Modelling solute dispersion in periodic heterogeneous porous media: model benchmarking against intermediate scale experiments

Samer Majdalani, Vincent Guinot, Carole Delenne, Hicham Gebran

Modelling solute dispersion in periodic heterogeneous porous media: model benchmarking against intermediate scale experiments Manuscript accepted for publication in Journal of Hydrology (Editor decision 9 March 2018)

Samer Majdalani[a], Vincent Guinot[a,b], Carole Delenne[a,b], Hicham Gebran[c]

[a] *Univ. Montpellier, Polytech Montpellier/HSM, France*
[b] *Inria Lemon, France*
[c] *Lebanese University Fanar, Department of Mathematics*

Abstract

This paper is devoted to theoretical and experimental investigations of solute dispersion in heterogeneous porous media. Dispersion in heterogenous porous media has been reported to be scale-dependent, a likely indication that the proposed dispersion models are incompletely formulated. A high quality experimental data set of breakthrough curves in periodic model heterogeneous porous media is presented. In contrast with most previously published experiments, the present experiments involve numerous replicates. This allows the statistical variability of experimental data to be accounted for. Several models are benchmarked against the data set: the Fickian-based advection-dispersion, mobile-immobile, multirate, multiple region advection dispersion models, and a newly proposed dispersion model based on pure advection. A salient property of the latter model is that its solutions exhibit a ballistic behaviour for small times, while tending to the Fickian behaviour for large time scales. Model performance is assessed using a novel objective function accounting for the statistical variability of the experimental data set, while putting equal emphasis on both small ad large time scale behaivours. Besides being as accurate as the other models, the new purely advective model has the advantages that (i) it does not exhibit the undesirable effects associated with the usual Fickian operator (namely the infinite solute front propagation speed), and (ii) it allows dispersive transport to be simulated on every heterogeneity scale using scale-independent parameters.

Keywords: Solute transport, heterogeneous porous media, intermediate scale, Fickian behaviour, dispersion modelling.

1. Introduction

In many circumstances, the classical Fickian operator fails to account correctly for the behaviour of solutes in heterogeneous porous media. The Advection-Dispersion (AD) model exhibits poor performance. Attempting to calibrate this model against field or laboratory data has been seen to lead to contradictory conclusions. Field scale dispersion data have been reported to yield a growing trend for the dispersion coefficient D with the scale of the experiment [25]. A number of laboratory

experiments, in contrast, indicate that no clear trend can be identified for the variations in D with experiment scale. For instance, [58] report an increasing trend for the dispersion coefficient. In [39], an increasing trend is found for $D(x)$, but the authors notice that this conclusion may be biased by experimental noise. In [61], identifying a trend for the variations of D with distance is found very difficult if not impossible. In [14], no scaling trend is identified for the dispersion coefficient, even over short distances. More recently, laboratory experiments carried out on an artificial, periodic porous medium [48] show that contradictory trends in $D(x)$ can easily be inferred if the breakthrough curves are not sampled with sufficient accuracy and the tracer experiments are not replicated a sufficient number of times. Several models with scale-dependent dispersion have been proposed in the literature [2, 40, 54, 55, 67, 68, 69, 70]. All these models have shown a good ability to reproduce field- or laboratory-obtained experimental breakthrough curves via a proper parameter tuning. This makes a benchmarking of their respective predictive capabilities very difficult [22]. The following models have been used extensively for benchmarking against experimental data sets.

The Fractional Advection-Dispersion (FAD) model builds up on the Continuous Time Random Walk (CTRW) formalism [43, 49, 51]. FAD occurs when the motion of the solute molecules is non-Brownian. Different behaviours may be obtained depending on the assumptions made on the characteristic times and lengths of molecule jumps [43, 44, 41, 42]. In the presence of trapping effects, an inverse power law asymptotic behaviour may be observed for the probability density function of solute residence time in the porous media. This results in subdiffusive dispersion processes, with a variance of molecule positions growing slower than time. Another type of non-Fickian behaviour is that of Levy motion, whereby the characteristic time for particle motion is finite, but the characteristic length of the jumps in molecule positions is infinite [6, 7]. The resulting behaviour is called superdiffusion, with a variance of molecule positions growing faster than time. All these models share the common feature that the governing equations incorporate fractional derivatives with respect to time and/or space, hence the term "fractional". FAD models have been tested against experimental data sets obtained from laboratory experiments [8, 13, 37, 45]. In [37] the best fit was obtained by making the dispersion parameters scale-dependent. In [62], a FAD model was tested against in situ data obtained from experiments at the scale of 1m to 1km. Comparisons with data observed at the metric scale [9] showed that time-varying fractional orders of differentiation were essential in reconstructing the heavy tailing in the observed breakthrough curves.

The Mobile-Immobile (MI) model [24, 64] is based on the assumption of a mobile region (where the solute obeys a standard AD model) exchanging with an immobile region. The MI formalism has been used to describe different physical settings. The simple structure of this model allows analytical solutions to be obtained for a number of configurations [17, 31, 53, 65]. Several versions of the MI model with a scale-dependent dispersion coefficient have been explored in [23]. The

best fit against the experimental laboratory results obtained in [38] was achieved for a dispersion coefficient varying exponentially with the travelled distance.

The Multiple Rate (MR) model [35] is a generalization of the MI model. Several immobile regions exchange with the mobile region according to different exchange rates. Increasing the number of regions and varying the exchange kinetics allows for anomalous diffusion processes to be reproduced via a proper distribution of the exchange rates between the mobile and immobile fractions [18].

Multiple Region Advection-Dispersion (MRAD) models have been proposed to account for the dispersion of solutes in heterogeneous soils in the presence of macropores, high- or low-permeability inclusions or several spatial scales of hydraulic heterogeneity. Note that the term MRAD is not the name given to these models by their authors but a term proposed by the authors of the present paper for the sake of terminology convenience. In these models, several different mobile regions, each having its own velocity fields and dispersion coefficient, exchange mass. Several closure models have been investigated for the exchange between the two regions. Although most applications include two mobile regions [1, 11, 12, 26, 27, 28, 34, 59], applications with three mobile regions have been reported [33]. Two region models have been tested against numerical experiments [11, 12, 16] and laboratory experiments [29, 30]. They are shown to become equivalent to a single region model with a Fickian behaviour (that is the AD model) in the limit of long times and travel distances [1, 16, 30]. Conversely, they are deemed more accurate than the AD model for small times and highly contrasted hydraulic properties [30].

All these models have shown a good ability to reproduce field- or laboratory-obtained experimental breakthrough curves via a proper parameter tuning. This makes a benchmarking of their respective predictive capabilities very difficult. As shown in [30], tracer tests involving a strong heterogeneity allow for a better model discrimination than tests involving weakly variable porous media. Moreover, pulse tracer tests are also deemed more discriminatory in terms of model response than step injection tests, especially for long time and/or travel distances [30]. However, most experiments report either step tracer tests [38, 39, 46, 52, 58, 61] or very long pulses that may be interpreted as a succession of two steps [56, 58, 30]. A few exceptions are reported in [30, 32, 63].

As shown in a previous publication [48], the AD, FAD and MI models with scale-independent parameters fail to account for the behaviour of experimental breakthrough curves at small space and time scales when the porous medium is strongly heterogeneous and periodic. Two main reasons were identified for this. Firstly, the size of the Representative Elementary Volume (REV) [5] is at least one order of magnitude larger than the spatial period of the Model Heterogeneous Porous Medium (MHPM). Dispersion models are not valid at spatial scales smaller than the REV size. Secondly, a Laplace analysis of the theoretical AD, FAD and MI modelled breakthrough curves [48]

shows that these models yields infinite signal propagation speeds. An infinite concentration wave speed is clearly physically unrealistic. Besides, the finite propagation speed of the concentration signal exerts a strong influence on the behaviour of the experimental breakthrough curves for small times and distances [48], which explains that the above three models are more inaccurate for small times and short distances than for long time and distances. That Fickian-based dispersion models only seem to become more accurate as the spatial scale increases is only due to the fact that the Peclet number increases with distance (therefore, dispersion, albeit modelled wrongly, has a decreasing importance in the modelled signal) [48]. These conclusions are to be extended to the FAD model with superdiffusive behaviour. Indeed, this model is obtained under the assumption of heavy-tailed PDFs for the particle jump length [49], thus allowing for infinite particle velocities. A conclusion of the study [48] is therefore that models where advective processes play a predominant role should be expected to give better results than AD- and FAD-based models at small scales.

The experimental results in [48] also indicate that previously identified scale dependency of the dispersion coefficient may easily be explained by the variability between the replicates of a same experiment.

The objectives of the present paper are the following.

(i) Build a high-quality experimental database for Intermediate Scale Experiments (ISE) of dispersion of tracers in heterogeneous porous media. In [48] it was chosen to build a periodic heterogeneous porous medium made of a series of 15 cm long columns enclosing high permeability conduits surrounded by single-sized glass beads. However, for a single period and two periods, the results were biased by the influence of the inlet and outlet boundary conditions. Consequently, experiments were meaningful for a minimum of three successive periods. In the present work, the experimental setup was revised so that the experiments be meaningful even for a single period.

(ii) Propose a model benchmarking methodology with an enhanced discriminatory power. Bearing in mind the conclusions in [30], the proposed methodology consists in realizing a step injection and using both the breakthrough curve and its time derivative to benchmark the various models. Moreover, for each model, a single parameter set is used to reproduce the experimental signal at all scales. This approach is retained because the ISE [48] shows that there exists a model with scale-independent coefficients that allows the breakthrough curves to be reproduced at all scales (although this model is unknown).

(iii) Benchmark the AD, MI, MR and MRAD models against the experimental breakthrough curves and their time derivatives.

(iv) Determine whether a purely advective multiregion model can provide a viable alternative to models embedding a Fickian or fractional Laplacian description of dispersion, with the

advantage that a purely advective model involves finite signal propagation speeds.

A brief description of the AD, MI, MR and MRAD models for the dispersion of tracers (that is, inert solutes not subjected to degradation, adsorption/desorption and the concentration of which does not influence te flow field) is given in Section 2 as well as the proposed Purely Advective Multiregion (PAMR) model. The experimental setup is described in Section 3. The model benchmarking procedure and experimental results are described respectively in Sections 4. Sections 5 and 6 are devoted to discussion and conclusions.

2. Models

2.1. The AD model

The simplest known model for passive solute transport in porous media is the Advection-Dispersion (AD) model. The governing equation is the following:

$$\partial_t c + u\partial_x c - D\partial_{xx} c = 0 \tag{1}$$

where c is the concentration, u is the flow velocity and D the dispersion coefficient. In the case of a constant input concentration at the upstream boundary, this model yields an S-shaped solution for the concentration and a gaussian-shaped for its derivative with respect to time.

2.2. The MI-MR model

The mobile-immobile (MI) model was first proposed in [64]. In what follows, owing to the assumption of passive transport, the adsorption/desorption terms are cancelled in the governing equations. Using the assumptions of constant water contents for the mobile and immobile regions, the governing equations are simplified into

$$\partial_t c_{\mathrm{m}} + u\partial_x c_{\mathrm{m}} - D\partial_{xx} c_{\mathrm{m}} = \frac{k}{\Theta}(c_{\mathrm{im}} - c_{\mathrm{m}}) \tag{2a}$$

$$\partial_t c_{\mathrm{im}} = \frac{k}{(1-\Theta)}(c_{\mathrm{m}} - c_{\mathrm{im}}) \tag{2b}$$

$$\Theta = \frac{\theta_{\mathrm{m}}}{\theta_{\mathrm{m}} + \theta_{\mathrm{im}}} \tag{2c}$$

where c_{m} and c_{im} are respectively the concentrations in the mobile and immobile regions, θ_{m} and θ_{im} are respectively the water contents of the mobile and immobile regions, Θ is the normalized water content of the mobile fraction, and k is the exchange rate constant between the mobile and immobile regions.

The MI model may be called a single rate model since it contains a single exchange rate constant k between mobile and immobile regions. A generalization of the MI model, called the Multiple Rate (MR) model, was later proposed by [35]. In this model the mobile region can exchange

with multiple immobile regions, each having its own exchange rate constant k_j. Using again the assumption of passive scalar transport, normalizing the water contents of the mobile and immobile regions lead to:

$$\partial_t c_{\mathrm{m}} + u\partial_x c_{\mathrm{m}} - D\partial_{xx} c_{\mathrm{m}} = \sum_{j=1}^{N} \frac{k_j}{\Theta}((c_{\mathrm{im}})_j - c_{\mathrm{m}}) \tag{3a}$$

$$\partial_t (c_{\mathrm{im}})_j = \frac{k_j}{\Theta_j}(c_{\mathrm{m}} - (c_{\mathrm{im}})_j) \;\; j = 1, \ldots, N \tag{3b}$$

Where N is the number of immobile regions with normalized water content Θ_j, and k_j ($j = 1$, ... ,N) are the multi-exchange rate constants between the mobile and the N immobile regions. For the sake of consistency, in this paper the MI and MR models designate the single-rate mobile-immobile model and the multi-rate mobile-immobile model respectively.

