Modelling solute dispersion in periodic heterogeneous porous media: model benchmarking against intermediate scale experiments
Samer Majdalani, Vincent Guinot, Carole Delenne, Hicham Gebran

To cite this version:
Abstract

This paper is devoted to theoretical and experimental investigations of solute dispersion in heterogeneous porous media. Dispersion in heterogeneous 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 and large time scale behaviours. 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 yield infinite signal propagation speeds. An infinite concentration 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 times 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 the 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:

$$\frac{\partial}{\partial t} c + u \frac{\partial}{\partial x} c - D \frac{\partial^2}{\partial x^2} c = 0$$

(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

$$\frac{\partial}{\partial t} c_m + u \frac{\partial}{\partial x} c_m - D \frac{\partial^2}{\partial x^2} c_m = \frac{k}{\Theta} (c_m - c_m)$$

(2a)

$$\frac{\partial}{\partial t} c_{im} = \frac{k}{(1 - \Theta)} (c_m - c_{im})$$

(2b)

$$\Theta = \frac{\theta_m}{\theta_m + \theta_{im}}$$

(2c)

where $c_m$ and $c_{im}$ are respectively the concentrations in the mobile and immobile regions, $\theta_m$ and $\theta_{im}$ are respectively the water contents of the mobile and immobile regions, $\Theta$ 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:

$$\partial_t c_m + u \partial_x c_m - D \partial_{xx} c_m = \sum_{j=1}^{N} \frac{k_j}{\Theta_j} ((c_{im})_j - c_m)$$

$$\partial_t (c_{im})_j = \frac{k_j}{\Theta_j} (c_m - (c_{im})_j), j = 1, \ldots, N$$

Where $N$ is the number of immobile regions with normalized water content $\Theta_j$, and $k_j$ ($j = 1, \ldots, N$) 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 adsorption/desorption terms are set to zero. Normalizing the water contents as in the previous subsection leads to the following governing equations:

$$\partial_t c_i + u_i \partial_x c_i - D_i \partial_{xx} c_i = \sum_{j \neq i} \frac{k_{ij}}{\Theta_j} (c_j - c_i), \ i = 1, \ldots, R$$

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

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

and the total concentration in the porous medium is defined as

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

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 Appendix A, the variance $c^{(2)}$ of the solute particle locations obeys the following equation

$$d_t c^{(2)} = \sum_{i=1}^{R} \Theta_i \left(v_i c_i^{(1)} + 2D_i\right), \ v_i \equiv u_i - \sum_{j=1}^{R} \Theta_j u_j$$
where \( c_i^{(1)} \) is the average abscissa of the particle locations in the region \( i \), obeying

\[
d_t c_i^{(1)} = v_i + \sum_{j \neq i} k_{ij} \frac{\Theta_j}{\Theta_i} \left( c_j^{(1)} - c_i^{(1)} \right)
\]  

(7)

As shown in Appendix B (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 \( v \equiv [v_1, \ldots, v_R] \) belongs to the range of the matrix \( M \) defined as

\[
M_{ij} = \begin{cases} 
-\sum_{p \neq i} k_{jp} \frac{\Theta_p}{\Theta_i} & \text{if } i = j \\
\frac{k_{ij} \Theta_i}{\Theta_j} & \text{if } i \neq j
\end{cases}
\]  

(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 (Appendix A)

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

(9)

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

A direct consequence of the above property is the following (Appendix B, 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 \( M \) 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 Advection 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 $\Theta_i$ 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:

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

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 $\frac{k_{ij}}{v_j}$ with $k_{ij} \Theta_i \Theta_j$. As shown by equation (9), a finite dispersion coefficient $D_\infty$ 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 $M$ (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 Appendix A, 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 $\Theta_1, \Theta_2$, 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 becomes a satisfactory approximation 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 MP3TM). 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 325TM) and saved on a data logger (Campbell CR1000TM) every 5 s. The cumulated outlet volume $V$ is measured by weighing the effluent every 5s 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.

