categories with volumetric proportions $p_{F_1}(\boldsymbol{x}) = 0.2$, $p_{F_2}(\boldsymbol{x}) = 0.5$ and $p_{F_3}(\boldsymbol{x}) = 0.3$ one can use equations 2.13

$$t_1 = G^{-1}[0.2] = -8.84 \tag{2.14}$$

$$t_2 = G^{-1}[0.2 + 0.5] = 0.524. \tag{2.15}$$

By doing so, the PDF of the Gaussian variables at each spatial location are divided into three intervals which honour the observed volumetric proportion.

2.1.4 Truncated pluri-Gaussian random fields

Sediments' deposits rarely show an orderly or stationary structure. When geological surveys suggest no natural sequence between facies that is, certain facies can be in contact with more than two facies, the TGS model is too restrictive. To reproduce non-stationary geometries, we need two or more GRFs. For this reason, a simulation derived from the truncation of multiple GRFs is called PGS. Three examples of PGS fields with Gaussian, exponential and Matérn covariance functions are illustrated in figure 2.5. The idea behind

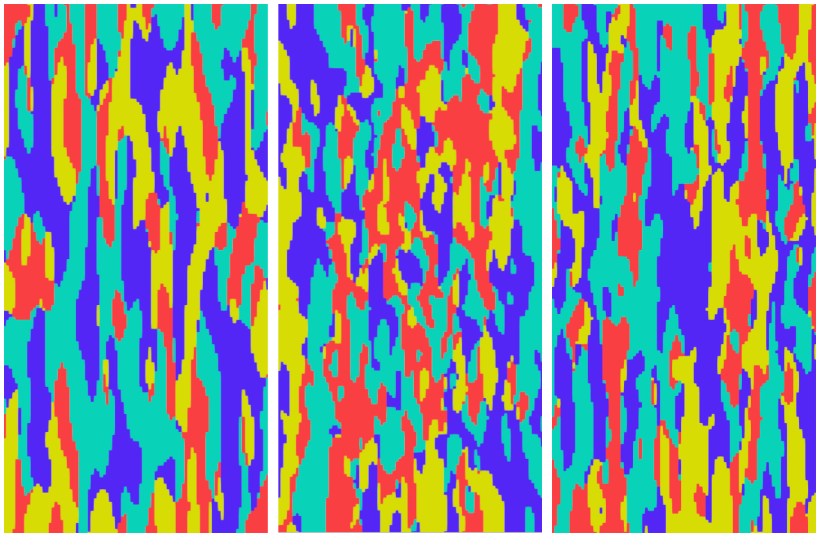

Figure 2.5: Bi-truncated GRFs characterised by three different covariances. From left to right: Gaussian, exponential and Matérn covariance.

the PGS method is to simulate two or more continuous GRFs on a spatial domain and compare them through a series of inequalities which allows us to assign a unique value to each cell (figure 3.1). In our case, this value is an indicator of the facies and the series of inequality is described by the *truncation rule* (figure 2.6). For example, assuming that the underlying GRFs of figure 2.5 are $Z_1(\boldsymbol{x})$ and $Z_2(\boldsymbol{x})$, the red cells in the domain satisfy the condition $t_1 < Z_1(\boldsymbol{x}) < t_3$ and $s_1 < Z_2(\boldsymbol{x}) < s_2$, as illustrated by the truncation rule (figure 2.6). The volumetric fraction of the facies F_i is equal to the probability p of finding F_i at the location $\boldsymbol{x}$ which is

$$p_{F_i}(\boldsymbol{x}) = \left[G_1(t_i) - G_1(t_{i-1})\right]\left[G_2(s_i) - G_2(s_{i-1})\right]. \tag{2.16}$$

The problem of defining the thresholds given the volume proportions (i.e. the probabilities) leads to the optimisation of i error functions. In other words,

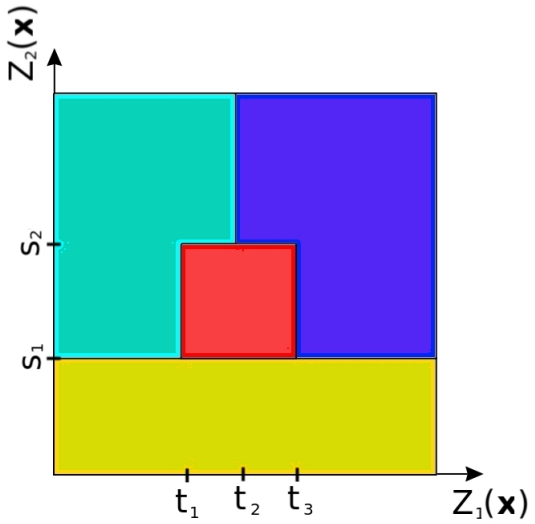

Figure 2.6: Qualitative truncation rule for figure 2.3. Note that the areas of the different facies do not correspond to the their respective proportions in the simulations, not because of the qualitative nature of the image, but because the underlying continuous variables are Gaussian and not uniformly distributed.

there are infinite threshold combinations which honour a given set of volume proportions and some constraints need to be defined beforehand to allow for a single solution. However, it is always possible to arbitrary fix thresholds and compute the corresponding set of proportions.

2.1.5 Correlation functions and variograms

Depending on geological factors such as the depositional history, the orogenesis or the tectonic deformations, the subsurface sediment structure might show strongly different patterns. To match these spatial characteristics, a large number of covariances are available in literature [32]. Some of the most popular covariances in the geostatistical practice together with their respective spectral representations, i.e. their form after the Fourier transform being applied to the covariance, are described by the expressions 2.17, 2.18 and 2.19. As illustrated in section 2.1.6, the spectral representation of the covariance is essential for the implementation of the `OpenFOAM`®geostatistical utility `setRandomField` that can be downloaded at [6].

Gaussian

$$C(d;\lambda) = \sigma^2\left(1 - e^{-\frac{\pi}{4}\left(\frac{d}{\lambda}\right)^2}\right) \quad \text{and} \quad S(\boldsymbol{a};\lambda) = \sigma^2\left(\frac{\lambda}{\pi}\right)^n e^{-\frac{1}{\pi}(\boldsymbol{a}\lambda)^2} \tag{2.17}$$

Exponential

$$C(d;\lambda) = \sigma^2\left(1 - e^{-\frac{|d|}{\lambda}}\right) \quad \text{and} \quad S(\boldsymbol{a};\lambda) = \sigma^2\lambda^n \frac{\Gamma\left(\frac{n+1}{2}\right)}{\left(\pi\left(1+\boldsymbol{a}^2\lambda^2\right)\right)^{\frac{n+1}{2}}} \tag{2.18}$$

Matérn [33]

$$\begin{aligned} C(d;\nu,\lambda) &= \sigma^2\left(1 - \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\sqrt{2\nu}\frac{d}{\lambda}\right)^{\nu} K_\nu\left(\sqrt{2\nu}\frac{d}{\lambda}\right)\right) \\ S(\boldsymbol{a};\nu,\lambda) &= \sigma^2\lambda^n \frac{\Gamma(\nu+\frac{n}{2})(2\nu)^\nu}{\Gamma(\nu)\pi^{\frac{n+1}{2}}\left(2\nu+\lambda^2\boldsymbol{a}^2\right)^{\nu+\frac{n}{2}}} \end{aligned} \tag{2.19}$$

The parameters in the expressions 2.17-2.19 are defined as:

- d is the distance which separates a pair of random variables;
- λ is the correlation length;
- $\boldsymbol{a}$ is the frequency vector;
- n is the number of dimensions;
- Γ is the gamma function;
- ν is a smoothness parameter;
- K_ν is the modified Bessel function of the second kind.

The exponential and Gaussian covariances are both special cases of the Matérn covariance. For $\nu = 1/2$ the Matérn covariance becomes the exponential while for $\nu \to \infty$ it becomes the Gaussian. As it can be noted from figure 2.2, the Gaussian covariance 2.17 generates smoother patterns than the exponential covariance 2.18. In terms of smoothness, the pattern generated by the Matérn covariance in figure 2.2 is between the Gaussian and the exponential with slight variations depending on the value of ν.

2.1.6 Geostatistical implementation

We leveraged on the well established use of geostatistics to reproduce geological properties distributions [8] to associate randomly generated truncated pluri-Gaussian fields with subsoil permeability patterns. Although the complexity of reproducing the subsurface sediment distribution has been tackled using different approaches [8], we chose the truncated pluri-Gaussian simulation method because it can reproduce the high contrast heterogeneity patterns characteristic of the subsoil with a high degree of realism.

Several well-known methods have been implemented in literature for generating GRFs [34, 35, 36, 37]. According to [38] a GRF can be represented using a stochastic Fourier integral:

$$Z(\boldsymbol{x}) = \int_{-\infty}^{+\infty} e^{-2\pi i \boldsymbol{a}\cdot\boldsymbol{x}} \sqrt{S(\boldsymbol{a})} dW(\boldsymbol{a}). \tag{2.20}$$

where $\boldsymbol{a}$ are frequencies, $dW(\boldsymbol{a})$ is a complex Gaussian random variable and $S(\boldsymbol{a})$ is the amplitude of the spectral measure. This can be rewritten as

$$Z(\boldsymbol{x}) = \int_{-\infty}^{+\infty} \cos(2\pi \boldsymbol{a}\cdot\boldsymbol{x}) \sqrt{S(\boldsymbol{a})} dW(\boldsymbol{a}) + i \int_{-\infty}^{+\infty} \sin(2\pi \boldsymbol{a}\cdot\boldsymbol{x}) \sqrt{S(\boldsymbol{a})} dW(\boldsymbol{a}). \tag{2.21}$$

that in its discrete form becomes

$$Z(\boldsymbol{x}) = \sum_{j=0}^{N_f} \cos(2\pi \boldsymbol{a}_j\cdot\boldsymbol{x}) \sqrt{S(\boldsymbol{a}_j)} W_j + i \sum_{j=0}^{N_f} \sin(2\pi \boldsymbol{a}_j\cdot\boldsymbol{x}) \sqrt{S(\boldsymbol{a}_j)} W'_j. \tag{2.22}$$

Equation 2.22 represents the core of the PGS algorithm implemented in the `setRandomFiled` geostatistical library as it is used to generate continuous GRFs. Several geostatistical features were implemented in the `setRandomFiled` library to enable a good match of the geometrical properties between the virtual geostatistical field and the real geological media. These parameters are illustrated in table 2.1 along with the accepted values. An accurate description of the PGS algorithm implementation for the exponential covariance can be found in section 3.3.1.

Geostatistical library parameters		
Feature	*Description*	*Values*
Field name	Name of the field	string of characters
Field type	Distribution of the random variable	uniform, gaussian, logNormal, truncated, bitruncated
Correlation function	Variable spatial dependence	exponential, Gaussian, Matérn
Function parameters	Mean, variance and correlation lengths	$\mathbb{R}$
Enable statistics	Metrics of the generated fields	true, false
Disable dimensions	1D, 2D or 3D	true, false
Periodicity	Periodic field	true, false
Frequencies	Number of frequencies	$\mathbb{N}$
Thresholds values	Categories' limits expressed as input values for the Gaussian	$\mathbb{R}$
Percentiles values	Categories' limits expressed as volume fraction of sediments	[0 1]
Field values	Numerical value assigned to each category	$\mathbb{R}$

Table 2.1: Most relevant geostatistical features available for the `OpenFOAM` library `setRandomField`. The third column collects the valid input for setting the features in the first column.

2.2 Flow and transport mathematical models

While in 1856 Henry Darcy was working on a freshwater filtering problem for the fountains of Dijon, he first proposed a relation between the flow rate and the pressure difference [14]:

$$q = -K\frac{h_2 - h_1}{l} \tag{2.23}$$

where q is the rate of water displacement per unit area $[LT^{-1}]$, l is the distance between the inlet and outlet sections of the porous media $[L]$, h_2 and h_1 are the heights $[L]$ of water measured in manometers upstream and downstream respectively and K is a factor of proportionality, namely hydraulic conductivity $[LT^{-1}]$. Another fundamental relationship to 2.23 for transport in fluids was derived by Adol Fick in 1855:

$$J = -D\frac{dc}{dx} \tag{2.24}$$

where J is the diffusive flux $[L^2T^{-1}]$, D is the molecular diffusion coefficient $[L^2T^{-1}]$ and c is the volume fraction of solute in fluid [-].

These groundbreaking empirical intuitions paved the way for the study and modelling of flow and transport in porous media. In the next sections, the mathematical description of the flow and transport mechanisms is illustrated.

2.2.1 Continuity equation

The continuity equation for most of the groundwater application purposes is a scalar equation where the unknown fluid density appears in the temporal derivative and in the divergence. The interpretation is that, for a fixed region in space, the temporal and spatial variation of one property equals the property variation produced by the sinks/sources (Eq. 2.25)

$$\frac{\partial(\rho\phi)}{\partial t} + \nabla \cdot (\rho \boldsymbol{V}) = \rho_f R_f \tag{2.25}$$

where ρ is the fluid density $[ML^{-3}]$, ϕ is the porosity $[-]$, $\boldsymbol{V}$ is the Darcy's velocity $[LT^{-1}]$ and R_f is a source or sink rate $[T^{-1}]$ of fluid with density ρ_f $[ML^{-3}]$. Expanding the derivatives we find

$$\rho\frac{\partial\phi}{\partial t} + \phi\frac{\partial\rho}{\partial t} + \rho\nabla \cdot \boldsymbol{V} + \boldsymbol{V} \cdot \nabla\rho = \rho_f R_f \tag{2.26}$$

that under the following hypothesis:

- constant porosity;
- constant and homogeneous density;
- no sources or sinks in the domain;

yields the incompressibility condition

$$\nabla \cdot \boldsymbol{V} = 0. \tag{2.27}$$

2.2.2 From Navier-Stokes to Darcy's law

We provide a short derivation of the Darcy law from the Navier-Stokes equation. We start by deriving Navier-Stokes momentum equation 2.34 from Newton's second law 2.28 and, under low Reynolds number assumption, simplifying it into the Stokes equation. Upscaling the Stokes equation 2.36 as illustrated by [21, 24] leads to the Darcy equation (e.g. [39]).

Cauchy momentum equation

Starting from Newton's second law

$$\frac{D\boldsymbol{p}}{Dt} = \boldsymbol{F}, \tag{2.28}$$

and assuming constant density, one can see the derivative of the momentum $\boldsymbol{p}$ at the left hand side as

$$\frac{D\boldsymbol{p}}{Dt} = \frac{D\boldsymbol{u}}{Dt}\rho dxdydz \tag{2.29}$$

and the resulting force $\boldsymbol{F}$ on the right hand side as the sum of body and surface forces

$$\boldsymbol{F} = \boldsymbol{F}_b + \boldsymbol{F}_s. \tag{2.30}$$

If the only body force considered is the gravitational force

$$\boldsymbol{F}_b = \boldsymbol{g}\rho dxdydz \tag{2.31}$$

and the surface forces are described by the the stress tensor $\boldsymbol{\sigma}$

$$\boldsymbol{F}_s = (\nabla \cdot \boldsymbol{\sigma})\, dxdydz \tag{2.32}$$

then Newton's second law 2.28 becomes Cauchy momentum equation:

$$\frac{D\boldsymbol{u}}{Dt} = \frac{1}{\rho}\nabla \cdot \boldsymbol{\sigma} + \boldsymbol{g}. \tag{2.33}$$

Navier-Stokes momentum equation - Lagrangian form

Navier-Stokes momentum equation can be derived as a particular case of Cauchy momentum equation 2.33 by setting the Cauchy stress tensor $\boldsymbol{\sigma}$ to be the sum of viscous (or deviatoric) stress matrix $\boldsymbol{\tau}$ and the pressure (or volumetric) stress $-p\mathbb{I}$.

$$\rho \frac{D\boldsymbol{u}}{Dt} = -\nabla p + \nabla \cdot \boldsymbol{\tau} + \rho \boldsymbol{g} \tag{2.34}$$

Navier-Stokes momentum equation - Eulerian form

Expanding the total derivative of equation 2.34, the Navier-Stokes momentum equation becomes

$$\rho \left(\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{v} \cdot \nabla \boldsymbol{u} \right) = -\nabla p + \nabla \cdot \boldsymbol{\tau} + \rho \boldsymbol{g} \tag{2.35}$$

where $\boldsymbol{u}$ is the velocity of a particle of fluid with respect to a Lagrangian reference system and $\boldsymbol{v}$ the Eulerian velocity of the fluid, i.e. with respect to a fixed reference frame system.

If lagrangian acceleration of particles flowing through a porous media are considered negligible the material derivative $D\boldsymbol{u}/Dt$ vanishes and the Navier-Stokes equation 2.34 simplifies into

$$0 = -\nabla p + \nabla \cdot \boldsymbol{\tau} + \rho \boldsymbol{g} \tag{2.36}$$

Assuming that the force resisting the motion of a particle at a point inside an elementary channel is proportional and opposite to the Darcy's velocity $\boldsymbol{V}$, the divergence of the stress tensor can be written as

$$\nabla \cdot \boldsymbol{\tau} = -\mu V_i k_{ji}^{-1} \quad i, j = 1, 2, 3 \tag{2.37}$$

where $\boldsymbol{k}$ is a function of the channel's geometry and μ is the dynamic viscosity of the fluid. Substituting equation 2.37 in equation 2.36 yields

$$0 = -\nabla p - \mu V_i k_{ji}^{-1} + \rho \boldsymbol{g} \tag{2.38}$$

which is the Darcy law

$$\boldsymbol{V} = -\frac{\boldsymbol{k}}{\mu} \left(\nabla p - \rho \boldsymbol{g} \right). \tag{2.39}$$

The Darcy velocity $\boldsymbol{V}$ is also called apparent velocity because it represents the velocity of the fluid as if the porous matrix did not occupy any space while the real velocity $\boldsymbol{v} = \boldsymbol{V}/\phi$ is the velocity that would be measured by a flow meter in the porous domain and it clearly exceeds the Darcy velocity.

2.2.3 Darcy's law

The constitutive equation for groundwater flows is Darcy's law (Eq. 2.39). The most relevant statements about this experimental law which has been rewritten multiple times using disparate notations, concern the hypothesis which go under each formulation [40, 41]. The formulation we adopt here (Eq. 3.5) is the most general, rich and complete, preparatory to the code implementation work

$$\boldsymbol{V} = -\frac{\boldsymbol{k}}{\mu}(\nabla p + \rho g \nabla z) \tag{2.40}$$

where $\boldsymbol{k}$ is the hydraulic permeability tensor $[L^2]$, μ is the dynamic viscosity $[ML^{-1}T^{-1}]$, p is the pressure $[MT^2L]$, g is the gravity constant $[LT^{-2}]$ and $\nabla z = (0, 0, 1)$ $[-]$. The original formulation (Eq. 2.23), although it suffers from a uniform density assumption which is rarely explicitly stated, is probably more suitable for direct application to practical hydrogeological problems where the uniform density hypothesis applies. The uniform density assumption ($\nabla \rho = 0$) is seldom explicitly stated for two main reasons:

- it is implicit in the more restrictive assumptions of incompressibility condition (Eq. 2.27) that is often assumed to close the system of three equations (2.40) in four unknowns ($\boldsymbol{V}$ and p);
- fluids for which *pressure compressibility* is generally negligible, show a non-negligible *chemical compressibility* behavior, i.e. relevant density changes. For instance, relevant quantity of salt can be dissolved into water before sensible volume variations can take place; as a result, we observe significant density increments. Apparently *pressure incompressible* liquids such as water suddenly become *chemically compressible*.