2.3. The MRAD model

The Multiple Region Advection-Dispersion (MRAD) model is based on the assumption of R regions flowing in parallel, exchanging mass according to linear kinetics [1, 11, 12, 26, 27, 28, 34, 59]. In each of these regions, the AD model is assumed valid. The original model allows for different heads in the various flow regions, thus allowing for water exchange between the regions in addition to solute exchange. In the present experiments, however, the upstream and downstream sections of each model column are connected to a single inflow and outflow pipe, thus making the head in all regions identical on the scale of the heterogeneity. Consequently, the hydraulic source term between the various regions is set to zero. Bearing in mind the assumption of passive scalar transport, any degradation or adosrption/desorption terms are set to sero. Normalizing the water contents as in the previous subsection leads to the following governing equations:

$$\partial_t c_i + u_i \partial_x c_i - D_i \partial_{xx} c_i = \sum_{j \neq i} \frac{k_{ij}}{\Theta_i} \left(c_j - c_i\right), \;\; i = 1, \ldots, R \tag{4a}$$

$$k_{ij} = k_{ji} \, \forall \, (i, j) \tag{4b}$$

$$\Theta_i = \frac{\theta_i}{\sum_{j=1}^{R} \theta_j}, \;\; \sum_{i=1}^{R} \Theta_i = 1 \tag{4c}$$

and the total concentration in the porous medium is defined as

$$c\left(x,t\right) \equiv \sum_{i=1}^{R} \Theta_i c_i \left(x,t\right) \tag{5}$$

As mentioned in the introduction, most applications involve $R = 2$ regions, with the exception of [33] where $R = 3$ and $k_{ij} \neq 0$ only for $j = i \pm 1$. As shown in AppendixA, the variance $c^{(2)}$ of the solute particle locations obeys the following equation

$$\mathrm{d}_t c^{(2)} = \sum_{i=1}^{R} \Theta_i \left(v_i c_i^{(1)} + 2D_i\right), \;\; v_i \equiv u_i - \sum_{j=1}^{R} \Theta_j u_j \tag{6}$$

where $c_i^{(1)}$ is the average abscissa of the particle locations in the region i, obeying

$$\mathrm{d}_t c_i^{(1)} = v_i + \sum_{j \neq i} \frac{k_{ij}}{\Theta_i} \left(c_j^{(1)} - c_i^{(1)} \right) \tag{7}$$

As shown in AppendixB (Result 3), all the $c_i^{(1)}(t)$ tend to a limit value $c_i^{(1,\infty)}$ as t tends to infinity provided that the vector $\mathbf{v} \equiv [v_1, \ldots, v_R]$ belongs to the range of the matrix $\mathbf{M}$ defined as

$$M_{ij} = \begin{cases} -\sum_{p \neq i} \frac{k_{ip}}{\Theta_i} & \text{if} \quad i = j \\ \frac{k_{ij}}{\Theta_i} & \text{if} \quad i \neq j \end{cases} \tag{8}$$

If this is the case, $d_t c^{(2)}$ tends to a finite, limit value for long times, which is a Fickian (or "normal") behaviour. The limit value of the dispersion coefficient is shown to be (AppendixA)

$$D_\infty = \sum_{i=1}^{R} \Theta_i \left(D_i + \frac{1}{2} v_i c_i^{(1,\infty)} \right) \tag{9}$$

If the vector $\mathbf{v}$ does not belong to the range of $\mathbf{M}$, the $c_i^{(1)}$ do not converge and a superdiffusive (anomalous) behaviour is obtained.

A direct consequence of the above property is the following (AppendixB, Result 4): if r regions exchange mass with each other but do not exchange mass with the $R - r$ remaining regions, all regions can be renumbered in such a way that $k_{ij} = 0$ for $(i, j) \in \{1, \ldots, r\} \times \{r + 1, \ldots, R\}$. Then, the matrix $\mathbf{M}$ is block-diagonal and a necessary condition for the asymptotic behaviour of the dispersion process to be Fickian is that the average speed of the first r regions be the same as that of the remaining $R - r$ regions. If $\sum_{i=1}^{r} \Theta_i u_i \neq \sum_{i=1}^{R} \Theta_i u_i$, the $c_i^{(1)}(t)$ are not all bounded and a superdiffusive asymptotic behaviour is obtained.

2.4. Purely Advective Multiple Region (PAMR) model

The PAMR model proposed in this paper is based on the consideration that dispersion is a purely advective process on the microscale. The Fickian model arises only as an asymptotic property of Brownian movement [19, 20, 21, 43, 49] that is valid for a large number of solute particle displacements (in other words, for large space and time scales compared to the typical duration and length of the Brownian motion jumps). Although the term "Brownian" was originally used to designate the movement of small particles suspended in fluids, its meaning has broadened with time. The adjective "Brownian" is widely used as an equivalent to "Wiener process", a continuous but non-differentiable [10] random process with zero mean and variance proportional to time. Generalizing the concept has led to that of fractional Brownian motion, a useful concept for anomalous diffusion modelling [49]. The application fields may be totally disconnected from physics, as in e.g. financial mathematics. The first mathematical formalization of Brownian motion is attributed to Bachelier, with his thesis on the mechanisms of stock exchange [3]. Bearing this in

mind, the term "Brownian" is used in the present paper to designate a Markovian random particle displacement process. The Brownian character of solute movement in random, heterogeneous media is easily justified by considering the individual trajectories of solute molecules, that are assumed to travel at the same velocity as the surrounding fluid molecules (Figure 1a). If the pores of the medium are assumed to have random and isotropic orientation, size and spacing, the velocity field can also be assumed random at the scale of the grains/pores. Consequently, the cumulated movement of the fluid molecules in a coordinate system moving at the average fluid velocity may be considered random at this scale (Figure 1). Given the random pore orientation, the direction of the movement of a particle is totally uncorrelated from one pore to the next. This corresponds to the mathematical definition of the standard Brownian motion, whereby the particle displacement is (i) an isotropic, random function of space (ii) a Markov process that assumes zero correlation between successive Brownian displacements. If the medium is not totally random, however, the Brownian assumption may not hold any more. A structured medium exhibits a certain degree of periodicity (Figure 1c). Such periodicity may be encountered at the scale of the pore. It may also arise at much larger scales, as in the case of e.g. low permeability lenses in an aquifer. In such a case, the flow field cannot be considered totally random. Therefore, the cumulated particle displacements in the coordinate system moving at the average flow velocity are also periodic to some extent (Figure 1d).

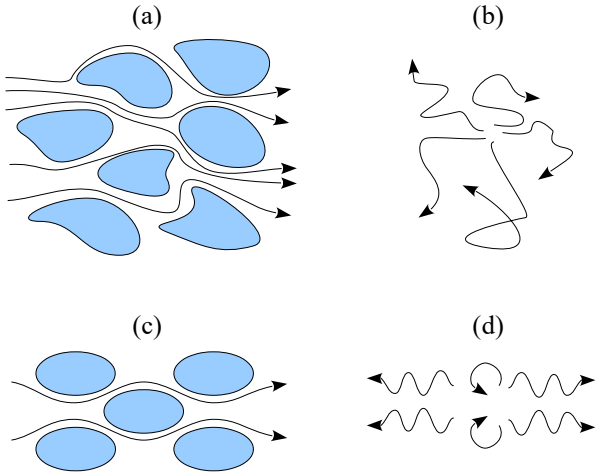


Figure 1: Solute dispersion as the result of a Brownian process. Random heterogeneous medium: (a) solute particle trajectories in the laboratory reference frame, (b) trajectories plotted in the coordinate system moving at the average solute speed. Periodic medium: (c) trajectories in the laboratory coordinate system, (d) trajectories in the coordinate system moving at the average solute speed.

In the extreme, totally periodic case of Figures 1c-d, the cumulated (relative) particle displacements are totally periodic, with a maximum vertical amplitude equal to one half of the vertical spatial period. Some particles travel to the left because they are slower than the mean flow velocity on the average. Some other particles travel to the right because their average velocity is larger than the mean flow velocity. However, despite their diverging character, the trajectories of the particles remain deterministic because the medium is strictly periodic. One may expect a certain

amount of randomness in these trajectories because of molecular diffusion, but a long time may be necessary for the random component of the movement to predominate over the deterministic component. Real-world situations, including the experiments reported in the present paper, lie somewhere between these two extreme configurations. In the light of the considerations above, a conceptually satisfactory scale-independent model for dispersion should be expected to satisfy the following two requirements.

(R1) No Fickian fluxes should be included in the governing equations. This requirement stems from the reasoning that (i) the simplest possible scale-independent dispersion model is sought, (ii) the Fickian model has been shown in [48] to yield undesirable behaviours for small times and travel distances, (iii) consequently, Fickian fluxes must be ruled out. As far as point (ii) is concerned, the Fickian model yields infinite wave propagation speeds for the modelled solute front, an unphysical behaviour yielding zero arrival times in solute breakthrough experiments. In contrast, the delay between the inlet and the outlet was pointed out as an essential feature of the experimental breakthrough curves reported in [48].

(R2) The model should yield the Fickian behaviour as a limit, asymptotic case for large times and distances. Such a behaviour has indeed been confirmed experimentally in the case of homogeneous media, as well as periodic heterogeneous media provided a sufficient number of periods is covered [48].

The conceptual model proposed hereafter aims to fulfill these two requirements, based on the following two assumptions.

(A1) The flow velocity within an averaging volume is partitioned into R regions over which the flow velocity is homogeneous. The flow region i has a normalized water content $\Theta_i, \sum_{i=1}^{R} \Theta_i = 1$. Under steady state flow conditions, the normalized water contents Θ_i and the flow velocities u_i are constant and uniform.

(A2) Two adjacent regions may exchange solute particles owing to the random velocity distribution within the porous medium (Figure 1b). Consequently, the concentrations within two adjacent regions tends to even out with time. The solute exchange rate between two regions is assumed proportional to the difference between the solute concentrations in the two regions.

Assumptions (A1-2) lead to the following governing equation:

$$\partial_t \left(\Theta_i c_i\right) + \partial_x \left(\Theta_i u_i c_i\right) = \sum_{j \neq i} k_{ij} \left(c_j - c_i\right), \; i = 1, \ldots, R \tag{10}$$

The advective part of the model (left-hand side of the equation) stems from assumption (A1). The source term (right-hand side of the equation) is the simplest possible formulation satisfying assumption (A2).

A first advantage of this model over those reported in the previous subsections is that it satisfies the two requirements (R1-R2). Indeed, the transport term in equation (10) is purely advective, thus satisfying (R1). Moreover, the Fickian model is obtained as a limit case for asymptotically long times and travel distances. This is easily shown by noticing that the model (10) is a particular case of the MRAD model (4a). Equation (10) is obtained by setting $D_i = 0$ in equation (4a) and replacing $\frac{k_{ij}}{\Theta_i}$ with $k_{ij}\Theta_i\Theta_j$. As shown by equation (9), a finite dispersion coefficient D_∞ may be obtained even though the coefficients D_i are all zero. A sufficient condition for this is that the vector v belong to the range of $\mathbf{M}$ (see Appendices A-B for the details of the derivation).

Another interesting feature of the PAMR model is that it allows for (ballistic) anomalous dispersion for small times, which is compatible with the apparent increase in the Fickian-based dispersion coefficient with observation scale reported in a number of studies.

Therefore, the PAMR model is considered more satisfactory from a conceptual point of view than the MRAD model because it does not have the drawbacks of Fickian dispersion models for small times and/or distances, while retaining its advantages for large times and/or distances. Nevertheless, the increased accuracy of the model for small times/distances is achieved at the expense of model parsimony. As shown in AppendixA, at least two regions are needed to obtain an asymptotic Fickian behaviour for dispersion. Then, the model has three independent parameters: one of the normalized water contents Θ_1, Θ_2, one of the flow velocities v_1, v_2 and the exchange parameter k_{12}. In contrast, the AD model requires only two parameters (the flow velocity v and the dispersion coefficient D), for the same asymptotic behaviour. The increased number of parameters was to be expected in that the Fickian behaviour is only an asymptotic property of the model. The additional parameters control the characteristic time/distance above which Fickian behaviour bceomes a satisfactory approcimation of the dispersion process.

3. Experimental setup and results

3.1. Experimental setup

The MHPM consists of a PVC column (10 cm in diameter, 15 cm in length) containing a cylindrical cavity (2.5 cm in diameter, 10 cm in length) placed in the centre of the column and surrounded by 1 mm glass spheres. For more details about the construction of the MHPM, please see [48]. The inflowing discharge is supplied using a peristaltic pump (Gilson MP3$^{\mathrm{TM}}$). Step tracing experiments are done by injecting salty water (deionised water + NaCl at $C_0 = 0.1$Mol/L) into the columns initially containing deionised water. A flow of 7.5 L/h salty water is induced into the study column until the outlet concentration c stabilizes to C_0. The outlet concentration is measured using a conductimeter (WTW TetraCon 325$^{\mathrm{TM}}$) and saved on a data logger (Campbell CR1000$^{\mathrm{TM}}$) every 5 s. The cumulated outlet volume V is measured by weighing the effluent every 5 s and saved on the data logger. The study column consists of a series of N connected MHPM.

In this paper, six series were investigated with $N = 1$ to 6. We used the 12 MHPM columns described in [48]. For each series, four replicates of the step tracing experiment were conducted for four 90° rotations of the study column. The purpose was to eliminate biases arising from possible asymmetry in the column geometry and density effects. The various column combinations used for the various experiments are summarized in Table 1. A mean breakthrough curve is deduced from all the replicates.

N	V_0 (L)	L (m)	Column groups	Total replicates
1	0.461	0.15	A, B, C, D, E, F, G, H, I, J, K, L	48
2	0.922	0.30	AB, CD, EF, GH, IJ, KL	24
3	1.383	0.45	ABC, DEF, GHI, JKL	16
4	1.844	0.60	ABCD, EFGH, IJKL	12
5	2.305	0.75	ABCDE, FGHIJ	8
6	2.766	0.90	ABCDEF, GHIJKL	8

Table 1: Experiment replicates. N is the number of columns, L is the total length of the porous medium, V_0 is the pore volume.

The main difference between the present experimental setup and that in [48] is the MHPM connection pattern (Figure 2). In [48], the first and last MHPM in the study column could not be considered as periodical heterogeneities because they had different flow inlet and outlet connections (Figure 2, top): the first MHPM had a divergent flow inlet and a parallel flow outlet while the last MHPM had a parallel flow inlet and a convergent flow outlet. In the present experiment, each MHPM of the study column can be considered as a single periodical heterogeneity because all MHPM have identical inlet and outlet connections (Figure 2 bottom): a divergent flow inlet and a convergent flow outlet. The advantage of the present setup is that the breakthrough curve can be obtained directly for a single heterogeneity ($N = 1$).

