Benchmarking numerical codes for tracer transport with the aid of laboratory-scale experiments in 2D heterogeneous porous media

Fadji Hassane Maina[1], Philippe Ackerer[1], Anis Younes[1], Alberto Guadagnini[2,3], Brian Berkowitz[4]

1. Hydrology and Geochemistry Laboratory – LHyGeS, Strasbourg University - CNRS, 67000 Strasbourg, France.

2. Dipartimento di Ingegneria Civile e Ambientale (DICA), Politecnico di Milano, Piazza L. Da Vinci, 32, 20133 Milano, Italy

3. Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, Arizona, USA

4. Department of Earth and Planetary Sciences, Weizmann Institute of Science, Rehovot 7610001 Israel

Accepted in: **Journal of Contaminant Hydrology**

ABSTRACT

We present a combined experimental and numerical modeling study that addresses two principal questions: (i) is any particular Eulerian-based method used to solve the classical advection-dispersion equation (ADE) clearly superior (relative to the others), in terms of yielding solutions that reproduce BTCs of the kind that are typically sampled at the outlet of a laboratory cell? and (ii) in the presence of matches of comparable quality against such BTCs, do any of these methods render different (or similar) numerical BTCs at locations within the domain? To address these questions, we obtained measurements from carefully controlled laboratory experiments, and employ them as a reference against which numerical results are benchmarked and compared. The experiments measure solute transport breakthrough curves (BTCs) through a square domain containing various configurations of coarse, medium, and fine quartz sand. The approaches to solve the ADE involve Eulerian-Lagrangian and Eulerian (finite volume, finite elements, mixed and discontinuous finite elements) numerical methods. Model calibration is not examined; permeability and porosity of each sand were determined previously through separate, standard laboratory tests, while dispersivities are assigned values proportional to mean grain size. We find that the spatial discretization of the flow field is of critical importance, due to the non-uniformity of the domain. Although simulated BTCs at the system outlet are observed to be very similar for these various numerical methods, computed local (point-wise, inside the domain) BTCs can be very different. We find that none of the numerical methods is able to fully reproduce the measured BTCs. The impact of model parameter uncertainty on the calculated BTCs is characterized through a set of numerical Monte Carlo simulations; in cases where the impact is significant, assessment of simulation matches to the experimental data can be ambiguous.

1. Introduction

Quantification of contaminant transport in porous media often relies on models based on various forms of the advection-dispersion equation (ADE) (e.g., Bear, 1972; Berkowitz et al., 2000). Careful assessments of this model, however, in the context of comparison to carefully controlled laboratory measurements, are surprisingly limited in the literature (e.g., Silliman and Simpson, 1987; Berkowitz et al., 2002; Levy and Berkowitz, 2003; Jose and Cirpka, 2004; Loyaux-Lowniczak et al., 2012). Historically, laboratory experiments on dispersion have focused on column experiments with one-dimensional (on average) flow, yielding temporal breakthrough curve (BTC) measurements at the column outlet (see for example Klotz et al., 1980 who reported experimental results for 4000 column tests). Significantly, too, these columns have traditionally contained macroscopically homogeneous packings of porous material.

Transport experiments in rectangular flow cells with two-dimensional flow fields and heterogeneous packing arrangements with different types of porous media are particularly scarce (e.g. Silliman et al., 1998; Levy and Berkowitz, 2003; Konz et al., 2009; Chiogna et al., 2010 and references therein). Controlled transport experiments in fully three-dimensional flow fields and domain are even scarcer (Danquigny et al., 2004; Oswald and Kinzelbach, 2004).

And yet, notwithstanding the paucity of such studies, the need for careful study of the ability of popular models to reproduce or match measurements from controlled experiments seems clear. In particular, such analyses should be a prerequisite to application of these models at larger field scales, for which information on the structural and hydraulic properties of the domain, and on initial and boundary conditions, is often scarce and/or uncertain. Moreover, it should be recognized that high-resolution, laboratory-scale experiments are particularly critical for numerical method verification when analytical solutions are not available.

Here, we present a combined experimental and numerical modelling study of transport in fully water-saturated, heterogeneous porous media. We address two principal questions: (i) is any particular Eulerian-based method (i.e., finite volume, mixed and discontinuous finite element, Eulerian-Lagrangian finite element) used to solve the classical advection-dispersion equation (ADE) clearly superior (relative to the others), in terms of yielding solutions that reproduce BTCs of the kind that are typically sampled at the outlet of a laboratory cell? and (ii) in the presence of matches of comparable quality against such BTCs, do any of these

methods render different (or similar) numerical BTCs at (point-wise) locations within the domain?

To address these questions, we obtained measurements from carefully controlled laboratory "benchmarking" experiments, and then examined the ability of various numerical solutions of the ADE to simulate the transport behavior measured in these domains. Experimental data are well suited for numerical code benchmarking if (i) the related uncertainties on experimental data are estimated, and (ii) simulations can be performed with limited calibration when model parameters have been estimated independently together with their associated uncertainty. Furthermore, experimental data are essential when analytical solutions are not available. Because field experiments cannot avoid model calibration due to unknown aquifer heterogeneity and the difficulty and uncertainty in handling boundary and initial conditions, laboratory scale experiments are a valuable tool for numerical benchmarking. This is true especially under non-uniform flow conditions, due to boundary conditions, structural heterogeneity of the porous medium and/or a combination thereof.

The experimental setup and the mathematical model and numerical methods are described in section 2. The benchmarking of the numerical codes is detailed in section 3 and the ability of the ADE to capture the salient features of the observed data, employing various solution techniques, is discussed in section 4.

2. Materials and Methods

2.1. Experimental set-ups

The experiments were conducted in flow cells of internal dimensions 24.9 × 24.9 × 1.5 cm. Two reservoirs contained two solutions: fresh water and saline water. The fresh solution contained 75% tap water and 25% double deionized water. The saline solute contained the same ratio of fresh and saline water with an additional 500 mg/L of NaCl; chloride was considered as the tracer to be measured.

Three grain sizes were used in this study, and the cell was packed three times with various configurations of these sands: once with medium grain sand (uniformly), and twice with various arrangements of fine, medium and large grain sizes (non-uniform). These sands are clean, sieved and well-rounded quartz sands (UNIMIN, USA) with minimal surface coatings (99.8% pure SiO_2, as reported by UNIMIN). The experiments were conducted with different porous medium structures; interfaces between the different sands were relatively discontinuous (sharp). The hydraulic conductivity and porosity of each of the three sands were estimated previously from constant head column experiments (Levy and Berkowitz,

2003). For each configuration, the flow cell was carefully packed with the sand being added under water, with shaking, stirring and pressing, to eliminate any bubbles and to achieve as uniform a packing as possible. This method reduces the possible occurrence of preferential pathways and channeling; the resulting flow patterns (see, for example, Fig. 4 below, with similar results for the other configurations) indicate that the packings were uniform.

The parameter set is listed in Table 1 and the different sand packing arrangements are shown in Fig. 1. Detailed geometries of the heterogeneous domain are given in Figs. 2 and 3, for sand packing arrangements 2 and 3, respectively. These figures also show the experimental set-up: the inlet and outlet for the fresh water and saline solutions are near the upper left and lower right sides, respectively, as indicated in Figs. 2 and 3. Visualization of a typical resulting flow pattern is shown in Fig. 4.

At initial time, injected saline solution replaced fresh water at the inlet point of the flow cell that contained fresh water with a constant discharge of 4 mL/min. Samples were collected from the outlet over set time intervals. The samples were then measured by an electrical conductivity (EC) meter (TWIN, Japan), with maximum measurement error of 1%, and then converted to tracer concentration values by using calibration curves. Each tracer experiment was repeated to confirm reproducibility and to obtain a first approximation of experimental errors. Experiments were performed in steady-state flow conditions without sink/source terms. Qualitative information is also provided by pictures taken at regular time intervals as shown in Fig. 4 for packing 3.

2.2. Mathematical model

Under steady-state flow conditions without sink/source terms, and assuming that the fluid density remains constant and that the porous medium can be considered as rigid, the flow is modeled by the equation:

$$\nabla \cdot (\mathbf{T} \nabla h) = 0 \tag{1}$$

where $\mathbf{T}$ is the transmissivity tensor (L^2/T) and h the hydraulic head (L). Transmissivity is computed by multiplying the hydraulic conductivity of the porous material by the flow cell thickness (1.5 cm). A constant flux (Neumann condition) was prescribed at the flow cell inlet whereas a constant hydraulic head (Dirichlet condition) was prescribed at the flow cell outlet.

The solute transport is modeled by the classical ADE, given by:

$$\frac{\partial C}{\partial t} + \nabla \cdot (\mathbf{u}C) - \nabla \cdot (\mathbf{D}\nabla C) = 0 \qquad (2)$$

where $C(\boldsymbol{x},t)$ [M/L^3] is the unknown concentration at vector location $\boldsymbol{x}$ and time t, and D [L²/T] is the dispersion tensor defined by:

$$D_{ij} = (\alpha_T |u| + D_e)\delta_{ij} + (\alpha_L - \alpha_T)\frac{u_i u_j}{|u|} \qquad \text{i,j=1,2} \qquad (3)$$

where u [L/T] is the pore water velocity of components u_i, α_L and α_T [L] are the longitudinal and transverse dispersivities, respectively, δ_{ij} is the Kronecker delta function, and D_e [L^2/T] is an effective molecular diffusion coefficient in the porous medium.

Solute concentration was set to zero in the domain, as the initial condition. The concentration was prescribed at the flow cell inlet and the usual dispersion-free boundary condition was applied at the flow cell outlet.

The dispersivity coefficients are often assumed to depend mainly on the grain size and on the uniformity coefficient of the grain size distribution (Klotz et al., 1980). The sand in the packing has a grain size distribution very similar to that of the sand used by Danquigny et al. (2004). Therefore, in all simulations considered in this study, longitudinal and transverse dispersivities are set to the average grain size and to one tenth of the average grain size for the longitudinal and transverse dispersivities, respectively (Table 1). The factor of about 1/10 between both dispersivities is a modeling choice adopted from a variety of practical applications. It has also been estimated from different 2D laboratory scale experiments (e.g., Ballarini et al., 2012).

2.2. Numerical methods and computational codes

Numerical solution of the transport equation, for advective-dominated transport, remains far from trivial. Numerical methods are still under development for accurate solutions that avoid artificial dispersion and/or numerical oscillation with reasonable space and time discretization, i.e., computational costs. Here, we examine different grid-based numerical methods to reproduce the experiments presented in section 2.1: Eulerian methods (finite difference, finite element, mixed hybrid and discontinuous finite element) and Eulerian-Lagrangian methods (characteristic methods and Eulerian-Lagrangian localized adjoint). Adaptive mesh techniques are not considered in the present analysis, although they were already applied to the homogeneous benchmark (packing 1) used here (Esfandiar et al., 2015).

We focus on Eulerian methods for two reasons: (i) they are very common and widely used in research and engineering, (ii) fully Lagrangian methods such as random walks are prone to fluctuations in concentration due to the computation of the dispersive term (Thomson et al., 1984) and are therefore not well suited for solving non-linear problems such as density driven flow and/or reactive transport.

Amongst the numerous computational codes based on the Eulerian approach, MODFLOW and FEFLOW are particularly popular. The numerical methods we consider are embodied in the codes listed below.