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

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 (\( N = 1 \)).

![Definition sketch for column connection. Top: experimental setup reported in [48]. Bottom: present experimental setup.](image)

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 $3 \times 10^{-3}$ ms$^{-1}$, while the speed of the slower one is approximately $7 \times 10^{-4}$ 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.
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 such a way that the CFL in the fastest flow region was equal to unity. For CFL = 1,
Numerical diffusion is known to be zero and the exact solution of 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 $L^p$-norm of the modeling 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

$$J_1 = \left( \frac{1}{T} \int_0^T |e(t)|^p \, dt \right)^{\frac{1}{p}}, \quad p > 0$$

**Modeling error definition.** The measure of the modeling error was defined on the basis of the following considerations: (i) the modeling error is zero whenever the modeled 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:

\[
e(t) = \begin{cases} 
\frac{\partial_t c_{\text{min}}(t) - \partial_t c_{\text{mod}}(t)}{\Delta \partial t c} & \text{if } \partial_t c_{\text{mod}}(t) < \partial_t c_{\text{min}}(t) \\
0 & \text{if } \partial_t c_{\text{min}}(t) \leq \partial_t c_{\text{mod}}(t) \leq \partial_t c_{\text{max}}(t) \\
\frac{\partial_t c_{\text{mod}}(t) - \partial_t c_{\text{max}}(t)}{\Delta \partial t c} & \text{if } \partial_t c_{\text{mod}}(t) > \partial_t c_{\text{max}}(t)
\end{cases}
\] (15a)

\[
\partial_t c_{\text{min}}(t) \equiv \min_r \partial_t c_{\text{expe},r}(t), \quad \partial_t c_{\text{max}}(t) \equiv \max_r \partial_t c_{\text{expe},r}(t), \quad \Delta \partial t c = \partial_t c_{\text{max}}(t) - \partial_t c_{\text{min}}(t) \quad (15b)
\]

where \(\partial_t c_{\text{expe},r}(t)\) is the derivative of the concentration measured at time \(t\) for the replicate number \(r\) of the experiment, \(c_{\text{mod}}(t)\) 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 artificially 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:

\[
e(t) = \begin{cases} 
\frac{\partial_t c_{\text{av}}(t) - \partial_t c_{\text{mod}}(t)}{\sigma(t)} & \text{if } \partial_t c_{\text{mod}}(t) \leq \partial_t c_{\text{av}}(t) \\
\frac{\partial_t c_{\text{mod}}(t) - \partial_t c_{\text{av}}(t)}{\sigma(t)} & \text{if } \partial_t c_{\text{mod}}(t) \geq \partial_t c_{\text{av}}(t)
\end{cases}
\] (16a)

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

**Calibration algorithm.** In the following, the calibration is performed on five experiments simultaneously \((N = 1 \ldots 5)\) 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 \(J_2\):

\[
J = \sum_{p=1}^{5} J_1(p) J_2(p) \quad (17)
\]

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 \(N_{\text{bits}} = 6\) bits = 64 values). As proposed in [47], the initial population is composed of \(N_{\text{pop}} = 8\) chromosomes,
each of which is the concatenation of the various genes. The population is thus represented by a
\( N_{\text{pop}} \times (N_p \cdot N_{\text{bits}}) \) 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.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(D)</td>
<td>Fickian dispersion coefficient</td>
<td>(4.041 \times 10^{-5} \text{ m}^2\text{s}^{-1})</td>
</tr>
<tr>
<td>(u)</td>
<td>Advection velocity</td>
<td>(7.193 \times 10^{-4} \text{ ms}^{-1})</td>
</tr>
<tr>
<td>(J)</td>
<td>Objective function</td>
<td>7.76</td>
</tr>
</tbody>
</table>

