In Memoriam Adelina Georgescu

GLOBAL RANDOM WALK SIMULATIONS FOR SENSITIVITY AND UNCERTAINTY ANALYSIS OF PASSIVE TRANSPORT MODELS[*]

Nicolae Suciu[†]   Călin Vamoş[‡]   Harry Vereecken[§]   Peter Knabner[¶]

Abstract

The Global Random Walk algorithm (GRW) performs a simultaneous tracking on a fixed grid of huge numbers of particles at costs comparable to those of a single-trajectory simulation by the traditional Particle Tracking (PT) approach. Statistical ensembles of GRW simulations of a typical advection-dispersion process in groundwater systems with randomly distributed spatial parameters are used to obtain reliable estimations of the input parameters for the upscaled transport model and of their correlations, input-output correlations, as well as full probability distributions of the input and output parameters.

MSC: 65M75, 82C70, 65C05

1 Introduction

Models of passive scalar transport in highly heterogeneous media, such as groundwater systems, turbulent atmosphere, or plasmas, are often based on a stochastic partial differential equation for the concentration field $c(\mathbf{x}, t)$,

$$\partial_t c + \mathbf{V}\nabla c = D\nabla^2 c, \tag{1}$$

with space variable drift $\mathbf{V}(\mathbf{x})$ which is a sample of a random velocity field, and a local diffusion coefficient D which is assumed constant [9, 10, 14, 15, 7]. The normalized concentration solving (1) for the initial condition $c(\mathbf{x}, 0) = \delta(\mathbf{x} - \mathbf{x}_0)$ is the probability density function of the diffusion process described by the Itô stochastic ordinary differential equation

$$X_i(t) = x_{0i} + \int_0^t V_i[\mathbf{X}(t')]dt' + W_i(t), \tag{2}$$

where $i = 1, 2, 3$, $x_{0i} = X_i(0)$ are deterministic initial positions and W_i are the components of a Wiener process of mean zero and variance $2Dt$ [5].

In this paper we consider contaminant transport in saturated groundwater systems. The time-stationary random velocity field $\mathbf{V}(\mathbf{x})$ is, in this case, the solution of the continuity and Darcy equations

$$\nabla \mathbf{V} = 0, \ \mathbf{V} = -K\nabla h, \tag{3}$$

where $K(\mathbf{x})$ is the hydraulic conductivity of the medium and h is the piezometric head [7]. Dirichlet boundary conditions, consisting of constant heads at the inlet and outlet boundaries of the domain, ensure the stationarity in time of the velocity field $\mathbf{V}$. The hydraulic conductivity K is supplied by various interpretations of field-scale measurements in form of a spatially distributed random parameter (random field) [2].

If the random velocity field, obtained by solving (3) for an ensemble of realizations of the K field, has a finite correlation range then it can be shown that, under certain conditions, the ensemble mean concentration is described asymptotically by an upscaled model of form (1), with drift coefficient given by the mean velocity and enhanced diffusion coefficients proportional with

the velocity correlation lengths [6, 4]. Under less restrictive conditions, with the only assumption that the first two spatial moments of the concentration are finite at finite times, the mean concentration can still be described by an equivalent Gaussian distribution with time variable diffusion coefficients [15], referred to as the "macrodispersion" model in the hydrological literature [2]. Root-mean-square deviations of the solutions to (1), for fixed realizations of the velocity field, from the predictions of the upscaled model are often used to quantify the uncertainty in stochastic modeling of transport in random environments [9, 12, 13, 14]. When the estimated mean-square uncertainty is acceptably small, one considers that "ergodic conditions" are met and the macrodispersion model can be successfully used to describe the transport in a single realization of the groundwater formation [9]. Nevertheless, for contamination risk assessments mean-square uncertainty assessments are not enough and extreme values of the stochastic predictions are also required. Such a task can be carried out by assessing the correlations and the full probability distributions of the input/output parameters [1].