MODFLOW (Harbaugh et al., 2000) is the U.S Geological Survey modular finite difference flow model. Finite differences are used to solve the flow equation, and the transport equation is solved by operator splitting, which is a common approach. The ADE is split into two equations: the first equation describes advection and the second, dispersion. This allows for use of a specialized scheme for solving the advective term and a conventional scheme for solving the dispersive term to improve the numerical accuracy. In MODLFOW, the advection term can be solved with Eulerian-Lagrangian methods (method of characteristics, modified characteristic method, and hybrid method of characteristics), the standard finite-difference method, and the third order of the total variation diminishing (TVD) scheme, while the dispersive part is computed by finite differences. The different schemes were tested during the simulation of the flow cell experiments outlined in section 2.1. Eulerian-Lagrangian methods produce oscillations whereas the upstream finite difference scheme leads to excessive numerical dispersion. Therefore, only results provided by the TVD scheme will be presented. The main idea of the TVD scheme is based on the so-called “ultimate algorithm” proposed by Leonard (1988). It assumes that the sum of concentration differences between adjacent nodes diminishes over successive transport steps. This scheme is mass conservative and generates limited numerical dispersion or unphysical oscillations (Zheng and Wang, 1999; Zhang and Shu, 2010). TVD schemes are generally much more accurate in solving advection-dominated problems, as compared to standard methods. In the ultimate scheme, the interface concentrations are determined through a third-order polynomial interpolation of nodal concentration, supplemented by a flux limiter to minimize unphysical oscillations which may occur if sharp concentrations fronts are involved.

FEFLOW (Finite Element subsurface FLOW) solves the governing flow, mass and heat transport equations in porous and fractured media using finite elements (Diersch, 2013). The advective part of the ADE can be solved using full upwinding, no upwind (concentration is computed by linear interpolation), least square upwinding, streamline upwinding, or shock

capturing. The two latter schemes add a second interpolation function to the standard Galerkin function to stabilize the solution in advection dominant problem. A set of preliminary runs revealed that the full upwinding scheme provided the most accurate match to the concentration data from the flow cell experiments. Therefore, the different results showed here have been obtained by this scheme.

TRACES (Transport of RadioACtive Elements in Subsurface, 2009) is a computer program for the simulation of flow and reactive transport in saturated porous media developed at Strasbourg University. The flow equation is solved by mixed hybrid finite elements. The transport equation is split into two parts: the advective part is solved by discontinuous finite elements (explicit scheme in time) and the dispersion by the mixed hybrid finite element method. Mixed finite elements are well suited for solving elliptic and parabolic partial differential equations, which are the mathematical representation of many problems, for instance groundwater flow and dispersion of solutes (Younes et al., 2010; Farhloul and Serghini Mounim, 2005). The Discontinuous Galerkin Finite Element method combined with a slope limiting procedure can solve advective-dominant transport without oscillations and with very limited numerical diffusion (Diaw et al., 2001). Discontinuous Galerkin Finite Elements allow for the computation of sharp fronts due to the discontinuity of the concentration function between adjacent elements. Mass conservation is ensured by the upwind approximation of the convective flux. A slope limiting procedure reduces the variation of concentration within an element to avoid, if necessary, oscillations. This procedure adds limited numerical dispersion at the element level. The time discretization scheme used is explicit, which requires strict adherence to the Courant criterion. This criterion is respected in regions with significant concentration gradients only. To avoid the use of small time steps due to high velocities (e.g., at the inflow or outflow), adhering to this criterion is irrelevant within regions where local concentrations do not vary in time.

ELLAM (Eulerian Lagrangian Localized Adjoint Method), first introduced by Celia et al. (1990), provides a methodology that maintains the accuracy and efficiency of Eulerian and Lagrangian methods, while also conserving mass and systematically treating any type of boundary condition (Russell and Celia, 2002). The approach is based on test functions (linear in this case) with full space time dependence that are solutions to the formal adjoint operator defined on local space time partitions of the domain. The definition of the adjoint operator avoids treating explicitly the hyperbolic (advective) part of the equation in an Eulerian framework. However, it requires the computation of some integrals along the characteristics which can introduce some numerical dispersion and/or oscillations. This drawback has been

reduced by a more accurate interpolation scheme (Younès et al., 2006). The integrals are computed along characteristics using particle tracking, and the other terms of the transport equation are solved by standard Galerkin finite elements. The method is not limited by restrictions on Courant or Peclet numbers and therefore should be useful for solution of advection-dominated transport problems which generally cannot be solved accurately by standard finite-difference and finite-element methods (Healy and Russell, 1993). In this code, the velocity field is calculated by the mixed hybrid finite element formulation of the flow equation which ensures accurate computation of the characteristics.

3. Benchmarking of numerical methods

The simulation of solute transport in porous media requires an appropriate spatial and temporal discretizations that allow computation of an accurate flow field and solute concentration map with negligible numerical diffusion and/or oscillations. Numerical diffusion and oscillations are often addressed in transport simulations through the grid Péclet number (e.g., Kinzelbach, 1986). The accuracy of the computed velocities related to the grid size employed for the solution of the groundwater flow problem in the presence of a given (deterministic) spatial distribution of hydraulic conductivity is addressed only in a limited number of studies, an example being the work of Cainelli et al. (2012).

The three laboratory experiments described in section 2.1 were simulated by the four computational codes described in section 2.2. A first set of simulations was performed to identify the required grid size for an accurate flow field computation. The accuracy of the numerical methods was then analyzed with (i) the BTCs at the experimental model outlet, and (ii) comparison of model simulations of BTCs at some specific locations inside the flow cell. First, we compared the computed flux-averaged BTCs to their measured counterparts taken at the outlet of the flow cell. Second, we compared computed local BTCs, representing resident concentrations, as an additional metric upon which we analyze the relative differences between the numerical methods. Finally, we examined the sensitivity of the model outputs to grid size, by comparing the spatial distribution of concentration inside the flow cell for packing 3, determined from the different numerical methods 50 min from tracer injection. We also provide some information on the code performance in terms of computation.

3.1 Optimization of the spatial and temporal discretizations

The accuracy of the velocity field is estimated through BTCs computed by advective transport only with different grid sizes. These simulations are based on particle tracking

displaced along streamlines, hereafter referred to as Advective Particle Tracking (APT). In each simulation, 100000 particles were initially distributed close to the inlet over a rectangular domain whose opposite corners have the following coordinates: (0.05, 21.0) cm and (0.50, 22.0) cm. The BTCs inside the domain were computed by counting the particle located within a circle of 0.5 cm radius and centered on the reference location (see Figs. 2 and 3 for each packing). Model parameters (hydraulic conductivity, effective porosity) and injection flow rate were derived from previous experiments (Table 1). The flow computation was performed with TRACES. Grid cells are squares with uniform sides $\Delta x = \Delta y$ so that mixed finite elements are equivalent to finite differences (Younès et al., 2010) and the following results can be extended to the flow field computation with MODFLOW. The flow field was computed with grid cells of 4 mm, 2 mm, 1 mm and 0.5 mm. The grid size was found to have no significant effects for the homogeneous test case (packing 1). For the two other test cases, the coarser mesh led to an apparent spreading of the solute as shown in Fig. 5 for packing 2 and Fig. 6 for packing 3. These different travel time distributions are related to the hydraulic head computations whose gradients are smoother for the coarse discretization (Δx = 4 mm) than for a finer discretization (Δx = 1 mm), as it shown in Fig. 7 for packing 2. Grid convergence was reached for a grid cell size of 1 mm; smaller grid sizes did not yield visibly different results. The corresponding grid Péclet numbers are Pe = 1.0, 1.9, and 4.3, respectively, for the fine, medium and coarse sand. Grid convergence is reached even if these grid Péclet numbers do not fulfill the traditional requirement, i.e., Pe << 1 (note that this condition is theoretically established only for one-dimensional uniform flow (Kinzelbach, 1986)).

Similar tests were performed with the triangular grids used for FEFLOW and ELLAM. The triangular grids are obtained from the rectangular grids by dividing each square into two triangles by the square diagonal. The diagonal orientation was changed from square to square, to avoid artificial anisotropy due to the mesh. The tests on triangular grids provide the same results as the rectangular grids. Therefore, the following computations were performed on grids with a spatial discretization of 1 mm size.

The temporal discretization was modified until no significant change appeared in simulated concentrations at the outlet. The maximum time step length was fixed to 0.25 min. In the following simulations, advective transport by APT is displayed also on the BTCs to highlight the contribution of dispersion on the simulated concentration. Finally, simulations with dispersivities 10 times greater than the values reported in Table 1 were performed to evaluate

the sensitivity of the simulated BTCs. No significant changes were observed because the shape of the BTCs is due mainly to advective transport.

3.2 Comparison of results between numerical methods and measured BTCs

The simulations were performed assuming that neither the boundary conditions (input flow rate at the inflow) nor the parameters (see Table 1) are subject to uncertainty. Because the objective of this section is to compare numerical results, the measured BTCs are used here as a reference solution to examine the quality of the numerical results.

For the first set of experiments (homogeneous medium), all computed BTCs compare to the measurements quite well (Fig. 8). From the APT simulation, we see that the shape of the BTC is due mainly to advective transport and that the contribution of the dispersion is negligible. Therefore, even if numerical diffusion exists, it cannot be highlighted by this kind of experiment. The four numerical methods provide very similar results except at the beginning of the BTC (around 50 minutes) where the finite element based flow models (FEFLOW, ELLAM) show an earlier breakthrough.

The simulations of the second set of experiments (packing 2) do not compare well to the experimental results (Fig. 9). Even the global shape of the measured BTC is not well reproduced by the numerical simulations, especially between 60 to 90 minutes after injection, even when dispersion/diffusion is neglected (APT). Dispersion/diffusion leads to a smoother evolution of the concentration versus time as expected. The BTCs computed by the codes are significantly different from each other as well. Results based on ELLAM show some oscillations at the end of the simulation period due to the numerous interpolations required by the method. Concentrations computed by MODFLOW are significantly different from the concentrations computed by the other computational approaches between 40 to 60 minutes after injection.

Similar results are obtained for packing 3, although the simulated solutions showed a better match to the experimental results (Fig. 10). The difference between advective transport (APT) and advective-dispersive transport is quite small, which indicates the predominant role of the advective transport for this test case. The transport computed by FEFLOW shows some excessive dispersion/diffusion and MODLOW computation shows some excessive dispersion between 60 and 100 minutes after injection. Both experiments were simulated using a finer mesh (Δx = 0.5 mm) without significant changes in the computed concentrations.

The results of the simulations for all three packing arrangements are summarized in Table 3, in terms of metrics that provide a quantitative comparison:

$$ME = \frac{1}{N}\sum_{i=1}^{N}\left(C_i^s - C_i^m\right); \quad AME = \frac{1}{N}\sum_{i=1}^{N}\left|C_i^s - C_i^m\right|; \qquad RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(C_i^s - C_i^m\right)^2} \tag{4}$$

where N is the number of measured concentrations, C_i^s and C_i^m respectively being the simulated and measured normalized concentrations. In (4), ME represents the average error and is a good indicator when discrepancies between computed and observed concentrations are due to a translation along the time axis; AME is the absolute mean error and RMSE is the root mean square error. Note that RMSE and AME provide the same type of information, RMSE being more sensitive to high differences between computed and measured normalized concentrations.

On average, we find that the APT simulations provide the highest errors, showing that dispersion, although quite small, improves the quality of the simulations. The values of ME are relatively small, showing that the average flow is well reproduced by the simulation; this finding might be expected given that the flow rate (at the inflow in our case) is prescribed. This result also shows that the various schemes tested (especially the upwind schemes) do not generate any artificial advective transport. The differences between AME and RMSE remain limited for all test cases because the numerical schemes analyzed here do not render outliers in concentration values.

All three metrics are quite small for packing 1 (less than 0.05 when dispersion is included) with some slightly higher values for FEFLOW. These metrics are the highest for packing 2 (close to 0.09 for MODFLOW), where all codes fail to reproduce the shape of the BTC (see Fig. 9). More significant differences among the codes are evidenced for packing 3, especially between FEFLOW and TRACES (see experiment 2).