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.
4.3. MI model

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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$</td>
<td>Fickian dispersion coefficient</td>
<td>$3.720 \times 10^{-7}$ m$^2$s$^{-1}$</td>
</tr>
<tr>
<td>$k$</td>
<td>Exchange coefficient</td>
<td>$7.751 \times 10^{-2}$ s$^{-1}$</td>
</tr>
<tr>
<td>$u$</td>
<td>Advection velocity</td>
<td>$4.471 \times 10^{-3}$ ms$^{-1}$</td>
</tr>
<tr>
<td>$\Theta_{im}$</td>
<td>Normalised immobile fraction</td>
<td>0.8333</td>
</tr>
<tr>
<td>$\Theta_m$</td>
<td>Normalised mobile fraction</td>
<td>0.1667</td>
</tr>
<tr>
<td>$J$</td>
<td>Objective function</td>
<td>3.505</td>
</tr>
</tbody>
</table>

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.
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.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$</td>
<td>Fickian dispersion coefficient</td>
<td>$6.200 \times 10^{-7} \text{m}^2 \text{s}^{-1}$</td>
</tr>
<tr>
<td>$k_1$</td>
<td>Exchange coefficient with region 1</td>
<td>$9.303 \times 10^{-1} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$k_2$</td>
<td>Exchange coefficient with region 2</td>
<td>$6.820 \times 10^{-2} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$k_3$</td>
<td>Exchange coefficient with region 2</td>
<td>$6.519 \times 10^{-2} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$u$</td>
<td>Advection velocity</td>
<td>$4.471 \times 10^{-3} \text{m} \text{s}^{-1}$</td>
</tr>
<tr>
<td>$\Theta_1$</td>
<td>Normalised immobile fraction 1</td>
<td>0.1017</td>
</tr>
<tr>
<td>$\Theta_2$</td>
<td>Normalised immobile fraction 2</td>
<td>0.1155</td>
</tr>
<tr>
<td>$\Theta_3$</td>
<td>Normalised immobile fraction 3</td>
<td>0.5775</td>
</tr>
<tr>
<td>$\Theta_m$</td>
<td>Normalised mobile water content</td>
<td>0.1658</td>
</tr>
<tr>
<td>$J$</td>
<td>Objective function</td>
<td>1.975</td>
</tr>
</tbody>
</table>

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.
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: $\Theta_1$, $k_{12}$, $u_1$, $u_2$, $D_1$ and $D_2$. The best parameter set (Figure 7) is given in Table 5.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1$</td>
<td>Fickian dispersion coefficient in region 1</td>
<td>$1.034 \times 10^{-5}$ m$^2$ s$^{-1}$</td>
</tr>
<tr>
<td>$D_2$</td>
<td>Fickian dispersion coefficient in region 2</td>
<td>$1.897 \times 10^{-5}$ m$^2$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_{12}$</td>
<td>Exchange coefficient</td>
<td>$1.570 \times 10^{-2}$ s$^{-1}$</td>
</tr>
<tr>
<td>$u_1$</td>
<td>Flow velocity in region 1</td>
<td>$2.406 \times 10^{-3}$ ms$^{-1}$</td>
</tr>
<tr>
<td>$u_2$</td>
<td>Flow velocity in region 2</td>
<td>$5.28 \times 10^{-4}$ ms$^{-1}$</td>
</tr>
<tr>
<td>$\Theta_1$</td>
<td>Normalised water content 1</td>
<td>0.124</td>
</tr>
<tr>
<td>$\Theta_2$</td>
<td>Normalised water content 2</td>
<td>0.876</td>
</tr>
<tr>
<td>$J$</td>
<td>Objective function</td>
<td>1.61</td>
</tr>
</tbody>
</table>