By means of a substitution as simple as

$$\boldsymbol{k} = \frac{\mu \boldsymbol{K}}{\rho g}, \tag{2.41}$$

and defining a scalar function called *hydraulic head* $h = \frac{p}{\rho g} + z$ $[L]$, equation 2.40 can be rewritten as

$$\boldsymbol{V} = -\boldsymbol{K} \nabla h. \tag{2.42}$$

where $\boldsymbol{K}$ is the hydraulic conductivity $[LT^{-1}]$. Given the popularity of equation 2.42, it is worth noting that, notwithstanding the similar notation, permeability $\boldsymbol{k}$ and hydraulic conductivity $\boldsymbol{K}$ are profoundly different: the first merely describes a properties of the porous medium, the second depends on

both, the properties of the porous medium *and* on the properties of the fluid itself (density and viscosity). As a consequence, $\boldsymbol{k}$ does not change with the fluid while $\boldsymbol{K}$ does. In brief:

- permeability $\boldsymbol{k}$ is the ability of a specific material to allow for any fluid to pass;
- hydraulic conductivity $\boldsymbol{K}$ describes the ease of flow through porous media related to a specific combination of flow and material.

Stated this necessary distinction, we will mostly use Darcy's law in the form (Eq. 2.40).

Groundwater flow

The combination of continuity equation (Eq. 2.25) with Darcy's law (Eq. 2.40) provides a system of four equations which contains all the six unknowns (p, ρ, ϕ and $\boldsymbol{V}$) which are physically relevant to fully describe groundwater flow. Clearly, this problem cannot be solved unless two constitutive equations for ρ and ϕ are provided or the number of unknowns is reduced. Under the assumptions of equation 2.27, ϕ and ρ become parameters so that the number of unknowns is reduced and the resulting system can be rewritten as the canonical equation for groundwater flow

$$\frac{\partial}{\partial x}\left(-\frac{k_x}{\mu}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(-\frac{k_y}{\mu}\frac{\partial p}{\partial y}\right)+\frac{\partial}{\partial z}\left(-\frac{k_z}{\mu}\left(\frac{\partial p}{\partial z}+\rho g\right)\right)=0. \qquad (2.43)$$

2.2.4 Advection-dispersion transport model

Classical transport theories identify the spread of a nonreactive solute in porous media as a consequence of three mechanisms:

a) molecular diffusion;

b) mechanical (or kinematic) dispersion (sometimes effective diffusion);

c) advection (often convection).

A fourth type of dispersion called Taylor dispersion [42, 43] describes the spread of the solute under the variation of the velocity profile through small tubes. Given the symmetry of such dispersion which moves with the mean speed of the flow, it has negligible effects from a meso-scale perspective. The combination of the first three mechanisms (a-c) is usually embedded in the

hydrodynamic dispersion matrix $\boldsymbol{D}$. Unfortunately, no consensus about terminology has been reached yet and the attribution of different meanings to the same word by distinct communities (e.g. geologists, engineers and mathematicians) could lead to misleading interpretations. Beyond the terminology, mechanism (a) indicates the motion of particles at molecular scale caused by their own random collisions, mechanism (b) describes velocity-driven processes where molecules disperse by tracking through the tortuous pathways of the medium because of the variability of the complex, microscopic velocities while mechanism (c) represents the way particles are carried by the bulk motion of the fluid [44]. Before introducing their mathematical representations, an important distinction is needed: when flow velocity is zero, the only flux is induced by molecular diffusion (a) while when flow velocity is non-zero solute spreading is induced by the simultaneous contributions of molecular diffusion (a), mechanical dispersion (b) and advection (c). Taking the spread of a contaminant in a homogeneous porous medium as an example, in case of no flow molecular diffusion (a) would generate an increasing spherical plume. However, if flow velocity is added, the plume would spread uniformly as it advected with the flow (c) but it also elongates since the spreading caused by mechanical dispersion (b) is generally observed to be greater in the direction of the flow rather than in the transverse direction. In case of fluid flow, the contribution of each mechanism is variable and the dominating mechanism depends on the Péclet number $[-]$

$$\boldsymbol{Pe} = \frac{\boldsymbol{V} l}{\phi D_{mol}} \tag{2.44}$$

where l $[L]$ is a characteristic length (e.g. the correlation length) and D_{mol} $[L^2T^{-1}]$ is the *molecular diffusion* coefficient. It has been observed from experiments that

- for $\boldsymbol{Pe} < 1$ molecular diffusion dominates mechanical dispersion;
- for $1 < \boldsymbol{Pe} < 10$ molecular diffusion an mechanical dispersion are comparable;
- for $\boldsymbol{Pe} > 10$ mechanical dispersion dominates molecular diffusion.

Fick's law: molecular diffusion and mechanical dispersion

We will now focus on case (a), hence no flow velocity. Detailed description of the collision between molecules at aquifer scale is still computationally prohibitive hence, predictions on the spread of a solute in a porous medium rely on Fick's law (Eq. 2.45). Similar to Darcy's law for hydraulic flow, Ohm's

law for charge transport and Fourier's law for heat transport, the constitutive equation for the diffusive flux per unit area $\boldsymbol{J_0}$ $[LT^{-1}]$ happens in the direction of the negative gradient

$$\boldsymbol{J_0} = -D_0 \nabla c. \tag{2.45}$$

The diffusion equation can be derived from the continuity equation 2.25 by substituting the density ρ, the Darcy flux $\rho \boldsymbol{V}$ and the fluid sink/source rate $\rho_f R_f$ respectively with the volume fraction of solute in fluid c $[-]$, the diffusion flux per unit area $\boldsymbol{J_0}$ and the sink/source rate $c_s R_s$ of the solute (no porosity ϕ is present because we are not considering the porous matrix yet)

$$\frac{\partial c}{\partial t} - \nabla \cdot (D_0 \nabla c) = c_s R_s \tag{2.46}$$

where D_0 is the molecular diffusion coefficient in the fluid $[L^2 T^{-1}]$. In common with other linear parabolic partial differential equations, equation 2.45 shows some basic properties:

- disturbances propagate at infinite speed;
- it smooths out roughness in the initial boundary data;
- there is a loss of information as signals propagate;
- there is a maximum principle.

The molecular diffusion and the mechanical dispersion fluxes in porous media are then given by

a) $\boldsymbol{J_{mol}} = -\phi D_{mol} \nabla c$;

b) $\boldsymbol{J_{mec}} = -\phi \boldsymbol{D_{mec}} \nabla c$;

where D_{mol} is the effective molecular diffusion coefficient in the porous medium $[L^2 T^{-1}]$ and $\boldsymbol{D_{mec}}$ is the mechanical dispersion matrix $[L^2 T^{-1}]$. While the use of Fick's law (Eq. 2.45) for case (a) is widely acknowledged, the Fickianity of the dispersion flux (case b) is still debated. However, since no alternative model is currently available, we will assume its validity for the next chapters. The effective molecular diffusion coefficient D_{mol} is usually smaller than the molecular diffusion coefficient D_0 because of the hinder effect of the solid boundaries of the porous structure where the liquid phase is enclosed. The range of the ratio D_{mol}/D_0 is usually $0.1 - 0.7$ and is called tortuosity.

Mechanical dispersion models

Mechanical dispersion needs to be taken into account when solute is transported through a porous matrix. Based on the analogy with the molecular diffusion, the mechanical dispersion flux is assumed to happen against the concentration gradient and proportionally to the mechanical dispersion coefficient however, the parallelism with molecular diffusion holds as long as channels and geometry of the porous matrix are uniformly distributed in space so that the overall effect of mechanical dispersion can be assimilated to the molecular agitation at the microscopic scale. Based on this parallelism, the mechanical dispersion coefficient $\boldsymbol{D_{mec}}$ has the same units as the molecular diffusion coefficient and it describes all the displacements that the porous matrix induces on the fluid and that are not represented neither by the advective or diffusive models, i.e. sharp velocity gradients and highly permeable channels. One constitutive model [15, 45] for mechanical dispersion is

$$\boldsymbol{D_{mec}} = \alpha_T |\boldsymbol{V}| \mathbb{I} + (\alpha_L - \alpha_T) \frac{\boldsymbol{V}^T \boldsymbol{V}}{|\boldsymbol{V}|} \tag{2.47}$$

where α_L and α_T are the longitudinal and transversal mechanical dispersion coefficients, $\boldsymbol{V}$ is the local Darcy velocity written as a row vector and $\mathbb{I}$ is the identity matrix. In a simplified case where the flow is only longitudinal, the Darcy velocity has one non-zero component $\boldsymbol{V} = (V_x, 0, 0)$ and the mechanical dispersion model 2.47 can be rewritten as

$$\boldsymbol{D_{mec}} = \begin{bmatrix} \alpha_L V_x & 0 & 0 \\ 0 & \alpha_T V_x & 0 \\ 0 & 0 & \alpha_T V_x \end{bmatrix} \tag{2.48}$$

Typically $\boldsymbol{\alpha}$ in the longitudinal direction is roughly taken one order of magnitude bigger than in the transversal direction ($\alpha_T \simeq 0.1 \alpha_L$).

Advection-dispersion equation

The mathematical representation of the advective flux is

c) $\boldsymbol{J_{adv}} = \boldsymbol{V} c$.

The net flux per unit area $\boldsymbol{J}$ $[LT^{-1}]$ is then given by the sum of the three fluxes (a - c)

$$\boldsymbol{J} = \boldsymbol{J_{adv}} + \boldsymbol{J_{mol}} + \boldsymbol{J_{mec}} = \boldsymbol{V} c - \phi(\mathbb{I} D_{mol} + \boldsymbol{D_{mec}}) \nabla c = \boldsymbol{V} c - \phi \boldsymbol{D} \nabla c \tag{2.49}$$

for which the sum of $\mathbb{I}D_{mol}$ and $\boldsymbol{D_{mec}}$ gives the hydrodynamic dispersion matrix $\boldsymbol{D}$. The advection-dispersion equation can be derived from the continuity equation 2.25 by substituting the concentration c to the density ρ and the net flux per unit area $\boldsymbol{J}$ 2.49 to the Darcy flux $\rho\boldsymbol{V}$. By doing so, the continuity equation 2.25 becomes

$$\frac{\partial(\phi c)}{\partial t} + \nabla \cdot (\boldsymbol{V}c) - \phi\nabla \cdot (\boldsymbol{D}\nabla c) = c_s R_s. \tag{2.50}$$

In our study no reactions have been considered so that, expanding the variables' dependency, equation (2.50) simplifies into the advection-dispersion equation

$$\frac{\partial c(\boldsymbol{x}, t)}{\partial t} + \nabla \cdot [\boldsymbol{v}(\boldsymbol{x})c(\boldsymbol{x}, t)] - \nabla \cdot [\boldsymbol{D}(\boldsymbol{x})\nabla c(\boldsymbol{x}, t)] = 0 \tag{2.51}$$

where $\boldsymbol{v}$ represents the real velocity of the flow through the channel throats.

2.2.5 Transport upscaling

Upscaling means to find analogies between phenomena which happen on separate scales and try to fit the smaller scale mathematical description to the larger scale mechanism. One example that has widely been tested is the adaptation of Fick's law initially developed for molecular diffusion (a) description to capture the effect of the mechanical dispersion (b) driven by the porous media channel structure through a redefinition of the molecular diffusion coefficient. The same way [1], the advection-dispersion equation 2.51 first formulated to model the transport at the meso-scale, is proposed to capture the effect of the sediment heterogeneity at the field scale. To do so, the hydrodynamic dispersion matrix $\boldsymbol{D}$ becomes the macrodispersion matrix $\boldsymbol{D_{mac}}$ which simulates the dispersion effect that the sediment regions with different properties have on the solute

$$\frac{\partial(\phi c)}{\partial t} + \nabla \cdot (\bar{\boldsymbol{V}}c) - \phi\nabla \cdot (\boldsymbol{D_{mac}}\nabla c) = 0 \tag{2.52}$$

where instead of $\boldsymbol{V}$ the spatially averaged Darcy velocity $\bar{\boldsymbol{V}}$ is adopted.

Advection-dispersion equation however, was conceived and experimentally verified in meso-scale porous domains where sediment properties were homogeneously distributed. Thus, it is understandable how its application to macro-scale transport problems characterised by sharp variations in geological properties might sometimes be affected by large uncertainties. Indeed, the attempts made to adapt the original hydrodynamic dispersion coefficient to account for macro-dispersion phenomena often results in large macro-dispersion values

affected by elevated uncertainties. Some macro-dispersion predictive and calibrated methods to estimate the macro-dispersion coefficient are illustrated in section 3.3.5.

2.2.6 Transport implementation

The in-house solver for scalar transport `scalarTransportFoam` was initially adopted for preliminary transport tests on PGS fields. As the name suggests, this solver allows for the solution of the classical advection-diffusion equation which is suitable to describe the transport of a scalar through a porous media. This initial version did not include any mechanical dispersion model that results in transport being driven by advection and isotropic molecular diffusion. The implementation of an adaptive time step together with a condition on the permitted concentration threshold on the outlet boundary as additional features of `scalarTransportFoam` yielded the `adaptiveScalarTransportFoam` solver. A further extension of the `adaptiveScalarTransportFoam` which foresees the aggregation of a mechanical dispersion model as an additional term to the diffusion eventually fully represents the model described in equation 2.51. At the current stage, the last version of the transport solver is represented by the `rhoDarcyFoam` utility which, among other features, allows the simulation of transport in compressible fluids. A comprehensive description of its potentiality can be found in [45] where the code along with some benchmark examples are presented in detail. The core of the paper [45] is also reported in the appendix **??** of the present work.

Chapter 3

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

This chapter was published as a paper on the GEM-International Journal of Geomathematics [26].

3.1 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`®[46], for i) generating random pluri-Gaussian porous media, ii) solving the steady state Darcy-scale flow, iii) solving the advection diffusion equation, iv) 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.

3.2 Introduction

Subsurface flow and solute transport modelling is used in several engineering and environmental fields (CO_2 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 [47, 48, 49], and the definition of appropriate models and their parameterization remains an open field of research [50, 51, 52, 53]. An important challenge is how to simulate non-Fickian behaviour, which originates mainly from physical heterogeneities emerging across multiple scales [5, 54, 1]. 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 [55].

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 [56, 57]. 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 [58], 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 [59, 60, 61]. 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 [61] and their spatial organisation. For example this latter point was recently addressed in [53], 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 [62], a feature that also influences reactive transport settings [63, 64].

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 [65]. Solute transport has been widely investigated in continuous Gaussian and non-Gaussian permeability fields [66, 67], and methods have also been proposed to handle non-continuous fields, suitable to reproduce geomaterials where property transition is marked by sharp interfaces [68]. 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 [69, 52]. Facies properties can be qualitatively linked to non-Fickian parameters for alluvial aquifers [70], 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 [71], 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 de-

viation 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`®[46], as part of the SECUReFoam library [6]. 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 [62, 72, 5]. 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 [34]. Several approaches are available to reproduce complex subsurface structures (sequential Gaussian simulations [73], Markov chain probability [74], Multiple-point statistics [75]) as well as a number of geostatistical open toolboxes (GSLib [36], T-PROGS [76]). Nevertheless, few open-source tools exist that provide integrated geostatistical, flow and transport simulation solvers (OpenGeoSys [77], porousMultiphaseFoam [78], DuMux [79]). This work is structured as follows: in Section 2 we give the mathematical overview of the problem, in Section 3 we describe the testcases and summarise the numerical methodologies. Numerical results and the post-processing are presented in Section 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.

3.3 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.

3.3.1 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) [65]. 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 [31]);
- high geological realism since the truncation rule allows simulating observed geometrical relations between geological facies [8, 80, 31].

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 [38, 34], a discrete-in-frequencies continuous-in-space representation of a complex GRF is therefore given by:

$$Z(\boldsymbol{x}) = \sum_{j=0}^{N_f} \cos(2\pi \boldsymbol{a}_j \cdot \boldsymbol{x}) \sqrt{S(\boldsymbol{a}_j)} W_j + i \sum_{j=0}^{N_f} \sin(2\pi \boldsymbol{a}_j \cdot \boldsymbol{x}) \sqrt{S(\boldsymbol{a}_j)} W'_j \quad (3.1)$$

where $\boldsymbol{x}$ is the position vector, $\boldsymbol{a}_j = (a_{x,j}, a_{y,j}, a_{z,j})$ is the j^{th} frequency vector, W_j and W'_j are independent complex Gaussian random variable and $S(\boldsymbol{a}_j)$ 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 $\gamma(\boldsymbol{r})$ between a pair of values of a random variable located at points separated by the distance $\boldsymbol{r}$. We denote the correlation function as $\rho(\boldsymbol{r})$ and the variance as σ^2 (where $\gamma(\boldsymbol{r}) = \sigma^2 \rho(\boldsymbol{r})$).

In this work, we assume an exponential correlation function

$$\rho(\boldsymbol{r}) = 1 - e^{-\sqrt{\frac{r_x^2}{\lambda_x^2}+\frac{r_y^2}{\lambda_y^2}+\frac{r_z^2}{\lambda_z^2}}} \tag{3.2}$$

with corresponding spectrum

$$S(\boldsymbol{a}) = \sigma^2 \|\boldsymbol{\lambda}\|^d \frac{\Gamma\left(\frac{d+1}{2}\right)}{\left(\pi\left(1 + a_x^2\lambda_x^2 + a_y^2\lambda_y^2 + a_z^2\lambda_z^2\right)\right)^{\frac{d+1}{2}}}, \tag{3.3}$$

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. 3.1), which bins continuous values into a set of categories. The smooth transition which characterises the GRF is

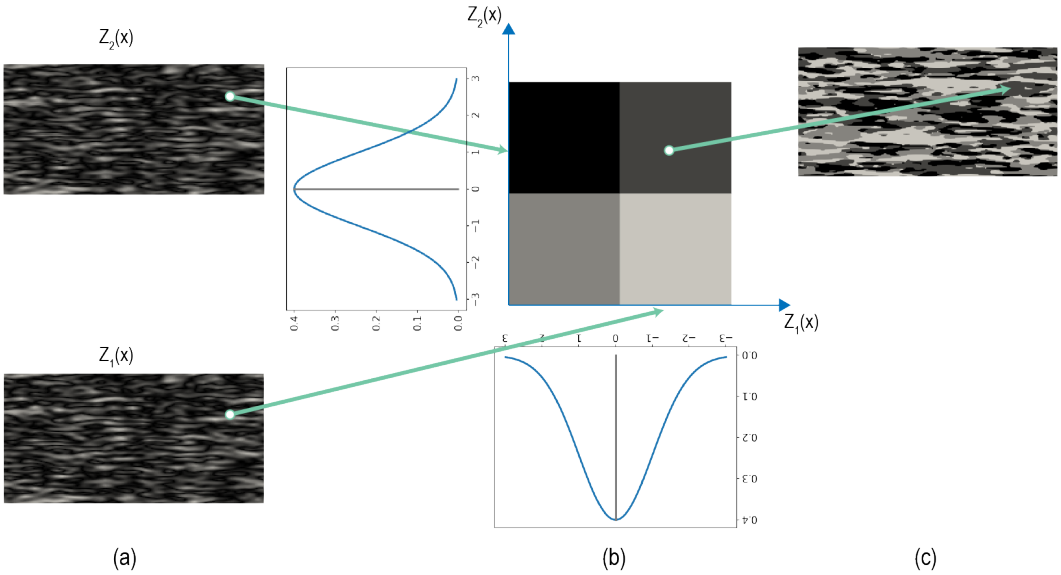

Figure 3.1: Truncated pluri-Gaussian simulation. a) Continuous multivariate Gaussian random fields Z_1 and Z_2 generation; b) truncation rule for four facies domain and its corresponding thresholds on the Gaussian distribution of the variables; c) sample of a two-dimensional truncated pluri-Gaussian random field. The arrows indicate the contribution of the two GRFs in assigning a given category at a selected location in space.

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 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

$$p_{\varphi_i}(\boldsymbol{x}) = [G(r_i) - G(r_{i-1})]\,[G(s_i) - G(s_{i-1})] \quad i = 1 \dots n \tag{3.4}$$

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 [29, 30] 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 [31]. In this work, we assume the single truncation diagram, in Fig. 3.1. In our simulations we vary the correlation lengths $\boldsymbol{\lambda}$ 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.