Note that we consider here the measured concentrations as a reference even though uncertainties exist in the parameter values and in the geometry of the block-homogeneous zones. These uncertainties do not allow for more detailed comparisons among the codes at this stage. Indeed, a numerical code with numerical dispersion may provide better results than a more accurate code where an underestimated dispersivity value is used.

As stated above, local BTCs (associated with resident concentration) were also computed to evaluate the relative differences among the numerical methods. To this end, the grid and time step sizes chosen after convergence at the outlet BTCs were used, i.e., Δx = 1 mm and Δt_{max} = 0.25 min. There are no measurements from within the laboratory flow cells, and the main purpose of these comparisons was to investigate how numerical methods describe local

BTCs at some strategic points in the domain. We do so by selecting points located at the boundaries of two or more zones in the heterogeneous structures (packing 2 and 3).

While the differences between the codes remain quite small for the local BTCs close to the inlet, the differences are very significant far from the outlet (Fig. 11 for packing 2 and Fig. 12 for packing 3). We recall here that the differences in the simulations rely on the numerical method and the grid. ELLAM, FEFLOW and MODFLOW suffer from numerical dispersion/diffusion (see BTCs at 'C' for both packings for FEFLOW and MODFLOW, BTC at 'C' for packing 2 for ELLAM). Because the flow fields of different codes are very similar, the difference in the computed concentrations may rely on (i) the discretization of the advective part of the transport equation, (ii) the handling of a full dispersion tensor (which might be a critical issue for a finite volume based approach such as embodied in MODFLOW) and/or (iii) the coupling between advection and dispersion when the numerical method used to solve advection and dispersion are different (which is the case for ELLAM and TRACES).

These comparisons show that the ability of a numerical model to reproduce a flux-averaged BTC (such as that sampled at a system outflow, or along a borehole under field conditions) does not necessarily imply that local concentrations are computed accurately.

3.3 Numerical method performance

The sensitivity to discretization is frequently studied in assessing the accuracy of numerical methods. The spatial discretization sensitivity of numerical solutions presented here was obtained after temporal convergence, i.e., when reducing the time step does not change the solution. We compare concentration distributions at 50 min after injection, for packing 3, obtained with two different grid sizes, for the four models (Fig. 13). The grid size Δx = 1 mm is taken as the reference solution, which allowed reasonable simulation of the BTCs at the flow cell outlet (results with the Δx = 0.5 mm grid did not differ but required increased computational time).

The change in grid size affects the discretization of the flow field and the potential numerical dispersion of the numerical method. We recall that, due to the mesh geometry (squares), MODFLOW and TRACES are based on the same velocity field. The effect of fine sand inclusions located close to the outlet (bottom right in Fig. 1, packing 3) is not properly taken into account for the coarser grid.

We note that all of the numerical methods tested are affected by numerical dispersion with the exception of the APT scheme. In this context, we found that FEFLOW and MODFLOW are more sensitive to spatial discretization than TRACES and ELLAM. Decreasing mesh size reduces greatly the numerical dispersion generated by all of these methods and improves significantly the accuracy of the numerical solutions. Standard methods embedded in MODFLOW and FEFLOW have to be applied under restricted conditions to limit problems with numerical stability and/or numerical dispersion. These concentration distributions provide a good illustration of the numerical dispersion generated by the standard methods, as well as their dependence to grid size.

Because the numerical methods are different, the computational costs are different for the same grid. Accuracy of computational codes should be linked to computational costs such as CPU time and required memory. CPU time is of course correlated to the number of unknowns, but also to local criteria such as the Courant number of the advective part of the computation. For example, TRACES uses an explicit scheme for the discontinuous finite element scheme. The number of unknowns is equal to the number of elements for MODFLOW, the number of nodes for FEFLOW and ELLAM, and the number of element edges for TRACES.

Table 2 summarizes the required computational needs for the same grid used for the simulation of the packing 2. However, computer time is not a proper criterion to compare computational codes, because it depends on many parameters including, e.g., the type of compiler, the type of solver, and programming skills, and is provided here only as additional information.

4. Reliability of the ADE in the context of the numerical methods tested

The simulations of the second and third sets of experiments (packing 2 and 3) do not yield complete matches to the experimental results (Figs. 9 and 10), irrespective of the numerical method employed to solve the ADE. Several reasons can be suggested:

(i) Uncertainty in data, including the idealized geometry of the heterogeneities in the simulations compared to the experimental setup, and/or uncertainty in the hydraulic conductivities and porosities due to sand packing as reported by, e.g., Ruch (1992).

(ii) Lack of knowledge regarding transport processes – and how to model them – at the interface between two porous media, as reported by, e.g., Berkowitz et al. (2009, and references therein) for a layered column. Other experimental results suggest that, at least for small variations in medium heterogeneity, the global transport in such a

configuration can be considered as the convolution of homogeneous blocks (Delay et al., 1997), without significant effects of mixing processes at the block interfaces.

(iii) The ADE may not be appropriate to fully characterize the transport, as demonstrated from repeated and diverse experiments demonstrating non-Fickian transport (e.g., Berkowitz et al., 2000, 2006).

We address here the first factor by incorporating parameter uncertainties in the numerical simulations. We assume therefore that the difference between computed and observed concentrations is due to the uncertainty in hydraulic conductivity and porosity due to packing as observed by Ruch (2002) who reported hydraulic conductivity differences of about 15% by repeating column experiments.

Uncertainties in parameters were simulated by (forward) Monte Carlo simulations assuming uniform distribution of the hydraulic conductivity over a given range for each sand. These ranges were fixed to $4 \times 10^{-5} – 5 \times 10^{-4}$, $9.0 \times 10^{-4} – 15 \times 10^{-4}$ and $5 \times 10^{-3} – 15 \times 10^{-3}$ m/s for the fine, medium and coarse sands, respectively. These values were estimated from other laboratory experiments performed with similar sand (same average grains size) by Ruch (1992), Schroth et al. (1996), Silliman and Caswell (1998), Chao et al. (2000), Levy and Berkowitz (2003) and Jose et al. (2004). Effective porosity is assumed to be distributed over [0.32 to 0.38] which represents 10% of the estimated value. Although dispersivity values were assumed on the basis of average grain size, their uncertainties were not taken into account. Numerical simulations with dispersivity values 10 times greater did not show significant differences in the BTCs. We also performed a set of MC simulations where (*a*) hydraulic conductivity was modeled as a spatially varying random process within each of the sands, and (*b*) conductivity values within each grid cell were randomly selected from the above mentioned uniform distributions. We found that the effects on the target BTCs due to this type of internal variability were not significant when compared to the effects of the conductivity contrasts between the sand types (details not reported). Given the objective of our study, we do not pursue this avenue of investigation here.

Because the objectives are not a detailed statistical analysis of the concentration distribution, we performed only 100 MC runs using code TRACES with the optimal spatial discretization of $\Delta x = 1$ mm (see section 3). These types of uncertainty quantification analyses could also be performed in the context of typical model calibration, where the uncertainty associated with estimated model parameters is conditional on available data as a result of an inverse modeling procedure. As stated in the Introduction, we emphasize that aspects of

model calibration in randomly heterogeneous porous media are not the subject of this investigation. Measured and 30 Monte Carlo computed BTCs chosen randomly amongst the 100 simulations are shown in Figs. 14 and 15. The observed data lie within the collection of computed BTCs. The uncertainty related to the parameters is a possible explanation of the differences between measured and computed concentrations (Fig. 9 and 10) with the initial dataset (Table 1). The impact of parameter uncertainties is significant for packing 2, mainly 60 to 80 minutes after injection where the shape of the BTC may be either convex or concave.

Given the analysis of these results, together with the recognition that hydraulic conductivity could also be modeled as a spatial random function within each type of sand, one can conclude that the significant impact of the hydraulic conductivity uncertainty on the BTCs does not allow unambiguous rejection of the classical ADE formulation to quantify solute transport in these experiments. Of course, this conclusion does not negate the other two factors noted above, as well.

5. Conclusions

Based on the results of our analysis, we conclude the following:

1. Simulation of tracer transport in heterogeneous media is far from being straightforward. Mesh discretization can impact considerably the results of computations of both flow and concentration fields. While numerous numerical works are focused on detailed analysis of artificial dispersion and/or oscillation due to the numerical solution of the transport equation, we show here also that the accuracy of the solution of the flow equation is a critical issue which needs to be properly addressed through a consistent and rigorous convergence analysis.
2. Enhanced numerical methods used in computational systems such as TRACES and ELLAM are more accurate and perform better than standard methods (finite volume, Galerkin finite elements) for solving the transport equation, at least for the experiments considered here. They provide fewer problems related to numerical dispersion and the way the full dispersion/diffusion tensor can be handled effectively in the numerical solution method.
3. Model verification based only on flux-averaged concentrations (of the kind detected at a flow cell outlet and/or pumping well) does not guarantee the numerical accuracy of the methodology underlying the modeling results. Although simulated BTCs at the outflow may be very close, computed local concentration values (inside the domain) can be very different, depending on the numerical

method employed. As a consequence, the choice of a numerical method can include an element of uncertainty to model results, which should be considered in combination with uncertainties associated with knowledge of model parameters.

4. Although the experiments were carried out with great care and accuracy, natural experimental uncertainties in the hydraulic conductivity and effective porosity values (eventually including detailed knowledge of the location of boundaries between different materials) can be significant and contributes to our inability to fully demonstrate (or reject) the capability of the ADE to interpret the available data, regardless of the numerical method employed to solve the flow and transport scenario.

Laboratory experiments are useful for model testing and numerical methods benchmarking. However, they required very detailed information on both concentrations and parameters. To fully answer the question of mathematical model validation, the presented setup still requires some modifications to reach the 'ideal' configuration. Due to the boundary condition at the inflow (prescribed flux), the average velocity will not change whatever the hydraulic conductivity in the domain. Therefore, it is better to use prescribed hydraulic heads and to measure the water flux at the outflow. It is also highly recommended to follow local concentrations over time inside the porous medium, as discussed in, e.g., Loyaux-Lawniczak et al. (2012) and Raveh-Rubin et al. (2015).

Acknowledgements

P.A. and B.B. gratefully acknowledge the financial support of the Ministry of Science and Technology, Israel and the Ministry of Research, France. B.B. was also supported by a research grant from Stanley Magidson and Darlene Switzer-Foster. B.B. holds the Sam Zuckerberg Professorial Chair in Hydrology. B.B. thanks Guy Katz for assistance with performing the laboratory experiments. A.G. and B.B. acknowledge funding from MIUR (Italian Ministry of Education, University and Research) and IMOE (Israel Ministry of National Infrastructures, Energy and Water Resources) in the context of the Water JPI, WaterWorks 2014, Project: WE-NEED - Water NEEDs, availability, quality and sustainability). The authors thank Wolfgang Nowak and two anonymous referees for constructive review comments.

References

Ballarini, E., Bauer, S., Eberhardt, C., Beyer, C., 2012. Evaluation of transverse dispersion effects in tank experiments by numerical modeling: Parameter estimation, sensitivity analysis and revision of experimental design. J. Contam. Hydrol. 134-135, 22-36, doi:10.1016/j.jconhyd.2012.04.001.

Bear, J., 1972. Dynamics of Fluids in Porous Media. American Elsevier Publishing, New York.

Berkowitz, B., Klafter, J., Metzler, R., Scher, H., 2002. Physical pictures of transport in heterogeneous media: advection-dispersion, random walk and fractional derivative formulations. Water Resour. Res. 38, 1191. doi:10.1029/2001WR001030.