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.
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: \( \Theta_1 \), \( \Theta_2 \), \( k_{12} \), \( k_{13} \), \( k_{23} \), \( u_1 \), \( u_2 \) 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 the 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).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k_{12} )</td>
<td>Exchange coefficient between regions 1 and 2</td>
<td>( 1.518 \times 10^{-4} ) s(^{-1} )</td>
</tr>
<tr>
<td>( k_{13} )</td>
<td>Exchange coefficient between regions 1 and 3</td>
<td>( 1.252 \times 10^{-3} ) s(^{-1} )</td>
</tr>
<tr>
<td>( k_{23} )</td>
<td>Exchange coefficient between regions 2 and 3</td>
<td>( 5.803 \times 10^{-3} ) s(^{-1} )</td>
</tr>
<tr>
<td>( u_1 )</td>
<td>Flow velocity in region 1</td>
<td>( 3.942 \times 10^{-3} ) ms(^{-1} )</td>
</tr>
<tr>
<td>( u_2 )</td>
<td>Flow velocity in region 2</td>
<td>( 8.668 \times 10^{-4} ) ms(^{-1} )</td>
</tr>
<tr>
<td>( u_3 )</td>
<td>Flow velocity in region 3</td>
<td>( 1.079 \times 10^{-4} ) ms(^{-1} )</td>
</tr>
<tr>
<td>( \Theta_1 )</td>
<td>Normalised water content 1</td>
<td>0.1389</td>
</tr>
<tr>
<td>( \Theta_2 )</td>
<td>Normalised water content 2</td>
<td>0.1430</td>
</tr>
<tr>
<td>( \Theta_3 )</td>
<td>Normalised water content 3</td>
<td>0.7271</td>
</tr>
<tr>
<td>( J )</td>
<td>Objective function</td>
<td>1.592</td>
</tr>
</tbody>
</table>

Table 6: Three region PAMR model. Calibration results.
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: $\Theta_1$, $k_{12}$, $k_{13}$, $k_{34}$, $u_1$, $u_2$, $u_3$ 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.

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

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.
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.

<table>
<thead>
<tr>
<th>Model</th>
<th>No. parameters</th>
<th>$J$ calibration eq.(16a)</th>
<th>$J$ validation eq.(15a)</th>
<th>$J$ validation eq.(16a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>2</td>
<td>7.750</td>
<td>2.63 x 10^{-2}</td>
<td>1.600</td>
</tr>
<tr>
<td>MI</td>
<td>4</td>
<td>3.505</td>
<td>3.45 x 10^{-2}</td>
<td>1.948</td>
</tr>
<tr>
<td>MR</td>
<td>8</td>
<td>1.975</td>
<td>2.63 x 10^{-2}</td>
<td>1.476</td>
</tr>
<tr>
<td>MRAD</td>
<td>6</td>
<td>1.610</td>
<td>6.35 x 10^{-2}</td>
<td>2.391</td>
</tr>
<tr>
<td>PAMR 3</td>
<td>8</td>
<td>1.592</td>
<td>3.40 x 10^{-2}</td>
<td>1.753</td>
</tr>
<tr>
<td>PAMR 4</td>
<td>8</td>
<td>1.608</td>
<td>3.43 x 10^{-2}</td>
<td>1.747</td>
</tr>
</tbody>
</table>

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.
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 increase 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 Appendix A and the stability analysis in Appendix B).

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 Appendix A.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 Appendix A.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 Appendix A.2.4 shows that \( D_{loc} \) converges faster to the asymptotic value than does \( D_{av} \). Indeed, the difference \( D_{\infty} - D_{loc} \) is a decreasing exponential function of time, while \( D_{\infty} - D_{av} \) 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 \( D_{\infty} \) 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 the 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 the 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 periodical 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.
Appendix A. Equations of moments

Appendix A.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 $u$:

$$\partial_t c - D \partial_{xx} c = 0$$  \hfill (A.1)

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

$$c^{(p)} = \int_{-\infty}^{+\infty} x^p c(x, t) \, dx$$  \hfill (A.2)