When solving advection-dominated transport problems associated to (1), like the one considered here, with Péclet numbers Pe= $U\lambda/D = 100$, where U is the amplitude of the mean velocity and λ a correlation length, the challenge is to ensure the stability of the solutions and to avoid the numerical diffusion [7]. Therefore, numerical solutions to the Itô equation (2), implemented in so called Particle Tracking (PT) algorithms, are often used to simulate trajectories of computational particles and to estimate concentrations by particles densities. PT methods are stable, free of numerical diffusion, thus suitable for advection-dominated transport problems. However, since the computational costs increase linearly with the number of particles, the estimated concentrations are too inaccurate for large-scale simulations of transport in groundwater. Overcoming the limitations of the sequential PT procedure, the Global Random Walk (GRW) has no limitations as concerning the number of particles [9, 16]. As shown in Sect. 2.2 below, GRW provides accurate simulations of the concentration field at costs comparable to those of a single-trajectory PT simulation.

The paper is organized as follows. After recalling basic notions about Euler schemes and PT methods in Section 2.1, we introduce in Section 2.2 the GRW algorithm as a weak numerical scheme for the Itô equation and in Section 2.3 we present a two-dimensional GRW algorithm. A Monte Carlo approach based on GRW is described in Section 3.1. Finally, in Section

3.2 we demonstrate the ability of the GRW approach to produce a detailed sensitivity and uncertainty numerical analysis of the macrodispersion model.

2   Numerical simulations of diffusion processes

2.1   Itô equation and Particle Tracking

Let us consider the one-dimensional Itô equation (2) and an equidistant time discretization $0 < \delta t < \cdots < k\delta t < \cdots < K\delta t = T$. In most of its implementations, the PT simulation of the particle's trajectory consists of an Euler approximation Y_t of the solution $X(t)$, which is a continuous time process satisfying the iterative scheme

$$Y_{k+1} = Y_k + V_k \delta t + \delta W_k, \tag{4}$$

where $Y_k = Y_{k\delta t}$, $V_k = V(Y_k)$, and $\delta W_k = W_{k+1} - W_k$ is the increment of the Wiener process. While the *strong convergence* of order $\beta > 0$ of the Euler scheme requires

$$\lim_{\delta t \longrightarrow 0} E\left(|X_t - Y_t|\right) \leq C\delta t^\beta,$$

where E denotes the expectation, for the *weak convergence* of order $\beta > 0$, it suffices that

$$\lim_{\delta t \longrightarrow 0} |E\left(g(X_t)\right) - E\left(g(Y_t)\right)| \leq C\delta t^\beta,$$

for some functionals $g(X_t)$ (e.g. moments $E(X_t^m)$, $m \geq 1$).

For strong pathwise convergence, the Euler scheme (4) has to consider the Wiener process specified in the Itô equation (2). For weak convergence, when only the probability distribution is approximated, the increments of the Wiener process can be replaced by random variables ξ with similar moments. For weak Euler scheme of order $\beta = 1$ the first three moments of ξ have to satisfy, for some constant M, the condition [5, Sect. 5.12]

$$|E(\xi)| + \left|E(\xi^3)\right| + \left|E(\xi^2) - \delta t\right| \leq M\delta t^2.$$

Easily generated noise increments satisfying the above condition are the two-states random variables

$$\xi : \Omega \longrightarrow \{-\sqrt{2D\delta t}, +\sqrt{2D\delta t}\}, \ P\{\xi = \pm\sqrt{2D\delta t}\} = \frac{1}{2}. \tag{5}$$

2.2   Global Random Walk

As far as one approximates probability distributions and their moments the trajectories of the weak Euler scheme are in fact not necessary. The probability distribution of the surrogate random increments of the Wiener process (5) is the limit over a large number of trials N of the relative frequency n/N of occurrence of n heads or tails of an unbiased coin. This can also be thought of as probability that a random walker takes unbiased left/right jumps of constant length $\delta x = \sqrt{2D\delta t}$ on a lattice,

$$P\{\leftarrow\} = P\{\rightarrow\} = \lim_{N \longrightarrow \infty} \frac{n^{\leftarrow}}{N} = \lim_{N \longrightarrow \infty} \frac{n^{\rightarrow}}{N} = \frac{1}{2}, \tag{6}$$

where $n^{\leftarrow}$ and $n^{\rightarrow}$ are the number of walkers jumping to the first-neighbor left site and to the first-neighbor right site, respectively.