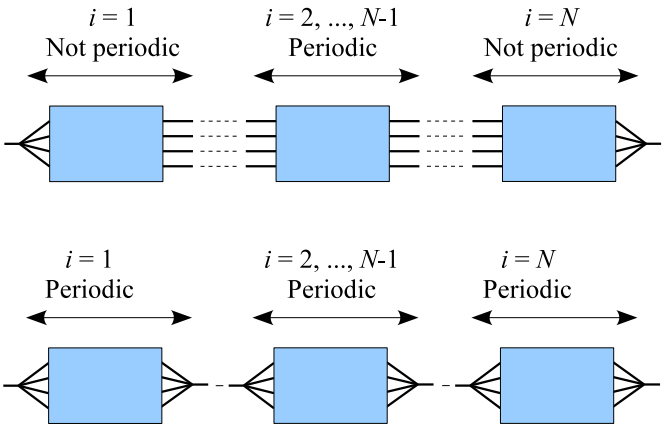


Figure 2: Definition sketch for column connection. Top: experimental setup reported in [48]. Bottom: present experimental setup.

3.2. Experimental breakthrough curves

Figure 3 shows the experimental breakthrough curves for $N = 1$ to $N = 6$ MHPM. As expected from [48], the breakthrough concentration signal tends to the classical S-shaped solution of the AD

model as N increases. In the present experiments, this behaviour is achieved for fewer MHPMs than in [48] ($N = 5$ instead of 10). This is attributed to the different injection geometry (compare Figure 2, top and Figure 2, bottom). Plotting the time derivative of the normalized concentration signal allows two main transport modes to be identified. From the location of the peaks on Figure 3 for $N = 1$, the faster mode travels at approximately 3×10^{-3} ms^{-1}, while the speed of the slower one is approximately 7×10^{-4} ms^{-1}. As time (and distance) grows, the relative amplitude of the faster peak decreases and only one peak can be detected in the $\partial c/\partial t$ signal for $N = 5, 6$. Note that the time derivatives of the concentration signal for N= 1 to 3 is strikingly similar to experimental propagators obtained in heterogeneous porous media [50, 57] and replicated by pore-scale modelling [10]. In contrast, the $c(t)$ signal makes these two modes more difficult to detect, even at early times. For this reason, the time derivative $\partial c/\partial t$ of the concentration signal is used all throughout this manuscript for model benchmarking.

The enhanced discriminatory power of the time derivative of the $c(t)$ signal over the signal itself should not come as a surprise. Since the injection signal is a concentration step, using its time derivative $\partial c/\partial t$ is equivalent to performing a breakthrough experiment using a Dirac (pulse) input signal. From the point of view of the frequency domain analysis, the Laplace/Fourier transforms of the Dirac signal gives an equal weight to all frequencies, while the Laplace/Fourier transform of the step function is the inverse of the frequency, thus giving less importance to higher frequencies. Since our earlier experiments [48] showed that high frequencies are essential in discriminating between models, the Dirac injection should be preferred. Such an input signal, however, is extremely difficult to generate with a good control on experimental conditions. Using the time derivative $\partial c/\partial t$ with a step injection is an efficient way of obviating this difficulty.

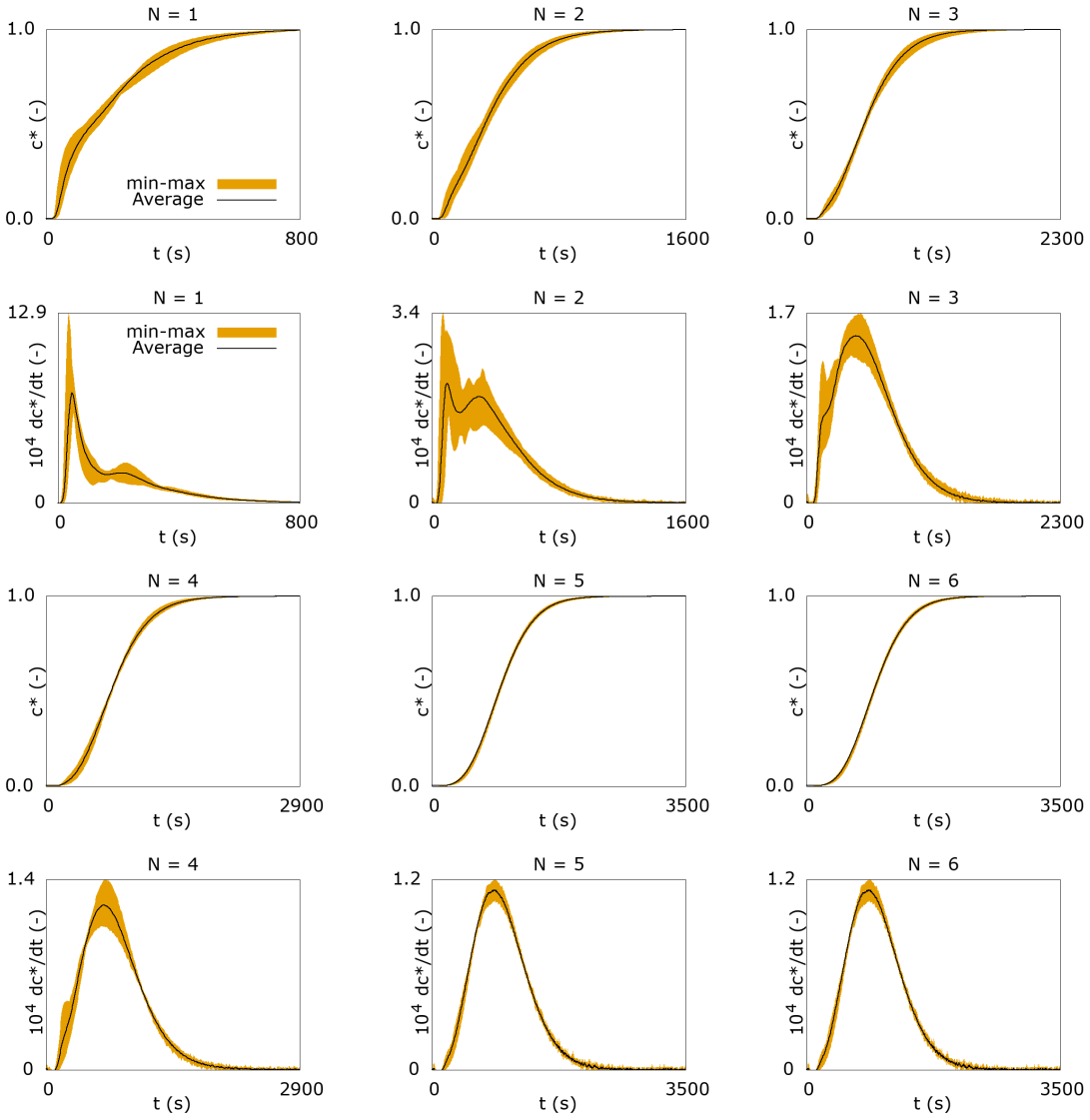


Figure 3: Experimental breakthrough curves and their time derivatives for $N = 1, \ldots, 6$ MHPM.

4. Model benchmarking

4.1. Calibration method

Objective function. As discussed in the introduction, a sound benchmarking should allow for an efficient discrimination of both the long-time and short-time behaviours of the various models against experimental datasets. In [48], the short-time behaviour was enhanced by using the Laplace transforms of the experimental breakthrough curves. The Laplace transform of the signal was used only because the unit transfer function of a single column was not accessible from the data. Since, in the present experiments, the $c(t)$ signal is available for $N = 1$, the Laplace transform can be avoided. The governing equations were solved in the (x, t) domain. A second-order, conservative finite volume method [60] was used. This method has proved less diffusive than the second-order-in-time MUSCL-Hancock method for small values of the Courant-Friedrichs-Lewy (CFL) number. In order to minimise numerical diffusion as much as possible, the computational time step was adapted in usch a way that the CFL in the fastest flow region was equatl to unity. For CFL $= 1$,

numerical diffusion is known to be zero and the exact solutionof the advection part of the equation is obtained.Moreover, different mesh sizes were tested. Varying the computational cell size from $\Delta x = 1.5 \times 10^{-5}$ m (i.e. 1000 cells per 15 cm long column) to $\Delta x = 1.5 \times 10^{-4}$ m (i.e. 100 cells per column) showed no noticeable difference between the numerical solutions, showing that numerical convergence was achieved.

Let $e(t)$ be the model error, that is, a measure of the difference between the model output and the experimentally measured concentration (see next paragraph for $e(t)$ definitions). The usual approach consists in computing the objective function as the Lp-norm of the modelling error over the time interval $[0, T]$ of interest. Dividing by the length of the time integration interval so as to remove the bias arising from the length of the time series yields the following objective function

$$J_1 = \left(\frac{1}{T} \int_0^T |e(t)|^p \, \mathrm{d}t \right)^{\frac{1}{p}}, \; p > 0 \tag{11}$$

In practice, the model output is discretized using a time step Δt over the interval $[0, T]$, with $T = n\Delta t$. In the case of a constant Δt, the trapezium rule leads to the following estimate:

$$J_1 = \left(\frac{1}{T} \sum_{j=1}^{n} \left(\frac{e_j + e_{j+1}}{2} \right)^p (t_{j+1} - t_j) \right)^{1/p} = \left(\frac{1}{n} \sum_{j=1}^{n} \left(\frac{e_j + e_{j+1}}{2} \right)^p \right)^{1/p} \tag{12}$$

where e_j is the error at time t_j. When the simulation is carried out over a long time, however, the objective function J_1 gives a small relative weight to the early times of the model response. In order to enhance the discriminatory power of early simulated times, it is proposed that the objective function be computed not by integrating with respect to time but with respect to frequency. The lower and upper bounds of the frequency are respectively $\nu_0 \equiv \frac{1}{T}$ and $\nu_1 \equiv \frac{1}{\Delta t}$. The frequency-based objective function is defined as

$$J_2 = \left(\frac{1}{\nu_1 - \nu_0} \int_{\nu_0}^{\nu_1} |e(\nu)|^p \, \mathrm{d}\nu \right)^{\frac{1}{p}} = \left(\frac{1}{\nu_1 - \nu_0} \int_{\Delta t}^{T} \frac{1}{t^2} |e(t)|^p \, \mathrm{d}t \right)^{\frac{1}{p}} \tag{13}$$

The factor $\frac{1}{t^2}$ illustrates the stronger weight given to small times. The trapezium rule leads to the following formula for J_2

$$J_2 = \left[\frac{1}{\nu_1 - \nu_0} \sum_{j=1}^{n-1} \left(\frac{e_j + e_{j+1}}{2} \right)^p \left(\frac{1}{t_j} - \frac{1}{t_{j+1}} \right) \right]^{1/p} = \left[\frac{n}{n-1} \sum_{j=1}^{n-1} \left(\frac{e_j + e_{j+1}}{2} \right)^p \left(\frac{1}{j} - \frac{1}{j+1} \right) \right]^{1/p} \tag{14}$$

Since the objective is to achieve a correct description of both short- and long-time behaviours, the final objective function is defined as the product $J_1 J_2$.

Modelling error definition. The measure of the modelling error was defined on the basis of the following considerations: (i) the modelling error is zero whenever the modelled variable is equal

to the measured one, (ii) the error must be normalized with the amplitude of the experimental variability, assessed from several replicates of the experiment. These two conditions allow at least two possible definitions to be proposed for the modelling error.

Definition 1: the error is zero when the modelled signal is within the min-max interval of the experimental signal. When the modelled signal is outside the experimental min-max interval, the error is taken equal to the distance to the closer interval bound:

$$e\left(t\right)=\begin{cases}\frac{\partial_t c_{\min}(t)-\partial_t c_{\mathrm{mod}}(t)}{\Delta\partial_t c} & \text{if} \quad \partial_t c_{\mathrm{mod}}(t)<\partial_t c_{\min}(t)\\ 0 & \text{if} \quad \partial_t c_{\min}(t)\leq\partial_t c_{\mathrm{mod}}(t)\leq\partial_t c_{\max}(t)\\ \frac{\partial_t c_{\mathrm{mod}}(t)-\partial_t c_{\max}(t)}{\Delta\partial_t c} & \text{if} \quad \partial_t c_{\mathrm{mod}}(t)>\partial_t c_{\max}(t)\end{cases} \tag{15a}$$

$$\partial_t c_{\min}\left(t\right)\equiv\min_r \partial_t c_{\mathrm{expe},r}\left(t\right),\ \ \partial_t c_{\max}\left(t\right)\equiv\max_r \partial_t c_{\mathrm{expe},r}\left(t\right),\ \ \Delta\partial_t c=\partial_t c_{\max}(t)-\partial_t c_{\min}(t) \tag{15b}$$

where $\partial_t c_{\mathrm{expe},r}\left(t\right)$ is the derivative of the concentration measured at time t for the replicate number r of the experiment, $c_{mod}\left(t\right)$ is the modelled concentration at time t. In this definition, the error is zero whenever the modelled signal can be explained by the experimental variability between the various replicates. Numerical experiments carried out by reducing articifially the min-max range showed very little sensitivity of the respective performance of the various models considered in this paper.