3.3.2 The flow model

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

$$\boldsymbol{V} = -\frac{\boldsymbol{k}}{\mu}(\nabla p + \rho g \nabla z), \tag{3.5}$$

where $\boldsymbol{V}$ is the Darcy velocity vector $[LT^{-1}]$, $\boldsymbol{k}$ 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`®[46] is based on equation (3.5) assuming an incompressible fluid. Therefore pressure can be computed according to a Poisson equation

$$\nabla \cdot \boldsymbol{V} = -\nabla \cdot \frac{\boldsymbol{k}}{\mu} \nabla p = 0 \tag{3.6}$$

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., $\boldsymbol{k} = k\mathbb{I}$, $\mathbb{I}$ 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.

3.3.3 Local transport model

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

$$\boldsymbol{J}_{adv} = \boldsymbol{V}c. \tag{3.7}$$

In line with previous work [62], we neglect mechanical dispersion and model the diffusive fluxes $\boldsymbol{J}_{mol}$ $[LT^{-1}]$ as

$$\boldsymbol{J}_{mol} = -\phi \boldsymbol{D}_{mol} \nabla c \tag{3.8}$$

where $\boldsymbol{D}_{mol}$ $[L^2T^{-1}]$ 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

$$\frac{\partial c}{\partial t} + \nabla \cdot (\boldsymbol{v}c) - \boldsymbol{D}_{mol} \nabla^2 c = 0 \tag{3.9}$$

where $\boldsymbol{v} \equiv \boldsymbol{V}/\phi$ is the fluid velocity, i.e. the velocity that would be measured by a flow meter in the porous domain and $\boldsymbol{D}_{mol} = D\mathbb{I}$. We impose a constant concentration on the whole inlet face of the domain (3.9) 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.

3.3.4 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:

$$\frac{\partial C}{\partial t} + \overline{v_x}\frac{\partial C}{\partial x} - D_{xx}^{mac}\frac{\partial^2 C}{\partial x^2} = 0, \tag{3.10}$$

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 (3.13) while inferred macrodispersion assessments can be performed using the moments' method (3.20) or applying the least square method to the breakthrough curve, as illustrated in section 3.3.5.

3.3.5 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 $\overline{c}$.

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

Péclet number

$$Pe_x\ [-]\ = \frac{\overline{v_x}\lambda_x}{D_{mol}}; \tag{3.11}$$

effective permeability

$$k_x^{eff}\ [m^2]\ = -\frac{\overline{v_x}\mu}{\frac{\partial p}{\partial x} - \rho g}; \tag{3.12}$$

nominal macrodispersion

$$D_{mac}^{ij}\ [m^2/s]\ = \phi \boldsymbol{\lambda}^T \boldsymbol{V} \tag{3.13}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{V}$ are typical lengths and velocity vectors. Equation (3.12) quantifies the effective permeability based on the average field velocity and the pressure gradient. In case of a two dimensional Gaussian permeability field, analytical solutions which allow the effective permeability to be computed only on the permeability field basis already exists in literature [15, 81, 82, 1]. How these analytical expression compares with the computed effective permeability in non-Gaussian permeability field remains an open question. Equation (3.13) allows the macrodispersion matrix to be approximated a priori starting from geostatistical (correlation length $\boldsymbol{\lambda}$) and flow (velocity $\boldsymbol{V}$) 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 subsection 3.3.5.

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. (3.10) 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

$$F(T;\mu_1,\nu) = \bar{c} = \Phi\left(\sqrt{\frac{\nu}{T}}\left(\frac{T}{\mu_1} - 1\right)\right) + \mathrm{e}^{\frac{2\nu}{\mu_1}}\Phi\left(-\sqrt{\frac{\nu}{T}}\left(\frac{T}{\mu_1} + 1\right)\right) \tag{3.14}$$

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 (3.14) and corresponds to the PDF of the solute arrival times. This PDF is expressed as [83]

$$f(T;\mu_1,\nu) = \frac{\partial \bar{c}}{\partial T} = \sqrt{\frac{\nu}{2\pi T^3} exp\left[-\frac{\nu(T-\mu_1)^2}{2\mu_1^2 T}\right]}. \tag{3.15}$$

Other analytical solutions are available for different boundary conditions on finite domains [84]. 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 [54] if

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

In the assumption of a Fickian model such as (3.15) arrival times display a sharp and exponential tail as $t \to \infty$. 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 [54, 62].

The moments' method Following [85, 86], the estimation of the statistical moments of the cumulative Inverse Gaussian is performed by approximating its parameters μ_1 and ν as

$$E[\bar{c}] = \mu_1 \tag{3.16}$$

$$Var[\bar{c}] = \mu_2 - \mu_1^2 = \frac{\mu_1^3}{\nu}. \tag{3.17}$$

To compute the first and second order moments we used

$$\begin{aligned}\mu_1 &= \int_0^{+\infty} fTdT = \int_0^{+\infty} F'TdT = -\int_0^{+\infty} FdT + [FT]_0^{+\infty} \\ &= -\sum_{i=0}^{+\infty} F_i \Delta T + F_{+\infty} T_{+\infty}, \end{aligned} \tag{3.18}$$

$$\begin{aligned}\mu_2 &= \int_0^{+\infty} fT^2dT = \int_0^{+\infty} F'T^2dT = -2\int_0^{+\infty} FTdT + \left[FT^2\right]_0^{+\infty} \\ &= -2\sum_{i=0}^{+\infty} F_i T_i \Delta T + F_{+\infty} T_{+\infty}^2. \end{aligned} \tag{3.19}$$

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

$$V_x = \frac{L_x}{\mu_1} \tag{3.20}$$

$$D_{xx}^{mac} = \frac{\mu_2 V^3}{2L_x} \tag{3.21}$$

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 Gaussian approximation, a normalised error e was defined as

$$e = \frac{||\bar{c}(T) - F(T)||}{||\bar{c}(T)||} \cdot 100. \tag{3.22}$$

Least squares estimation Parameter estimation is performed by minimising the least squared error between numerical data and the models (3.14)-(3.15). 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 equation (3.15) while for the second case the analytical function is equation (3.14). 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. This method was applied to the low and high contrast scenarios described in section 3.3.10 and the corresponding parameter values are reported in tables 3.6 and 3.7.

3.3.6 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 $2 \cdot 10^6$.

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 equation (3.2), the number of permeability zones, as well as the volumetric proportion for each of the facies (see table 3.1). 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 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} m^2 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.

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

Table 3.1: Parameters kept constant throughout the simulations.

The simulation workflow is divided into three steps

- geostatistics: the permeability domain is generated using the truncated pluri-Gaussian algorithm;
- flow: equation 3.5 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 [6] 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 subsection-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.

3.3.7 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 (section 3.3.9). Finally, we analyse the impact of three physical parameters on arrival time PDFs, namely permeability contrast (section

3.3.10), the longitudinal correlation length used to generate the fields (section 3.3.11) and Péclet number (section 3.3.12). By testing transport behaviour in settings with increasing longitudinal correlation lengths and permeability contrast between facies, we aim to:

- assess the sensitivity of transport behaviour and the onset of non-Fickian transport with respect to the correlation length and permeability constast of the generated permeability fields;
- quantify the discrepancy from Fickian behaviour as a function of both correlation lengths and permeability contrast, considered separately.

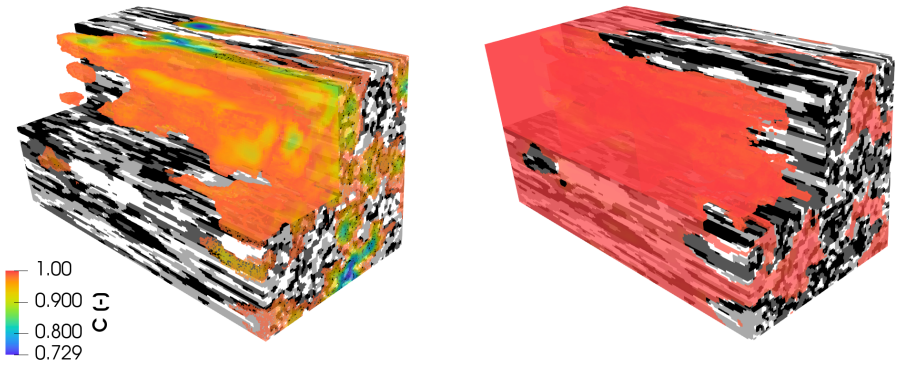

Figure 3.2: Solute plume distribution in high contrast permeability domain at late time. Domain sizes are 2x1x1 and correlation lengths along the three directions are set to (0.8, 0.1, 0.1). On the left panel it is possible to observe how low solute concentration values (0 blue - 0.99 red) are confined to low permeability regions (10^{-12} $[m^2]$ dark grey - 10^{-13} $[m^2]$ 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} $[m^2]$ dark grey - 10^{-11} $[m^2]$ black).

Before moving to these detailed analyses, figure 3.2 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 3.2 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.

3.3.8 Velocity PDFs

The velocity PDF has a direct influence on non-Fickian transport features [87] as large differences between permeability values may induce bi- or multi-modal velocity distributions [53]. The PDFs of point velocity values with increasing longitudinal correlation lengths λ_x are shown in figure 3.3. The longitudinal velocity V_x normalised with the average longitudinal velocity is reported on the horizontal axis of figures 3.3 and 3.4 as V_x^*. The vertical axis of figure 3.3 reports the probability density distribution as a function of the longitudinal velocity $p(V_x^*)$. As expected, the velocity distributions in figure 3.3 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 (figure 3.4), similar to what is shown for a bimodal permeability distribution by [53]. 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 figure 3.3 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.

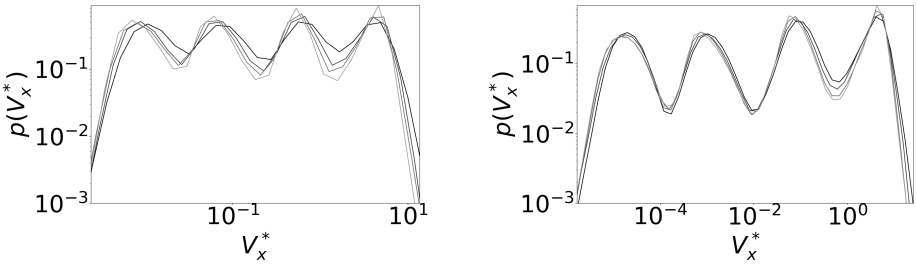

Figure 3.3: PDFs of the longitudinal velocity component for low and high contrast. Velocity PDFs are shown for low (left) and high (right) permeability contrast. The correlation lengths λ_x span between 0.4 (darker lines) and 1.0 (lighter lines).

Figure 3.4 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. 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.

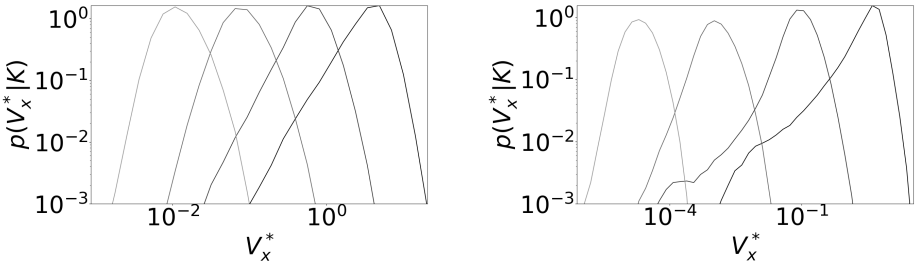

Figure 3.4: Conditional PDFs of the longitudinal velocity values in low (left) and high (right) permeability contrast. Results are shown for correlation lengths $\lambda_x = 0.4$. The curves are shown with different colours depending on the facies permeability, i.e., lighter colours correspond to low permeability and darker colours to high permeability media.

3.3.9 Variability of transport behaviour across multiple realizations

Figure 3.5 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' heterogenities as the solute fill the whole domain. Local injections in high/low conductivity have been considered in previous works [71] 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) per-

meability 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 Figure 3.5). 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.

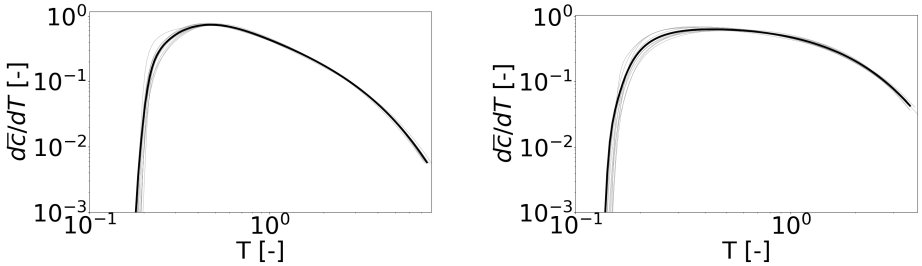

Figure 3.5: PDFs of solute arrival times associated with 10 realisations with the same correlation length ($\lambda_x = 0.8m$). Low and high permeability contrast on the left and right side of the panel respectively.

3.3.10 Effect of permeability contrast

Figure 3.6 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 figure 3.6 are shown in table 3.2.

The simulated BTCs (i.e., CDFs of the solute arrival times) are shown in the top left panel of figure 3.6 while on top right panel of figure 3.6 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 figure 3.6 are computed applying the moment's method and the least square method illustrated in 3.3.5 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 figure 3.6, bottom left and the first error column in table 3.3). 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 equation 3.14 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 [84]. 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 (figure 3.6, bottom right), regardless of the method used to estimate its parameters (Least Square or Moments method).

In summary, figure 3.6 suggests that as the permeability contrast increases, the evolution of the solute concentration shows significant departure from the Fickian model.

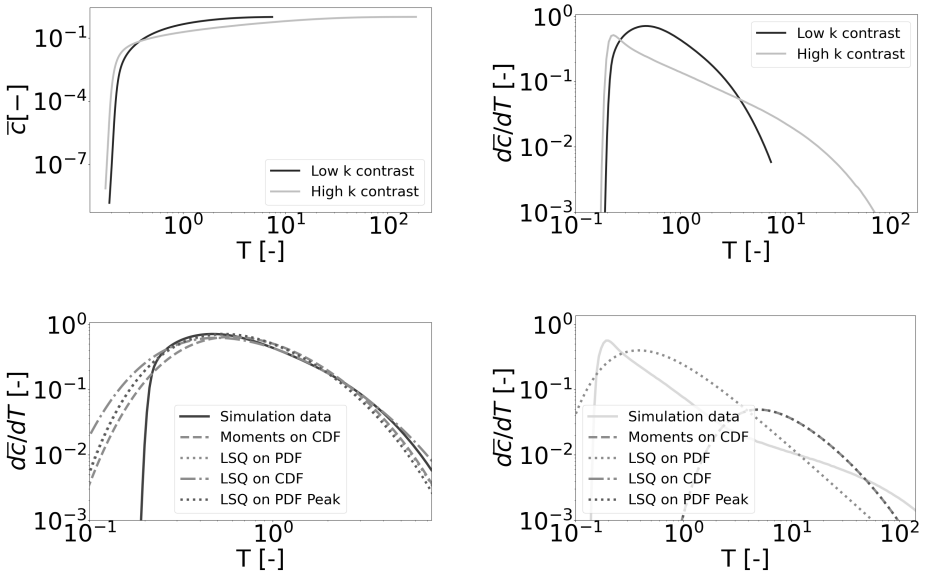

Figure 3.6: Top panels: breakthrough curves (left) and their time derivatives (right) simulated on identical geological structure with low and high permeability contrast between the facies. Bottom panels: the curves are overlapped with the corresponding Inverse Gaussian approximations via least square (LSQ) or moments' method estimation in low (left) and high (right) permeability contrast case.

A detailed analysis of these results is shown in Table 3.3. Our results suggest that the level of accuracy of macrodispersion models in capturing transport behaviour decreases with the permeability contrast. Table 3.3 reports the

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

Table 3.2: Geostatistical and flow parameters for the low and high permeability contrast fields used to generate results reported in Figure 3.6. The flow parameters were defined by equations (3.11) and (3.12). Considering that the permeability interval in the high contrast permeability case ranges between 10^{-9} and 10^{-15}, the variance of the natural logarithm of the permeability fields where the four sediment categories are equally probable can reach values up to 26.

values of macrodispersion parameters computed through approximation (3.13) (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) thus the estimated values cannot be considered as reliable.

3.3.11 Effect of spatial correlation

Transport simulations are performed on PGS domains sharing comparable geostatistical parameters (table 3.4) while increasing longitudinal correlation lengths (figure 3.7). These provide interesting insights into the transition from Fickian to anomalous transport in relation to the connectivity degree of the sediment structure (figure 3.8). 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 the conductivity fields (see Figure 3.1). This length can be interpreted as the characteristic length over which facies' transitions are observed.

Figure 3.8 displays the PDFs of the solute arrival times obtained for a range of values assigned to λ_x. As the longitudinal correlation length increases,

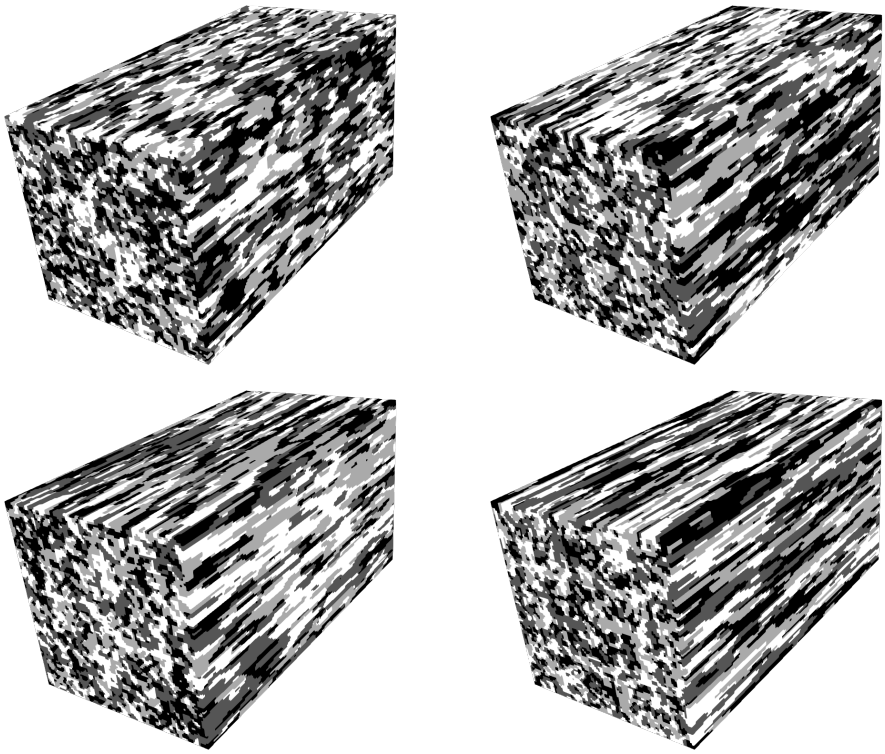

Figure 3.7: Truncated pluri-Gaussian permeability fields with increasing longitudinal correlation length λ_x. Clockwise order from top left panel: $\lambda_x = 0.4$, $\lambda_x = 0.6$, $\lambda_x = 0.8$, $\lambda_x = 1.0$.

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

Table 3.3: 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. The *reference solution* values represent the average longitudinal velocity and the nominal macrodispersion as from equation (3.13). *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 ($\sigma > 10^6$).

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 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 figure 3.8) 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.