The evaluation of the moments $E(X_t^m)$ within the numerical implementation of the weak Euler scheme consists of an arithmetic average, over an ensemble of trajectories (4), of the position of the particles at a given time, which approximates the stochastic average with respect to the probability distribution, $E(X_t) = \int x^m P(t, dx)$. The latter average can also be estimated by discretizing the integral on a regular grid of length L and space step δx as a sum $\sum_{i=1}^{L} (i\delta x)^m P(i\delta x)$, where the probability distribution at a fixed time $P(i\delta x)$ can be approximated by the relative frequency of occupation of the i-th lattice site, n_i/N. Since, according to (5), the walkers cannot be trapped at lattice sites, the occupancy number n_i is the sum of numbers of walkers reaching the site i from the left, $n_i^{\rightarrow}$, and from the right, $n_i^{\leftarrow}$, i.e. $n_i = n_i^{\rightarrow} + n_i^{\leftarrow}$. One obtains thus the estimation of the m-th order moment of X_t given by

$$E(X_t^m) = \sum_{i=1}^{L} (i\delta x)^m \left(\frac{n_i^{\rightarrow}}{N} + \frac{n_i^{\leftarrow}}{N} \right). \tag{7}$$

For large N, the random variables $n_i^{\rightarrow}$ and $n_i^{\leftarrow}$ occurring in (6-7) can be well approximated as follows. If the number n_i of walkers at the grid site i is even then half of them jump to the left and half to the right, $n_i^{\leftarrow} = n_i^{\rightarrow} = n/2$. If n_i is odd then one walker is allocated to either $n_i^{\leftarrow}$ or to $n_i^{\rightarrow}$ with the same probability, $P\{\leftarrow\} = P\{\rightarrow\} = 1/2$. One obtains in this way a GRW algorithm for the Wiener process, described by equation (2) without drift term [16]. Figure 1 illustrates the evolution of the number n_i of random walkers over the first three simulation steps, obtained with a straightforward

MATLAB implementation of the above one-dimensional GRW algorithm. The concentration at a given time (solution of (1)) can be simply estimated as $c(i\delta x) = n_i/\delta x$.
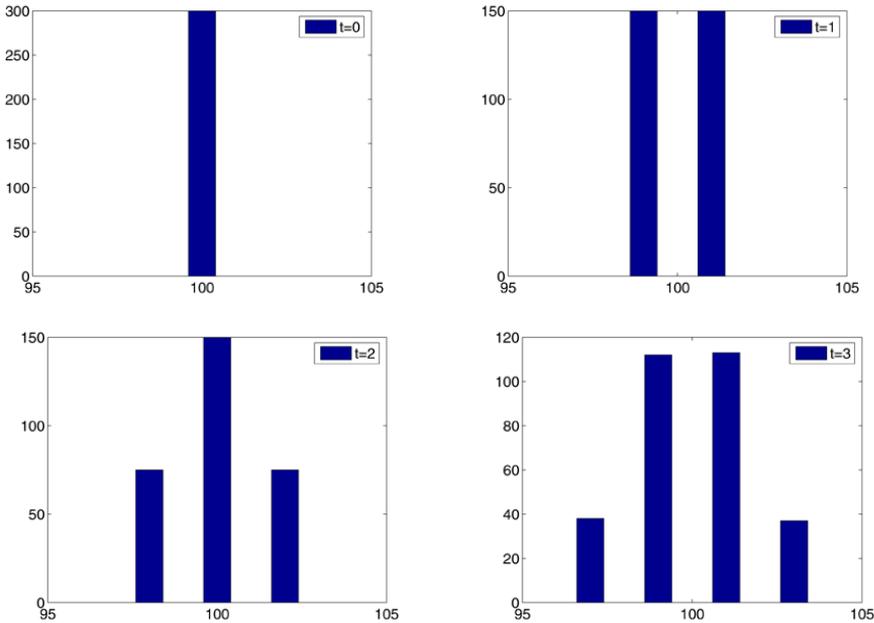

Figure 1: Distribution of $N = 300$ random walkers after the first three time steps of the GRW simulation.

Unlike the discrete-time grid-free weak Euler scheme, the GRW algorithm is a discrete time-space stochastic scheme. As follows from (5) the constant amplitude δx of the random jumps ξ is related to the time step δt and the diffusion coefficient D by

$$D = \frac{\delta x^2}{2\delta t}. \tag{8}$$

Since the numerical scheme is constrained by the relation (8), GRW is not affected by numerical diffusion. GRW is also stable because the number of random walkers N is conserved. Figure 2 shows the estimated mean $M = E(X_t)$ and diffusion coefficient $D = [E(X_t^2) - E(X_t)^2]/(2t)$, computed according to (7), as well as the final distribution of n_i for a diffusion process with $D = 1$ resulted from a GRW simulation with $\delta x = 1$ and $\delta t = 0.5$.