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

$$\int_{-\infty}^{+\infty} x^p (\partial_t c - D \partial_{xx} c) \, dx = 0$$  \hfill (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

$$\frac{dc^{(0)}}{dt} = 0 \Rightarrow c^{(0)}(t) = 1$$  \hfill (A.4a)

$$\frac{dc^{(1)}}{dt} = u \Rightarrow c^{(1)}(t) = 0$$  \hfill (A.4b)

$$\frac{dc^{(2)}}{dt} - 2D = 0 \Rightarrow c^{(2)}(t) = 2Dt$$  \hfill (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 ($D_{\text{loc}}$) and an average dispersion ($D_{\text{av}}$):

$$D_{\text{loc}} = \frac{1}{2} \frac{dc^{(2)}}{dt}$$  \hfill (A.5a)

$$D_{\text{av}} = \frac{1}{2t} c^{(2)}$$  \hfill (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...
Appendix A.2. Spatial moments for the MRAD model

Appendix A.2.1. Governing equations

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

\[
\Theta_i \partial_t c_i + \Theta_i v_i \partial_x c_i - \Theta_i D_i \partial_{xx} c_i = \sum_{j \neq i} k_{ij} (c_j - c_i) \quad (A.6a)
\]

\[
k_{ij} = k_{ji}, \quad \sum_{i=1}^{R} \Theta_i v_i = 0 \quad (A.6b)
\]

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

\[
d_t c_i^{(0)} = \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} (c_j^{(0)} - c_i^{(0)}) \quad (A.7)
\]

Since \( c_i^{(0)} (t = 0) = 1 \forall i \), \( d_t c_i^{(0)} \) remains identically zero at all times, which results in

\[
c_i^{(0)} (t) = 1 \forall t \geq 0 \quad (A.8)
\]

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

\[
d_t c_i^{(1)} = v_i c_i^{(0)} + \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} (c_j^{(1)} - c_i^{(1)}) = v_i + \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} (c_j^{(1)} - c_i^{(1)}) \quad (A.9)
\]

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

\[
d_t c_i^{(2)} = v_i c_i^{(1)} + 2D_i c_i^{(0)} + \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} (c_j^{(2)} - c_i^{(2)}) = v_i c_i^{(1)} + 2D_i + \sum_{j \neq i}^{R} \frac{k_{ij}}{\Theta_i} (c_j^{(2)} - c_i^{(2)}) \quad (A.10)
\]

Multiplying by \( \Theta_i \) and summing over \( i = 1, \ldots, R \), using the property \( \sum_{i=1}^{R} \Theta_i v_i = 0 \) yields equation (6).
Appendix A.2.2. Small time behaviour

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

$$\left( d_t c_i^{(1)} \right)_{t=0} = v_i$$  \hspace{1cm} (A.11)

Consequently the following equivalence holds

$$c_i^{(1)}(t) \sim v_it$$  \hspace{1cm} (A.12)

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

$$d_t c_i^{(2)} = R \sum_{i=1}^{R} \Theta_i (v_i^2 + 2D_i)$$  \hspace{1cm} (A.13)

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

$$D_{\text{loc}} = R \sum_{i=1}^{R} \Theta_i (v_i^2 t + D_i)$$  \hspace{1cm} (A.14a)
$$D_{\text{av}} = R \sum_{i=1}^{R} \Theta_i \left( \frac{1}{2} v_i^2 t + D_i \right)$$  \hspace{1cm} (A.14b)

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

Appendix A.2.3. Long time behaviour

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

$$c_i^{(1)} \rightarrow c_i^{(1,\infty)} \forall i$$  \hspace{1cm} (A.15)

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

$$d_t c_i^{(2)} = R \sum_{i=1}^{R} \Theta_i \left( v_i c_i^{(1,\infty)} + 2D_i \right)$$  \hspace{1cm} (A.16)