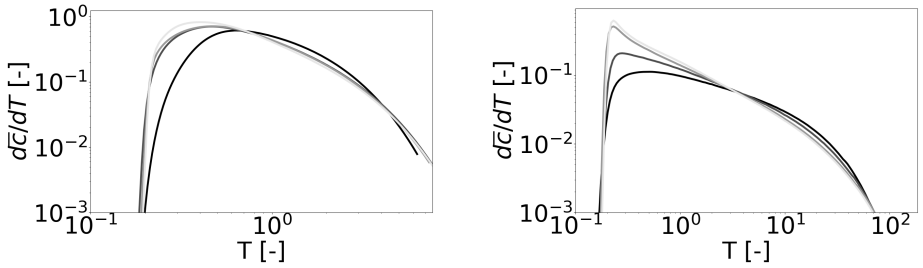

Figure 3.8: Arrival time PDFs computed as $d\bar{c}/dT$ for low (left panel) and high (right panel) permeability contrast for as a function of the longitudinal correlation length λ_x. From the darkest to the lightest curve the longitudinal correlation length $\lambda_x = 1$ increases with evenly spaced interval from 0.4 to 1.

Tables 3.4 and 3.5 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 3.6 and 3.7. As a result of the emergence of preferential flow-paths, the effective permeability shows a positive trend for increasing correlation lengths.

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

Table 3.4: Geostatistical and flow parameters for low permeability contrast simulations. The flow parameters are defined by equations (3.11) and (3.12).

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

Table 3.5: Geostatistical and flow parameters' for high permeability contrast simulations. The flow parameters were defined by equations (3.11) and (3.12) and are computed in the longitudinal direction of the flow. Longitudinal correlation length increases from 0.4 to 1.0.

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 figure 3.9). 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.

Low contrast

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

Table 3.6: 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. *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.

High contrast

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

Table 3.7: 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. *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.

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

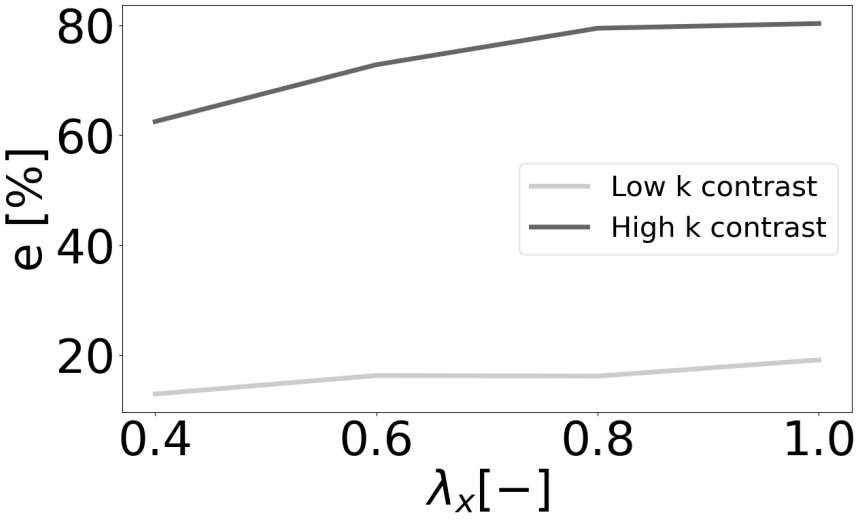

Figure 3.9: Longitudinal correlation length vs relative error computed as the departure from Fickian model (eq. 3.22) with parameters estimated with the moments method (Method 1).

3.3.12 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 section 3.3.10) or correlation lengths (section 3.3.11) 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, from 10^{-9} m^2/s to 10^{-11} m^2/s. This corresponds 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 figure 3.10. This result is in agreement with [71] 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 figure 3.10 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 figure 3.10, 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 section 3.3.11: 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.

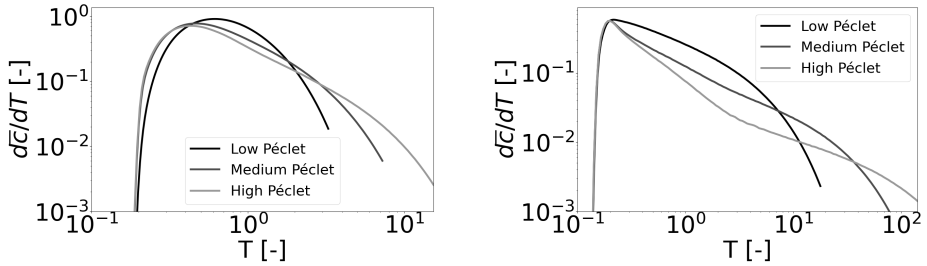

Figure 3.10: Arrival times PDFs in low (left) and high (right) permeability contrast domains sharing the same geological structure ($\lambda_x = 0.8$m) while characterised by a different Péclet. In case of low permeability contrast, the Péclet number ranges between $8 \cdot 10^2$ and $8 \cdot 10^4$ while for high permeability contrast the Péclet ranges between $6 \cdot 10^3$ and $6 \cdot 10^5$. Significant Péclet variation for these simulations was obtained by changing the molecular diffusion coefficient.

Figure 3.11 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 resuting from the moments' method (M.1 in tables 3.6 and 3.7).

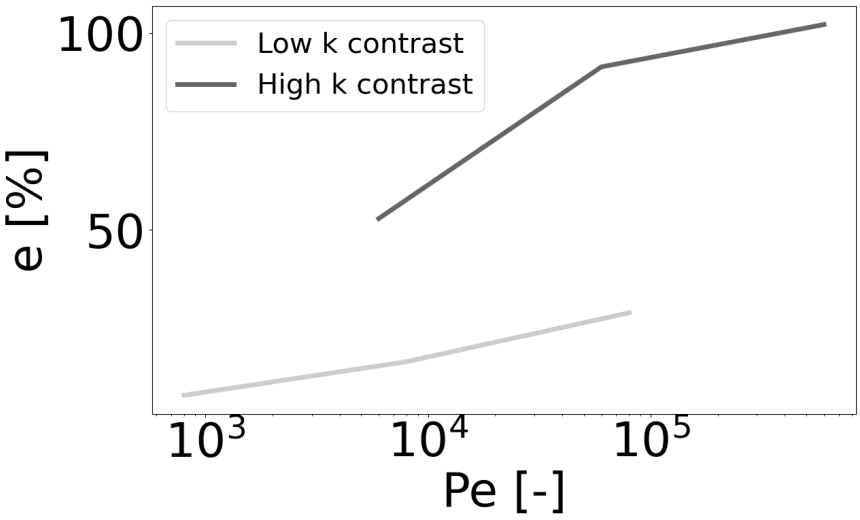

Figure 3.11: Péclet number vs relative error computed as in eq. (3.22). The interplay between increasing Péclet numbers, permeability contrast and relative error is qualitatively similar to the one exhibited by the increasing longitudinal correlation length in fig. 3.8.

3.4 Conclusions

In this work we have explored the impact of permeability contrast and, similarly to [70, 71, 53], correlation length and Péclet 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 [88] and Generalised Multi-Rate Transfer equations [89].

Chapter 4

Geostatistical and flow modelling case study: the Herten site

The study in the previous chapter is based on idealised fully stationary random fields. With the purpose of extending some of the conclusions to real applications, in this chapter we consider a real field site, namely the Hertn site [90, 91]. The switch from idealised fully stationary random fields to a real field site required a software that could handle conditional geostatistical simulations. For this reason, our geostatistical utility [6] adopted in the previous chapter is replaced by a more advanced commercial software that performs PGS simulations with conditioning data.

The objectives of this chapter are twofold: to recount the steps followed to reconstruct a 3D geological domain conditional on boreholes data (section 4.2) using the PGS plug-in of a commercial software (Petrel), and to perform flow simulations on a real geological domain 4.3.

The first objective aims to provide a universal procedure for applying the PGS algorithm to a realistic case study. This procedure was elaborated and tested during a two-month internship at Ephesia, a Swiss-French company that developed a PGS plug-in for a well-known commercial geological modelling software produced by Schlumberger (Petrel). The idea behind it is to virtually reproduce a geological domain applying a PGS algorithm to hard data that mimic the ones usually available on field (boreholes) and to compare them with a more complete reconstruction obtained with a different algorithm [92] applied to cross sectional data [90], seldom available for large field sites. The second objective is to compute the flow field on this domain and show the effects of the geostatistical approximation on the flow. A more detailed anal-

ysis of the transport is instead the focus of the next chapter. We test these PGS algorithms on a geological domain representing a high resolution aquifer consisting of an outcrop of a late Pleistocene unconsolidated fluvio-glacial and fluvial-braided river sediments in the Rhine basin [90]. The 3-D domain was reconstructed from a set of detailed 2D cross sectional data as illustrated in [91].

Having access to geological data surveyed with high resolution, it is possible to benchmark the PGS geostatistical algorithm and its effect on flow and transport against it. [ADD The second part of this chapter looks at the flow field in the Herten reconstructed using these geostatical algorithms, while a more detailed study of the transport is the focus of the next chapter]

4.1 Site description

An analysis of several open geological data sets (Norne [93], Sleipner [94], Herten and Descalvado [95], Volve [96], etc..), indicated a 1120 m^3 depositional formation in the upper Rhine valley of southern Germany near the town of Herten, as the most suitable domain for our study purposes. The 16x10x7 metres quarry was excavated and surveyed in 1999 and based on our knowledge, Herten and Descalvado are the only sites where geological properties of 2D cross sections were directly measured. This characteristic, together with the cubical spatial shape and the number of geological properties surveyed at high spatial resolution were our drivers in the choice for test case domain. The data analog, consisting of hydrogeological properties measured at six parallel 2D sections, were released in 2006 along with a detailed sedimentological interpretation [90] and its three-dimensional reconstruction [91, 95, 27]. According to the hydro-facies classification suggested by [90], the sediments can be divided into ten categories whose spatial locations are indicated in figure 4.1. The fluvio-glacial river sediments show a high degree of heterogeneity with permeability values that span approximately six orders of magnitude [90].

The three-dimensional permeability field of figure 4.1 was obtained by [92] applying the sequential 2-D simulation with conditioning data [92] to the six longitudinal cross-sections excavated from the Herten field site 4.2. In more details, the three-dimensional reconstruction of the geological domain that will be taken as reference for performing flow and transport simulations in the next sections, is the combined result of direct observations and geostatistical simulation [92]: after multiple cross sections were mapped (figure 4.3) and digitized during the excavation process, the 2D analogs were used as training images to perform multiple point statistics (MPS) simulations between cross

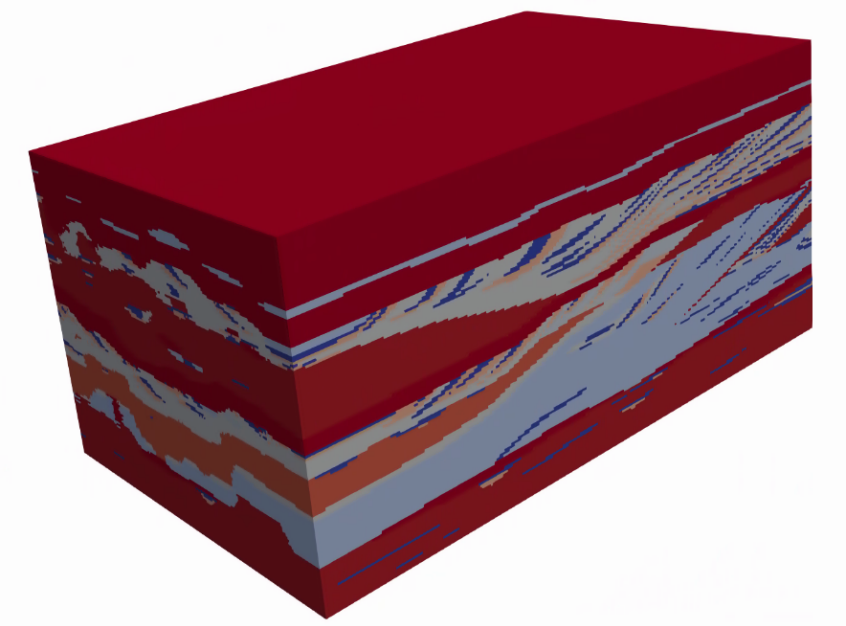

Figure 4.1: The ten different colours represent the spatial distribution of the ten facies which characterise Herten. It shows the results of the application of *sequential 2-D simulation with conditioning data* algorithm [92] to the six cross sections of hard data collected at Herten by the authors of [95]. The algorithm was applied 100 times to the same 2D cross sections and the results published at [27]. This image was obtained by the authors selecting at each cell the sediment category that appears with the highest frequency over the ensemble of 100 realizations.

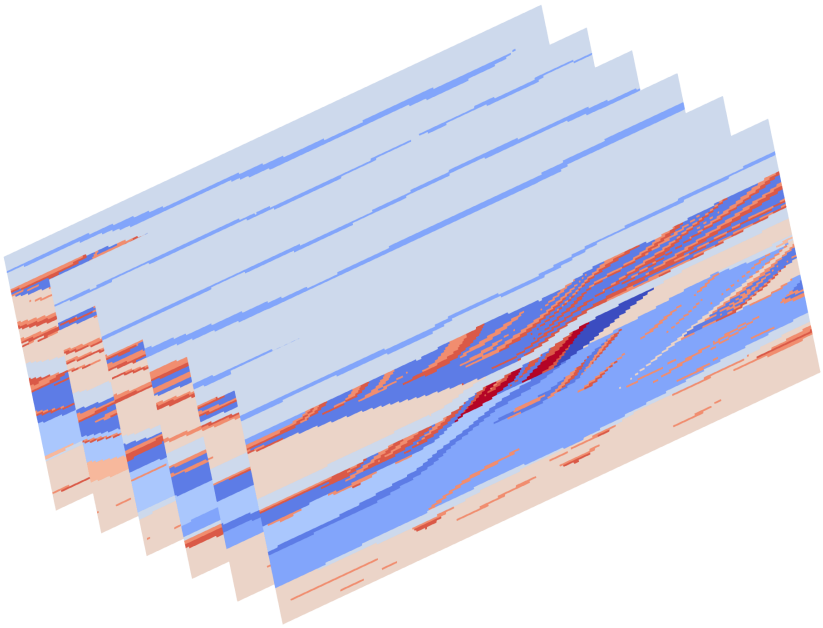

Figure 4.2: The digital reconstruction of the original six longitudinal cross sections surveyed at Herten field site [90] map the sediment categories in space. Except for the colour scheme which changes from a visualisation software to another (e.g. Paraview and Petrel), the external sections show a perfect matching with the the boundaries of the 3D reconstruction of Herten in figure 4.1.

sections. Usually, the number of dimensions between input and output data in

Figure 4.3: One of the six high resolution photo of Herten cross sections surveyed by [95]. Different colours indicate the presence of sediments with different geological properties for example permeability of porosity. The reconstruction of the permeability field in figure 4.1 show a good geometrical affinity with the shape of in this photo.

classical MPS algorithms has to be consistent. However, thanks to a custom method called *sequential 2D simulation with conditioning data* [92], this hypothesis was relaxed and an ensemble of 100 three-dimensional realisations of the simulated domain are provided at [27] alongside the 2D surveyed analogs. Out of this ensemble, in this study a new map is adopted by selecting at each cell the most probable sediment category that is, the one which occurs with the highest frequency over the ensemble of 100 realizations. The conditional PGS simulation illustrated in section 4.2 was performed on a subset of the original 2D cross sectional data. The idea behind it, consists of conditioning a simulation on a hard data set arranged in a format more commonly available at the field scale than 2D cross sectional data. In this sense, the hard data adopted in section 4.2 and plotted in figure 4.4, represent boreholes or core samples data.

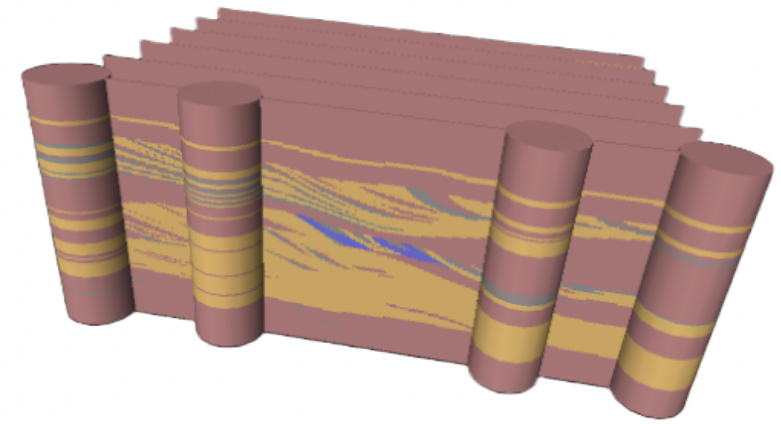

Figure 4.4: 3D visualisation of the 2D cross sections surveyed at Herten field site and described in [90]. The cylindrical shapes mimic the geometry of boreholes which is the typical geometry of the data available at field scale. These vertical data are used as boreholes to perform conditional simulations in section 4.2. The slightly different patterns from figure 4.2 are the results of lumping ten sediment categories into four classes, each characterised by sediments with similar permeability.

4.2 Conditioned PGS application

The procedure to perform PGS simulations starting from boreholes data offers an original and universal framework to develop PGS realisations conditional on real data. In this context we adopt the geostatistical terminology where the meaning of simulation is to reconstruct the permeability field and not solving partial differential equations. The main goal of the pre-processing stage is to determine the truncation or lithotype rule. This can be achieved in several ways, i.e. analogy with sites having the same geological setting, direct observation of outcrops or analysis of cores and well-logs. In this work, since a larger amount of information is available, we decided to draw it based on transition probability and bivariate transition probability matrices outcome. These matrices, described more in details in the next paragraph, are respectively indicators of conditional and joint probability of random variables. In this case, the random variable is represented by the sediment category values that can be taken by each cell in the spatial domain. The hard data for the conditional modelling stage were extracted from the reference data set (figure 4.1) as if they were the outcome of a survey campaign where several equally spaced boreholes were drilled (some of them are visible in figure 4.4).

The procedure can be subdivided into three main stages, each formed of minor successive steps:

- pre-processing (Matlab)

    1. best proportion trend map: from the 100 equally probable Herten reconstruction available at [27], a new three-dimensional map is adopted by selecting at each cell the sediment category that appears with the highest frequency and, based on the similarity in permeability values, the original ten sediment categories individuated by [27] are lumped into four permeability homogeneous categories;
    2. transition probability matrix and bivariate transition probability matrix: the probability of finding sediment categories A and B at the two sides of a cell border and the probability of a cell being associated with sediment B given the neighbouring cell of sediment A, respectively;
    3. truncation or lithotype rule that represents the contact pattern of the sediments, as illustrated in 2.1.4;

- simulation (Petrel and PGS plug in)

4. hard data (core samples, boreholes, wells) preparation to be imported into Petrel;
5. proportion trend map for each sediment category: Kriging interpolation of hard data provides the spatial distribution of each sediment category;
6. dip angle map: the observed sediments' inclination is one of the simulation's inputs that allows to better reproduce the target geological structure;
7. conditional simulation: hard data, Kriging and dip angle are used to reconstruct the geological data set with an elevate realism degree;

- post-processing (`OpenFOAM`®)

8. polyMesh grid generation: a grid with the same sizes and the same resolution of the available permeability field is created;
9. permeability field import: manipulation of the permeability field to make them readable from `OpenFOAM`®;
10. flow and transport simulation: Darcy and advection-dispersion equations are solved by means of `OpenFOAM`®libraries over the conditional permeability field created in the previous steps. The results are illustrated in section 5.