It is also possible to simplify the GRW algorithm by completely removing the randomness from the scheme. This is done by setting $n_i^{\leftarrow}$ and $n_i^{\rightarrow}$ to the exact value of $n/2$. In this case N has no longer the meaning of a number of random walkers and can be taken as an arbitrary positive real number, for instance equal to 1. This deterministic GRW scheme is equivalent to the finite-difference scheme for the heat equation and converges as δx^2 for $\delta x \to 0$ [16]. Since according to relation (8) $\delta x^2 \sim \delta t$, the deterministic GRW has the same order of convergence with the time step as the weak Euler scheme of order $\beta = 1$. The convergence of the stochastic GRW simulation reaches the same order only if the number of random walkers N is large enough to smooth out the random fluctuations of n_i. Figure 3 shows the dependence on N of the absolute error $eD(t) = |D_{grw}(t) - D|$ and the convergence of the norm $\|D_{grw} - D\|$ defined by

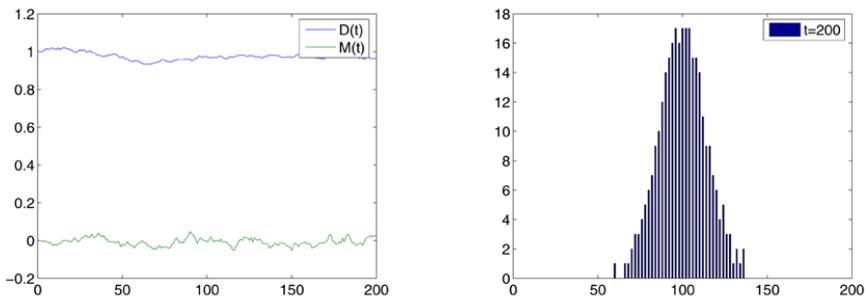$$\|D_{grw} - D\|^2 = \sum_{k=1}^{T/\delta t} [D_{grw}(k\delta t) - D]^2 .$$

Figure 2: Estimation of the diffusion coefficient $D(t)$ and of the mean $M(t)$ (left) and distribution of $N = 300$ random walkers after 200 time steps in the GRW simulation (right).

Note that the GRW scheme described above is practically insensitive to the number of random walkers N. Assuming that all L grid points contain random walkers at all the computation time steps, one needs LT calls of a uniformly-distributed random-numbers generator for the entire simulation. Hence, the total computation time is of the order of that for the simulation of a single trajectory of the Itô process by the weak Euler scheme. Since for $N =$
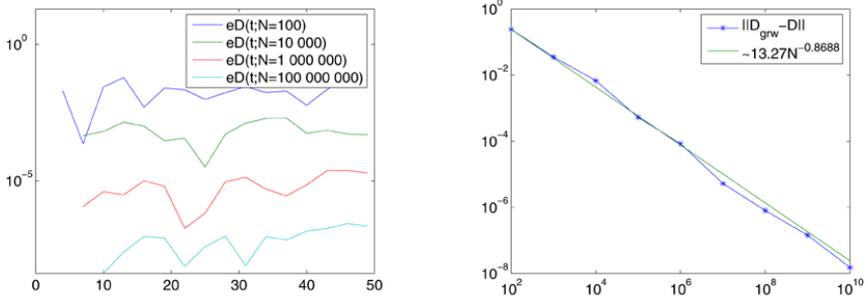
Figure 3: Errors for the estimations of diffusion coefficients for increasing N (left) and the convergence of the error norm (right).

1 the output of the simulation is the trajectory of a single random walker, GRW can be thought of as a superposition of particle tracking procedures for arbitrary large numbers of particles. Since the computational cost of a simulation for N trajectories with the Euler scheme is of the order of NT, the GRW algorithm achieves a speed-up of computations, with respect to PT, of the order N/L. For example, while the convergence investigations with GRW presented in Figure 3 were performed in about one second, similar investigations with the Euler scheme required several minutes on the same computer. In case of realistic simulations of diffusion processes, when very large numbers of particles should be considered, e.g. $N = 10^{24}$ (Avogadro's number), as well as large grids of the order of $L = 10^6$ nodes, a huge speed-up of computations by a factor of 10^{18} can be achieved by using the GRW algorithm.