Definition 2: the error is defined as the distance from the average measurement. In this definition the scaling factor is $\sigma(t)$, the standard deviation of $\partial_t c(t)$ between the various replicates:

$$e\left(t\right)=\begin{cases}\frac{\partial_t c_{\mathrm{av}}(t)-\partial_t c_{\mathrm{mod}}(t)}{\sigma(t)} & \text{if} \quad \partial_t c_{\mathrm{mod}}(t)\leq\partial_t c_{\mathrm{av}}(t)\\ \frac{\partial_t c_{\mathrm{mod}}(t)-\partial_t c_{\mathrm{av}}}{\sigma(t)} & \text{if} \quad \partial_t c_{\mathrm{mod}}(t)\geq\partial_t c_{\mathrm{av}}(t)\end{cases} \tag{16a}$$

$$\partial_t c_{\mathrm{av}}(t)\equiv\frac{1}{M}\sum_{r=1}^{M}\partial_t c_{\mathrm{expe},r}(t),\ \ \sigma(t)\equiv\left[\frac{1}{M-1}\sum_{r=1}^{M}\left(\partial_t c_{\mathrm{expe},r}(t)-\partial_t c_{\mathrm{av}}(t)\right)^2\right]^{\frac{1}{2}} \tag{16b}$$

Calibration algorithm. In the following, the calibration is performed on five experiments simultaneously ($N=1\ldots5$) using the definition (16a) for the error. The final objective function is taken as the sum for each experiment of the product of the two functions J_1 and J_2:

$$J=\sum_{p=1}^{5}J_1(p)J_2(p) \tag{17}$$

A binary Genetic Algorithm (GA) [36] is used to calibrate the different models. In this kind of algorithm, the set of parameters is represented by a "chromosome" in which each parameter is a "gene". The gene is described by a min/max interval and a binary encoding of a given length that determine the number of possible values taken by each of the N_p parameters (e.g N_{bits}= 6 bits = 64 values). As proposed in [47], the initial population is composed of $N_{\mathrm{pop}}=8$ chromosomes,

each of which is the concatenation of the various genes. The population is thus represented by a $N_{\text{pop}} \times (N_p N_{\text{bits}})$ matrix initially filled with randomly generated one and zero bits. A simulation is run for each chromosome and the results are compared in term of objective function. Half of the chromosomes that yield to the poorest results are discarded through natural selection and replaced with new offsprings. The best chromosome is kept intact, and the others are subjected to mutation (see [47] for a detailed description).

A first set of parameter bounds is defined, with physically permissible values according to the experiments. The genetic algorithm searches for the best parameter values within 64 possibilities. To ensure a wide coverage of the parameter space, the GA is run during 50,000 generations. Accuracy is improved by running the algorithm a second time over a narrower interval defined about the result of the first run, with a width about 15% that of the initial interval.

The models are calibrated using the five experiments ($N = 1$ to 5) and validated using the sixth experiment ($N = 6$). For the sake of conciseness, the results are displayed for $N = 1$, 3 and 5 only.

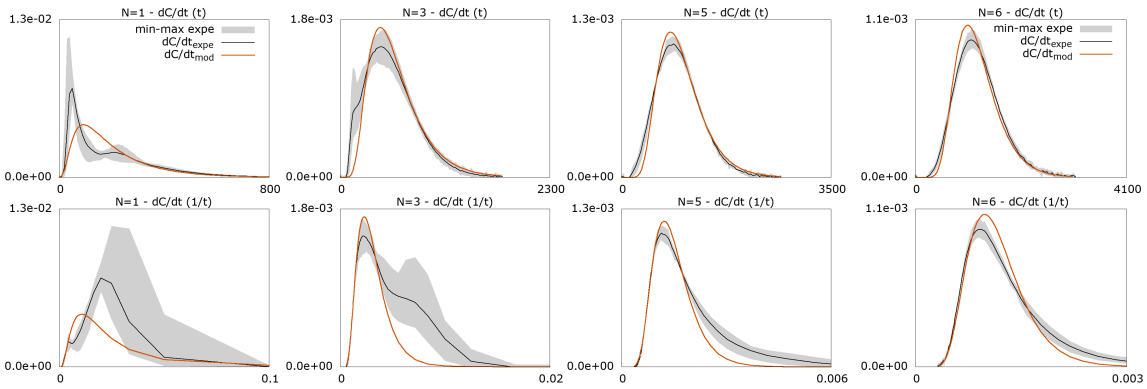
4.2. AD model

The best parameter values obtained for the AD model (Figure 4) are given in Table 2.

Parameter	Meaning	Numerical value
D	Fickian dispersion coefficient	4.041×10^{-5} m^{2}s^{-1}
u	Advection velocity	7.193×10^{-4} ms^{-1}
J	Objective function	7.76

Table 2: AD model. Calibration results.

As expected, the AD model is not able to reproduce the experimental results for $N < 5$. The modelling results improve as N increases. Owing to the single advection velocity, the model fails to represent the main two transport models at small times and distances. This is particularly visible in the frequency domain (Figure 4, bottom), where the higher frequencies are inaccurately accounted for. While the model seems to perform correctly for $N = 5$, the plot in the frequency domain shows that high frequencies remain underestimated by the model, since the model signal is not within the min/max confidence interval of the experimental curve. As N increases, however, the discrepancy between the modelled and experimental signals becomes smaller. This was expected from the analysis presented in the Appendices, because the Fickian dispersion model is more appropriate for large travel times and distances than for short ones.



(a) Calibration (b) Validation

Figure 4: a) Calibration of the AD model. Top: time derivative of the concentration with respect to time. Bottom: time derivative of the concentration as a function of frequency. From left to right: best result for each experiment, obtained with the same set of parameters. b) Model validation against the 6-MHPM experiment.

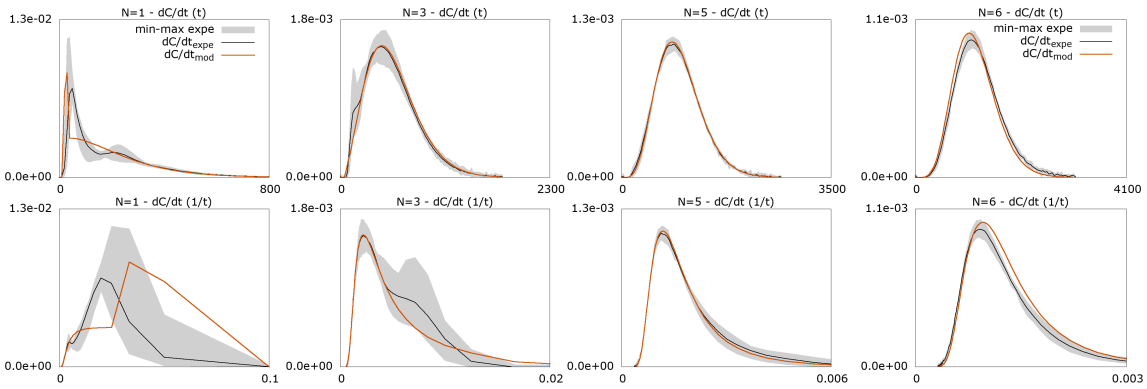
4.3. MI model

The four parameters to be calibrated for the MI model (2a, 2b) are Θ_{m}, k, u and D. The remaining parameter is determined automatically, $\Theta_{\mathrm{im}} = 1 - \Theta_{\mathrm{m}}$. The simulation results after calibration (Figure 5), are obtained with the parameter set in Table 3.

Parameter	Meaning	Numerical value
D	Fickian dispersion coefficient	3.720×10^{-7} m^{2}s^{-1}
k	Exchange coefficient	7.751×10^{-2} s^{-1}
u	Advection velocity	4.471×10^{-3} ms^{-1}
Θ_{im}	Normalised immobile fraction	0.8333
Θ_{m}	Normalised mobile fraction	0.1667
J	Objective function	3.505

Table 3: MI model. Calibration results.

In contrast with the AD model, the MI model produces too early a peak time. The inflection points in the time signal for $N = 1$ and in the frequency signal for $N = 3$ are missed. For $N = 5$, the MI model results are within the bounds of the min/max confidence interval, which is an improvement over the AD model (compare Figs. 4, 5, second column from the right). The validation simulation (Figure 5, rightmost column) also shows a better agreement with the experimental curve than that of the AD model.



(a) Calibration (b) Validation

Figure 5: a) Calibration of the MI model. Top: time derivative of the concentration with respect to time. Bottom: time derivative of the concentration as a function of frequency. From left to right: best result for each experiment, obtained with the same set of parameters. b) Model validation against the 6-MHPM experiment.

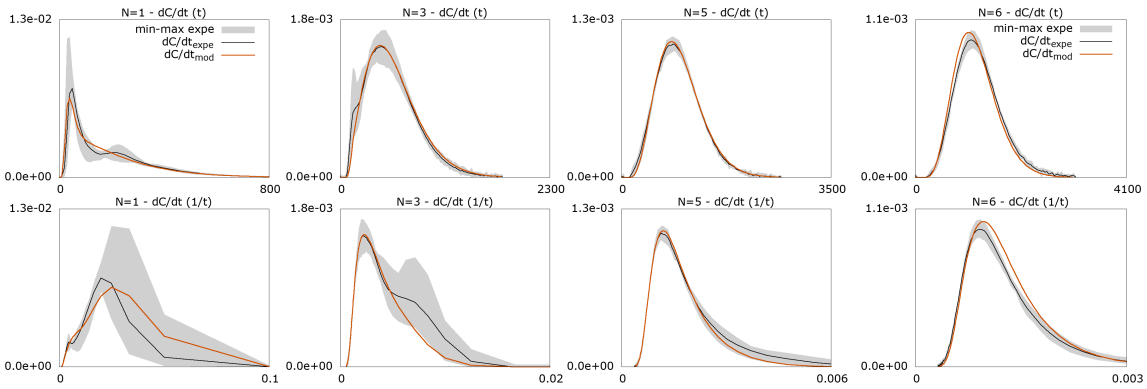
4.4. MR model

Calibrating the MR model (3a, 3b) with three immobile regions gives the values in Table 4. Figure 6 shows the corresponding breakthrough curves.

Parameter	Meaning	Numerical value
D	Fickian dispersion coefficient	$6.200 \times 10^{-7}\mathrm{m^2s^{-1}}$
k_1	Exchange coefficient with region 1	9.303×10^{-1} $\mathrm{s^{-1}}$
k_2	Exchange coefficient with region 2	6.820×10^{-2} $\mathrm{s^{-1}}$
k_3	Exchange coefficient with region 2	6.519×10^{-2} $\mathrm{s^{-1}}$
u	Advection velocity	4.471×10^{-3} $\mathrm{ms^{-1}}$
Θ_1	Normalised immobile fraction 1	0.1017
Θ_2	Normalised immobile fraction 2	0.1155
Θ_3	Normalised immobile fraction 3	0.5775
Θ_m	Normalised mobile water content	0.1658
J	Objective function	1.975

Table 4: RI model. Calibration results.

While the model does not allow the inflection points in the time- and frequency-domain signals to be reconstructed for $N = 1, 3$, the modelled signal lies within the min/max confidence interval for all calibration runs $N = 1, \ldots, 5$. The validation run ($N = 6$) produces slightly too early a signal, just as the MI model.



(a) Calibration (b) Validation

Figure 6: a) Calibration of the MR model with three immobile regions. Top: time derivative of the concentration with respect to time. Bottom: time derivative of the concentration as a function of frequency. From left to right: best result for each experiment, obtained with the same set of parameters. b) Model validation against the 6-MHPM experiment.

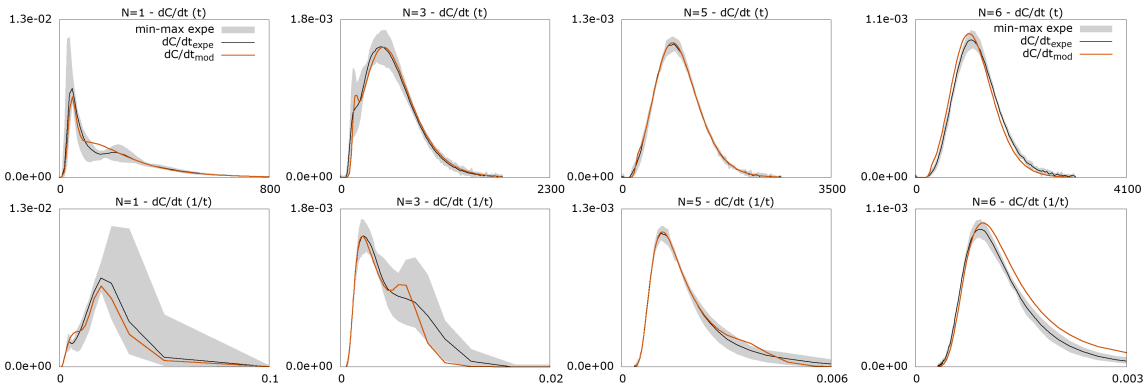
4.5. Two region MRAD Model

The simplest possible version of the MRAD model, with two mobile regions, is used. This model has eight parameters in total, but only six of them are independent and must be calibrated: Θ_1, k_{12}, u_1, u_2, D_1 and D_2. The best parameter set (Figure 7) is given in Table 5.

Parameter	Meaning	Numerical value
D_1	Fickian dispersion coefficient in region 1	1.034×10^{-5} m^2 s^{-1}
D_2	Fickian dispersion coefficient in region 2	1.897×10^{-5} m^{2}s^{-1}
k_{12}	Exchange coefficient	1.570×10^{-2} s^{-1}
u_1	Flow velocity in region 1	2.406×10^{-3} ms^{-1}
u_2	Flow velocity in region 2	5.28×10^{-4} ms^{-1}
Θ_1	Normalised water content 1	0.124
Θ_2	Normalised water content 2	0.876
J	Objective function	1.61

Table 5: MRAD model. Calibration results.

Unlike the AD, MI and MR models, the MRAD model allows the inflection points in the experimental signals to be reproduced. Although the slower peak is too early for $N = 1$ (Figure 7, leftmost column), its location is correct for $N = 3$. In contrast, the faster peak is overestimated for $N = 3$. The small time/high frequency behaviour of the MRAD validation run ($N = 6$, rightmost column on Figure 7) is slightly less accurate than that of the MR model.



(a) Calibration (b) Validation

Figure 7: a) Calibration of the MRAD model with two mobile regions. Top: time derivative of the concentration with respect to time. Bottom: time derivative of the concentration as a function of frequency. From left to right: best result for each experiment, obtained with the same set of parameters. b) Model validation against the 6-MHPM experiment.