Berkowitz, B., Scher, H., Silliman, S.E., 2000. Anomalous transport in laboratory-scale, heterogeneous porous media. Water Resour. Res. 36, 149-158, doi:10.1029/1999WR900295.

Berkowitz, B., Cortis, A., Dentz, M., Scher, H., 2006. Modeling non-Fickian transport in geological formations as a continuous time random walk. Rev. Geophys. 44, RG2003. doi:10.1029/2005RG000178.

Berkowitz B., Cortis A., Dror I., Scher H. 2009. Laboratory experiments on dispersive transport across interfaces: The role of flow direction. Water Resour. Res., 45, W02201, doi:10.1029/2008WR007342, 2009

Cainelli, O., Bellin, A., & Putti, M. (2012). On the accuracy of classic numerical schemes for modeling flow in saturated heterogeneous formations. Adv. Water Resour., 47, 43-55.

Celia, M.A., Russell, T.F., Herrera, I., Ewing, R.E., 1990. An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation. Adv. Water Resour. 13, 187-206. doi:10.1016/0309-1708(90)90041-2.

Chao, H.C., Rajaram, H., 2000. Intermediate-scale experiments and numerical simulation of transport under radial flow in a two-dimensional heterogeneous porous medium. Water Resour. Res. 36 (10), 2869-2884.

Chiogna, G., Eberhardt, C., Grathwohl, P., Cirpka, O.A., Rolle, M., 2010. Evidence of compound-dependent hydrodynamic and mechanical transverse dispersion by multitracer laboratory experiments. Environ. Sci. Technol. 44, 688-693.

Danquigny, C., Ackerer, P., Carlier, J.P., 2004. Laboratory tracer tests on three-dimensional reconstructed heterogeneous porous media. J. Hydrol. 294, 196-212.

Delay, F., Ackerer, P., Danquigny, C., 2005. Simulating solute transport in porous or fractured formations using random walk particle tracking: a review. Vadose Zone J. 4, 360-379.

Delay F., Porel G., de Marsily G., 1997. Predicting solute transport in heterogeneous media from results obtained in homogeneous ones: an experimental approach. J. Contam. Hydrol., 25, 63-84.

Diaw, E.B., Lehmann, F., Ackerer, P., 2001. One-dimensional simulation of solute transfer in saturated–unsaturated porous media using the discontinuous finite elements method. J. Contam. Hydrol. 51, 197-213. doi:10.1016/S0169-7722(01)00129-2.

Diersch, H.J., 2013. FEFLOW: Finite Element Modeling of Flow, Mass and Heat Transport in Porous and Fractured Media. Springer Science & Business Media.

Esfandiar, B., Porta, G., Perotto, S., Guadagnini, A., 2015. Impact of space-time mesh adaptation on solute transport modeling in porous media. Water Resour. Res., 51, 1315-1332. doi:10.1002/2014WR016569.

Farhloul, M., Serghini Mounim, A., 2005. A mixed-hybrid finite element method for convection–diffusion problems. Appl. Math. Comput. 171, 1037-47. doi:10.1016/j.amc.2005.01.101.

Harbaugh, A.W., Banta, E.R., Hill, M.C., McDonald, M.G., 2000. MODFLOW-2000, the US Geological Survey Modular Ground-Water Model: User Guide to Modularization Concepts and the Ground-Water Flow Process. US Geological Survey Reston, VA. http://funnel.sfsu.edu/students/dotsona/geosci/courses/G700/Documents-Manuals-PDFs/DOC3_MODFLOW2000_ModConcepts_GWFlowProcess_ofr00-92.pdf.

Healy, R.W., Russell, T.F., 1998. Solution of the advection-dispersion equation in two dimensions by a finite-volume Eulerian-Lagrangian localized adjoint method. Adv. Water Resour. 21, 11-26. doi:10.1016/S0309-1708(96)00033-4.

Jose S., Rahman M., Cirpka O., 2004. Large-scale sandbox experiment on longitudinal effective dispersion in heterogeneous porous media. Water Resour. Res., 40, W12415, doi:10.1029/2004WR003363, 2004

Kinzelbach W., 1986. Groundwater Modelling An Introduction with Sample Programs in BASIC. Developments in Water Science, Elsevier, 25, 333p.ISBN: 978-0-444-42582-9

Klotz, D., Seiler, K.P., Moser, J., Neumaier, F., 1980. Dispersivity and velocity relationship from laboratory and field experiments. J. Hydrol. 45, 169-184.

Konz, M., Ackerer, P., Younes, A., Huggenberger, P., Zechner, E., 2009. 2D stable layered laboratory-scale experiments for testing density-coupled flow models. Water Resour. Res. 45, doi:10.1029/2008WR007118.

Leonard, B.P., 1988. Universal Limiter for Transient Interpolation Modeling of the Advective Transport Equations: The ULTIMATE Conservative Difference Scheme. NASA Technical Memorandum 100916: 115.

Levy, M., Berkowitz, B., 2003. Measurement and analysis of non-Fickian dispersion in heterogeneous porous media. J. Contam. Hydrol. 64, 203-226. doi:10.1016/S0169-7722(02)00204-8.

Loyaux-Lawniczak S., Lehmann F., Ackerer P. (2012). Acid/base front propagation in saturated porous media: 2D laboratory experiments and modeling. J. Contam. Hydrol., 138-139, 15-21.

Oswald, S., Kinzelbach W., 2004. Three-dimensional physical benchmark experiments to test variable-density flow models, J. Hydrol. 290, 22-42.

Raveh-Rubin, S., Y. Edery, I. Dror and B. Berkowitz (2015). Nickel migration and retention dynamics in natural soil columns, Water Resour. Res., 51, 7702-7722, doi:10.1002/2015WR016913.

Ruch, M., 1992. Transfert d'un pollutant non réactif en milieu poreux saturé: etude sur modèles physiques tridimensionnels et simulations. PhD Thesis, Strasbourg University (F), 162 pp.

Russell, T.F., Celia, M.A., 2002. An overview of research on Eulerian-Lagrangian localized adjoint methods (ELLAM). Adv. Water Resour. 25, 1215-1231. doi:10.1016/S0309-1708(02)00104-5.

Schroth, M.H., Ahearn, S.J., Selker, J.S., Istok, J.D., 1996. Characterization of Miller-similar silica sands for laboratory hydrologic studies. Soil Sci. Soc. Am. J. 60, 1331-1339.

Silliman, S.E., Caswell, S., 1998. Observations of measured hydraulic conductivity in two artificial, confined aquifers with boundaries. Water Resour. Res. 34 (9), 2203-2213.

Silliman, S.E., Simpson, E.S., 1987. Laboratory evidence of the scale effect in dispersion of solutes in porous media. Water Resour. Res. 23, 1667-1673.

Silliman, S.E., Zheng, L., Conwell, P., 1998. The use of laboratory experiments for the study of conservative solute transport in heterogeneous porous media. Hydrogeol. J. 6, 166-177.

Thomson, N. R., Sykes, J. F., Lennox, W. C., 1984, A Lagrangian porous media mass transport model, Water Resour. Res., 20(3), 391-399.

2009. TRACES (Transport RéActif de Contaminants dans les Eaux Souterraines - Transport of RadioACtive Elements in Subsurface). User's Guide. LHyGeS, Strasbourg University, 70p.

Younes, A., Ackerer, P., Lehmann, F., 2006. A new efficient Eulerian–Lagrangian localized adjoint method for solving the advection–dispersion equation on unstructured meshes. Adv. Water Resour. 29, 1056-1074. doi:10.1016/j.advwatres.2005.09.003.

Younes, A., Ackerer, P., Delay, F., 2010. Mixed finite elements for solving 2-D diffusion-type equations. Rev. Geophys. 48, RG1004. doi:10.1029/2008RG000277.

Zhang, X., Shu, C., 2010. A genuinely high order total variation diminishing scheme for one-dimensional scalar conservation laws. SIAM J. Numer. Anal. 48, 772-795. doi:10.1137/090764384.

Zheng, C., Wang, P.P., 1999. MT3DMS: A Modular Three-dimensional Multispecies Transport Model for Simulation of Advection, Dispersion, and Chemical Reactions of Contaminants in Groundwater Systems. Documentation and user's guide. U.S. Army Engineer Research and Development Center No. SERDP-99-1, Vicksburg, MS.

Table 1. Porous medium properties and assumed dispersivities based on average grain size

Sand type:	"Coarse"	"Medium"	"Fine"
Average grain size (m)	1.1×10^{-3}	0.53×10^{-3}	0.23×10^{-3}
Hydraulic conductivity (m/s)	5.0×10^{-3}	1.5×10^{-3}	0.20×10^{-3}
Porosity (-)	0.35	0.35	0.35
Longitudinal dispersivity (m)	1.1×10^{-3}	0.53×10^{-3}	0.23×10^{-3}
Transverse dispersivity (m)	1.1×10^{-4}	0.53×10^{-4}	0.23×10^{-4}

Table 2. Computational time for the numerical codes for grid size of Δx = 1 mm and Δt = 1 minute.

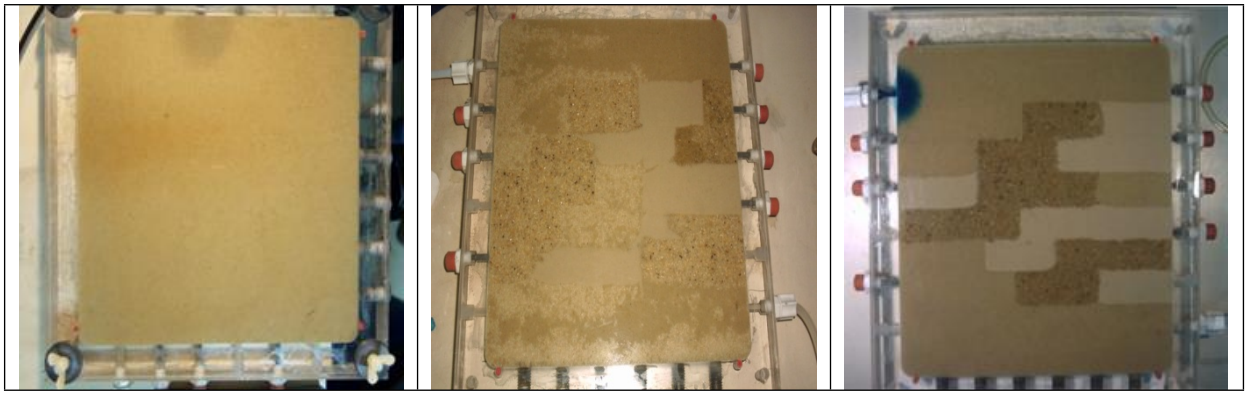
	MODFLOW	FEFLOW	TRACES	ELLAM
No. cells	249×249	249×249×2	249 × 249	249×249×2
No. unknowns	62001	62500	124500	62500
CPU time (s)	5028	431	3806	423

Table 3. Some metrics of the simulations to assess discrepancies between models and data.