The pre-processing for this study starts from the sequential 2-D simulations performed by [92] with six surveyed cross sections as conditioning data. A hundred equally probable realizations of the excavated geological field site were published at [27] by the same authors who surveyed the six cross sections. Our *best proportion trend map* (figure 4.1) maps the cells' values according to the sediment category mode that is, the highest occurrence frequency throughout the whole ensemble of simulations. Given the similarity in permeability values, the original ten sediment categories individuated by [27] have been lumped into four permeability homogeneous categories (1). This also facilitates the evaluation of transition probability and bivariate transition probability matrices for the sketching of the lithotype rule. The permeability values of these four sediments classes range from $1.84 \cdot 10^{-9}$ m^2 to $3.30 \cdot 10^{-15}$ m^2 and they are obtained as weighted averages of the permeability values of the facies and their volumetric fractions. Taking the best proportion trend map as reference, its

sediment transition probability matrix 4.1

$$\begin{bmatrix} 9.65 \cdot 10^{-1} & 1.52 \cdot 10^{-3} & 3.10 \cdot 10^{-2} & 2.82 \cdot 10^{-3} \\ 1.04 \cdot 10^{-1} & 4.31 \cdot 10^{-1} & 3.29 \cdot 10^{-1} & 1.36 \cdot 10^{-1} \\ 4.92 \cdot 10^{-2} & 7.60 \cdot 10^{-3} & 9.37 \cdot 10^{-1} & 6.44 \cdot 10^{-3} \\ 1.92 \cdot 10^{-1} & 1.35 \cdot 10^{-1} & 2.77 \cdot 10^{-1} & 3.97 \cdot 10^{-1} \end{bmatrix} \tag{4.1}$$

and the bivariate transition probability matrix 4.2 (2)

$$\begin{bmatrix} 5.82 \cdot 10^{-1} & 9.14 \cdot 10^{-4} & 1.87 \cdot 10^{-2} & 1.70 \cdot 10^{-3} \\ 9.14 \cdot 10^{-4} & 3.77 \cdot 10^{-3} & 2.88 \cdot 10^{-3} & 1.19 \cdot 10^{-3} \\ 1.87 \cdot 10^{-2} & 2.88 \cdot 10^{-3} & 3.55 \cdot 10^{-1} & 2.45 \cdot 10^{-3} \\ 1.70 \cdot 10^{-3} & 1.19 \cdot 10^{-3} & 2.45 \cdot 10^{-3} & 3.49 \cdot 10^{-3} \end{bmatrix} \tag{4.2}$$

are computed [29, 97]. As their names suggest, each element of the transition and the bivariate transition probability matrix describes a transition probability from a facies to another. Transition and bivariate transition probability matrices are squared matrix with the number of rows (or columns) equal to the number of one element possible transitions, i.e. the number of facies. The f_j elements on the i-th row of the transition probability matrix describe the conditional probability of finding the element j at the position $x+1$ given the element i at the location x. In other words, this can be the probability of a cell being associated with sediment B given the neighbouring cell of sediment A, that is $P(B \mid A)$. For the definition of probability, the sum along the rows of the transition probability matrix is one. It is also interesting to notice that the highest values for both matrices are on the diagonals which is in line with what one would expect as it is always likelier to find the same sediment while moving from one cell to the next. In this context, bivariate transition probability is defined as the probability of finding sediment categories A and B at the two sides of a cell that is, $P(A \cup B)$. The b_{ij} values within the bivariate transition matrix differ from the elements inside the transition probability matrices as b_{ij} describes the joint probability of finding the facies i and j at two neighbouring cells. For this reason the bivariate transition probability matrix is symmetrical. As it is possible to observe, none of these probabilities is zero which means that all the facies are in contact with all the facies. This essential information is conveyed by these matrices and allows to define the contact pattern between sediments accordingly, i.e. by drawing a truncation rule where the perimeter of each category is in contact with all other categories. To simulate a contact pattern where all the facies are in contact, the truncation rule in figure 4.5 was sketched and used as input for the PGS plug-in in Petrel. Whether the contact pattern could have been honoured with

a different combination of thresholds, an arbitrary choice needs to be taken as the impact of the thresholds on transport goes beyond the scope of the present work. As explained in section 2.1.3, the inclusion (4) and later Krig-

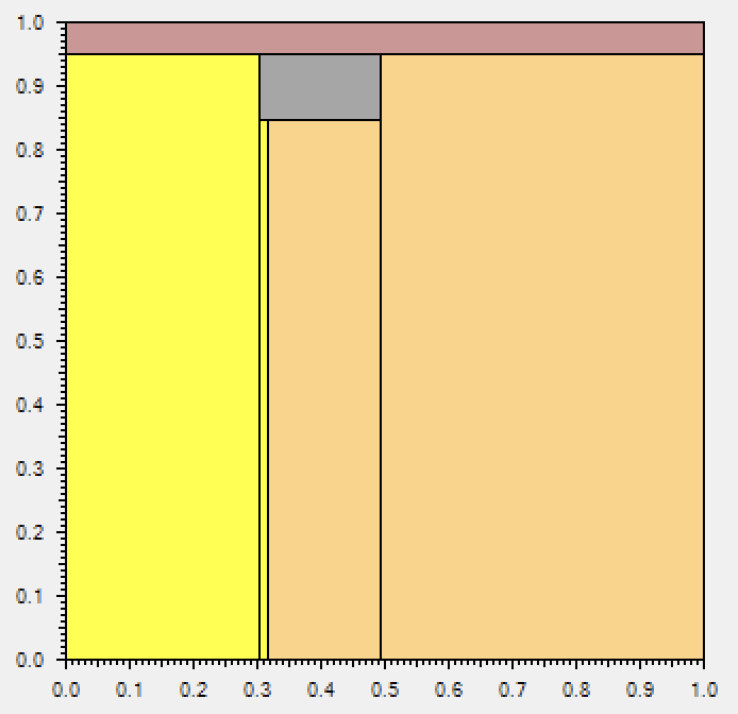

Figure 4.5: Herten lithotype rule characterised by the mutual contact between four homogeneous facies. The contact pattern is inferred by the transition and the bivariate transition matrices.

ing interpolation (5) of hard data (boreholes) extracted from the 2D sections would lead to smooth transition pattern between facies if no truncation rule was established. To increase the realism of the final geostatistical domain by matching the local inclination of the facies, a dip angle inclination map was imposed (6). Variogram or correlation function parameters required for the conditional simulation to be performed, were estimated on a trial-and-error basis. Additional requirements by Petrel to perform the PGS algorithm with conditioning data are the continuous distribution of each of the sediment over the domain, namely the proportion trend maps. These sediment maps can be achieved through the Kriging interpolation of the sediment categories that characterise the hard data found in the boreholes (7). The transport simulations illustrated in chapter 5 are performed in `OpenFOAM` on a regular grid (8) with the same dimensions as the one simulated in Petrel. This grid was

populated with permeability values by matching sediment categories and corresponding geological properties (9), as measured by [90]. Flow and transport simulations were then possible to be run (10) on the simulated permeability field (figure 4.6).

The increasing complexity of the simulation steps (1-7) reflects an increasing degree of realism. Notwithstanding the objectiveness of the geostatistical tools available, their application to a real data set like Herten requires some subjective choices that, although legitimate, need to be clearly stated. As a result of the trial and error process (6), these choices are reflected by the following parameters:

- facies correlation lengths in x, y and z direction are respectively set to 10, 2 and 0.1 meters. Although correlation lengths are essential parameters for variograms or correlation functions to be computed, a precise evaluation from field data is challenging while a good estimation could be reached by a visual analysis;

- the nugget value describes the short scale randomness implicit to spatial variables and it is defined as the intercept of the variogram at a distance that goes to zero. In theses cases it set to 0.0001 which means that short scale permeability randomness is quite small;

- Kriging search radius is the number of cells considered for the application of the Kriging algorithm and is set to 10.

These parameters are shared by the simulations in figure 4.6, which represent how the final reconstructions of the original data set with three different variograms look like (figure 4.1). As a posterior assessment of the geostatistical similarity, the transition and bivariate transition probability matrices could have been computed for the reconstructed permeability fields and compared with 4.1 and 4.2.

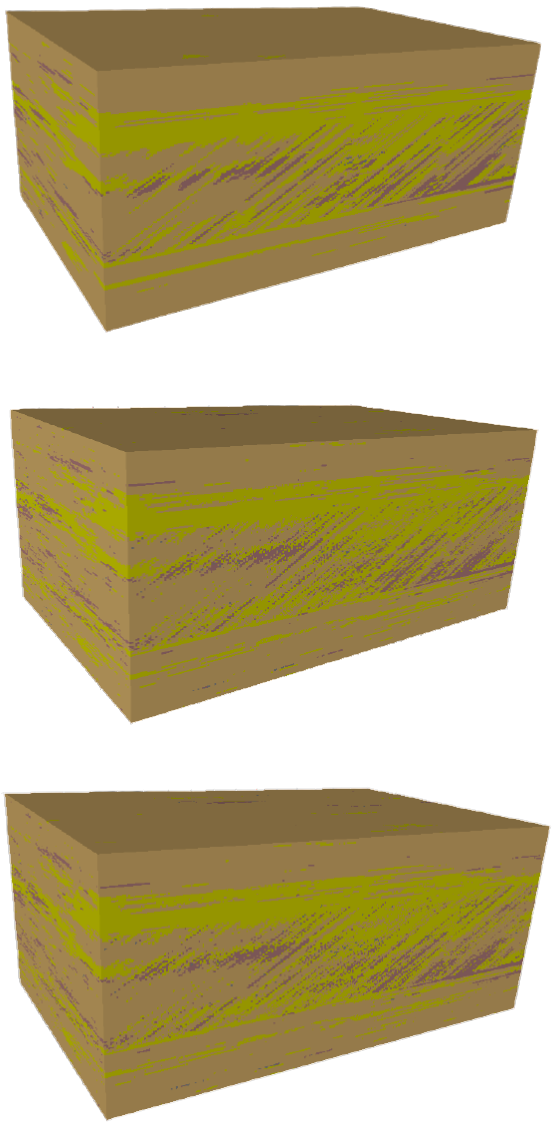

Figure 4.6: Herten PGS with (a) Gaussian (b) exponential and (c) spherical variogram. By means of the geostatistical techniques, which include the conditioning on boreholes, these simulations try to match the sediment structure of figure 4.1.

4.3 Spatial velocity analysis

In this section, a detailed analysis of the velocity field for a steady state flow simulation through the permeability field of Herten reconstructed by [91] is provided. To enable the comparison with the results from the transport analysis discussed in chapter 5 and obtained for the benchmark permeability field, the velocity analysis is performed on the benchmark permeability field instead of the reconstructed ones. It is worth to be noticed that while the reconstructed permeability field is characterised by just 4 facies, the benchmark permeability field shows 10 different permeability values. The PDF of the velocity field contains important information on the heterogeneity of velocities that are a key factor to non-Fickian transport as large differences between probability values indicate multi-modal velocity distribution which in turn might indicate the presence of preferential flow paths. The velocity field represents the solution of the Darcy's flow equation by means of the `OpenFOAM`®libraries available at [6]. The 16x10x7 meters dimensions of the domain are discretized using a 5 cm spatial resolution which makes the number of total cells equal to 8960000 equal. Boundary conditions for the pressure are zero gradient along the lateral sides while a constant gradient of 50 Pa/m along the longitudinal direction. To solve the advection-dispersion equation in the next section, the concentration gradient is zero everywhere except for the inlet boundary where the concentration is kept constant to one. While the number of cells is almost four times higher than the one used to discretise the domain in [26], pressure and concentration boundary conditions are maintained the same in both studies. To solve the Darcy's flow equation over the permeability domain, the same `OpenFOAM`®solver *simpleDarcyFoam* is adopted in chapter 3 and 4, while to solve the advection-dispersion equation two different versions of the modified *scalarTransportFoam* are adopted. In terms of computational time, although the larger number of cells makes the flow and transport simulations through Herten case study more expensive, the increase in computational time does not exactly scale with the number of cells and this is probably due to the improvements implemented in *macroScalarTransport*. Pressure and velocity fields for Herten permeability domain are visualised in figure 4.7. On the left panel, the pressure drop between inlet and outlet shows a smooth pattern compared to the velocity field values on the right. The sharp transitions in the velocity fields are in agreement with what expected from a high contrast permeability fields such as the one presented by Herten field site. The values of the longitudinal Darcy velocity normalised with the mean longitudinal velocity are reported on the horizontal axis of figures 4.8. The vertical axis on the left panel of figure 4.8 reports the probability density distribution as a function of the normalised

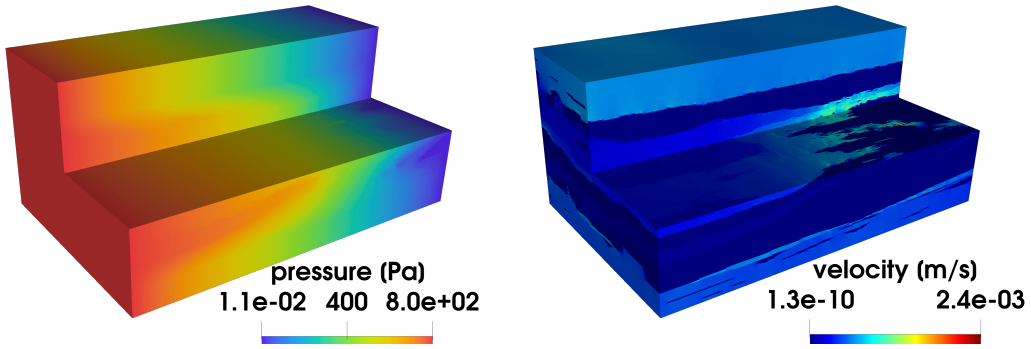

Figure 4.7: Clip of Herten pressure and velocity fields. Smooth transition pattern between permeability values on the left panel indicates a gradual drop in pressure while the discontinuities highlighted in the velocity field on the right panel are characteristics of high contrast permeability fields.

longitudinal Darcy velocity while on the vertical axis of the right panel of figure 4.8, the joint probability of $p(V_x^*, k = k_i)$ is reported. In other words, each of the curves on the right panel of figure 4.8 represents only longitudinal velocity values computed for cells associated with a given sediment category ($i = 1..10$). For both panels all the axes are plotted in logarithmic scale. Con-

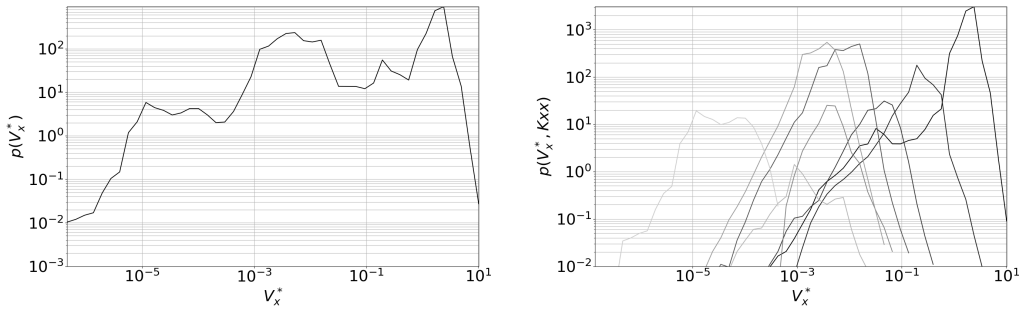

Figure 4.8: Probability density function (left) and joint probability density function (right) of normalised longitudinal Darcy velocity V_x^* field computed on the benchmark Herten domain. The greyscale adopted for the curves in the right panel associates bright colours to low permeabilities facies, e.g. fine sand, and dark colours to high permeability facies, e.g. gravel.

trary to what shown in figure 3.3, the velocity distribution is quite scatter and do not peak around the permeability values of each sediment category. This

is likely due to the higher number of sediment categories (10) in the reference field 4.1 than in the field adopted in section 3.3.8 where four equally spaced permeability facies with the same volume fraction populate the domain. The joint velocity distribution in figure 4.8 illustrates the distribution of velocities across the 10 sediment categories. The brighter lines corresponding to lower permeability values are quite symmetric and have their peaks around lower velocity values meaning that, as expected, it is likely to find low velocities in low permeability regions. Conversely, the fact that higher permeability values (darker lines) show an asymmetric velocity distribution with leftward tail, entails a larger distribution of velocities in highly permeable sediments meaning that, high permeable regions are characterised by low and high velocities while low permeable regions mainly present low velocities. This is in agreement to what was found in figure 3.4 and the conclusion that *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* is not only limited to synthetic domains but can be extended to real domains as well.

The parity diagram in figure 4.9 shows on the horizontal axis the normalised Darcy velocity lumped into 50 categories while on the vertical axis the longitudinal permeability of the sediments grouped into 50 categories. The diameter

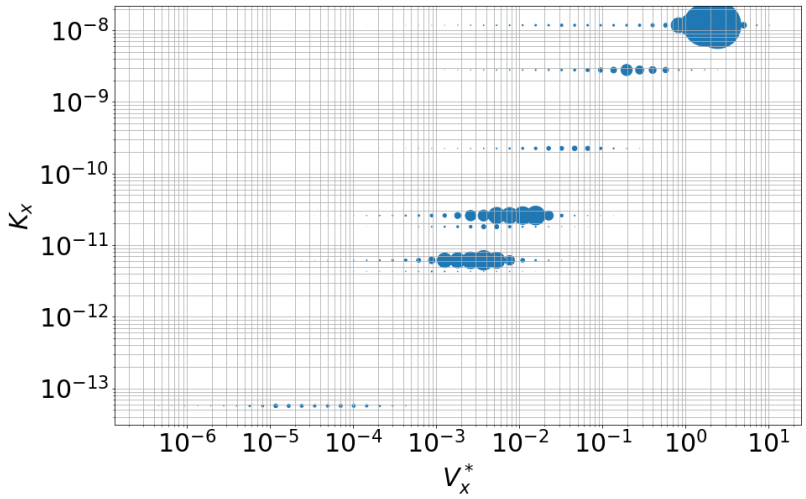

Figure 4.9: Parity diagram of longitudinal velocity and permeability. Permeabilities K_x and normalised velocities V_x^* were divided into categories and the diameters of the circles are proportional to the number of cells which fall into the permeability-velocity category identified by the values on the axes.

of the circles is proportional to the joint probability of each category of the respective velocity-permeability coordinates. The alignment along the diagonal is expected but it is interesting to notice the different spreading of the various facies which represent a different velocity variance in each facies.

4.3.1 Non-conditional and benchmark flow field spatial velocity analysis

In this section we anticipate the velocity fields that, together with 4.7, will be used to study the effect of the permeability field realism degree in section 5.2. At first sight the striking difference between the three velocity distributions in figures 4.8 and 4.8 clearly emerges. These relevant differences

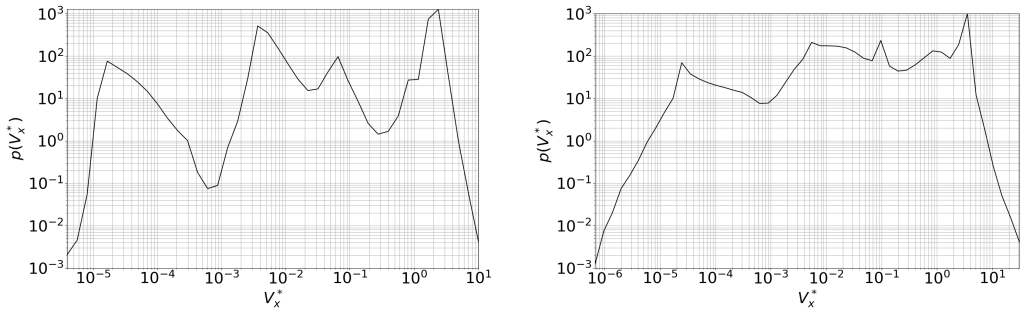

Figure 4.10: Spatial pdf of the velocity field obtained in the the unconditional (on the left) and conditional (on the right) reconstruction of Herten permeability field.

are indicators of substantial differences between the permeability fields reconstructed with different techniques. The permeability field on the left of figure 4.10 procedes from an unconditional simulation of Herten performed with our `setRandomField` `OpenFOAM`®library while the permeability field at the base of the right panel in figure 4.10 is obtained as described in section 4.2. This is quite noticeable since some important geostatistical parameters such as the permeability values of the sediments, their volumetric fractions and the correlation lengths between the unconditional permeability field and the conditional permeability field at the base of the velocity distributions in figures 4.10 are kept constant. Overall, the same qualitative differences can be observed between any of the two panels in figure 4.10 and the velocity distribution illustrated in figure 4.8 obtained from the benchmark permeability.