4.6. Three region PAMR model

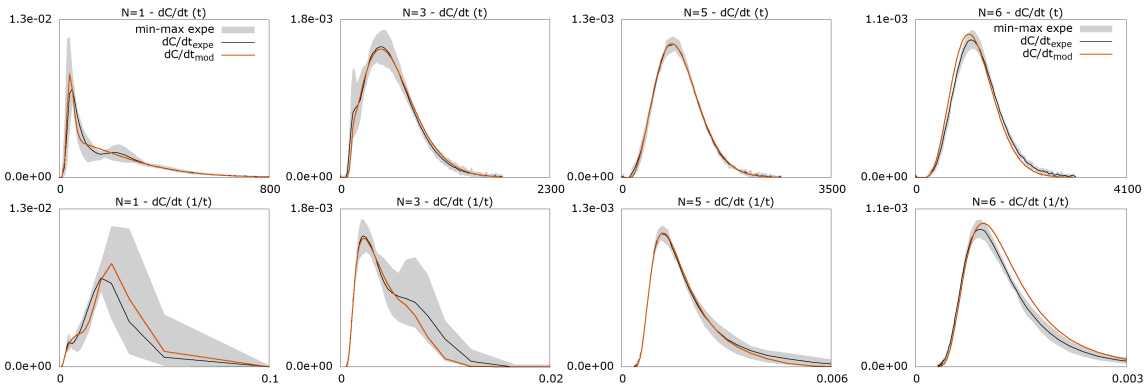
As far as the PAMR model is concerned, three regions represent the minimal level of complexity for which the model can be expected to bring an improvement over the previously tested models. This is because a two region PAMR model would be only a particular case of the two region MRAD model. In the previous paragraph, the optimal MRAD model is shown to have non-zero dispersion coefficients D_1 and D_2. Consequently, a two region PAMR model (with null dispersion) can only perform worse than the two region MRAD. Proposing a PAMR model is thus meaningful only for $R > 2$.

The three region PAMR model requires the calibration of 6 independent parameters: Θ_1, Θ_2, k_{12}, k_{13}, k_{23}, u_1, u_2 and u_3. The best parameter set (Fig 8) is given in Table 6.

As far as the calibration phase is concerned, the three region PAMR model is the only one to fit within the min/max experimental confidence intervals (albeit passing very near the upper limit for $N = 1$). For $N = 3$, the modelled signal is improved significantly over that of the MRAD model in both the time and frequency domains. As with he previous models, the modelled validation signal is slightly too early compared to the experimental one. The large frequency behaviour of the PAMR model is better than that of the MRAD model (compare Figs. 7, 8, rightmost columns).

Parameter	Meaning	Numerical value
k_{12}	Exchange coefficient between regions 1 and 2	1.518×10^{-2} s^{-1}
k_{13}	Exchange coefficient between regions 1 and 3	1.252×10^{-3} s^{-1}
k_{23}	Exchange coefficient between regions 2 and 3	5.803×10^{-3} s^{-1}
u_1	Flow velocity in region 1	3.942×10^{-3} ms^{-1}
u_2	Flow velocity in region 2	8.668×10^{-4} ms^{-1}
u_3	Flow velocity in region 3	1.079×10^{-4} ms^{-1}
Θ_1	Normalised water content 1	0.1389
Θ_2	Normalised water content 2	0.1430
Θ_3	Normalised water content 3	0.7271
J	Objective function	1.592

Table 6: Three region PAMR model. Calibration results.



(a) Calibration (b) Validation

Figure 8: a) Calibration of the PAMR model with three mobile regions. Top: time derivative of the concentration with respect to time. Bottom: time derivative of the concentration as a function of frequency. From left to right: best result for each experiment, obtained with the same set of parameters. b) Model validation against the 6-MHPM experiment.

4.7. Four region PAMR model

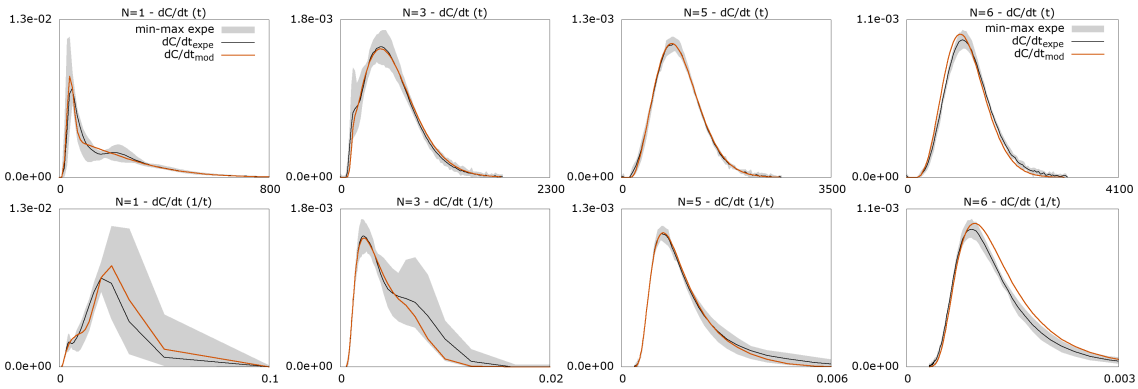
A four region PAMR model is also tested, with the following two objectives: (i) determine whether adding extra degrees of freedom allow for better calibration/validation results, (ii) investigate whether each of the two dispersion coefficients in the two region MRAD model could not be represented more efficiently using two regions exchanging mass. The rationale for this is presented in Subsection AppendixA.2.4. In this subsection, an asymptotic Fickian behaviour is shown to be obtained from two regions with identical fractions $\Theta_p (p = 1, 2)$ exchanging mass. The consequence is that each of the two regions in the MRAD model might be subdivided into two regions of equal size, resulting in a four region PAMR model. In this version, eight independent parameters must be calibrated: Θ_1, k_{12}, k_{13}, k_{34}, u_1, u_2, u_3 and u_4. The remaining coefficients are inferred automatically because $\Theta_2 = \Theta_1$, $\Theta_3 = \Theta_4 = \frac{1-\Theta_1}{2}$, and $k_{ij} = k_{ji} \forall (i, j)$.

The best result (Fig. 9) is obtained for the parameter set in Table 7.

Parameter	Meaning	Numerical value
k_{12}	Exchange coefficient between regions 1 and 2	$2.180\ \mathrm{s}^{-2}$
k_{13}	Exchange coefficient between regions 1 and 3	$1.561 \times 10^{-3}\ \mathrm{s}^{-1}$
k_{34}	Exchange coefficient between regions 3 and 4	$2.190 \times 10^{-1}\ \mathrm{s}^{-1}$
u_1	Flow velocity in region 1	$4.474 \times 10^{-3}\ \mathrm{ms}^{-1}$
u_2	Flow velocity in region 2	$7.624 \times 10^{-4}\ \mathrm{ms}^{-1}$
u_3	Flow velocity in region 3	$2.228 \times 10^{-5}\ \mathrm{ms}^{-1}$
u_4	Flow velocity in region 4	$1.448 \times 10^{-5}\ \mathrm{ms}^{-1}$
$\Theta_1 = \Theta_2$	Normalised water contents 1 and 2	0.1277
$\Theta_3 = \Theta_4$	Normalised water contents 3 and 4	0.3723
J	Objective function	1.608

Table 7: Four region PAMR model. Calibration results.

Comparing Figures 8-9 shows that the three and four region PAMR models give almost indiscernible results. The objective functions are also extremely similar.



(a) Calibration (b) Validation

Figure 9: a) Calibration of the PAMR model with four mobile regions. Top: time derivative of the concentration with respect to time. Bottom: time derivative of the concentration as a function of frequency. From left to right: best result for each experiment, obtained with the same set of parameters. b) Model validation against the 6-MHPM experiment.

5. Discussion

Table 8 summarizes the results obtained after model calibration and validation. The two error definitions (15a, 16a) presented in this paper are used for the validation.

Model	No. parameters	J calibration eq.(16a)	J validation eq.(15a)	J validation eq.(16a)
AD	2	7.750	2.63×10^{-2}	1.660
MI	4	3.505	3.45×10^{-2}	1.948
MR	8	1.975	2.63×10^{-2}	1.476
MRAD	6	1.610	6.35×10^{-2}	2.391
PAMR 3	8	1.592	3.40×10^{-2}	1.753
PAMR 4	8	1.608	3.43×10^{-2}	1.747

Table 8: Summary of the calibration and validation results for the 6 models. PAMR 3 and PAMR 4 correspond respectively to the three and four region PAMR models.

The following conclusions may be drawn.

Firstly, comparing the third and fifth columns in the table allows the respective predictive power of the various models at large scales to be assessed. Indeed, these two columns in the Table use the same definition (16a) for the objective function. While the AD and MI models perform poorly compared to the MR, MRAD and PAMR models in the calibration phase, their validation performance is similar in the validation phase. This may be explained by the fact that the Fickian model (towards which all models are asymptotically equivalent) becomes more valid as N increases. The MRAD model is the only one exhibiting a significantly decreasing performance in the validation phase compared to the calibration phase.

Secondly, the respective performance of the various models is the same for both model error formulae (15a, 16a). The MR and AD models, that have the smaller validation error (15a), also have the smaller validation error (16a). The MI, PAMR3 and PAMR4 models, that have intermediate validation error values with formula (15a), also have intermediate validation error values with formula (16a). Lastly, the MRAD model consistently has the larger validation error

values, be it with the error definition (15a) or (16a).

Thirdly, the error formula (15a) gives objective function values consistently 50 times as small as those obtained with the error definition (16a). These two modelling error definitions serve different purposes. The modelling error definition (15a) may be seen as a measure of the "plausibility" of a model. With this definition, two different models giving output signals within the min/max experimental error range will yield identically $J = 0$, even if they depart significantly from the mean experimental signal. With this definition, $J = 0$ means that the difference between the mean experimental measurement and the model output can be fully explained by experimental imprecision. These two models may be deemed equally plausible. In contrast, the error formula (16a) is a measure of the "accuracy" of a model. With this modelling error definition, the smaller J, the better the model output fits the average experimental response. In comparison with other studies, the present work features a large number of replicates. Even though a model porous media is dealt with, the replicates show significant statistical variations in the experimental breakthrough curves, be it for a given MHPM or for the ensemble of the 12 MHPM used in this study. These statistical variations are illustrated by the min-max interval on Figure 3. The interval is the widest for $N = 1$ MHPM (single period study column) and becomes narrower as N increases. Clearly, the statistical variations are stronger at the heterogeneity scale ($N = 1$) and are damped as the porous media tends to homogeneity ($N = 6$). This shows the importance of replicating experiments when heterogeneous porous media are involved. Most previously reported transport experiments use a limited number of replicates (a single experiment in many cases). This does not allow the experimental variability of the data set to be assessed. Given the width of the min-max cloud on Figure 3, many candidate models might fit within the experimental confidence interval, and if they do, they can all be considered as good candidate models.

In order to compare several modelling approaches, using high quality experimental data proves essential. A large number of replicates allows the statistical variations of the experimental data to be accounted for in the benchmarking process. Another issue in the calibration and validation process is the definition of the objective function. As shown in [48], many models provide an accurate description of the long-time behaviour of the breakthrough curves. In contrast, examining the short time behaviour of the experimental data sets leads to rule out a number of models [48]. Consequently, defining an objective function that gives equal weights to small and large time behaviours, as done in the present study, is deemed essential to an efficient model assessment. To our best knowledge, this approach has been little used in transport model benchmarking. Using the time derivative of the experimental signals in the calibration and validation processes for a more efficient discrimination between models has also been little reported in the literature.

In this study, we compare five modelling approaches: four existing models (AD, MI, MR, MRAD) and a new model (PAMR). The originality of the PAMR model is that transport is

modelled on a purely advection basis, without the need for dispersion terms. The four existing, dispersion-based models use a scale-independent dispersion coefficient. We insist on this choice because as shown in [48], (i) there is no experimental evidence that dispersion varies with time and distance, (ii) there exists a model with scale-independent coefficients that allows the MHPM experiments to be reproduced. Moreover, as mentioned in [38], while some studies suggest that dispersion may increases with time and distance, other studies suggest that dispersion may eventually tend to a constant asymptotic value.

If the response of the heterogeneous medium departs from the classical Fickian behaviour, we consider proposing a non-Fickian transport more appropriate than making the Fickian transport model scale-dependent (as in e.g. [23, 55]). Note that fractional advection-dispersion models were developed to eliminate the scale-dependent dispersion [6], albeit with limited success [22]. In the authors' view, the challenge tackled by the proposed PAMR model is not only to reproduce the non-Fickian transport process on the heterogeneity scale, but also to tend asymptotically towards a Fickian transport process (classical AD model) on an observation scale that is much larger than the heterogeneity scale.

Moreover, the PAMR model does not have the undesirable effects of the Fickian operator. As shown by Einstein [19, 20, 21], the Fickian model is valid only above a given time and space scale, under the assumption that a sufficient number of realizations of the Brownian motion have been realized and that the movement obeys a Markovian process (that is, no memory of the previous displacements is kept after a change in particle position and velocity). If the scale under consideration is such that a time or space correlation exists between the successive displacements of the solute particles (as is the case when a single MHPM is concerned), the Fickian model should be expected to be invalid. Along the same line, Benson et al. [6] insist on the fact that AD model is based on the divergence of a vector field which is evaluated by the limit of the flux of this vector on the surface of an enclosed volume when the volume shrinks towards zero. To quote these authors: "This is valid only if the flux is indeed a point vector quantity relative to the scale of observation, for example, heat flow in homogeneous material. Then the limit exists. Solute dispersion is a counterexample since it is primarily due to velocity fluctuations that arise only as an observation space grows larger, invalidating the limit".