2.3 Two-dimensional GRW algorithm

For a two-dimensional transport problem, the solution of the parabolic equation (1) is simulated with N particles undergoing advective displacements and diffusive jumps according to the random walk law on a regular grid. The concentration at a given time $t = k\delta t$ and a point $(x_1, x_2) = (i_1\delta x_1, i_2\delta x_2)$ is given by

$$c(x_1, x_2, t) = \frac{1}{N\Delta_1\Delta_2} \sum_{i_1'=-s_1}^{s_1} \sum_{i_2'=-s_2}^{s_2} n(i_1 + i_1', i_2 + i_2', k), \qquad (9)$$

where $\Delta_l = 2s_l \delta x_l$, $l = 1, 2$, are the lengths of the symmetrical intervals centered at x_l and $n(i_1, i_2, k)$ is the number of particles which at the time step k lie at the grid point (i_1, i_2).

For constant diffusion coefficient D, the two-dimensional simulation consists of repeating the one-dimensional procedure on each of the two spatial directions [16, 11]. The one-dimensional GRW algorithm, which generalizes the algorithm presented in Section 2.2 to account for advective displacements, describes the scattering of the $n(i, k)$ particles from (x_i, t_k) by

$$n(j, k) = \delta n(j, j + v_j, k) + \delta n(j + v_j - d, j, k) + \delta n(j + v_j + d, j, k), \quad (10)$$

where $v_j = V_j \delta t / \delta x$ are discrete displacements produced by the velocity field and d describes the diffusive jumps. The quantities δn introduced in (10) are Bernoulli random variables and describe respectively, the number of particles which remain at the same grid site after an advective displacement and the number of particles jumping to the left and to the right of the advected position $j + v_j$. The distribution of the particles at the next time $(k + 1)\delta t$ is given by

$$n(i, k + 1) = \sum_j \delta n(i, j, k).$$

The average number of particles undergoing diffusive jumps and the average number of particles remaining at the same node after the displacement v_j are given by the relations

$$\overline{\delta n(j + v_j \pm d, j, k)} = \frac{1}{2} r \, \overline{n(j, k)},$$

$$\overline{\delta n(j, j + v_j, k)} = (1 - r) \, \overline{n(j, k)},$$

where $0 \leq r \leq 1$. The diffusion coefficient D is related to the grid steps by the relation

$$D = r \frac{(d\delta x)^2}{2\delta t},$$

which generalizes (8) and ensures that the scheme does not produce numerical diffusion.

Particularizing the above one-dimensional GRW algorithm for genuine diffusion, i.e. letting $v_j = 0$ in (10), one can easily see that the evolution of the mean number of particles is described by

$$\overline{n(i, k + 1)} = \frac{r}{2} \overline{n(i + d, k)} + (1 - r)\overline{n(i, k)} + \frac{r}{2} \overline{n(i - d, k)}. \quad (11)$$

which has the form of the explicit scheme for the heat equation. Since the scheme (11) is consistent and, by condition $r \leq 1$ (von Neumann's criterion), it is also stable, it converges with the order $O(\delta x^2)$. Moreover, as demonstrated numerically in [16], the un-averaged GRW solution $n(i, k)$ converges as $\mathcal{O}(\delta x^2) + \mathcal{O}(N^{-1/2})$. Thus, for sufficiently large numbers of particles GRW has the same order of convergence as the stable finite differences scheme.

It is worth noting that while for constant drift coefficients V_j the GRW algorithm is still equivalent to a finite difference scheme, the equivalence fails for space variable V_j. Indeed, in the latter case to the site i contribute not only particles jumping from two symmetrical left and right sites, like in the finite difference scheme (11), but also particles coming from distances $v_j \pm d$ which depend on the variable drift coefficient V_j. However, GRW remains equivalent to a superposition of many PT schemes and this makes it suitable for simulating advection-diffusion processes described by the parabolic equation (1). In fact, as shown in Section 2.2 above, GRW is a weak scheme for solving Itô equations, which approximates the true probability distribution (concentration) at all grid points and time steps, without solving for individual trajectories. This is the essential feature which considerably increases the performance of the GRW algorithm with respect to PT, where, after the sequential simulation of particles trajectories, a post-processing is required to count the contribution of the computational particles to the concentration, estimated at given points in space and time steps.