Chapter 5

Transport modelling study case: the Herten site

In [26] we established a possible hierarchy of non-Fickian triggering factors which can serve as a guideline for setting the limits of macro-scale transport models applicability. Although certainly relevant, the three parameters investigated in [26], namely permeability contrast, correlation length and Péclet number, are just few of the variables which play a relevant role in mono-phase non-reactive transport simulations through heterogeneous porous media. For example, it is essential to assess the relevance of the injection source shape on transport through permeability fields characterised by high contrast sediment as local or well shaped sources limit the range of sediments experienced by the solute. As evidenced in section 4, another relevant factor is the realism degree of the reconstructed geological domain. This is crucial for identifying the preferential flow paths and the velocity distributions that characterise a subsurface domain. As part of the novelties of this chapter, the simulation of more complex meso-scale transport dynamics including the effect of dispersion, is made possible by the developments to the `OpenFOAM`®library [6] which, among other features, introduces models for longitudinal and transversal mechanical dispersion. The main new feature introduced from *adaptiveScalarTransportFoam* used in chapter 3, to *macroScalarTransport* used in this chapter, is the anisotropic mechanical dispersion. At the present stage, our `OpenFOAM`®library [6] allows to perform transport simulations which include all the parameters that describe the classical meso-scale transport model described in section 2.2.4. Multi-phase, compressibility, variable density and other features are not used here as the focus is to study solute transport.

The Eulerian approach to hydrogeology constitutes a new paradigm for flow and transport studies where the impact of several geological and hydraulic

parameters on transport is assessed by using OpenFOAM®to solve partial differential equations. In this chapter, the effect of the following parameters on transport through the geostatistical reconstruction of Herten is assessed:

- injection area;
- geological realism;
- isotropic mechanical dispersion;
- anisotropic mechanical dispersion;
- Péclet number through changes to the molecular diffusion.

The first set of simulations is designed to assess the impact of the solute injection geometry on the overall transport. As highlighted by [53], the geometry of the injection source might be a key parameter in triggering anomalous transport. This is particularly relevant when permeability contrast is high, as relatively small fraction of low or high permeability sediment create flow barriers or highly connected regions which in turn can hinder the transport effects or act as preferential flow paths. Qualitative considerations are drawn by looking at the relative distance between breakthrough curves. Depending on the geometry of the injection source, breakthrough curves might show single or multi-modal distributions. Although local injections might better address the effect of anisotropic mechanical dispersion, due to two main reasons the injection source will be the whole inlet boundary from here onward: firstly the consistency with our previous study [11] where molecular diffusion is always isotropic and secondly wide injection areas allow a greater number of permeabilities to be explored by the solute while travelling through highly heterogeneous porous media.

An increasing degree of realism characterises the permeability fields of the second set of transport simulations. The purpose of these simulations is to assess the relevance of the geological reconstruction realism degree on transport mechanism. Of the three permeability fields, the first one aims at reproducing the Herten field site by means of the PGS algorithm implemented in OpenFOAM®and available at [6], the second one is conditionally generated on a subset of boreholes extracted from the Herten analog (figure 4.4) while the third one represents the best proportion trend map obtained as the mode of the hundred realisations available at [27]. The latter being graphically represented by the sediment distribution in figure 4.1, it is also taken as the reference permeability distribution for the simulations performed in all the other sections of the present chapter.

The third set of simulations is intended to describe the impact that isotropic mechanical dispersion has on solute transport. While in chapter 3 molecular diffusion was the only mechanism implemented in the transport simulation solver, the last version of the code [6] allows for the mechanical dispersion model 2.2.4 to be added to the simulation process. Thanks to this improvement, a more accurate description of the complex meso-scale transport mechanisms is provided which entails a more accurate description of macro-scale solute transport.

The fourth set of simulations evaluates the influence of anisotropic mechanical dispersion on transport. Although certainly relevant, the relation between longitudinal and transversal dispersion is not clear. As evidenced by [98], several empirical studies that attempted to approximate the longitudinal to transversal mechanical dispersion relation provided significantly different results. For this reason, a simple relation where the ratio between α_L and α_T is $10 : 1$ was deemed representative of anisotropic dispersion mechanism through a wide range of porous media.

The fifth set of simulations aims to test the effect of transport at different Péclet numbers on the breakthrough curve. This set of numerical experiments is substantially different from results presented in section 3.3.12 where simulations with a similar setup are performed without including the effect of mechanical dispersion. Although a positive relation between Péclet number and channeling effect was expected, at high Péclet the effect of mechanical dispersion smoothed out the effects of molecular diffusion producing an overlap of the breakthrough curves.

The effects of these five parameters on transport are assessed by varying one parameter value at time over three or four values while keeping the other constants. As a result, seventeen scenarios are proposed with the parameters' values specified in table 5.1. For all these five cases, the same flow and transport initial and boundary conditions adopted in section 3.3.12 and detailed in table 3.1 were selected. It is worth noting that, except for section 5.2, geostatistical parameters like correlation lengths and permeability values cannot be arbitrarily set as the permeability field provided by [95] is adopted as reference.

	S#1	*S#2*	*S#3*	*S#4*	*S#5*	*S#6*	*S#7*	*S#8*	*S#9*	*S#10*	*S#11*	*S#12*	*S#13*	*S#14*	*S#15*	*S#16*	*S#17*
geological realism	S	C	R	R	R	R	R	R	R	R	R	R	R R	R	R	R	
isotropic mec. disp.	10^{-3}	10^{-3}	10^{-3}	1	10^{-1}	10^{-2}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}
anisotropic mec. dispe.	10^{-3}	10^{-3}	10^{-3}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-1}	10^{-2}	10^{-3}	10^{-4}
Péclet number (diffusion)	10^{-10}	10^{-10}	10^{-10}	10^{-10}	10^{-10}	10^{-10}	10^{-10}	10^{-9}	10^{-10}	10^{-11}	10^{-10}	10^{-10}	10^{-10}	10^{-10}	10^{-10}	10^{-10}	10^{-10}
injection area	F	F	F	F	F	F	F	F	F	F	L	W	F	F	F	F	F

Table 5.1: The table summarises the setup of seventeen simulations where the influence of five main characteristics (geological realism, isotropic and anisotropic mechanical dispersion, Péclet number and injection area) on transport is assessed by testing three or four alternatives for each scenario. For the geological realism case the alternatives are: stochastic (S), conditional (C) or real (R) permeability field. The isotropic mechanical dispersion alternatives cover four values of the dispersion coefficient α_L 2.48, from 1 to 10^{-3} m^2/s. The anisotropic dispersion simulation set is characterised by four values of the transversal dispersion coefficient α_T, from 10^{-1} to 10^{-4} for a given ratio $\alpha_L/\alpha_T = 10$. The changes in Péclet number 2.44 are achieved by varying the molecular diffusion coefficient by three magnitude orders, from 10^{-9} to 10^{-11} m^2/s. The injection area scenarios cover three injection geometries local (L), well (W) and face (F) injection.

5.1 Effect of injection area

In this section we analyse the results of injection areas characterised by three different types of geometry on transport (figure 5.1). For the *local* injection

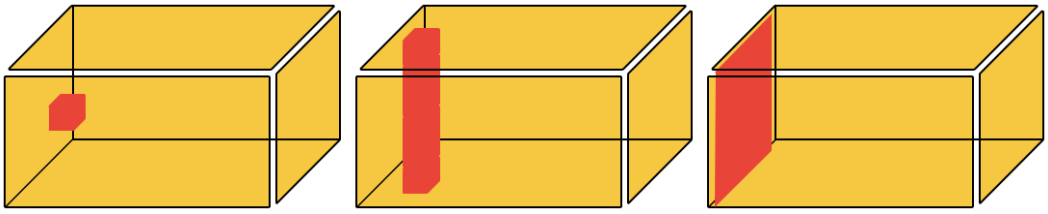

Figure 5.1: The three injection shapes tested in this section: local (left), well or borehole (center) and wall (right).

the solute is released through a cube with one meter side placed at the center of the injection face and one meter away from the high pressure boundary. The *well* injection is characterised by a parallelepiped of the same height of the Herten domain and a surface area of 0.1x0.1 m. In the case of *wall* injection the solute is released through the whole inlet boundary. While real tracer tests are typically characterised by localised or vertically shaped injection geometries, numerically simulated wall injection sources allow the solute to experience a wider range of sediments. It is interesting to observe the peaks of the temporal derivative of the BTCs in the panel on the right side of figure 5.3: the local injection (left panel in figure 5.2) shows a unimodal distribution which becomes tri-modal in the case of well injection (right panel in figure 5.2) and returns bi-modal for the wall-injection. As highlighted in [53], the geometry of the injection source plays a central role in determining the onset of anomalous transport as the release of solute through wide injection volumes allows for the concentration to be evenly distributed across the whole range of sediments while small injection volumes tend to limit the sediments that can be explored by the solute.

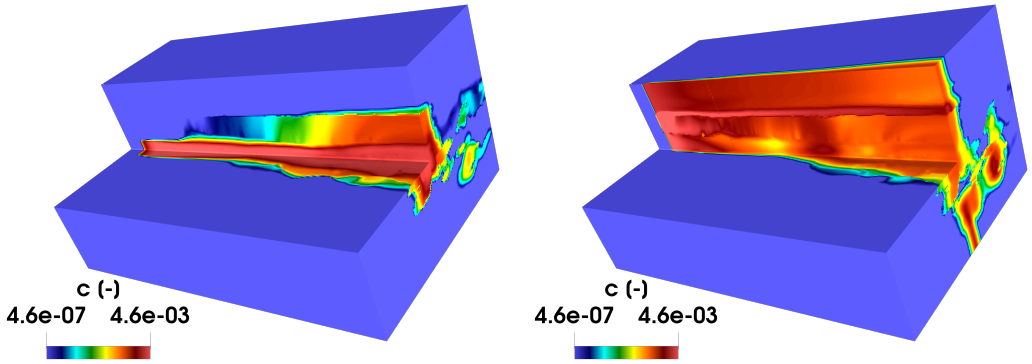

Figure 5.2: Concentration plume at the last time step of the simulation. On the left panel the the local injection is visualised while on the right panel the borehole injection scenario is represented.

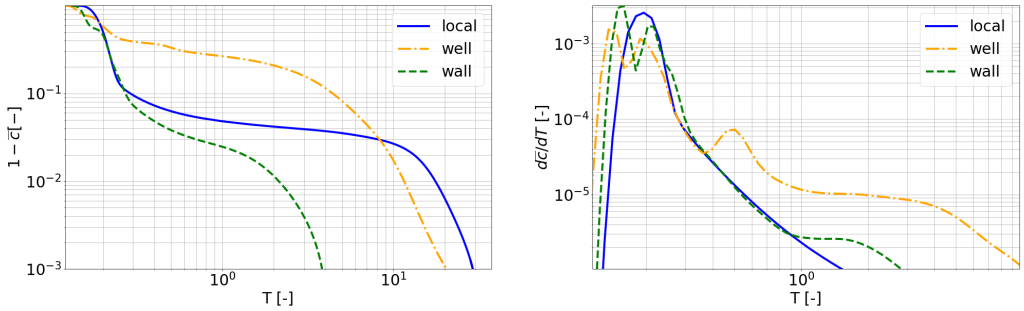

Figure 5.3: BTCs characterised by different shapes of the injection volume. The BTC obtained in the local injection setup is characterised by unimodal distribution while vertical and wall injections are characterised by multi-modal distributions. In case of the local and borehole injection the BTCs peaks correspond to the number of preferential facies through which the solute is flowing. The BTC distribution for the wall injection is the result of more complex dynamics as the overall mixing is enhanced by the elevate number of facies experienced by the solute.

5.2 Effect of permeability field realism degree

Given the natural inaccessibility of the subsurface, its virtual reconstruction is a challenging task which usually requires the application of geostatistical techniques. Geostatistical algorithms can be conditional or unconditional: the firsts preserve the values of surveyed or measured geological properties (hard data) at specified locations and attempt to reconstruct the unsurveyed regions by means of spatial interpolation. The seconds aim at reproducing geological domains by honouring geological parameters such as correlation length or variogram (soft data) throughout the domain. Due to the fundamental role played by the permeability field on transport simulations, it is crucial to understand the effect that these two types of geostatistical algorithms have on transport. In this section we run transport simulations on conditional and unconditional permeability fields and, by assuming that conditional simulations are characterised by a higher realism with respect to the unconditional ones, we ultimately assess the influence of the geological realism degree on transport. While an individual component is inherently involved in the assessment of the geological realism degree, in this study the geological realism degree can be interpreted as degree of similarity between the simulated and the benchmark field. Transport through three permeability fields is simulated:

- unconditional permeability field: it is generated using a PGS algorithm implemented in `OpenFOAM`®and available at [6]. The details about the unconditional PGS algorithm are provided in section 2.1.4 and the result is illustrated in the left panel of figure 5.4.

- conditional permeability field: it is the result of a geostatistical reconstruction which uses boreholes data extracted from the original bidimensional data analog [90] as hard data. More details on the conditional reconstruction process are available in section 4 and the result is presented on the right panel of figure 5.4.

- benchmark for the transport simulations in unconditional and conditional permeability fields: it is based on the detailed geological reconstruction of the Herten field site from bidimensional data available in [27]. The process followed to obtain the reference permeability field is described in section 4.1 and its spatial representation is available in figure 4.1.

It is worth noting the dip angle of the facies at the centre of the permeability field on the right of figure 5.4 which is applied on top of the conditional

algorithm. This dip angle map together with the conditional algorithm contributes to substantially increase the difference between the facies spatial distribution in the conditional and unconditional permeability fields 5.4. As a

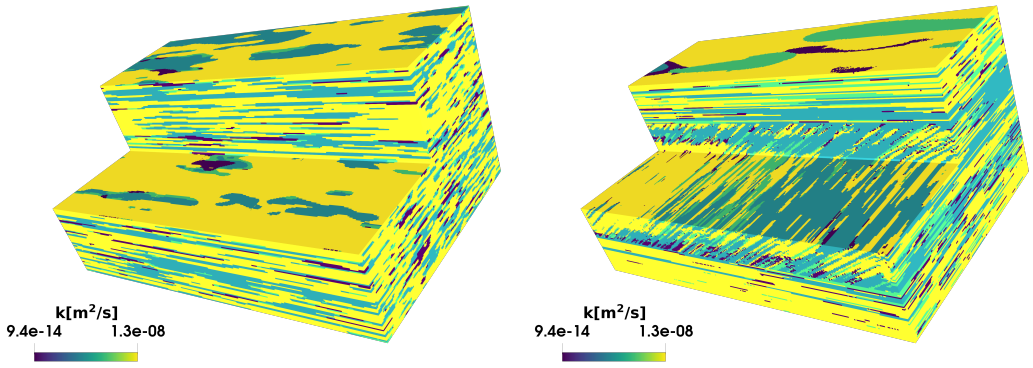

Figure 5.4: Herten permeability field clipped using along the longitudinal axis using a cubical shape. On the left panel the unconditional permeability field is represented while on the right panel the conditional permeability field is shown.

direct consequence, the concentration plumes observed at the same time for the unconditional and the conditional scenarios of figure 5.5 show profoundly different shapes. Although the tilt angle of the facies in the central region of the conditional domain makes it look more similar to the benchmark permeability field than the unconditional one (figure 5.4 vs figure 4.1), it also generates a flow barrier which smooths the peak of the plume (figure 5.5 left panel vs right panel). This effect is clearly visible for BTCs plotted against dimensional times (figure 5.6).

A relevant consequence of altering the permeability field is the modification of the average travelling time that is used to normalise the horizontal axis in non-dimensional of figure 5.7. When breakthrough curves are plotted against the dimensional times, a clear mismatch between benchmark and transport simulations in both conditional and unconditional permeability fields emerges. This discrepancy is representative of the velocity differences that could arise whenever synthetic permeability fields are used to estimate real world transport experiments at field scale using different levels of conditioning.

When a different timescale is used, a net shift between the breakthrough curve peaks emerge: while for BTCs normalised with the average travel time the first peak is the one of the conditional simulation, for BTCs plotted against dimensional times the peak's arrival time is the same. Ultimately, only the BTC obtained for the unconditional normalised simulation matches one of the

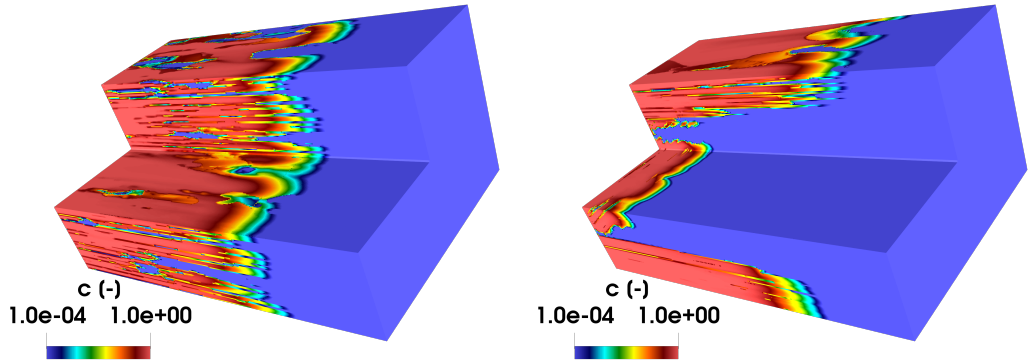

Figure 5.5: Concentration plume at around 8000 seconds. On the left panel the plume is travelling through the unconditional permeability domain while on the right panel the represented plume is flowing through the conditional permeability field.

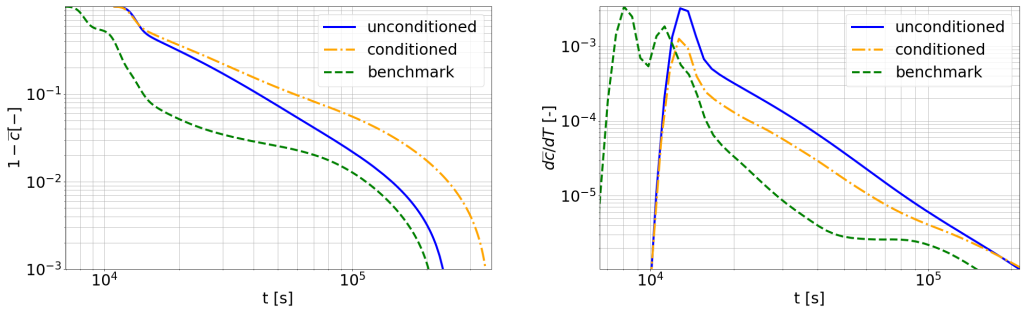

Figure 5.6: Dimensional time BTCs associated with different permeability fields: the *unconditional* BTC corresponds to a transport simulation through a domain generated in OpenFOAM using a PGS random algorithm, the *conditional* BTC is the output of a transport simulation where the permeability field is conditional on core samples randomly extracted from the 2D sections surveyed by [90] and the *benchmark* BTC represents the result of a transport simulation through a geological domain reconstructed in [91].