This explains why the AD model fails to reproduce the experimental results at small scales (Figure 4a and Table 8). As the size of the domain increases ($N = 6$ in Figure 4b), the AD model becomes better adapted to reproduce the experimental behaviour because the scale of observation becomes larger than the scale of one heterogeneity. The objective function for the AD model is $J = 7.76$. The MI model (one mobile region and one immobile region) gives better results than the AD model especially for $N \geq 3$ (Figure 5) but it still fails to reproduce the experimental behaviour for $N = 1$ (scale of one heterogeneity). The objective function becomes $J = 3.505$ with an improvement of 55% in comparison to AD model. The MR model, which is a generalization

of the MI model, gives better results than the MI model with an objective function $J = 1.975$ and an improvement of 44% over the MI model (Figure 6). The MR used in this study has one mobile region and three immobile regions and it starts to reproduce the experimental behaviour at the scale of one heterogeneity (Figure 6a). The MRAD model gives better results than the MR model (Figure 7), with an objective function of $J = 1.61$ (an improvement of 18% in comparison to MR). Concerning the PAMR model proposed in this study, we first used a simple version with three mobile regions. The simulations reproduce fairly well the experimental behaviour especially at the scale of one heterogeneity (Figure 8). With an objective function $J = 1.592$, the accuracy of the PAMR model is nearly identical to that of the MRAD model. A four region version of the PAMR gave similar simulation results (Figure 9), with $J = 1.608$. Adding a fourth region thus brings no improvement over the three region PAMR. This means that the porous media of this study (MHPM) can be well reproduced with a three region PAMR, and that a more complex multiple region PAMR will not lead to a better simulation. We consider the PAMR simulations very satisfactory in that (i) the PAMR model performs as satisfactorily as other existing models, (ii) it describes quite well the experimental behaviour at the scale of both a single and several heterogeneities, (iii) this is achieved using scale-independent parameters. Although the PAMR model is calibrated on a small observation scale (one to five heterogeneities), it asymptotically tends to a Fickian behaviour on a large observation scale (6 heterogeneities) with a fairly good simulation (Table 8). This asymptotic Fickian behaviour of the PAMR model was to be expected because the model is known to lead to the Fickian model for large scales under conditions that are met here (see AppendixA and the stability analysis in AppendixB).

As far as the reported, experimental scale dependence of the dispersion coefficient is concerned, the PAMR model allows such dependence to be explained, as shown in Subsection AppendixA.2.4. Solving the two region PAMR equations for a Dirac initial condition shows that the apparent dispersion coefficient obeys an exponential function of time. This is precisely the behaviour inferred in [23], where the best fit to the experimental data set was obtained by fitting an exponential function $D(x)$. The developments in Subsection AppendixA.2.4 also indicate that there exist two different ways of fitting the dispersion coefficient from field measurements. The first consists in adjusting D so as to fit the average spreading rate of the contaminant plume over time, yielding a first coefficient D_{av}. The second consists in adjusting D so as to fit the time derivative of the plume variance, yielding a local estimate D_{loc}. The variation analysis in AppendixA.2.4 shows that D_{loc} converges faster to the asymptotic value than does D_{av}. Indeed, the difference $D_\infty - D_{\mathrm{loc}}$ is a decreasing exponential function of time, while $D_\infty - D_{\mathrm{av}}$ is proportional to the inverse of time. This was to be expected in that D_{av} incorporates the effect of all times on the dispersion behaviour, including the times for which the dispersion behaviour is strongly non-Fickian. D_{loc} therefore provides a more accurate estimate of D_∞ than D_{av}. However, in practice, deriving D_{av} is easier than D_{loc} because the determination of D_{loc} requires many more sampling times and points

than that of D_{av} .

Lastly, it is worth pointing out that all multi-region models (including the PAMR model proposed in this paper) exhibit a certain degree of non-local behaviour, in contrast with the AD model that is purely local. Indeed, specifying te initial and boundary conditions for any of the MI, MR, MRAD and PAMR models requires that the initial and boundary conditions be specified for each of the flow regions. In practical applications, however, this is not possible. Only the average concentration $c = \sum_{i=1}^{R} \theta_i c_i$ within he bulk porous medium is known, and the space-time distribution $c_i(x,t)$ in every region i is unknown. A straightforward solution would be to set the concentrations in all regions to the same average concentration value. However, in real-world applications, such an initial state is likely to be most inaccurate. This is because, by definition of intermediate time and space scales, the concentrations in all flow regions have not reached equilibrium and can therefore not be assumed to be identical. Consequently, the distributions $c_i(x,t_0)$ to be used for an initial condition at a given time t_0 are functions of the past behaviours of the fields $c_i(x,t),\ t < t_0$ and the way the various flow regions have been exchanging solute. This is typically a non-local behaviour, a common feature shared by all widely admitted anomalous transport models.

6. Conclusions

In this paper, solute dispersion in model periodical heterogeneous media is studied from both an experimental and modelling perspective. The following, main results are found.

A large number of solute transport experiments through periodic heterogeneous porous media is presented. The significant number of experiment replicates yield high quality breakthrough curves, while allowing the experimental uncertainty to be characterized accurately. The authors are willing to make the data sets available to the scientific community.

A calibration procedure putting the emphasis on both small time and large time behaviours is set up. It allows for the benchmarking of several transport and dispersion models: the Advection-Dispersion (AD), Mobile-Immobile (MI), Multi-Rate (MR), Multi-Region Advection-Dispersion (MRAD) and Purely Advective Multi Region (PAMR) models.

Applying the AD model shows that, at the scale of a single heterogeneity, the dispersion process is non-Fickian. It tends asymptotically to a Fickian behaviour for an increasing number of heterogeneity periods. Even without a Fickian dispersion term, the PAMR model is seen to perform as well as the AD, MI, MR and MRAD models. The PAMR is able to simulate transport on every heterogeneity scale: on the single heterogeneity scale, PAMR can reproduce a non-Fickian behaviour while it tends to the observed, classical Fickian behaviour on the scale of several heterogeneities. Another advantage of the PAMR model is the absence of the undesirable effects of the Fickian term, such as the physically unsound infinite propagation speed of the solutions of the diffusion equation.

AppendixA. Equations of moments

AppendixA.1. Spatial moments for the AD model

The present subsection is devoted to the development of the equations for the spatial moments of the propagator in the AD model. The propagator is the solution $c(x,t)$ of the AD equation (1) for the initial condition $c(x,0) = \delta(x)$. The governing equation is first rewritten in the coordinate system moving at speed u:

$$\partial_t c - D\partial_{xx} c = 0 \tag{A.1}$$

Denoting by $c^{(p)}$ the pth-order spatial moment of the concentration

$$c^{(p)} \equiv \int_{-\infty}^{+\infty} x^p c(x,t)\,\mathrm{d}x \tag{A.2}$$

the governing equations for the moments are obtained by multiplying equation (A.1) by x^p and integrating over the real axis:

$$\int_{-\infty}^{+\infty} x^p \left(\partial_t c - D\partial_{xx} c\right)\mathrm{d}x = 0 \tag{A.3}$$

Using integration by parts to eliminate the higher-order derivatives, using the property $\lim_{x\to\pm\infty} c(x,t) = 0$, the following equations are obtained

$$\mathrm{d}_t c^{(0)} = 0 \Rightarrow c^{(0)}(t) = 1 \tag{A.4a}$$

$$\mathrm{d}_t c^{(1)} = u \Rightarrow c^{(1)}(t) = 0 \tag{A.4b}$$

$$\mathrm{d}_t c^{(2)} - 2D = 0 \Rightarrow c^{(2)}(t) = 2Dt \tag{A.4c}$$

The well-known property of a variance of particle locations proportional to time is retrieved. Note however that equation (A.4c) leads to two expressions for the dispersion coefficient, a local dispersion (D_{loc}) and an average dispersion (D_{av}):

$$D_{\mathrm{loc}} = \frac{1}{2}\mathrm{d}_t c^{(2)} \tag{A.5a}$$

$$D_{\mathrm{av}} = \frac{1}{2t} c^{(2)} \tag{A.5b}$$

The first is obtained from the differential equation (A.4c), the second is obtained from its solution under the assumption of a constant D. Eq. (A.5a) reflects a local behaviour at time t while Eq. (A.5b) reflects an average behaviour over the time interval $[0,t]$. Both expressions are equivalent when D is constant. When D is not constant (as e.g. in the MRAD model considered in Section

A.2), D_{av} incorporates the influence of early times when the behaviour is non-Fickian and can be expected to take a longer time to converge to the asymptotic value than the local coefficient D_{loc}.

AppendixA.2. Spatial moments for the MRAD model

AppendixA.2.1. Governing equations

For the sake of simplicity, the behaviour is analyzed in the coordinate system that moves at the average speed $\overline{u} = \frac{\sum_{i=1}^{R} \Theta_i u_i}{\sum_{i=1}^{R} \Theta_i}$. In this coordinate system, the governing equations are

$$\Theta_i \partial_t c_i + \Theta_i v_i \partial_x c_i - \Theta_i D_i \partial_{xx} c_i = \sum_{j \neq i} k_{ij} \left(c_j - c_i \right) \tag{A.6a}$$

$$k_{ij} = k_{ji}, \ \sum_{i=1}^{R} \Theta_i v_i = 0 \tag{A.6b}$$

where the v_i are defined as in (6). Integrating (A.6a) with respect to x over $(-\infty, +\infty)$, using the property $\lim_{x \to \pm\infty} c\left(x,t\right) = \lim_{x \to \pm\infty} \partial_x c\left(x,t\right) = 0$ yields

$$\mathrm{d}_t c_i^{(0)} = \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} \left(c_j^{(0)} - c_i^{(0)} \right) \tag{A.7}$$

Since $c_i^{(0)}\left(t=0\right) = 1 \, \forall i$, $\mathrm{d}_t c_i^{(0)}$ remains identically zero at all times, which results in

$$c_i^{(0)}\left(t\right) = 1 \forall \, t \geq 0 \tag{A.8}$$

Multiplying equation (A.6a) by x, integrating with respect to x over $(-\infty, +\infty)$, using integration by parts and the property $\lim_{x \to \pm\infty} c\left(x,t\right) = \lim_{x \to \pm\infty} \partial_x c\left(x,t\right) = 0$ yields

$$\mathrm{d}_t c_i^{(1)} = v_i c_i^{(0)} + \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} \left(c_j^{(1)} - c_i^{(1)} \right) = v_i + \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} \left(c_j^{(1)} - c_i^{(1)} \right) \tag{A.9}$$

Multiplying equation (A.6a) by x^2, integrating with respect to x over $(-\infty, +\infty)$, using integration by parts and the property $\lim_{x \to \pm\infty} c\left(x,t\right) = \lim_{x \to \pm\infty} \partial_x c\left(x,t\right) = \lim_{x \to \pm\infty} \partial_{xx} c\left(x,t\right) = 0$ yields

$$\mathrm{d}_t c_i^{(2)} = v_i c_i^{(1)} + 2 D_i c_i^{(0)} + \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} \left(c_j^{(2)} - c_i^{(2)} \right) = v_i c_i^{(1)} + 2 D_i + \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} \left(c_j^{(2)} - c_i^{(2)} \right) \tag{A.10}$$

Multiplying by Θ_i and summing over $i = 1, \ldots, R$, using the property $\sum_{i=1}^{R} \Theta_i v_i = 0$ yields equation (6).

679 *AppendixA.2.2. Small time behaviour*

680 At $t = 0$, $c_i^{(1)} = 0 \, \forall i$ and the governing equation (A.9) simplifies to

$$\left(\mathrm{d}_t c_i^{(1)}\right)_{t=0} = v_i \tag{A.11}$$

681 Consequently the following equivalence holds

$$c_i^{(1)}(t) \underset{0}{\sim} v_i t \tag{A.12}$$

682 Moreover, $c_i^{(2)}(t = 0) = 0 \; \forall \, i$ and equation (6) becomes

$$\mathrm{d}_t c^{(2)} = \sum_{i=1}^{R} \Theta_i \left(v_i^2 + 2D_i\right) \tag{A.13}$$

683 with an equivalent dispersion coefficient obtained from (A.5a, A.5b):

$$D_{\mathrm{loc}} = \sum_{i=1}^{R} \Theta_i \left(v_i^2 t + D_i\right) \tag{A.14a}$$

684

$$D_{\mathrm{av}} = \sum_{i=1}^{R} \Theta_i \left(\frac{1}{2} v_i^2 t + D_i\right) \tag{A.14b}$$

685 For small times, the equivalent dispersion coefficient is observed to increase linearly with distance.
686 However the growth rate of the average dispersion coefficient is smaller than that of the local one.

687 *AppendixA.2.3. Long time behaviour*

688 The solution is asymptotically stable (see AppendixB), consequently there exists a set of finite
689 asymptotic values $c_i^{(1,\infty)}$ such that

$$c_i^{(1)} \underset{t \to +\infty}{\longrightarrow} c_i^{(1,\infty)} \forall i \tag{A.15}$$

690 Substituting the property (A.15) into equation (A.10) gives the following long time, asymptotic
691 behaviour

$$\mathrm{d}_t c^{(2)} = \sum_{i=1}^{R} \Theta_i \left(v_i c_i^{(1,\infty)} + 2D_i\right) \tag{A.16}$$

692 The time derivative of the variance of particle locations becomes constant and a classical Fickian
693 behaviour is achieved, with (A.5a) simplifying into Equation (9).