The "reduced fluctuations" GRW algorithm generalizes the simple procedure described in Section 2.2 by

$$\delta n(j + v_j - d, j, k) = \left\{ \begin{array}{cc} n/2 & \text{if } n \text{ is even} \\ [n/2] + \theta & \text{if } n \text{ is odd,} \end{array} \right.$$

where $n = n(j, k) - \delta n(j, j + v_j, k)$, $[n/2]$ is the integer part of $n/2$ and θ is a variable taking the values 0 and 1 with probability $1/2$. Further, the number of particles jumping in the opposite direction, $\delta n(j, j + v_j + d, k)$ is determined by (10). This algorithm is appropriate for large scale problems, for two reasons. Firstly, the diffusion front does not extend beyond the limit concentration defined by one particle at a grid point, keeping a physical significant shape (unlike in finite differences schemes, where a pure diffusion front has a cubic shape of side $\sim \sqrt{2Dt}$). Secondly, the reduced fluctuations algorithm requires only a minimum number of calls of the random number generator.

A comparison with a PT code (done for the diffusion over ten time steps of N particle starting at the center of a cubic grid) shows that while for the GRW algorithm there were practically no limitations concerning the total number of particles and the computation time was of about one second, PT simulations for $N = 10^9$ particles already required a computing time of about one hour and 256 processors on a CRAY T3E parallel machine [16].

To compute moments, as for instance the variance of particle displacements $s_{ll}^2 = E(X_l^2) - E(X_l)^2$, $l = 1, 2$, a more accurate result is obtained if instead of the concentration (9) one uses the point density of the number of particles $n(i_1, i_2, k)$:

$$\frac{1}{(\delta x)^2} s_{ll}^2(k\delta t) = \frac{1}{N} \sum_{i_1, i_2} i_l^2 n(i_1, i_2, k) - \left[\frac{1}{N} \sum_{i_1, i_2} i_l \, n(i_1, i_2, k) \right]^2 .$$

With this, the effective diffusion coefficients will be computed as

$$D_{ll}^{eff}(k\delta t) = s_{ll}^2/(2k\delta t). \tag{12}$$

Let us consider N_{X_0} points uniformly distributed inside the initial plume, N/N_{X_0} particles at each initial point and let $n(i_1, i_2, k; i_{01}, i_{02})$ be the distribution of particles at the time step k given by the GRW procedure for a diffusion process starting at $(i_{01}\delta x_1, i_{02}\delta x_2)$. Writing the distribution for the extended plume as

$$n(i_1, i_2, k) = \sum_{i_{01}, i_{02}} n(i_1, i_2, k; i_{01}, i_{02}),$$

the averages defining the first two moments can be rewritten in the form

$$\frac{1}{N} \sum_{i_1, i_2} \alpha n(i_1, i_2, k) = \frac{1}{N_{X_0}} \sum_{i_{01}, i_{02}} \left(\frac{N_{X_0}}{N} \sum_{i_1, i_2} \alpha n(i_1, i_2, k; i_{01}, i_{02}) \right), \tag{13}$$

where α stands for i_l and i_l^2 respectively. As follows from (13), the first two moments $E(X_l)$, and $E(X_l^2)$, as well as the effective diffusion coefficients (12) are averages over the trajectories of the diffusion process starting at given initial positions and over the distribution of the initial positions.

3 Sensitivity and uncertainty analysis

3.1 Monte Carlo simulations

To enable the simulation of large ensembles of transport realizations, a linearization of the flow equation (3) was considered and the velocity samples were generated, for given statistics of the hydraulic conductivity K, by the Kraichnan's randomization method [8], which has been successfully used in numerical investigations on large scale behavior of the passive transport in aquifers [3, 9, 10]. We considered a log-normally distributed conductivity K, i.e. a normal lnK field with variance σ^2 and exponential isotropic correlation $\rho(|\mathbf{x}_1 - \mathbf{x}_2|) = \sigma^2 exp(-|\mathbf{x}_1 - \mathbf{x}_2|/\lambda)$, where λ is the correlation length. For a given pressure gradient between the inlet and outlet boundaries, which fixes the value of the ensemble mean velocity $U = |\langle \mathbf{V} \rangle|$, the incompressible Darcy flow, solution of equations (3), was approximated by a superposition of N_p periodic modes

$$V_i(\mathbf{x}) = U\delta_{i1} + U\sigma\sqrt{\frac{2}{N_p}} \sum_{l=1}^{N_p} p_i(\mathbf{q}_l) \sin(\mathbf{q}_l \cdot \mathbf{x} + \alpha_l). \qquad (14)$$