Packing 1	Experiment 1			Experiment 2		
	ME	**AME**	**RMSE**	**ME**	**AME**	**RMSE**
Modflow	0.0015	0.0133	0.0172	-0.0038	0.0071	0.0089
Feflow	-0.0177	0.0264	0.0370	-0.0190	0.0316	0.0384
APT	-0.0031	0.0375	0.0508	-0.0092	0.0239	0.0324
ELLAM	-0.0102	0.0184	0.0281	-0.0139	0.0187	0.0242
Traces	0.0035	0.0122	0.0178	-0.0001	0.0132	0.0189

Packing 2	Experiment 1			Experiment 2		
	ME	**AME**	**RMSE**	**ME**	**AME**	**RMSE**
Modflow	-0.0321	0.0697	0.0800	-0.0272	0.0739	0.0870
Feflow	0.0103	0.0456	0.0596	0.0135	0.0564	0.0730
APT	-0.0021	0.0571	0.0708	0.0007	0.0636	0.0824
ELLAM	-0.0179	0.0499	0.0614	-0.0146	0.0573	0.0746
Traces	-0.0127	0.0455	0.0523	-0.0084	0.0510	0.0630

Packing 3	Experiment 1			Experiment 2		
	ME	**AME**	**RMSE**	**ME**	**AME**	**RMSE**
Modflow	-0.0251	0.0471	0.0619	-0.0012	0.0469	0.0597
Feflow	0.0405	0.0414	0.0587	0.0688	0.0689	0.0849
APT	0.0201	0.0345	0.0414	0.0443	0.0500	0.0634
ELLAM	0.0164	0.0259	0.0324	0.0456	0.0478	0.0586
Traces	-0.0215	0.0276	0.0330	0.0053	0.0231	0.0281



Packing 1 Packing 2 Packing 3

Fig. 1. Sand packing arrangements used for the laboratory experiments. Packing 1: homogeneous packing with medium sand. Packing 2: heterogeneous packing arrangement with coarse, medium and fine sand. Packing 3: heterogeneous packing arrangement with coarse, medium and fine sand.

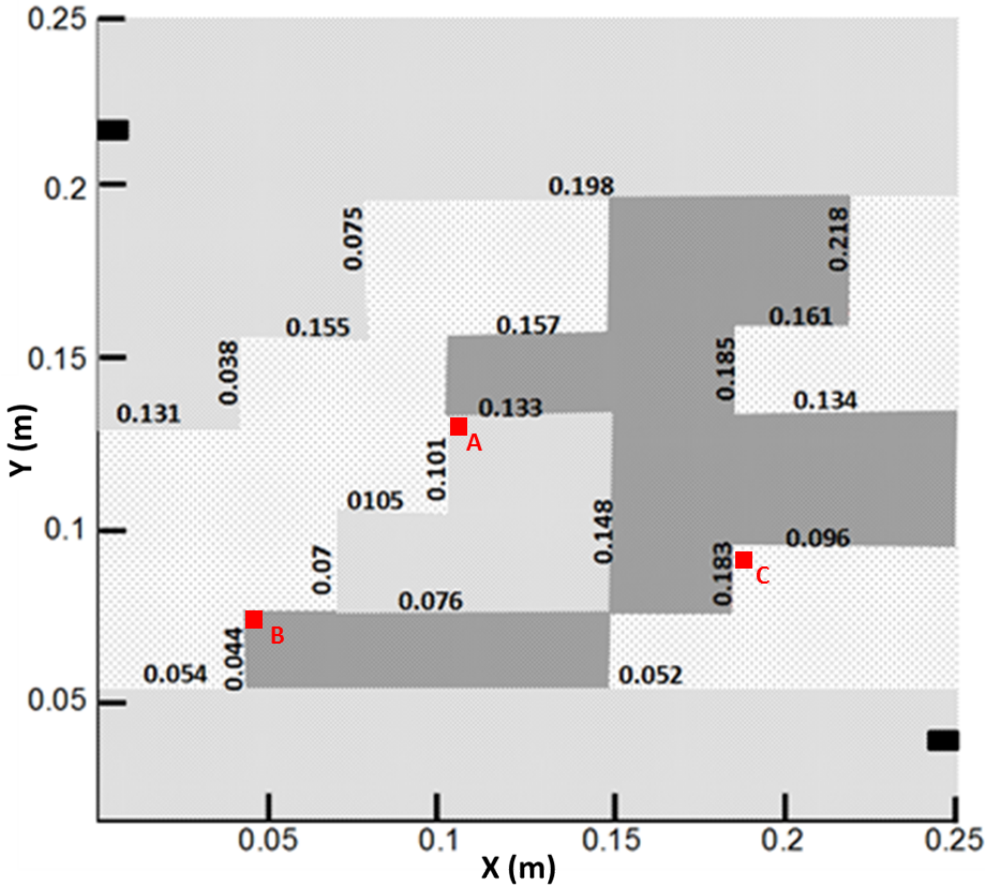


Fig. 2. Disposition of coarse (light grey), medium (grey) and fine (dark grey) sands in the Packing 2. Decimal values within the domains denote the coordinates in meters of each boundary (*x* and *y* dimensions are written vertically and horizontally, respectively). Inlet and outlet in each domain are represented by the black rectangles at the upper left and lower right sides, respectively. Red squares represent the locations where internal BTCs are computed.

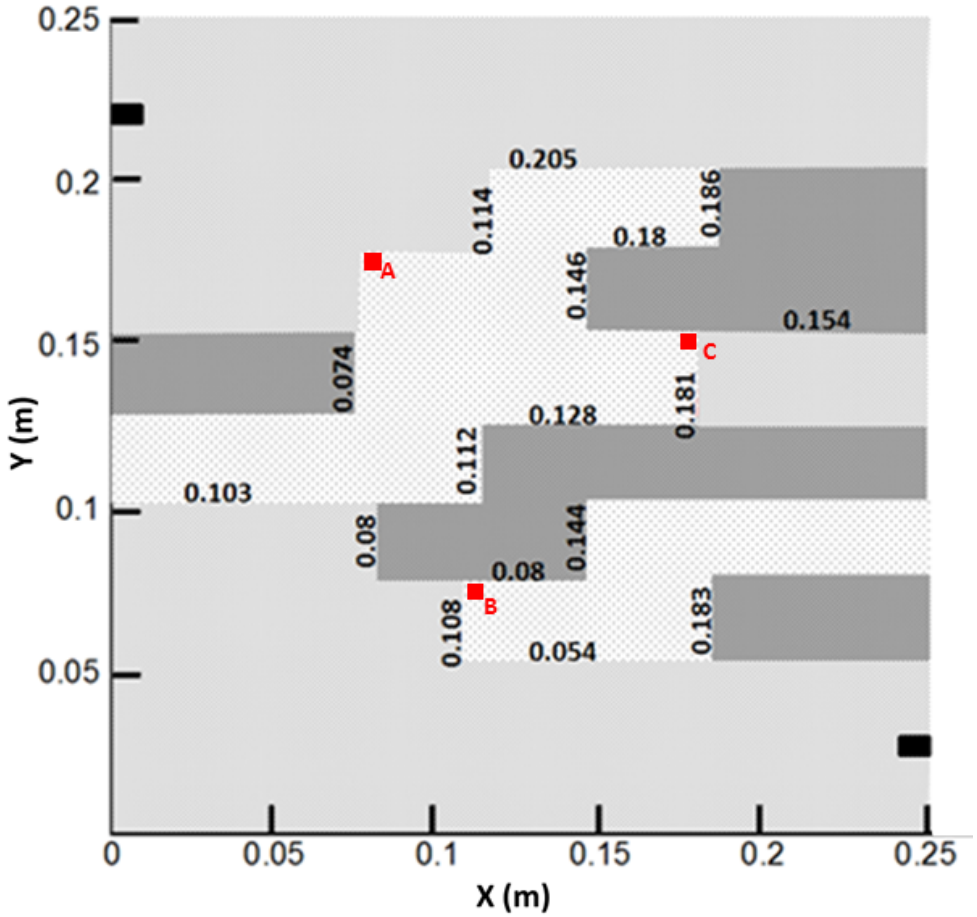


Fig. 3. Disposition of coarse (light grey), medium (grey) and fine (dark grey) sands in the Packing 3. Decimal values within the domains denote the coordinates in meters of each boundary (*x* and *y* dimensions are written vertically and horizontally, respectively). Inlet and outlet in each domain are represented by the black rectangles at the upper left and lower right sides, respectively. Red squares represent the locations where internal BTCs are computed.

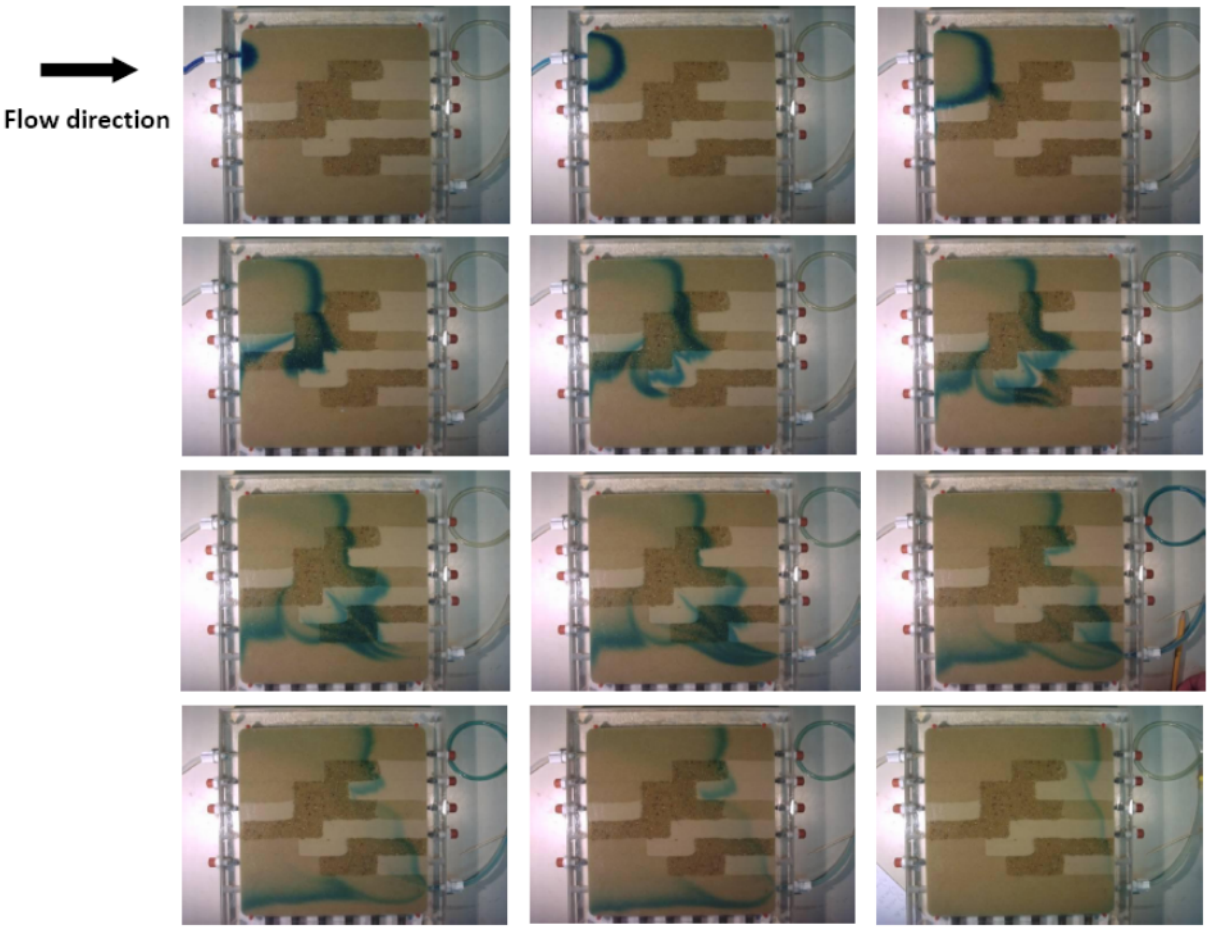


Fig. 4. Spatial evolution over time of tracer (blue dye) injected into Packing 3. Photographs are taken at approximately uniform time intervals over approximately one hour.