694 *AppendixA.2.4. A particular case: the two region model*

695 The simplest possible model consists of two regions :

$$R = 2, \; \Theta_2 = 1 - \Theta_1, \; v_2 = -\frac{\Theta_1}{\Theta_2} v_1 \tag{A.17}$$

Solving equations (A.9, 6) under assumptions (A.17) yields the following solution

$$c_i^{(1)} = \frac{\Theta_1\Theta_2}{k_{12}}\left(1-\exp\left(-\frac{k_{12}}{\Theta_1\Theta_2}t\right)\right)v_i,\ i=1,2 \tag{A.18a}$$

$$c^{(2)} = \frac{\Theta_1\Theta_2}{k_{12}}\left(\Theta_1 v_1^2+\Theta_2 v_2^2\right)\left(t+\frac{\Theta_1\Theta_2}{k_{12}}\left[\exp\left(-\frac{k_{12}}{\Theta_1\Theta_2}t\right)-1\right]\right)+2\left(\Theta_1 D_1+\Theta_2 D_2\right)t \tag{A.18b}$$

which yields

$$D_{\mathrm{loc}} = \frac{1}{2}\frac{\Theta_1\Theta_2}{k_{12}}\left(\Theta_1 v_1^2+\Theta_2 v_2^2\right)\left(1-\exp\left(-\frac{k_{12}}{\Theta_1\Theta_2}t\right)\right)+\Theta_1 D_1+\Theta_2 D_2 \tag{A.19a}$$

$$D_{\mathrm{av}} = \frac{1}{2}\frac{\Theta_1\Theta_2}{k_{12}}\left(\Theta_1 v_1^2+\Theta_2 v_2^2\right)\left(1+\Theta_1\Theta_2\frac{\exp\left(-\frac{k_{12}}{\Theta_1\Theta_2}t\right)-1}{kt}\right)+\Theta_1 D_1+\Theta_2 D_2 \tag{A.19b}$$

with the following limit behaviours

$$D_{\mathrm{loc}} \underset{0}{\sim} \Theta_1 D_1+\Theta_2 D_2+\frac{1}{2}\Theta_1\Theta_2\left(\Theta_1 v_1^2+\Theta_2 v_2^2\right)t \tag{A.20a}$$

$$D_{\mathrm{av}} \underset{0}{\sim} \Theta_1 D_1+\Theta_2 D_2+\frac{1}{4}\Theta_1\Theta_2\left(\Theta_1 v_1^2+\Theta_2 v_2^2\right)t \tag{A.20b}$$

$$D_{\mathrm{loc}} \underset{\infty}{\sim} D_{\mathrm{av}} \underset{\infty}{\sim} D_\infty = \Theta_1 D_1+\Theta_2 D_2+\frac{1}{2}\frac{\Theta_1\Theta_2}{k_{12}}\left(\Theta_1 v_1^2+\Theta_2 v_2^2\right) \tag{A.20c}$$

Long time and small time behaviours. A further restriction of the model (A.17) is obtained under the particular assumption $D_1 = D_2 = 0, \Theta_1 = \Theta_2 = \frac{1}{2}, v_1 = -v_2 = -a$. This leads to a particular case of the telegraph equation [4]. Its first mention as a model for turbulent dispersion is attributed to Davydov [15]. It gives a formula similar to the well-known formula established by Einstein in its publications on Brownian movement [19, 20, 21]:

$$D_\infty = \frac{a^2}{8k_{12}} \tag{A.21}$$

with $k_{12} = \frac{1}{4\tau}$, τ being the time scale of the Brownian movement. For small times, the advective process is predominant, with a so-called ballistic behaviour (a variance growing proportionally to the square of time). For large times, an asymptotic Fickian regime is reached. The Fickian behaviour is reached after a few times k_{12}^{-1}. This shows that a Fickian behaviour can be obtained from a purely advective model beyond a certain time scale. At smaller times, the dispersion process is anomalous. The advantage of the purely advective model over the Fickian model is that it implicitly rules out infinite signal propagation speeds.

Convergence rate to the asymptotic value D_∞. Equations (A.19a, A.19b) can be rewritten as

$$D_\infty - D_{\mathrm{av}} = A\frac{\exp(-Kt)}{Kt} \underset{t\to\infty}{\sim} \frac{A}{Kt} \tag{A.22a}$$

$$D_\infty - D_{\text{loc}} = A\exp(-Kt) \tag{A.22b}$$

$$A = \frac{K}{2}(\Theta_1 v_1^2 + \Theta_2 v_2^2), K = \frac{k_{12}}{\Theta_1\Theta_2} \tag{A.22c}$$

Comparing equations (A.22a, A.22b) shows that the local estimate D_{loc} of the dispersion coefficient tends faster to the asymptotic value D_∞ than does the average estimate D_{av}.

AppendixB. Stability of the solution

The purpose of this Appendix is to study the stability properties of the system (A.9). Note first that this system can be written in vector form as

$$\mathrm{d}_t\mathbf{x} = \mathbf{v} + \mathbf{M}\mathbf{x} \tag{B.1a}$$

$$\mathbf{x} \equiv \left[\mathbf{c}_1^{(1)}, \ldots, \mathbf{c}_R^{(1)}\right]^T, \ \mathbf{v} \equiv \left[\mathbf{v}_1, \ldots, \mathbf{v}_R\right]^T, \ \mathbf{M} = [M_{ij}] \tag{B.1b}$$

where the elements M_{ij} are defined as in (8). As shown in AppendixA, if the solutions of (B.1a) are stable, a Fickian behaviour leading to normal diffusion is obtained in the limit of long times.

It is first noticed that the matrix $\mathbf{M}$ can be written in the form $\mathbf{M} = -\mathbf{SD}$, with

$$\mathbf{D} = \mathrm{diag}\left[\Theta_i\right], \ S_{ij} = -\frac{M_{ij}}{\Theta_j} = \begin{cases} \sum_{p\neq i} \frac{k_{ip}}{\Theta_i\Theta_j} & \text{if} \quad i = j \\ -\frac{k_{ij}}{\Theta_i\Theta_j} & \text{if} \quad i \neq j \end{cases} \tag{B.2}$$

Consequently, the matrix $\mathbf{S}$ is symmetric.

Result 1. The matrix $\mathbf{S}$ is positive semidefinite.

Proof. For any vector $\mathbf{x} = [x_1, \ldots, x_R]^T$, rearranging and noting that $k_{ij} \geq 0 \, \forall \, (i,j)$ gives

$$\mathbf{x}^T\mathbf{S}\mathbf{x} = \sum_{i=1}^{R}\sum_{j>i} k_{ij}\left(\frac{x_i}{\Theta_i} - \frac{x_j}{\Theta_j}\right)^2 \geq 0 \, \forall \mathbf{x} \tag{B.3}$$

□

Corollary 1. The eigenvalues of $\mathbf{S}$ are all positive.

Proof. Let λ be an eigenvalue of $\mathbf{S}$ and $\mathbf{x}$ a corresponding eigenvector. Then

$$\mathbf{x}^T\mathbf{S}\mathbf{x} = \mathbf{x}^T\lambda\mathbf{x} = \lambda\mathbf{x}^T\mathbf{x} = \lambda\left\|\mathbf{x}\right\|^2 \geq 0 \tag{B.4}$$

□

Theorem 1. Let $\mathbf{S}$ and $\mathbf{D}$ be symmetric and positive definite matrices. Then the eigenvalues of $\mathbf{SD}$ are all strictly positive.

Proof. There exists an orthogonal matrix $\mathbf{C}$ such that $\mathbf{S} = \mathbf{C}\mathrm{diag}\left[\lambda_i\right]\mathbf{C}^T$. Defining $\mathbf{S}^{1/2} \equiv \mathbf{C}\mathrm{diag}\left[\sqrt{\lambda_i}\right]\mathbf{C}^T$, it is noted that the matrices $\mathbf{S}^{1/2}\mathbf{D}\mathbf{S}^{1/2}$ and $\mathbf{S}\mathbf{D}$ are similar because $\mathbf{S}\mathbf{D} = \mathbf{S}^{1/2}\mathbf{S}^{1/2}\mathbf{D}\mathbf{S}^{1/2}\mathbf{S}^{-1/2}$ (note that $\mathbf{S}^{-1/2}$ exists because the $\sqrt{\lambda_i}$ are all nonzero). Therefore the eigenvalues of $\mathbf{S}\mathbf{D}$ are the same as the eigenvalues of $\mathbf{S}^{1/2}\mathbf{D}\mathbf{S}^{1/2}$. Moreover, the matrix $\mathbf{S}^{1/2}\mathbf{D}\mathbf{S}^{1/2}$ is symmetric and positive definite because $\mathbf{D}$ is symmetric, positive definite and $\mathbf{S}^{1/2}$ is symmetric. Consequently, the eigenvalues of $\mathbf{S}\mathbf{D}$ are all striclty positive. □

Corollary 2. The eigenvalues of $-\mathbf{S}\mathbf{D}$ are all strictly negative.

Theorem 2. Let $\mathbf{S}$ and $\mathbf{D}$ be symmetric and positive semidefinite matrices. Then the eigenvalues of $\mathbf{S}\mathbf{D}$ are all positive.

Proof. Let $\lambda_1, \ldots, \lambda_R$ be the (positive) eigenvalues of $\mathbf{S}$. There exists an orthogonal matrix $\mathbf{C}$ such that $\mathbf{S} = \mathbf{C}\mathrm{diag}\left[\lambda_i\right]\mathbf{C}^T$. Defining $\mathbf{S}_\varepsilon = \mathbf{C}\mathrm{diag}\left[\lambda_i + \epsilon\right]\mathbf{C}^T, \epsilon > 0$. Following the reasoning of Theorem 1, $\mathbf{S}_\epsilon^{-1/2}$ exists and the eigenvalues of $\mathbf{S}_\epsilon\mathbf{D}$ are the same as the eigenvalues of $\mathbf{S}_\epsilon^{1/2}\mathbf{D}\mathbf{S}_\epsilon^{1/2}$. Consequently, they are all positive. Since $\mathbf{S}_\epsilon\mathbf{D} \to \mathbf{S}\mathbf{D}$ as $\epsilon \to 0$, it follows from the continuity of the spectrum of a matrix that the eigenvalues of $\mathbf{S}\mathbf{D}$ are all positive.

Corollary 3. The eigenvalues of $\mathbf{M} = -\mathbf{S}\mathbf{D}$ are all negative.

Result 2. The solutions of the differential system $\mathrm{d}_t\mathbf{x} = \mathbf{M}\mathbf{x}$ converge to an equilibrium solution.

Proof. Since $\mathbf{M}$ and $\mathbf{M}^T$ have the same spectrum, their eigenvalues are all negative. Therefore, the solutions of $\mathrm{d}_t\mathbf{x} = \mathbf{M}\mathbf{x}$ and $\mathrm{d}_t\mathbf{x} = \mathbf{M}^T\mathbf{x}$ have the same asymptotic behaviour. Let $\mathbf{x}$ be the solution of $\mathrm{d}_t\mathbf{x} = \mathbf{M}^T\mathbf{x}$. Defining $Q \equiv \frac{1}{2}\sum_{i=1}^{R}\frac{1}{\Theta_i}x_i^2$, one has

$$\mathrm{d}_t Q = \sum_{i=1}^{R}\frac{1}{\Theta_i}x_i\mathrm{d}_t x_i = \left(\mathbf{D}^{-1}\mathbf{x}\right)^T\mathrm{d}_t\mathbf{x} = \left(\mathbf{D}^{-1}\mathbf{x}\right)^T\mathbf{M}^T\mathbf{x} = \mathbf{x}^T\mathbf{D}^{-1}\mathbf{M}^T\mathbf{x} = -\mathbf{x}^T\mathbf{S}\mathbf{x} \tag{B.5}$$

From Result 1, the matrix $-\mathbf{S}$ is negative, semidefinite. Consequently, $\mathrm{d}_t Q \leq 0$ and $0 \leq Q\left(t\right) \leq Q\left(0\right)$ for $t > 0$. Consequently, the solution $\mathbf{x}\left(t\right)$ is bounded. Since the eigenvalues of $\mathbf{M}^T$ are real and negative (from Corollary 3), it follows from the theory of linear differential equations that there exists a vector $\mathbf{x}_\infty$ such that $\mathbf{x} \underset{t\to\infty}{\longrightarrow} \mathbf{x}_\infty$.□

Result 3. If $\mathbf{v} \in \mathrm{rge}\mathbf{M}$, the solution of the non-homogeneous system $\mathrm{d}_t\mathbf{x} = \mathbf{v} + \mathbf{M}\mathbf{x}$ converge to an equilibrium solution.

Proof. By assumption, there exists $\mathbf{a}$ such that $\mathbf{M}\mathbf{a} = \mathbf{v}$. Then $\mathrm{d}_t\mathbf{x} = \mathbf{v} + \mathbf{M}\mathbf{x} \Longleftrightarrow \mathrm{d}_t\left(\mathbf{a} + \mathbf{x}\right) = \mathbf{v} + \mathbf{M}\mathbf{x} = \mathbf{M}\left(\mathbf{a} + \mathbf{x}\right)$. Consequently, $\mathbf{a} + \mathbf{x}$ satisfies the homogeneous system $\mathrm{d}_t\mathbf{x} = \mathbf{M}\mathbf{x}$ and converges to an equilibrium solution, hence the result.

Result 4. If $\mathbf{M}$ is block-diagonal, $\mathbf{M} = \begin{bmatrix} \begin{bmatrix} M_{11} & \cdots & M_{1r} \\ \vdots & & \vdots \\ M_{r1} & \cdots & M_{rr} \end{bmatrix} & 0 \\ 0 & \mathbf{A} \end{bmatrix}$, a necessary condition for $\mathbf{v} = [v_1, \ldots, v_R]^T$ to belong to rge$\mathbf{M}$ is

$$\sum_{i=1}^{r} \Theta_i v_i = 0 \tag{B.6}$$

Proof. If $\mathbf{v}$ belongs to the range of $\mathbf{M}$, there exists a vector $\mathbf{a}$ such that $\mathbf{v} = \mathbf{M}\mathbf{a}$. Then, using the symmetry property $k_{ij} = k_{ji} \forall (i,j)$, one has

$$\sum_{i=1}^{r} \Theta_i v_i = \sum_{i=1}^{r} \Theta_i \sum_{j=1}^{r} M_{ij} a_j = \sum_{i=1}^{r} \Theta_i \sum_{\substack{j=1 \\ j \neq i}}^{r} \frac{k_{ij}}{\Theta_i} (a_j - a_i) = \sum_{i=1}^{r} \sum_{\substack{j=1 \\ j \neq i}}^{r} k_{ij} (a_j - a_i) = 0 \tag{B.7}$$

□

Corollary 4. The same result holds for the $R - r$ remaining elements of $\mathbf{v}$, $\sum_{i=r+1}^{R} \Theta_i v_i = 0$.