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

Appendix A.2.4. A particular case: the two region model

The simplest possible model consists of two regions:

$$R = 2, \Theta_2 = 1 - \Theta_1, v_2 = -\frac{\Theta_1}{\Theta_2}v_1$$  \hspace{1cm} (A.17)
Solving equations (A.9, 6) under assumptions (A.17) yields the following solution

\[ c_i^{(1)} = \frac{\Theta_1 \Theta_2}{k_{12}} \left( 1 - \exp \left( -\frac{k_{12}}{\Theta_1 \Theta_2} t \right) \right) v_i, \quad i = 1, 2 \]  

\[ c_i^{(2)} = \frac{\Theta_1 \Theta_2}{k_{12}} \left( \Theta_1 v_1^2 + \Theta_2 v_2^2 \right) \left( t + \frac{\Theta_1 \Theta_2}{k_{12}} \left( \exp \left( -\frac{k_{12}}{\Theta_1 \Theta_2} t \right) - 1 \right) \right) + 2 (\Theta_1 D_1 + \Theta_2 D_2) t \]  

which yields

\[ D_{\text{loc}} = \frac{1}{2} \frac{\Theta_1 \Theta_2}{k_{12}} \left( \Theta_1 v_1^2 + \Theta_2 v_2^2 \right) \left( 1 - \exp \left( -\frac{k_{12}}{\Theta_1 \Theta_2} t \right) \right) + \Theta_1 D_1 + \Theta_2 D_2 \]  

\[ D_{\text{av}} = \frac{1}{2} \frac{\Theta_1 \Theta_2}{k_{12}} \left( \Theta_1 v_1^2 + \Theta_2 v_2^2 \right) \left( 1 + \Theta_1 \Theta_2 \frac{\exp \left( -\frac{k_{12}}{\Theta_1 \Theta_2} t \right) - 1}{kt} \right) + \Theta_1 D_1 + \Theta_2 D_2 \]

with the following limit behaviours

\[ D_{\text{loc}} \sim \Theta_1 D_1 + \Theta_2 D_2 + \frac{1}{2} \Theta_1 \Theta_2 \left( \Theta_1 v_1^2 + \Theta_2 v_2^2 \right) t \]  

\[ D_{\text{av}} \sim \Theta_1 D_1 + \Theta_2 D_2 + \frac{1}{4} \Theta_1 \Theta_2 \left( \Theta_1 v_1^2 + \Theta_2 v_2^2 \right) t \]  

\[ D_{\text{loc}} \sim D_{\text{av}} \sim D_{\infty} = \Theta_1 D_1 + \Theta_2 D_2 + \frac{1}{2} \frac{\Theta_1 \Theta_2}{k_{12}} \left( \Theta_1 v_1^2 + \Theta_2 v_2^2 \right) \]  

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]:

\[ D_{\infty} = \frac{a^2}{8k_{12}} \]  

with \( k_{12} = \frac{1}{4\tau} \), \( \tau \) 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 \( D_{\infty} \). Equations (A.19a, A.19b) can be rewritten as

\[ D_{\infty} - D_{\text{av}} = A \frac{\exp(-Kt)}{Kt} \sim \frac{A}{Kt} \]  

(A.22a)
\[ D_\infty - D_{\text{loc}} = A \exp(-Kt) \]  
(A.22b)

\[ A = \frac{K}{2} (\Theta_1 \nu_1^2 + \Theta_2 \nu_2^2), \quad K = \frac{k_{12}}{\Theta_1 \Theta_2} \]  
(A.22c)

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

**Appendix B. 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

\[ \frac{dx}{dt} = v + Mx \]  
(B.1a)

\[ x \equiv [c_1^{(1)}, \ldots, c_R^{(1)}]^T, \quad v \equiv [v_1, \ldots, v_R]^T, \quad M = [M_{ij}] \]  
(B.1b)

where the elements \( M_{ij} \) are defined as in (8). As shown in Appendix A, 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 \( M \) can be written in the form \( M = -SD \), with

\[ D = \text{diag} [\Theta_i], \quad S_{ij} = -\frac{M_{ij}}{\Theta_j} = \begin{cases} \sum_{p \neq i} \frac{k_{ip}}{\Theta_i \Theta_p} & \text{if } i = j \\ -\frac{k_{ij}}{\Theta_i \Theta_j} & \text{if } i \neq j \end{cases} \]  
(B.2)

Consequently, the matrix \( S \) is symmetric.