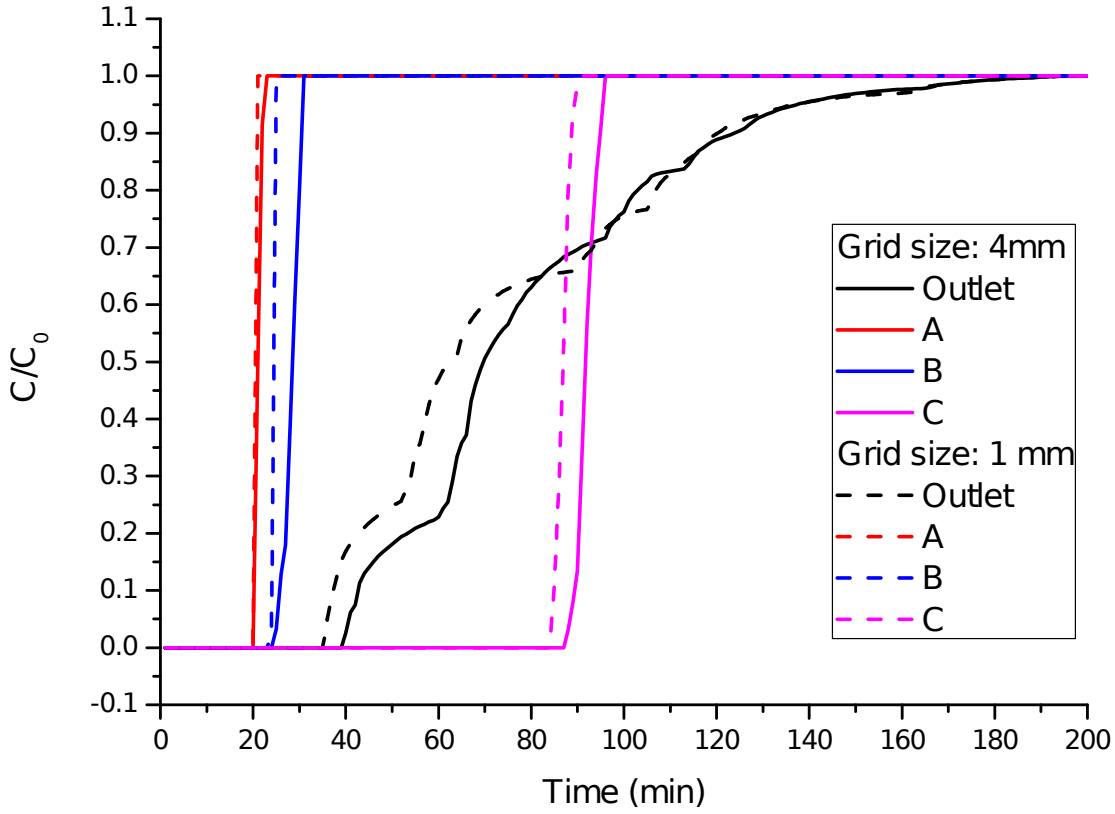


Fig. 5. Simulated breakthrough curves for Packing 2 considering advective transport only at points located in A (10.2, 13.2), B (4.5, 7.5) and C (18.5, 9.5); see Fig. 2.

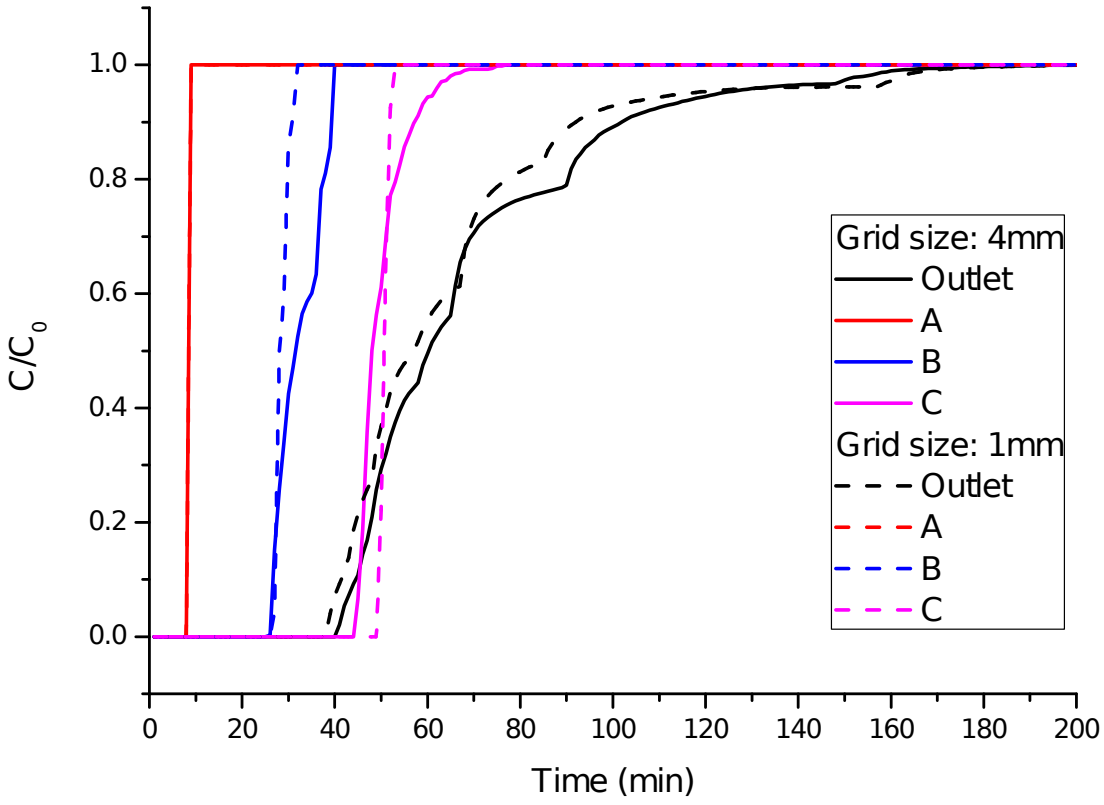
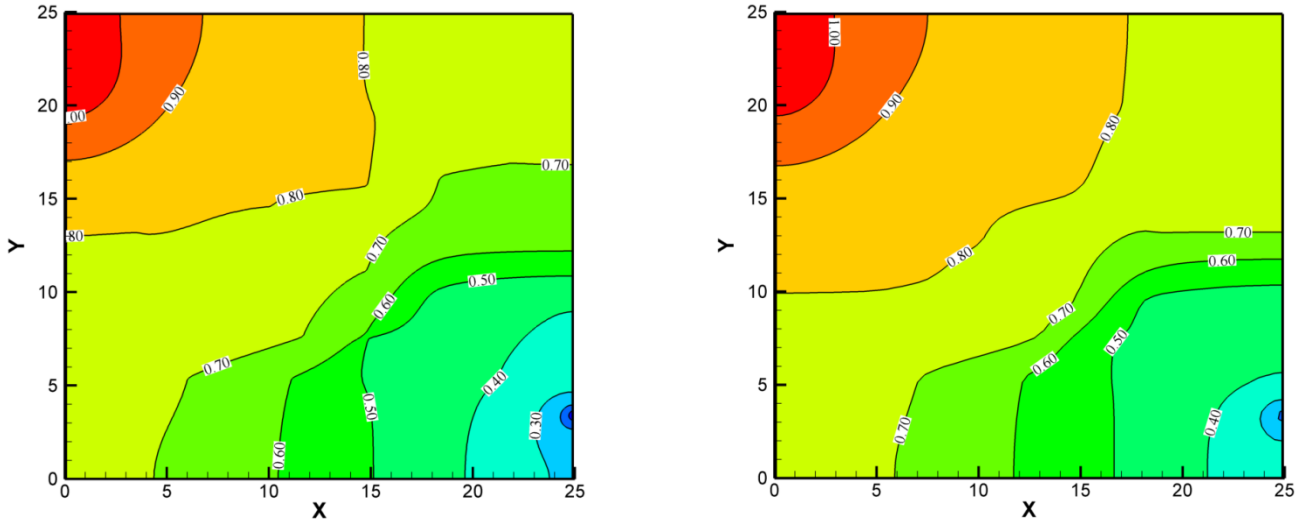


Fig. 6. Simulated breakthrough curves for Packing 3 considering advective transport only at points located in A (8.1, 17.3), B (11.3, 7.3) and C (17.6, 14.9); see Fig. 3.



(a) (b)

Fig. 7. Pressure head distribution for Packing 2 and spatial discretization of (a) Δx=1 mm and (b) Δx=4 mm.

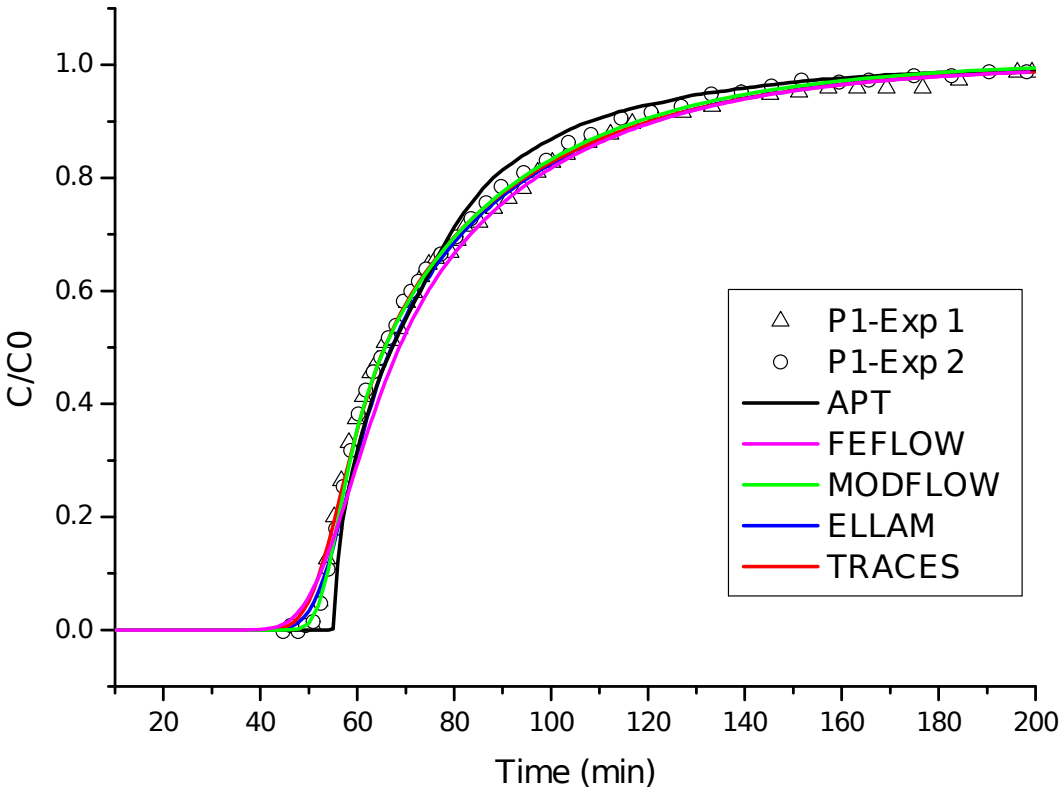


Fig. 8. Measured and simulated breakthrough curves at outlet for Packing 1 (P1).