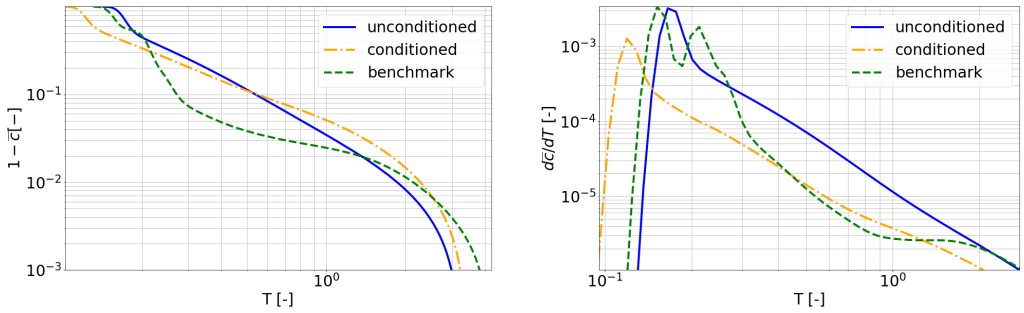

Figure 5.7: BTCs corresponding to transport simulations through different permeability fields. The same legend of figure 5.6 applies. It is worth noting that the timescale for these breakthrough curves is scaled on the average travelling time which varies between one realisation and another.

peaks which characterise the bimodal BTC distribution for the benchmark case.

The fact that the peak of the normalised unconditional breakthrough curve matches quite accurately the first peak of the normalised benchmark breakthrough curve both, in terms of time and magnitude might suggest that the input geological parameters for the unconditional simulation of Herten permeability field such as the correlation length and the variogram provide a geological reconstruction characterised by sediment spatial distribution that well captures the salient characteristics of the permeability field such as mean facies length, connectivity and geometry. Quite surprisingly, the additional pieces of information embedded in the conditional simulation produce a shift of the peak towards earlier arrival times which decreases the accuracy of the transport description. This does not compromise the effectiveness of the conditioning on borehole data while it stresses the central role of the conceptualisation step and the relevance in the choice of the geostatistical parameters set.

5.3 Effect of isotropic mechanical dispersion

In this section the effect of different isotropic mechanical dispersion coefficients on transport is tested. The mechanical dispersion flux, as explained in section 2.2.4, is one of the two fluxes which contributes to the hydrodynamic flux in the advection-dispersion model.

Starting from the assumption that the effect of randomly distributed pores

is qualitatively comparable to the one produced by the molecular agitation, the effect of porous media matrix channels on dispersion flux is mathematically modelled following a parallelism with the molecular diffusion that is, the product between the concentration gradient and the mechanical dispersion matrix 2.48. Being the hydrodynamic dispersion flux the sum on different scales of two mathematically similar terms, i.e. the isotropic molecular diffusion and the mechanical dispersion fluxes, the overall effect is a dispersive behaviour where molecular diffusion becomes relevant for low Péclet numbers while for advection dominated transport, the dispersion is driven by the mechanical dispersion. The mechanical dispersion implementation in the present work relies on the model described by equation 2.47. Isotropic mechanical dispersion is used to describe a comparable effect of the porous media matrix on the solute dispersion in the three directions of the space. In other words, the longitudinal α_L and transversal α_T dispersion coefficient of the mechanical dispersion matrix 2.48 take the same value α.

Having this mechanism included in the numerical model used for the transport simulation of this chapter, allows to include an additional level of physical realism that was not considered in the results of the simulations presented in [26]. Contrary to what expected, for a fixed value of the molecular diffusion

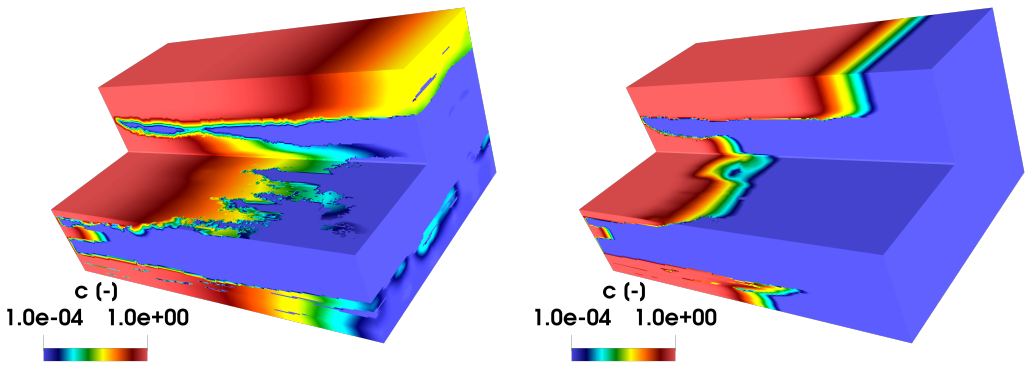

Figure 5.8: Concentration plumes at around 5000 seconds. The value of isotropic mechanical dispersion coefficient which is adopted for the simulation of plume on the left panel is $\alpha = 1\ m$ while on the right panel is $\alpha = 10^{-3}\ m$. The effect of the molecular diffusion on the concentration front is visible only for the low value of the mechanical dispersion coefficient.

coefficient when the mechanical dispersion is high, the front of the plume appears to be more spatially extended (left panel of figure 5.8) than in the case of low mechanical dispersion (right panel of figure 5.8). However this effect is valid on a local scale, where molecular diffusion is relevant.

On a global scale the overall effect of the mechanical dispersion is comparable to that of the molecular diffusion at a local scale: the larger the mechanical dispersion coefficient, the more homogeneous becomes the propagation front of the solute. This effect appears quite clearly looking at the BTCs in figure 5.9 where the unimodal BTC is associated with high value of the mechanical dispersion coefficient while for lower values of mechanical dispersion the bimodal distribution becomes more evident.

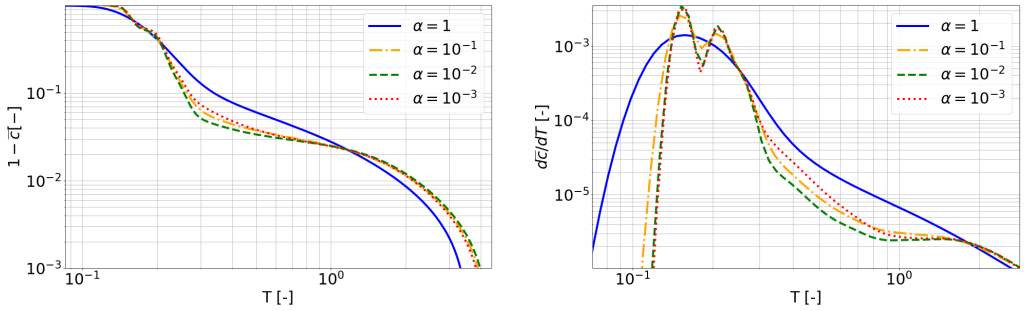

Figure 5.9: BTCs characterised by decreasing values of the mechanical dispersion coefficient. Longitudinal α_L and transversal α_T mechanical dispersion have the same value α. As the mechanical dispersion coefficient decreases the smoothing effect at heterogeneity scale becomes weaker and the advective fluxes give rise to a bimodal distribution.

5.4 Effect of anisotropic mechanical dispersion

In this section the effect of anisotropic mechanical dispersion coefficient on transport through the reconstructed permeability field of Herten is illustrated. Field observations and laboratory experiments often highlight the existence of anisotropic mechanical dispersion behaviour. As illustrated in [98], the measurements of longitudinal and transversal mechanical dispersion often yield significantly different values [98]. Although the physical mechanisms behind this behaviour are very complex and need to be studied at the pore-scale, being able to numerically reproduce it is an essential pre-requirement to tackle the understanding of non-Fickian transport in real macroscale applications. An anisotropic mechanical dispersion model which follows the one presented in 2.48 was implemented in [6] and its results presented in this section. Thanks to this improvement to the code, the longitudinal α_L and transversal α_T mechanical dispersion values can be independently set. Among the several empirical

models that describe the longitudinal to transversal ratio of the mechanical dispersion coefficient, one of the most popular relation estimates the longitudinal dispersion as ten times the transversal dispersion. Notwithstanding the longitudinal to transversal mechanical dispersion ratio adopted, the differences between figure 5.8 and figure 5.10 are negligible. Little differences also emerge

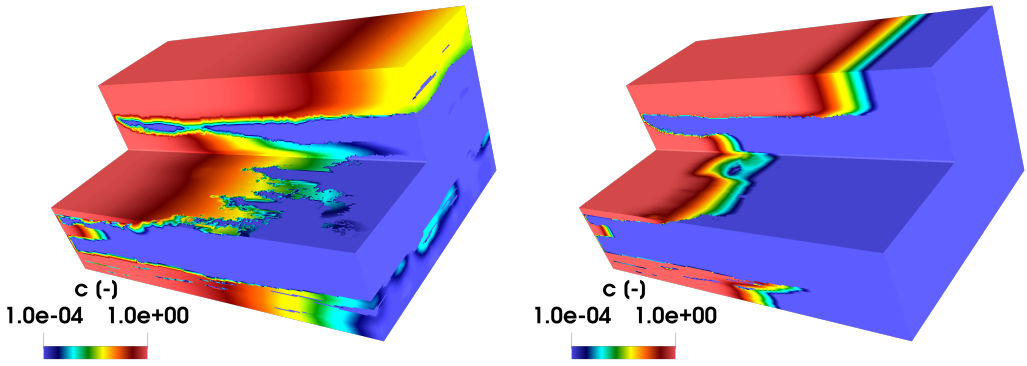

Figure 5.10: Concentration plumes at around 5000 seconds. The value of the longitudinal mechanical dispersion coefficient which is adopted for the simulation of plume on the left panel is $\alpha_L = 1m$ while on the right panel $\alpha_L = 10^{-3}m$. The ratio between longitudinal and transversal mechanical dispersion coefficient is $\alpha_L : \alpha_T = 10 : 1$

by comparing the BTCs in figure 5.11 and 5.9. Although these results suggest

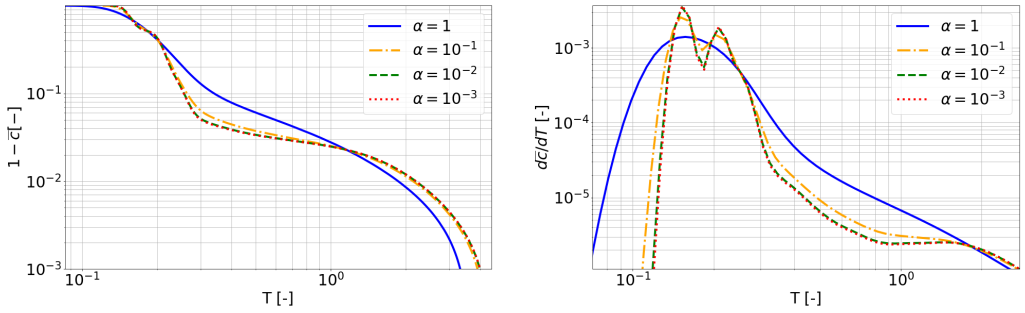

Figure 5.11: BTCs characterised by decreasing values of the longitudinal and transversal mechanical dispersion coefficients α_L and α_T. The ratio between the longitudinal and transversal values is to 10:1 for all the four cases. This enhances the spread in the longitudinal direction.

a negligible impact of the transversal mechanical dispersion on transport, different conclusions might be derived when transport simulations are performed

using different injection geometries. While solute in wall injections tends to fill the whole domain from the start, for local or borehole injections anisotropic mechanical dispersion might have a stronger impact on the solute plume as the transversal dispersion affects the lateral flow paths.

5.5 Effect of Péclet number

In this section the effect of Péclet number on transport through the reconstructed permeability field of Herten is presented. As already defined in 3.11, the Péclet number is computed using the molecular diffusion but it does not account for the mechanical dispersion. For a fixed set of initial and boundary conditions the molecular diffusion coefficient was varied of one magnitude order at each simulation. In this way, since characteristic length and Darcy velocity are kept constant, an increase of the molecular diffusion coefficient corresponds to an equivalent decrease in Péclet number and viceversa. Following the results from section 3.3.12, we decided to observe the transport behaviour under a set of similar Péclet numbers. For the transport results to be consistently compared, the characteristic length for this case is the estimated sediment correlation length and initial conditions replicates the setup adopted in section 3.3.12. The main difference between the simulations presented in section 3.3.12 and the one performed in the present section is represented by the isotropic mechanical dispersion. While in section 3.3.12 no mechanical dispersion is considered, for the results presented in this section the hydrodynamic dispersive flux is the result of the combined effect of molecular diffusion and isotropic mechanical dispersion. Contrary to what expected, no significant

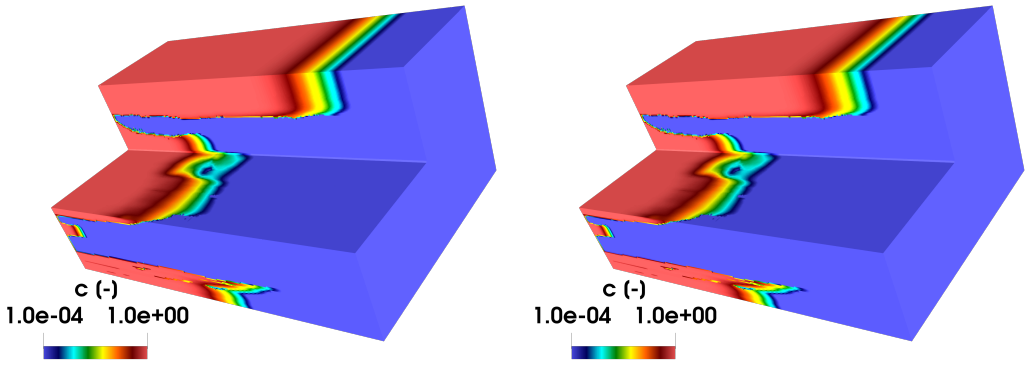

Figure 5.12: Concentration plumes at around 5000 seconds. The plume on the left panel is characterised by a low Péclet number while the one on the right has a high Péclet number.

differences can be appreciated for the solute plumes in figure 5.12 which represent the low and high Péclet number simulations. The reason behind these results lays in the mechanical dispersion effect which dominates and tends to smooth the concentration front of the plume. The value of the isotropic me-

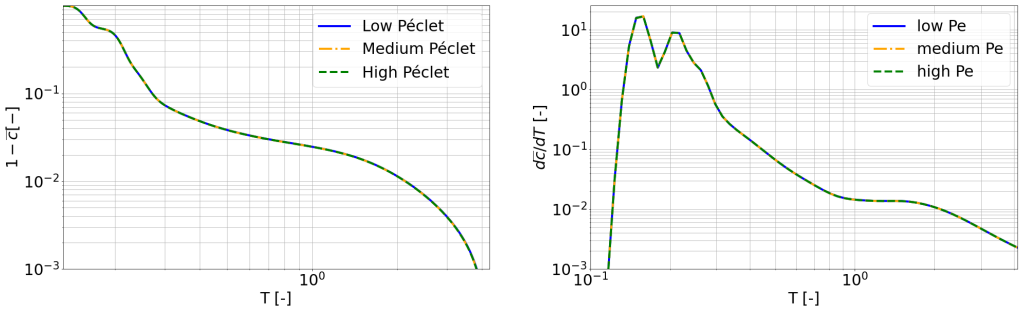

Figure 5.13: BTCs characterised by different values of the molecular diffusion coefficient D_{mol}. As the molecular diffusion decreases the effect of advection becomes predominant and characterised by higher Péclet.

chanical dispersion coefficient adopted for the simulations which yielded the BTCs in figure 5.13 is $\alpha = 10^{-3}m$ and the Péclet number evaluated using the estimated longitudinal correlation length as characteristic dimension ranges from $2.73 \cdot 10^{6}$ for the simulation characterised by low Péclet and $2.73 \cdot 10^{8}$ for the simulation characterised by high Péclet.

Chapter 6

Conclusions

In this thesis, flow (Darcy) and transport (advection-dispersion) meso-scale mathematical models have been discussed and applied to heterogeneous unconditional and conditional permeability fields by means of numerical simulations. The results from these numerical simulations aim to substitute empirical tracer tests, otherwise expensive and time consuming to be collected. Based on macro-scale transport model fitting of the meso-scale data, chapter 3 provides a quantitative estimation of the discrepancy between the macro-scale prediction and the numerical meso-scale results. The application of this procedure to unconditional permeability fields with different geostatistical and flow parameters allowed the compilation of a hierarchical list of non-Fickian transport triggering factors. Chapter 4 focuses on the process to be followed to condition the reconstruction of an aquifer on boreholes data. The result is compared against the aquifer reconstruction conditioned on several sections surveyed at a high resolution level. Flow simulations through the aquifer are run on the reconstructed permeability field of the aquifer. Eventually, in chapter 5 transport behaviour through highly heterogeneous porous medium is analysed for the unconditional and the conditional aquifer reconstructions.

The most relevant results of this work can be summarised as follow:

- chapter 3

    1. for the settings specified in chapter 3, the macrodispersion model can be confidently applied with good accuracy in low permeability contrast scenarios. In ergodic domains with permeability contrast up to one order of magnitude, the error which characterises the approximated solution of the macrodispersion model is in the order of 10-20%;

2. for ergodic domains where the solute is injected on a surface which encompasses all the sediments of the domain, the permeability contrast between facies is the most relevant factor in determining the onset non-Fickian transport. Although Péclet number and correlation length influence the fate of the solute, their role in triggering anomalous transport behaviour has secondary importance;

- chapter 4

3. the spatial distribution of fluid velocity is mainly determined by permeability contrast between facies and secondary by the correlation lengths. Since the velocity field controls the transport mechanisms, this result confirms the hierarchy of the non-Fickian triggering factors described in the previous point;
4. subsurface reconstruction is a challenging task that, in case of high contrast and sharp separation between facies can be effectively tackled using the PGS algorithm. The procedure described in section 4.2 provides some guidelines to perform conditional PGS simulations starting from boreholes data. Given the extreme complexity which characterises the subsurface geological systems, the relevant simplifications which constrain the numerical model inputs might compromise the overall simulation. Consequently, considerable time should be invested in the conceptualisation step and sensitive analysis might be used to individuate the best simplification procedure;

- chapter 5

5. the solute release location and the shape of the volume involved exert a strong control on the overall transport dynamics. Depending on the sediments where solute is initially injected, its spreading patterns deeply affect the BTCs. Consequently, depending on the experiment or simulation purpose, great attention is needed especially when local or borehole injection are performed. As a rule of thumb for BTCs close to local injection sources, the number of peaks correspond to the number of facies within the injection volume;
6. unconditional and conditional permeability fields which share the correlation lengths and the sediment proportions do not necessarily present the same effective permeability therefore substantial transport differences might arise from different velocity fields. Geostatistical parameters such as Kriging search radius and the dip angle

maps might be combined with conditional simulations and heavily control the pattern of the permeability field therefore rigorous procedures should be adopted before setting theirs values;

7. Péclet number provides useful qualitative indications on the expected trend of breakthrough curves for transport simulations where the only dispersive mechanism involved in the advection-dispersion equation is molecular diffusion. For transport models where mechanical dispersion is accounted for, Péclet number variations effectively characterise only low advection velocity scenarios where the mechanical dispersion effects are negligible;

8. in high advection velocity scenarios, mechanical dispersion plays a critical role in the solute transport process therefore, the mechanical dispersion model should be carefully selected. The difference between isotropic and anisotropic mechanical dispersion models in injection scenarios characterised by wide surface injection area is negligible. Nevertheless, the impact of these two mechanical dispersion models on transport should be assessed on local injection scenarios;

Beyond these conclusions, we would like to stress the relevance of the methods first proposed in this study and later expanded in [45]. The shift of perspective from the traditional Lagrangian approach to the Eulerian perspective frees the hydrogeological community from computational limitations which have often constrained the simulations to 2-D particle tracking experiments by proposing the numerical solutions of partial differential equations which describe the complex subsurface mechanisms, such as the advection-dispersion equation. This paradigm shift was enabled by the extension of `OpenFOAM`®libraries to hydrogeological problems. These libraries can be freely accessed at [6].