**Result 1.** The matrix \( S \) is positive semidefinite.

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

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

\[ \square \]

**Corollary 1.** The eigenvalues of \( S \) are all positive.

**Proof.** Let \( \lambda \) be an eigenvalue of \( S \) and \( x \) a corresponding eigenvector. Then

\[ x^T S x = x^T \lambda x = \lambda x^T x = \lambda \| x \|^2 \geq 0 \]  
(B.4)

\[ \square \]

**Theorem 1.** Let \( S \) and \( D \) be symmetric and positive definite matrices. Then the eigenvalues of \( SD \) are all strictly positive.
**Proof.** There exists an orthogonal matrix $C$ such that $S = Cd[\lambda_i]C^T$. Defining $S^{1/2} \equiv Cd[\sqrt{\lambda_i}]C^T$, it is noted that the matrices $S^{1/2}DS^{1/2}$ and $SD$ are similar because $SD = S^{1/2}S^{1/2}DS^{1/2} - S^{-1/2}$ (note that $S^{-1/2}$ exists because the $\sqrt{\lambda_i}$ are all nonzero). Therefore the eigenvalues of $SD$ are the same as the eigenvalues of $S^{1/2}DS^{1/2}$. Moreover, the matrix $S^{1/2}DS^{1/2}$ is symmetric and positive definite because $D$ is symmetric, positive definite and $S^{1/2}$ is symmetric. Consequently, the eigenvalues of $SD$ are all strictly positive. $\square$

**Corollary 2.** The eigenvalues of $-SD$ are all strictly negative.

**Theorem 2.** Let $S$ and $D$ be symmetric and positive semidefinite matrices. Then the eigenvalues of $SD$ are all positive.

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

**Corollary 3.** The eigenvalues of $M = -SD$ are all negative.

**Result 2.** The solutions of the differential system $d_x = Mx$ converge to an equilibrium solution.

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

$$d_t Q = \sum_{i=1}^{R} \frac{1}{\delta_i} x_i d_t x_i = (D^{-1}x)^T d_t x = (D^{-1}x)^T M^T x = x^T D^{-1} M^T x = -x^T Sx \quad (B.5)$$

From Result 1, the matrix $-S$ is negative, semidefinite. Consequently, $d_t Q \leq 0$ and $0 \leq Q(t) \leq Q(0)$ for $t > 0$. Consequently, the solution $x(t)$ is bounded. Since the eigenvalues of $M^T$ are real and negative (from Corollary 3), it follows from the theory of linear differential equations that there exists a vector $x_\infty$ such that $x \rightarrow x_\infty$. $\square$

**Result 3.** If $v \in \text{rge} M$, the solution of the non-homogeneous system $d_t x = v + Mx$ converge to an equilibrium solution.

**Proof.** By assumption, there exists a such that $Ma = v$. Then $d_t x = v + Mx \iff d_t (a + x) = v + Mx = M(a + x)$. Consequently, $a + x$ satisfies the homogeneous system $d_t x = Mx$ and converges to an equilibrium solution, hence the result.
Result 4. If $M$ is block-diagonal, $M = \begin{bmatrix} M_{11} & \cdots & M_{1r} \\ \vdots & \ddots & \vdots \\ M_{r1} & \cdots & M_{rr} \end{bmatrix}$, a necessary condition for $v = [v_1, \ldots, v_R]^T$ to belong to the range of $M$ is

$$\sum_{i=1}^{r} \Theta_i v_i = 0 \quad \text{(B.6)}$$

Proof. If $v$ belongs to the range of $M$, there exists a vector $a$ such that $v = Ma$. Then, using the symmetry property $k_{ij} = k_{ji}$, one has

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

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