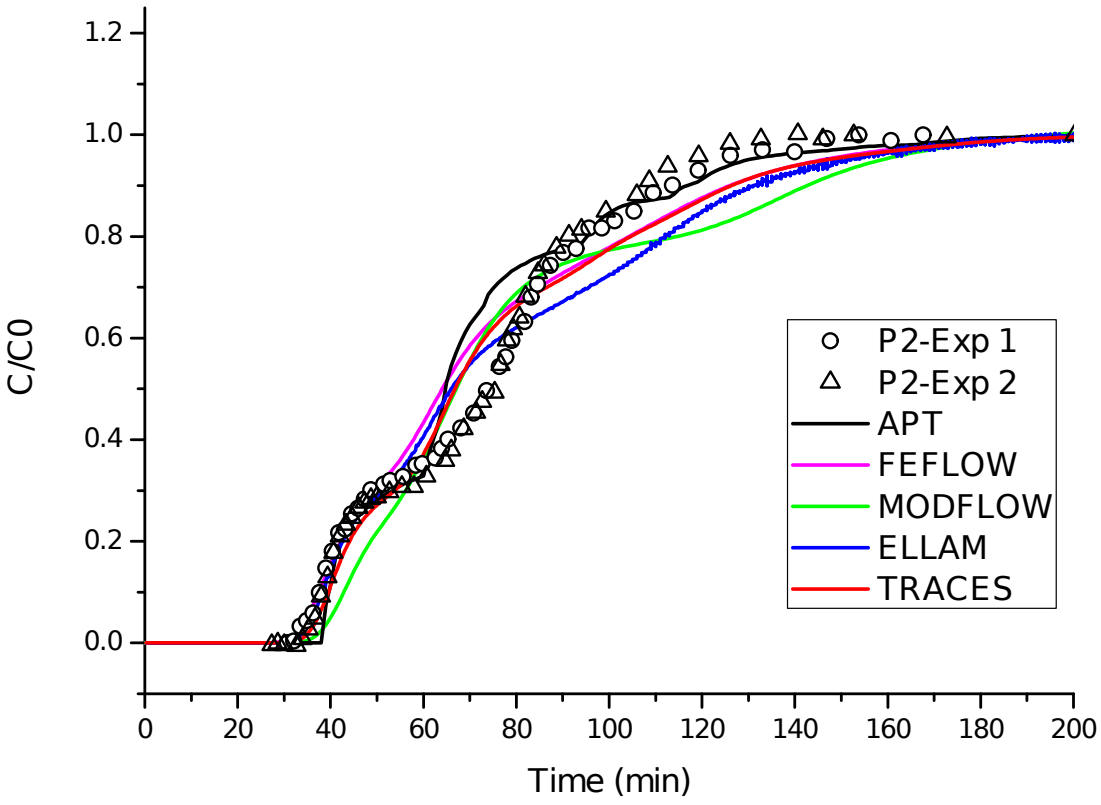


Fig. 9. Measured and simulated breakthrough curves at outlet for Packing 2 (P2).

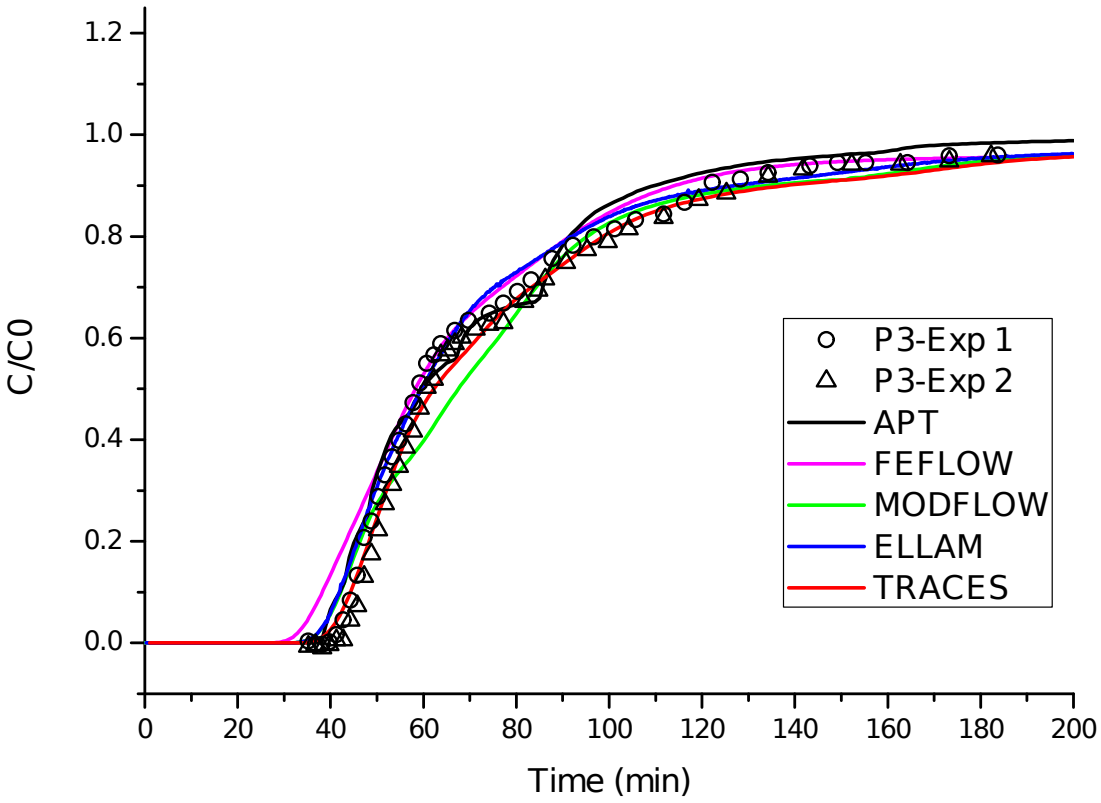


Fig. 10. Measured and simulated breakthrough curves at outlet for Packing 3 (P3).

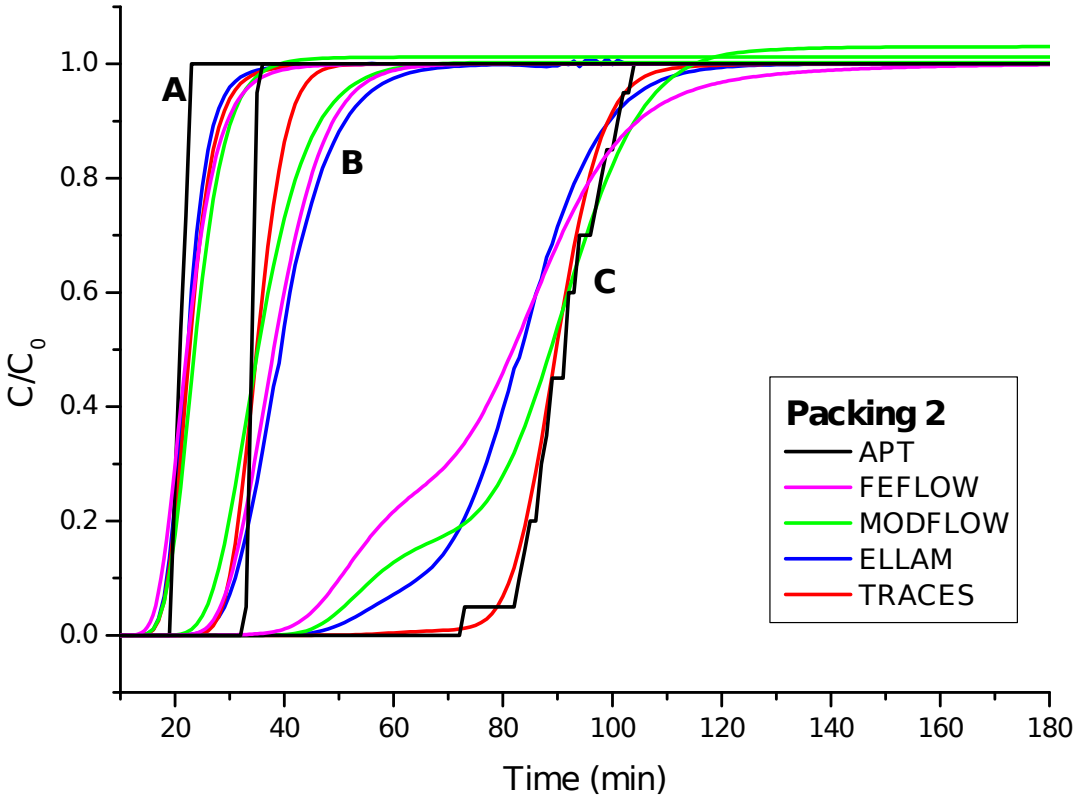


Fig. 11. Simulated breakthrough curves for Packing 2 at points located in A (10.2, 13.2), B (4.5, 7.5) and C (18.5, 9.5); see Fig. 2.

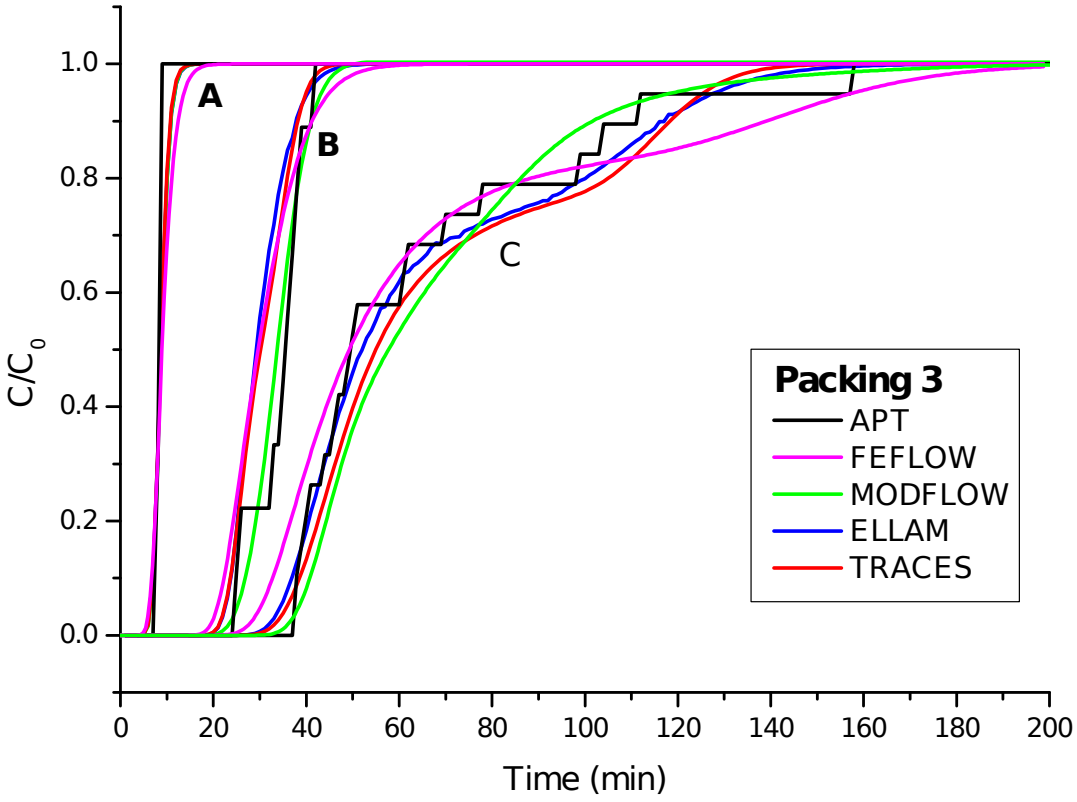


Fig. 12. Simulated breakthrough curves for Packing 3 at points located in A (8.1, 17.3), B (11.3, 7.3) and C (17.6, 14.9); see Fig. 3.