The wave vectors $\mathbf{q}_l$ are mutually independent random variables, with probability distribution proportional with the spectral density of the lnK field, and the phases α_l are random variables uniformly distributed in the interval $[0, 2\pi]$. The functions p_l are projectors which ensure the incompressibility of the flow. It has been shown that V_i tends to a Gaussian random field when $N_p \to \infty$ [8]. It was also found that $N_p = 6400$, which we fix in the following, provides reliable approximations of the velocity field at the problem's spatial scale considered here [9, 3].

The mean velocity occurring in (14), which can be freely chosen, was set to a typical value of $U = 1$ m/day. We also have chosen a typical local-scale diffusion coefficient in (1), $D = 0.01$ m^2/day, and $\lambda = 1$ m for the correlation length of the lnK field, so that the Péclet number was set to Pe= $U\lambda/D = 100$. We conducted Monte Carlo simulations for two cases, corresponding to two extreme degrees of heterogeneity: $\sigma^2 = 0.1$, for which the approximation (14) of the velocity field is accurate and the macrodispersion model is expected to provide a reliable description of the mean behavior of the transport process, and $\sigma^2 = 6$, an extremely large value, for which (14) is no longer close to the true solution of flow equations (3) but can however

serve to illustrate the situation when the macrodispersion model might be inadequate.

The behavior of a passive tracer, initially uniformly distributed in slabs of dimensions $100\lambda \times \lambda$ perpendicular to the mean flow direction, was simulated over 2000 days for the low heterogeneity case $\sigma^2 = 0.1$, in 1024 realizations of the random field (14), and over 300 days, in 100 realizations in the highly heterogeneous case $\sigma^2 = 6$. The plume's shapes in the two extreme cases are compared in Figure 4. (Note that the spatial simulation domain was, in all cases, large enough to avoid the influence of the boundaries.)
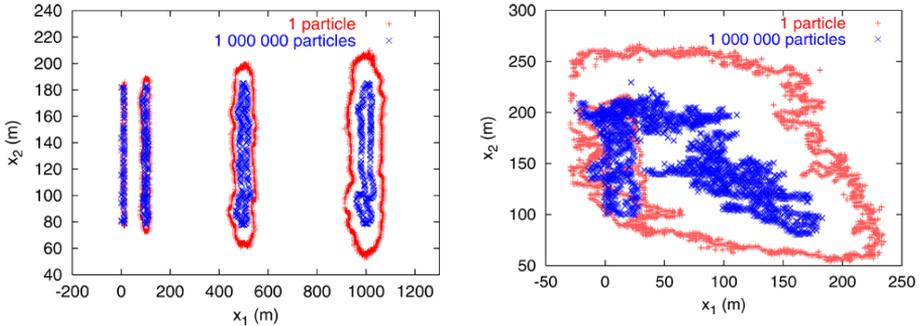

Figure 4: Plume contours for $\sigma^2 = 0.1$ at $t = 0$, 100, 500 and 1000 days (left panel) and for $\sigma^2 = 6$ at $t = 0$, 10, and 100 days (right panel).

Monte Carlo estimates, by equal-weight (arithmetic) averages over the corresponding ensembles of realizations, hereafter denoted by $\langle \cdots \rangle$, were computed for the set of input parameters of the macrodispersion model, consisting of longitudinal $u = E(X_1)/t$ and transverse $v = E(X_2)/t$ components of the center of mass velocity, longitudinal $D_x = D_{11}^{eff}$ and transverse $D_y = D_{22}^{eff}$ effective diffusion coefficients (12), for the only output parameter considered here, consisting of the cross-section space average concentration at the center of mass (hereafter denoted by c), as well as for their cross-correlations, $\langle uv \rangle$, $\langle uD_x \rangle$, $\langle uD_y \rangle$, $\langle vD_x \rangle$, $\langle vD_y \rangle$, $\langle D_xD_y \rangle$, $\langle uc \rangle$, $\langle vc \rangle$, $\langle D_xc \rangle$, and $\langle D_yc \rangle$. Probability densities of the parameters, approximated by histograms, were summed-up to estimate cumulative probability distributions.