Future works which might conveniently be developed within the computational framework illustrated in this work include the assessment of the lithotype thresholds for PGS permeability fields on transport and the fitting of more complex non-Markovian macro-models. Another interesting future study might involve the assessment of the contribution of variable density fluids to the onset of non-Fickian transport together with the study of multiphase transport in porous media. These are hot topics that can be effectively tackled with the latest version of our meso-scale transport simulation utility freely available at [6] with some benchmark example described in [45].

Bibliography

[1] Lynn W Gelhar and Carl L Axness. Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research*, 19(1):161–180, 1983.

[2] Roy Haggerty and Steven M Gorelick. Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity. *Water Resources Research*, 31(10):2383–2400, 1995.

[3] Elliott W Montroll and George H Weiss. Random walks on lattices. ii. *Journal of Mathematical Physics*, 6(2):167–181, 1965.

[4] Shlomo P Neuman. Eulerian-lagrangian theory of transport in space-time nonstationary velocity fields: Exact nonlocal formalism by conditional moments and weak approximation. *Water Resources Research*, 29(3):633–645, 1993.

[5] Marco Dentz, Tanguy Le Borgne, Andreas Englert, and Branko Bijeljic. Mixing, spreading and reaction in heterogeneous media: A brief review. *Journal of contaminant hydrology*, 120:1–17, 2011.

[6] Federico Municchi, Eugenio Pescimoro, Juan Hidalgo, and Matteo Icardi. Securefoam v0.1, aug 2022.

[7] Georges Matheron. Principles of geostatistics. *Economic geology*, 58(8):1246–1266, 1963.

[8] Christine E Koltermann and Steven M Gorelick. Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research*, 32(9):2617–2658, 1996.

[9] Marco Bianchi. Accounting for geological heterogeneity in groundwater flow models: a review of approaches. In *British Geological Survey internal report*, pages 1–44. British Geological Service, 2019.

[10] Hélène Beucher and Didier Renard. Truncated gaussian and derived methods. *Comptes Rendus Geoscience*, 348(7):510–519, 2016.

[11] Eugenio Pescimoro, Fulvio Boano, Audrey H Sawyer, and Mohamad Reza Soltanian. Modeling influence of sediment heterogeneity on nutrient cycling in streambeds. *Water Resources Research*, 55(5):4082–4095, 2019.

[12] Xiaoying Zhang, Funing Ma, Shangxian Yin, Corey D Wallace, Mohamad Reza Soltanian, Zhenxue Dai, Robert W Ritzi, Ziqi Ma, Chuanjun Zhan, and Xiaoshu Lü. Application of upscaling methods for fluid flow and mass transport in multi-scale heterogeneous media: A critical review. *Applied Energy*, 303:117603, 2021.

[13] Matteo Icardi, Gianluca Boccardo, Daniele L Marchisio, Tiziana Tosco, and Rajandrea Sethi. Pore-scale simulation of fluid flow and solute dispersion in three-dimensional porous media. *Physical review E*, 90(1):013032, 2014.

[14] M King Hubbert. Darcy's law and the field equations of the flow of underground fluids. *Transactions of the AIME*, 207(01):222–239, 1956.

[15] Federico Toschi and Marcello Sega. *Flowing matter*. Springer Nature, 2019.

[16] Branko Bijeljic, Peyman Mostaghimi, and Martin J Blunt. Signature of non-fickian solute transport in complex heterogeneous porous media. *Physical review letters*, 107(20):204502, 2011.

[17] Veronica L Morales, Marco Dentz, Matthias Willmann, and Markus Holzner. Stochastic dynamics of intermittent pore-scale particle motion in three-dimensional porous media: Experiments and theory. *Geophysical Research Letters*, 44(18):9361–9371, 2017.

[18] Eleonora Crevacore, Tiziana Tosco, Rajandrea Sethi, Gianluca Boccardo, and Daniele L Marchisio. Recirculation zones induce non-fickian transport in three-dimensional periodic porous media. *Physical Review E*, 94(5):053118, 2016.

[19] Roy Haggerty, Sean A McKenna, and Lucy C Meigs. On the late-time behavior of tracer test breakthrough curves. *Water Resources Research*, 36(12):3467–3479, 2000.

[20] Peter K Kang, Tanguy Le Borgne, Marco Dentz, Olivier Bour, and Ruben Juanes. Impact of velocity correlation and distribution on transport in fractured media: Field evidence and theoretical model. *Water Resources Research*, 51(2):940–959, 2015.

[21] Stephen Whitaker. Flow in porous media i: A theoretical derivation of darcy's law. *Transport in porous media*, 1(1):3–25, 1986.

[22] William G Gray and Kevin O'Neill. On the general equations for flow in porous media and their reduction to darcy's law. *Water Resources Research*, 12(2):148–154, 1976.

[23] Guillermo A Narsilio, Olivier Buzzi, Stephen Fityus, Tae Sup Yun, and David W Smith. Upscaling of navier–stokes equations in porous media: Theoretical, numerical and experimental approach. *Computers and Geotechnics*, 36(7):1200–1206, 2009.

[24] Ulrich Hornung. *Homogenization and porous media*, volume 6. Springer Science & Business Media, 1996.

[25] Stephen Whitaker. *The method of volume averaging*, volume 13. Springer Science & Business Media, 2013.

[26] Eugenio Pescimoro, Matteo Icardi, Giovanni Porta, and Marco Bianchi. Emergence of non-fickian transport in truncated pluri-gaussian permeability fields. *GEM-International Journal on Geomathematics*, 13(1):1–27, 2022.

[27] P Bayer, A Comunian, D Höyng, and G Mariethoz. Physicochemical properties and 3d geostatistical simulations of the herten and the descalvado aquifer analogs. 2015.

[28] Catherine E Powell et al. Generating realisations of stationary gaussian random fields by circulant embedding. *matrix*, 2(2):1, 2014.

[29] Steven F Carle and Graham E Fogg. Transition probability-based indicator geostatistics. *Mathematical geology*, 28(4):453–476, 1996.

[30] Gary S Weissmann, Steven F Carle, and Graham E Fogg. Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics. *Water Resources Research*, 35(6):1761–1770, 1999.

[31] Margaret Armstrong, Alain Galli, Hélène Beucher, Gaelle Loc'h, Didier Renard, Brigitte Doligez, Rémi Eschard, and Francois Geffroy. *Plurigaussian simulations in geosciences*. Springer Science & Business Media, 2011.

[32] Christian Lantuéjoul. *Geostatistical simulation: models and algorithms*. Number 1139. Springer Science & Business Media, 2001.

[33] Hans Wackernagel. Examples of covariance functions. In *Multivariate Geostatistics*, pages 55–59. Springer, 1998.

[34] Falk Heße, Vladyslav Prykhodko, Steffen Schlüter, and Sabine Attinger. Generating random fields with a truncated power-law variogram: A comparison of several numerical methods. *Environmental modelling & software*, 55:32–48, 2014.

[35] Yong Zhang, David A Benson, Mark M Meerschaert, Eric M LaBolle, and Hans-Peter Scheffler. Random walk approximation of fractional-order multiscaling anomalous diffusion. *Physical Review E*, 74(2):026706, 2006.

[36] Clayton V Deutsch and André G Journel. Gslib. *Geostatistical software library and user's guide*, 369, 1998.

[37] Feng Ruan and Dennis McLaughlin. An efficient multivariate random field generator using the fast fourier transform. *Advances in water resources*, 21(5):385–399, 1998.

[38] Benoit B Mandelbrot and John W Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437, 1968.

[39] Jacob Bear. *Dynamics of fluids in porous media*. Courier Corporation, 1988.

[40] M Dentz, DM Tartakovsky, E Abarca, Alberto Guadagnini, Xavier Sánchez-Vila, and Jesus Carrera. Variable-density flow in porous media. *Journal of fluid mechanics*, 561:209–235, 2006.

[41] Clifford I Voss and William R Souza. Variable density flow and solute transport simulation of regional aquifers containing a narrow freshwater-saltwater transition zone. *Water Resources Research*, 23(10):1851–1866, 1987.

[42] Geoffrey Ingram Taylor. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 219(1137):186–203, 1953.

[43] Rutherford Aris. On the dispersion of a solute in a fluid flowing through a tube. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 235(1200):67–77, 1956.

[44] J David Logan. *Transport modeling in hydrogeochemical systems*, volume 15. Springer Science & Business Media, 2001.

[45] Matteo Icardi, Eugenio Pescimoro, Federico Municchi, and Juan H Hidalgo. Computational framework for complex flow and transport in heterogeneous porous media. *arXiv preprint arXiv:2212.10961*, 2022.

[46] Henry G Weller, Gavin Tabor, Hrvoje Jasak, and Christer Fureby. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computers in physics*, 12(6):620–631, 1998.

[47] E Eric Adams and Lynn W Gelhar. Field study of dispersion in a heterogeneous aquifer: 2. spatial moments analysis. *Water Resources Research*, 28(12):3293–3307, 1992.

[48] Heidi Christiansen Barlebo, Mary C Hill, and Dan Rosbjerg. Investigating the macrodispersion experiment (made) site in columbus, mississippi, using a three-dimensional inverse flow and transport model. *Water Resources Research*, 40(4), 2004.

[49] A Fiori, Vladimir Cvetkovic, G Dagan, S Attinger, A Bellin, P Dietrich, A Zech, and G Teutsch. 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 Resources Research*, 52(12):9228–9234, 2016.

[50] Brendan Zinn and Charles F Harvey. 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 Resources Research*, 39(3), 2003.

[51] I Jankovic, M Maghrebi, A Fiori, and G Dagan. When good statistical models of aquifer heterogeneity go right: The impact of aquifer permeability structures on 3d flow and transport. *Advances in water resources*, 100:199–211, 2017.

[52] Marco Bianchi and Chunmiao Zheng. A lithofacies approach for modeling non-fickian solute transport in a heterogeneous alluvial aquifer. *Water Resources Research*, 52(1):552–565, 2016.

[53] Maosheng Yin, Yong Zhang, Rui Ma, Geoffrey R Tick, Marco Bianchi, Chunmiao Zheng, Wei Wei, Song Wei, and Xiaoting Liu. Super-diffusion affected by hydrofacies mean length and source geometry in alluvial settings. *Journal of Hydrology*, 582:124515, 2020.

[54] Brian Berkowitz, Andrea Cortis, Marco Dentz, and Harvey Scher. Modeling non-fickian transport in geological formations as a continuous time random walk. *Reviews of Geophysics*, 44(2), 2006.

[55] Jacob Bear. *Hydraulics of groundwater*. Courier Corporation, 2012.

[56] Felipe P.J. de Barros, Alberto Guadagnini, and Monica Riva. Features of transport in non-gaussian random porous systems. *International Journal of Heat and Mass Transfer*, 184:122244, 2022.

[57] M. Riva, A. Guadagnini, D. Fernandez-Garcia, X. Sanchez-Vila, and T. Ptak. Relative importance of geostatistical and transport models in describing heavily tailed breakthrough curves at the lauswiesen site. *Journal of Contaminant Hydrology*, 101(1-4):1–13, 2008.

[58] Gedeon Dagan. *Flow and transport in porous formations*. Springer Science & Business Media, 2012.

[59] Scott K. Hansen, Claus P. Haslauer, Olaf A. Cirpka, and Velimir V. Vesselinov. Direct breakthrough curve prediction from statistics of heterogeneous conductivity fields. *Water Resources Research*, 54(1):271–285, 2018.

[60] Alraune Zech, Sabine Attinger, Alberto Bellin, Vladimir Cvetkovic, Gedeon Dagan, Marco Dentz, Peter Dietrich, Aldo Fiori, and Georg Teutsch. A comparison of six transport models of the made-1 experiment implemented with different types of hydraulic data. *Water Resources Research*, 57(5):e2020WR028672, 2021.

[61] Shlomo P Neuman and Daniel M Tartakovsky. Perspective on theories of non-fickian transport in heterogeneous media. *Advances in Water Resources*, 32(5):670–680, 2009.

[62] Yaniv Edery, Alberto Guadagnini, Harvey Scher, and Brian Berkowitz. Origins of anomalous transport in heterogeneous media: Structural and dynamic controls. *Water Resources Research*, 50(2):1490–1505, 2014.

[63] Y. Edery, G.M. Porta, A. Guadagnini, H. Scher, and B. Berkowitz. Characterization of bimolecular reactive transport in heterogeneous porous media. *Transport in Porous Media*, 115(2):291–310, 2016.

[64] Y. Edery, M. Stolar, G. Porta, and A. Guadagnini. Feedback mechanisms between precipitation and dissolution reactions across randomly heterogeneous conductivity fields. *Hydrology and Earth System Sciences*, 25(11):5905–5915, 2021.

[65] Grégoire Mariethoz, Philippe Renard, Fabien Cornaton, and Olivier Jaquet. Truncated plurigaussian simulations to characterize aquifer heterogeneity. *Groundwater*, 47(1):13–24, 2009.

[66] Hrvoje Gotovac, Vladimir Cvetkovic, and Roko Andricevic. Flow and travel time statistics in highly heterogeneous porous media. *Water resources research*, 45(7), 2009.

[67] Guillem Sole-Mari, Monica Riva, Daniel Fernàndez-Garcia, Xavier Sanchez-Vila, and Alberto Guadagnini. Solute transport in bounded porous media characterized by generalized sub-gaussian log-conductivity distributions. *Advances in Water Resources*, 147:103812, 2021.

[68] Marco Bianchi and Daniele Pedretti. Geological entropy and solute transport in heterogeneous porous media. *Water Resources Research*, 53(6):4691–4708, 2017.

[69] Yong Zhang, Christopher T Green, and Graham E Fogg. The impact of medium architecture of alluvial settings on non-fickian transport. *Advances in water resources*, 54:78–99, 2013.

[70] Yong Zhang, Christopher T Green, and Boris Baeumer. Linking aquifer spatial properties and non-fickian transport in mobile–immobile like alluvial settings. *Journal of Hydrology*, 512:315–331, 2014.

[71] Yong Zhang, Christopher T Green, and Geoffrey R Tick. Peclet number as affected by molecular diffusion controls transient anomalous transport in alluvial aquifer–aquitard complexes. *Journal of contaminant hydrology*, 177:220–238, 2015.

[72] Heather Savoy, Thomas Kalbacher, Peter Dietrich, and Yoram Rubin. Geological heterogeneity: Goal-oriented simplification of structure and characterization needs. *Advances in Water Resources*, 109:1–13, 2017.

[73] Roussos Dimitrakopoulos and Xiaochun Luo. Generalized sequential gaussian simulation on group size ν and screen-effect approximations for large field simulations. *Mathematical Geology*, 36(5):567–591, 2004.

[74] Steven F Carle and Graham E Fogg. Modeling spatial variability with one and multidimensional continuous-lag markov chains. *Mathematical Geology*, 29(7):891–918, 1997.

[75] Sebastien Strebelle. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical geology*, 34(1):1–21, 2002.

[76] Steven F Carle. T-progs: Transition probability geostatistical software. *University of California, Davis, CA*, 84, 1999.

[77] Olaf Kolditz, Sebastian Bauer, Lars Bilke, Norbert Böttcher, Jens-Olaf Delfs, Thomas Fischer, Uwe J Görke, Thomas Kalbacher, Georg Kosakowski, CI McDermott, et al. Opengeosys: an open-source initiative for numerical simulation of thermo-hydro-mechanical/chemical (thm/c) processes in porous media. *Environmental Earth Sciences*, 67(2):589–599, 2012.

[78] Pierre Horgue, Cyprien Soulaine, Jacques Franc, Romain Guibert, and Gérald Debenest. An open-source toolbox for multiphase flow in porous media. *Computer Physics Communications*, 187:217–226, 2015.

[79] Bernd Flemisch, Melanie Darcis, K Erbertseder, B Faigle, A Lauser, Klaus Mosthaf, S Müthing, Philipp Nuske, A Tatomir, M Wolff, et al. Dumux: Dune for multi-{phase, component, scale, physics,...} flow and transport in porous media. *Advances in Water Resources*, 34(9):1102–1112, 2011.

[80] Niklas Linde, Philippe Renard, Tapan Mukerji, and Jef Caers. Geological realism in hydrogeological and geophysical inverse modeling: A review. *Advances in Water Resources*, 86:86–101, 2015.

[81] Joseph B Keller. A theorem on the conductivity of a composite medium. *Journal of Mathematical Physics*, 5(4):548–549, 1964.

[82] G Matheron. Composition des permeabilites en milieu poreux heterogene. methode de schwydler et regles de ponderation. 1967.

[83] Daniel M Tartakovsky and Marco Dentz. Diffusion in porous media: phenomena and mechanisms. *Transport in Porous Media*, 130(1):105–127, 2019.

[84] Martinus Theodorus Van Genuchten. *Analytical solutions of the one-dimensional convective-dispersive solute transport equation*, volume 1661. US Department of Agriculture, Agricultural Research Service, 1982.

[85] C Yu, AW Warrick, and MH Conklin. A moment method for analyzing breakthrough curves of step inputs. *Water Resources Research*, 35(11):3567–3572, 1999.

[86] A Kreft and A Zuber. On the physical meaning of the dispersion equation and its solutions for different initial and boundary conditions. *Chemical Engineering Science*, 33(11):1471–1480, 1978.

[87] Alessandro Comolli, Vivien Hakoun, and Marco Dentz. Mechanisms, upscaling, and prediction of anomalous dispersion in heterogeneous porous media. *Water Resources Research*, 55(10):8197–8222, 2019.

[88] Thomas Sherman, Nicholas B Engdahl, Giovanni Porta, and Diogo Bolster. A review of spatial markov models for predicting pre-asymptotic and anomalous transport in porous and fractured media. *Journal of Contaminant Hydrology*, 236:103734, 2021.

[89] Federico Municchi and Matteo Icardi. Generalized multirate models for conjugate transfer in heterogeneous materials. *Physical Review Research*, 2(1):013041, 2020.

[90] P Bayer, P Huggenberger, Philippe Renard, and Alessandro Comunian. Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: Field study. *Journal of Hydrology*, 405(1-2):1–9, 2011.

[91] Alessandro Comunian, Philippe Renard, Julien Straubhaar, and Peter Bayer. Three-dimensional high resolution fluvio-glacial aquifer analog–part 2: Geostatistical modeling. *Journal of hydrology*, 405(1-2):10–23, 2011.

[92] Alessandro Comunian, Philippe Renard, and Julien Straubhaar. 3d multiple-point statistics simulation using 2d training images. *Computers & Geosciences*, 40:49–65, 2012.

[93] Richard Wilfred Rwechungura, Eka Suwartadi, Mohsen Dadashpour, Jon Kleppe, Bjarne A Foss, et al. The norne field case-a unique comparative case study. In *SPE Intelligent Energy Conference and Exhibition*. Society of Petroleum Engineers, 2010.

[94] Tore A Torp and John Gale. Demonstrating storage of co2 in geological reservoirs: The sleipner and sacs projects. *Energy*, 29(9-10):1361–1369, 2004.

[95] Peter Bayer, Alessandro Comunian, Dominik Höyng, and Gregoire Mariethoz. High resolution multi-facies realizations of sedimentary reservoir and aquifer analogs. *Scientific data*, 2(1):1–10, 2015.

[96] TJ Szydlik, Simon Way, Patrick Smith, Lars Aamodt, and Christina Friedrich. 3d pp/ps prestack depth migration on the volve field. In *68th EAGE Conference and Exhibition incorporating SPE EUROPEC 2006*, pages cp–2. European Association of Geoscientists & Engineers, 2006.

[97] STEVEN F Carle, ERIC M Labolle, GARY S Weissmann, DAVID Van Brocklin, and Graham E Fogg. Conditional simulation of hydrofacies architecture: a transition probability/markov approach. *Hydrogeologic models of sedimentary aquifers, concepts in hydrogeology and environmental geology*, 1(1):147–170, 1998.

[98] JMPQ Delgado. Longitudinal and transverse dispersion in porous media. *Chemical Engineering Research and Design*, 85(9):1245–1252, 2007.