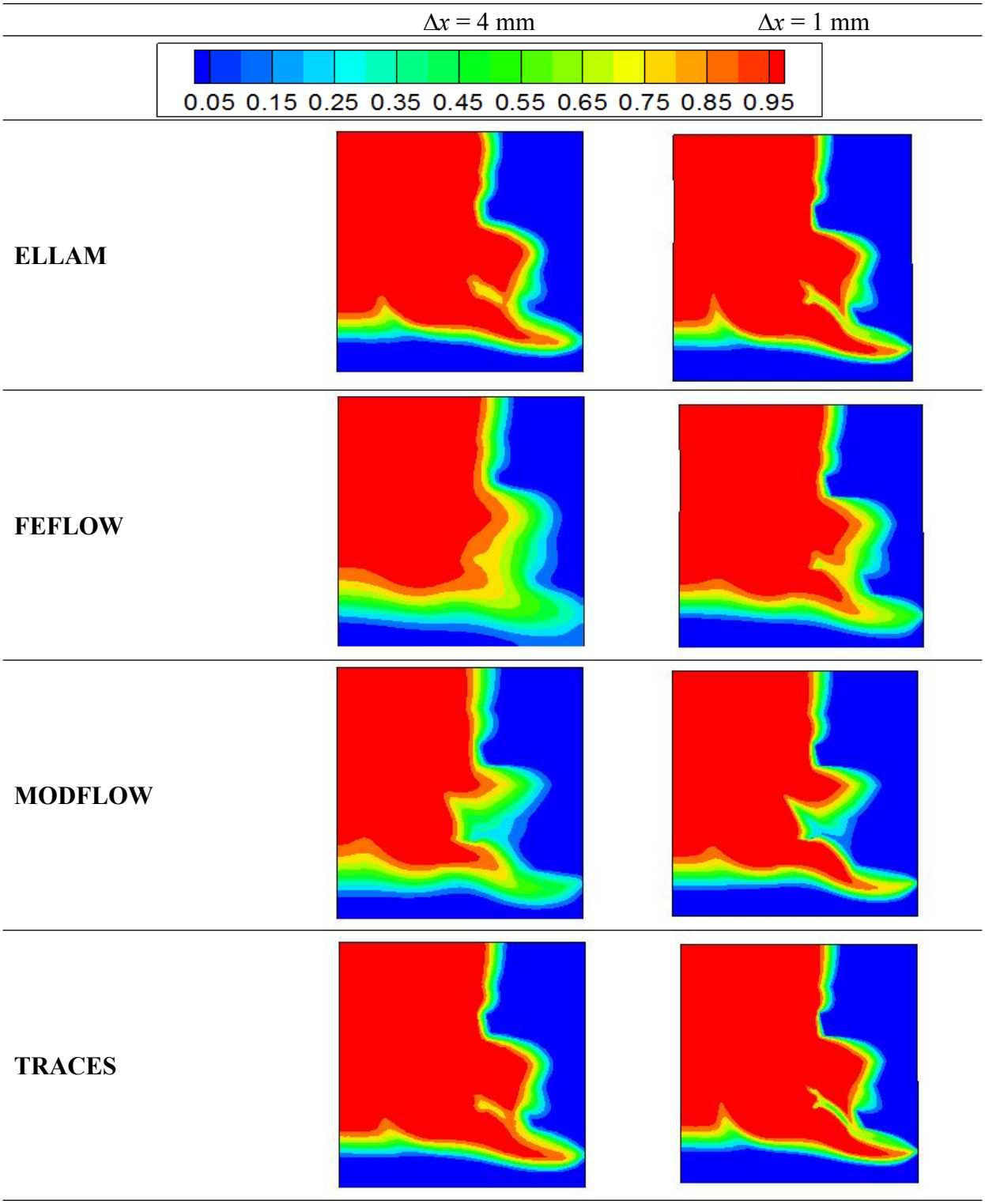


Fig. 13. Simulated concentration distributions 50 minutes after injection (Packing 3). Color bar indicates relative concentrations.

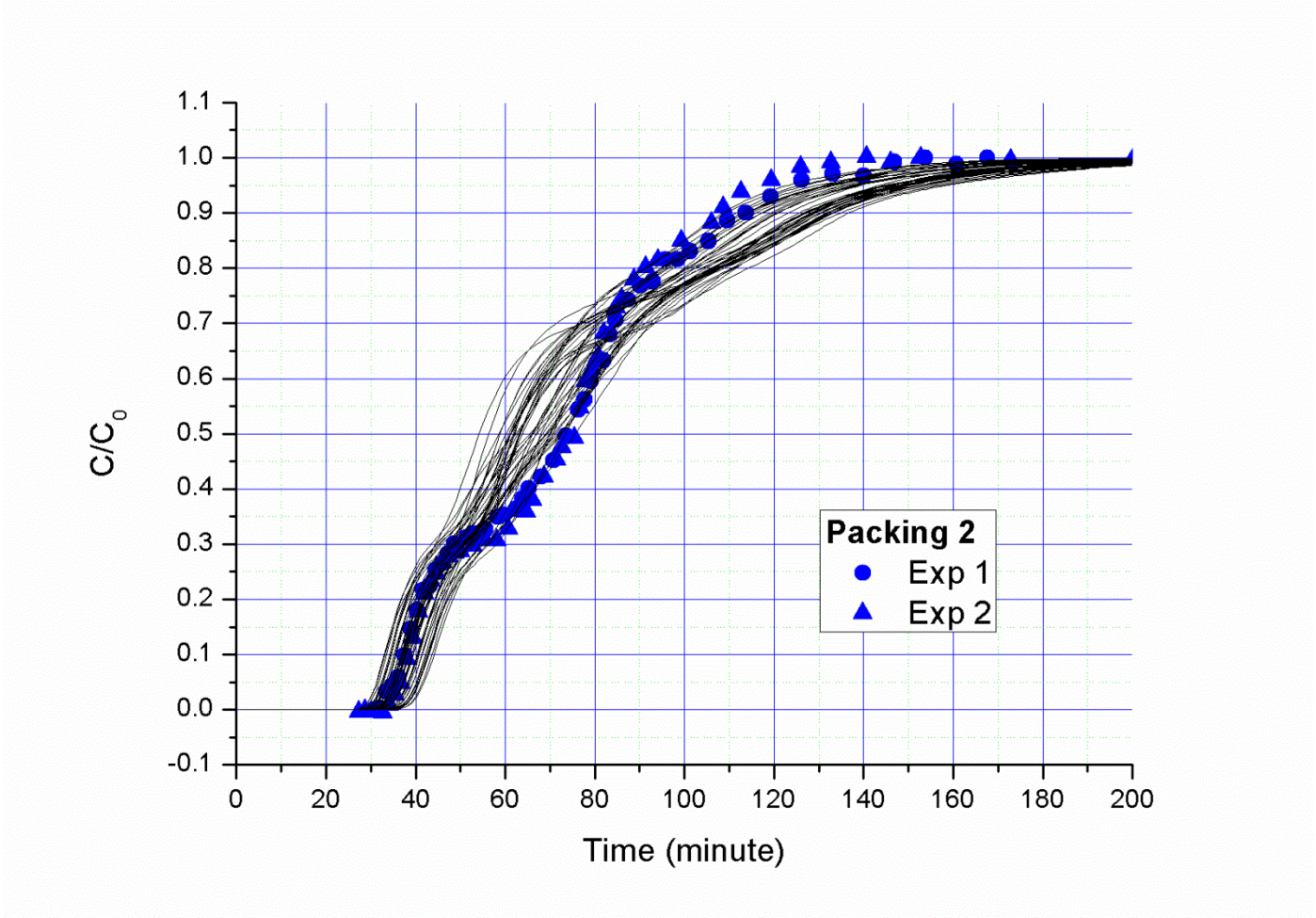


Fig. 14. Comparison between measured BTC and BTCs from 30 Monte Carlo simulations (Packing 2).

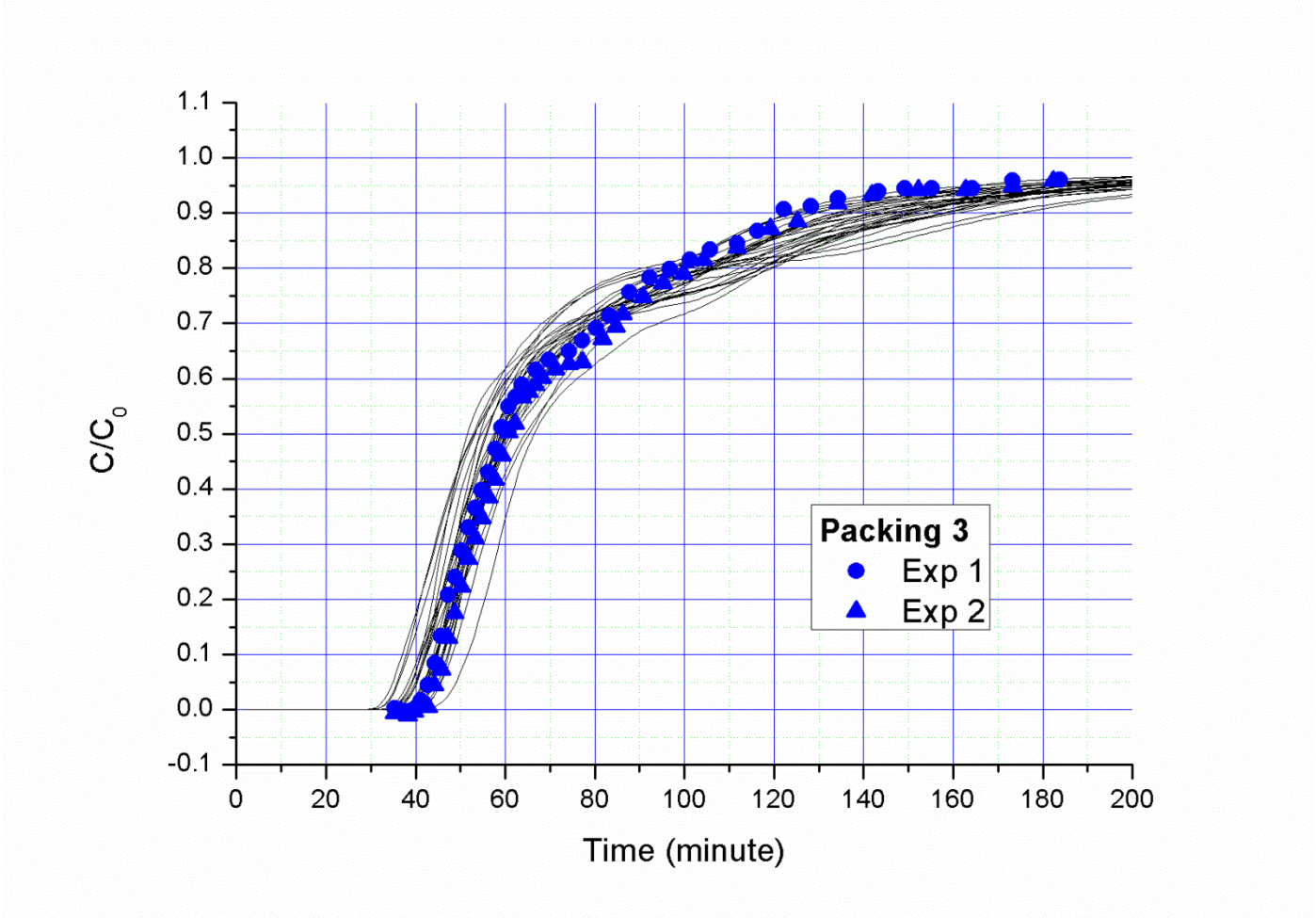


Fig. 15. Comparison between measured BTC and BTCs from 30 Monte Carlo simulations (Packing 3).