[1] Ahmadi, A., Quintard, M., Whitaker, 1998. Transport in chemically and mechanically heterogeneous porous media V. Two-equation model for solute transport with adsorption. Advances in Water Resources, 22, 59–86.

[2] Aral, M.M., Liao, B., 1996. Analytical solutions for two-dimensional transport equation with time-dependent dispersion coefficients. J. Hydrol. Eng. 1 (1), 20–32.

[3] . Bachelier L., 1900. Theorie de la speculation. Annales Scientiques de l'ENS, Ser. 3, 17, 21–86.

[4] Bakunin, OG., 2008. Turbulence and diffusion. Scaling versus equations. Springer-Verlag.

[5] Bear, J., 1972. Dynamics of Fluids in Porous Media. Dover, NY.

[6] Benson, D.A., Wheatcraft, S.W., Meerschaert, M.M., 2000a. The fractional-order governing equation of Lévy motion. Water Resour. Res. 36 (6), 1413–1423.

[7] Benson, D.A., Wheatcraft, S.W., Meerschaert, M.M., 2000b. Application of a fractional advective–dispersion equation. Water Resour. Res. 36 (6), 1403–1412.

[8] Berkowitz, B., Scher, H., Silliman, S.E., 2000. Anomalous transport in laboratory-scale, heterogeneous porous media. Water Resour. Res. 36 (1), 149–158.

[9] Berkowitz, B., Emmanuel, S., Scher, H., 2008. Non-Fickian transport and multiplerate mass transfer in porous media. Water Resour. Res. 44, W03402.

[10] Blunt, M.J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S. Mostaghimi, P., Paluszny, A., Pentland, C., 2013. Pore-scale imaging and modelling. Advances in Water Resources 51, 197-216.

[11] Cherblanc, F., Ahmadi, A., Quintard, M., 2003. Two-medium description of dispersion in heterogeneous porous media: calculation of macroscopic properties. Water Resour. Res. 39(6):1154–73.

[12] Cherblanc, F., Ahmadi, A., Quintard, M., 2007. Two-domain description of solute transport in heterogeneous porous media: Comparison between theoretical predictions and numerical experiments. Advances in Water Resources 30, 1127–1143.

[13] Cortis, A., Berkowitz, B., 2004. Anomalous transport in classical soil and sand columns. Soil Sci. Soc. Am. J. 68 (5), 1539–1548.

[14] Danquigny, C., Ackerer, P., Carlier, J.P., 2004. Laboratory tracer tests on threedimensional reconstructed heterogeneous porous media. J. Hydrol. 294, 196–212.

[15] Davydov, BJ., 1934. Diffusion equation with the inclusion of molecular velocity. Doklady Akademii Nauk SSSR, 2, 474.

[16] Davit, Y., Quintard, M., Debenest, G., 2010. Equivalence between volume averaging and moments matching techniques for mass transport models in porous media. International Journal of Heat and Mass Transfer, 53, Issues 21–22, 4985–4993.

[17] De Smedt, F., and P. J. Wierenga, 1979. A generalized solution for solute flow in soils with mobile and immobile water, Water Resour. Res., 15, 1137–114.

[18] Dentz, M., Berkowitz, B., 2003. Transport behavior of a passive solute in continuous time random walks and multirate mass transfer. Water Resources Research 39 (5), 1111.

[19] Einstein, A., 1905. On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. Annalen der Physik, 17, 549-560.

[20] Einstein, A., 1906. On the theory of the Brownian movement. Annalen der Physik, 19, 371-381.

[21] Einstein, A., 1908. The elementary theory of the Brownian motion. Zeitung f??r Elektrochemie, 14, 235-239.

[22] Gao, G., Zhan, H., Feng, S., Huang, G., Mao, X., 2009. Comparison of alternative models for simulating anomalous solute transport in a large heterogeneous soil column. Journal of Hydrology 377, 391-404.

[23] Gao, G., Zhan, H., Feng, S., Fu, B., Ma, Y., Huang, G., 2010. A new mobile-immobile model for reactive solute transport with scale-dependent dispersion, Water Resour. Res., 46, W08533.

[24] Gaudet, J. P., Jégat,H., Vachaud, G., Wierenga, P. J., 1977. Solute transfer, with exchange between mobile and stagnant water, through unsaturated sand, Soil Sci. Soc. Am. J., 41, 665–671.

[25] Gelhar, L.W., Welty, C., Rehfeldt, K.R., 1992. A critical review of data on field scale dispersion in aquifers. Water Resour. Res. 28 (7), 1955–1974.

[26] Gerke, H. H., Van Genuchten, M. T., 1993. A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media, Water Resour. Res., 29, 305–319.

[27] Gerke, H. H., Van Genuchten, M. T.1993. Evaluation of a first-order water transfer term for variably saturated dual-porosity flow models, Water Resour. Res., 29, 1225–1238.

[28] Gerke, H. H., Van Genuchten, M. T., 1996. Macroscopic representation of structural geometry for simulating water and solute movement in dualporosity media, Adv. Water Resour., 19, 343–357.

[29] Golfier, F., Quintard, M., Cherblanc, F., Zinn, BA., Wood, BD. 2007. Comparison of theory and experiment for solute transport in highly heterogeneous porous medium. Advances in Water Resources 30, 2235–2261.

[30] Golfier, F., Quintard, M., Wood, BD., 2011. Comparison of theory and experiment for solute transport in weakly heterogeneous bimodal porous media. Advances in Water Resources 34, 899–914.

[31] Goltz, M. N., Roberts, P. V., 1988. Simulations of physical nonequilibrium solute transport models: Application to a large-scale field experiment, J. Contam. Hydrol., 3, 37–63.

[32] Greiner, A., Schreiber, W., Brix, G., Kinzelbach, W., 1997. Magnetic resonance imaging of paramagnetic tracers in porous media: quantification of flow and transport parameters. Water Resour. Res. 33 (6), 1461–1473.

[33] Gwo, J., Jardine, P., Wilson, G., Yeh, G., 1996. Using a multiregion model to study the effects of advective and diffusive mass transfer on local physical non-equilibrium and solute mobility in a structured soil, Water Resour. Res., 32, 561–570.

[34] Gwo, J.-P., O'Brien, R., Jardine, P.M., 1998. Mass transfer in structured porous media: Embedding mesoscale structure and microscale hydrodynamics in a two-region model, J. Hydrol., 208, 204–222.

[35] Haggerty, R., Gorelick, S.M., 1995. Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity. Water Resour. Res. 31, 2383–2400.

[36] Haupt, R.L., Haupt, S.E., 2004. {Practical Genetic Algorithms, second ed., Wiley.

[37] Huang, G., Huang, Q., Zhan, H., 2006. Evidence of one-dimensional scale-dependent fractional advection–dispersion. Journal of Contaminant Hydrology, 85, 53-71.

[38] Huang, K., Toride, N., Van Genuchten, M. T., 1995. Experimental investigation of solute transport in large, homogeneous and heterogeneous, saturated soil columns, Transp. Porous Media, 18(3), 283–302.

[39] Irwin, N.C., Botz, M.M., Greenkorn, R.A., 1996. Experimental investigation of characteristic length scale in periodic heterogeneous porous media. Transp. Porous Media 25, 235–246.

[40] Jayawardena, A.A., Lui, P.H., 1984. Numerical solution of the dispersion equation using a variable dispersion coefficient: method and applications. Hydrol. Sci. J. 29 (3), 293–309.

[41] Kavvas, M.L., Kim, S., Ercan, A., 2015. Fractional Ensemble Average Governing Equations of Transport by Time-Space Nonstationary Stochastic Fractional Advective Velocity and Fractional Dispersion. I: Theory. Journal of Hydrologic Engineering, 20(2)

[42] Kavvas, M.L., Ercan, A., Polsinelli, J., 2017. Governing equations of transient soil water flow and soil water flux in multi-dimensional fractional anisotropic media and fractional time. Hydrol. Earth Syst. Sci., 21, 1547-1557.

[43] Klafter, J., Blumen, A., Shlesinger, M.F.. Stochastic pathway to anomalous diffusion. Phys. Rev. A 35, 3081–3085, 1987.

[44] Kumar, N., Harbola, U., Lindenberg, K., 2010. Memory-induced anomalous dynamics: Emergence of diffusion, subdiffusion, and superdiffusion from a single random walk model. Physical Review E, 82, 021101.

[45] Lévy, M., Berkowitz, B., 2003. Measurement and analysis of non-Fickian dispersion in heterogeneous porous media. J. Contam. Hydrol. 64, 203–226.

[46] Li, L., Barry, D.A., Culligan-Hensley, P.J., Bajracharya, K., 1994. Mass transfer in soils with local stratification of hydraulic conductivity. Water Resour. Res. 30 (11), 2891–2900.

[47] Majdalani, S., Angulo-Jaramillo, R., Di Pietro, L, 2008. Estimating preferential water flow parameters using a binary genetic algorithm inverse method. Environmental Modelling and Software 23, 950-956.

[48] Majdalani, S., Delenne, C., Chazarin J.P., Guinot, V., 2015. Solute transport in periodical heterogeneous porous media: Importance of observation scale and experimental sampling. Journal of Hydrology, 520, 52-60.

[49] Metzler, R., Klafter, J., 2000. The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys. Rep. 339, 1–77.

[50] Mitchell, J., Graf von der Schulenburg, D.A., Holland, D.J., Fordham, E.J., Johns, M.L., Gladden, L.F., 2008. Determining NMR flow propagator moments in porous rocks without the influence of relaxation. Journal of Magnetic Resonance 193, 218-225.

[51] Montroll, E.W., Weiss, G.H., 1965. Random walk on lattices. J. Math. Phys. 6, 167-181.

[52] Niehren, S., Kinzelbach, W., 1998. Artificial colloid tracer tests: development of a compact on-line microsphere counter and application to soil column experiments. Journal of Contaminant Hydrology, 35, 249-259.

[53] Parker, J. C., Valocchi, A.J., 1986. Constraints on the validity of equilibrium and first-order kinetic transport models in structured soils, Water Resour. Res., 22, 399–407.

[54] Pickens, J.F., Grisak, G.E., 1981a. Scale-dependent dispersion in a stratified granular aquifer. Water Resour. Res. 17 (4), 1191–1211.

[55] Pickens, J.F., Grisak, G.E., 1981b. Modeling of scale-dependent dispersion in hydrogeologic systems. Water Resour. Res. 17 (6), 1701–1711.

[56] Saiers, J.E., Hornberger, G.M., Hervey, C., 1994. Colloidal silica transport through structured, heterogeneous porous media. J. Hydrol. 163, 271–288.

[57] Scheven, U.M., Verganelakis, D., Harris, R., Johns, M.L., Gladden, L.F., 2005. Quantitative nuclear magnetic resonance measurements of preasymptotic dispersion in flow through porous media. Phys Fluids, 17, 117107

[58] Silliman, S.E., Simpson, E.S., 1987. Laboratory evidence of the scale effect in dispersion of solutes in porous media. Water Resour. Res. 23 (8), 1667–1673.

[59] Skopp, J., Gardner, W. R., Tyler, E. J., 1981. Solute movement in structured soils: Two-region model with small interaction, Soil Sci. Soc. Am. J., 45, 837–842.

[60] Soares-Frazao, S., Guinot, V., 2007.A second-order semi-implicit hybrid scheme for one-dimensional Boussinesq-type waves in rectangular channels. International Journal for Numerical Methods in Fluids, 58, 237–261.

[61] Sternberg, S.P.K., Cushman, J., Greenkorn, R.A., 1996. Laboratory observation of nonlocal dispersion. Transp. Porous Media 23, 135–151.

[62] Sun, H., Zhang, Y., Chen, W., Reeves, DM., 2014. Use of a variable-index fractional-derivative model to capture transient dispersion in heterogeneous media. Journal of Contaminant Hydrology, 157, 47-58.

[63] Tran Ngoc, T.D., Lewandowska, J., Vauclin, M., Bertin, H., 2011. Two-scale modeling of solute dispersion in unsaturated double-porosity media: homogenization and experimental validation. Int. J. Numer. Anal. Meth. Geomech. 35, 1536–1559.

[64] Van Genuchten, M.T., Wierenga, P.G., 1977. Mass transfer studies in sorbing porous media: II. Experimental evaluation with tritium (3H2O). Soil Sci. Soc. Am. J. 41, 272–278.

[65] Van Genuchten, M. T., Tang,D. H., Guennelon, R., 1984. Some exact solutions for solute transport through soils containing large cylindrical macropores, Water Resour. Res., 20, 335–346.

[66] Wiener, N., 1923. Differential space. Journal of Mathematics and Physics, MIT, 2, 131–174.

[67] Yates, S.R., 1990. An analytical solution for one-dimension transport in heterogeneous porous media. Water Resour. Res. 26 (10), 2331–2338.

[68] Yates, S.R., 1992. An analytical solution for one-dimensional transport in porous media with an exponential dispersion function. Water Resour. Res. 28 (8), 2149–2154.

[69] Zhang, R., Huang, K., Xiang, J., 1994. Solute movement through homogeneous and heterogeneous soil columns. Adv. Water Resour. 17 (5), 317–324.

[70] Zhou, L., Selim, H.M., 2002. A conceptual fractal model for describing time-dependent dispersivity. Soil Sci. 167 (3), 173–183.