3.2   Results

The left panel of Figure 5 shows that for low heterogeneity ($\sigma^2 = 0.1$) the only input-input relevant correlation is that between the longitudinal velocity of the center of mass and the transverse effective diffusion coefficient. The sensitivity of the transverse dispersion to the mean longitudinal flow indicates the increased role of the transverse dispersion for small mean flow velocity. The results for the highly heterogeneous case ($\sigma^2 = 6$) from the right panel of Figure 5 show stronger correlations between the input parameters, which are expected to facilitate the uncertainty propagation and to reduce the reliability of the macrodispersion model.
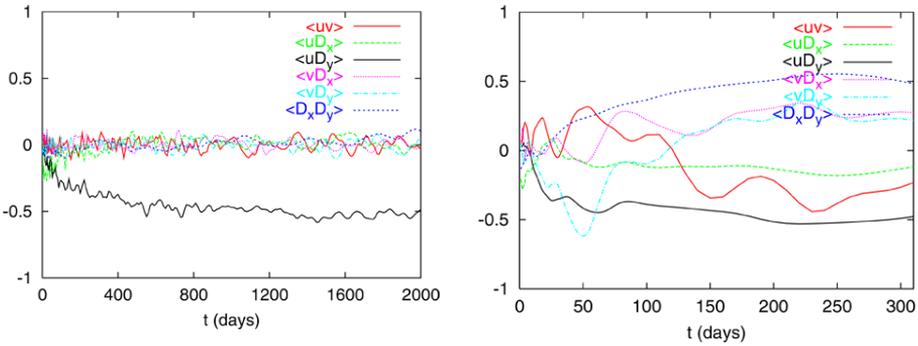

Figure 5: Correlations between input parameters of the macrodispersion model (velocity components of center of mass, u and v, and dispersion coefficients, D_x and D_y) for $\sigma^2 = 0.1$ (left panel) and $\sigma^2 = 6$ (right panel).

As expected, for low heterogeneity (left panel of Figure 6) there is a strong correlation between the longitudinal effective diffusion coefficient and the cross-section averaged concentration. This suggests that, when the only output parameter of interest is the cross-section concentration, the macrodispersion model can be trusted as reliable for single-realizations of the transport process, in agreement with other observations that the cross-section concentration can be modeled as an one-dimensional advection-diffusion process governed by the longitudinal effective diffusion coefficient [9]. The situation is different for high heterogeneity (right panel of Figure 6), where the cross-section concentration is also strongly correlated with the transverse effective diffusion coefficient. Again, this result renders questionable the applicability of the macrodispersion model to highly heterogeneous media.
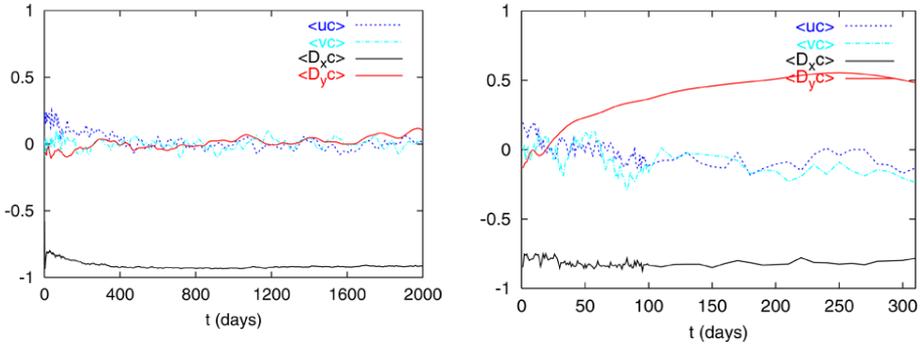
Figure 6: Correlations between input parameters u, v, D_x, and D_y, and the output parameter c (the cross-section space average concentration at the center of mass) for $\sigma^2 = 0.1$ (left panel) and $\sigma^2 = 6$ (right panel).

To illustrate the capability of the Monte Carlo approach based on GRW simulations to produce a full statistical description of the transport process, we present in Figure 7 the estimated cumulative probability distributions of the cross section concentration at the plumes center of mass and of the longitudinal velocity of the center of mass. In a forthcoming work, these probability distributions will be used as reference data in developing a probability density function method similar to those used in modeling turbulent transport [1]. The novelty of the new approach will consists of a three-dimensional GRW solution of the equations governing the evolution of the concentration probability density in the cartesian product between the physical space and the concentration domain.
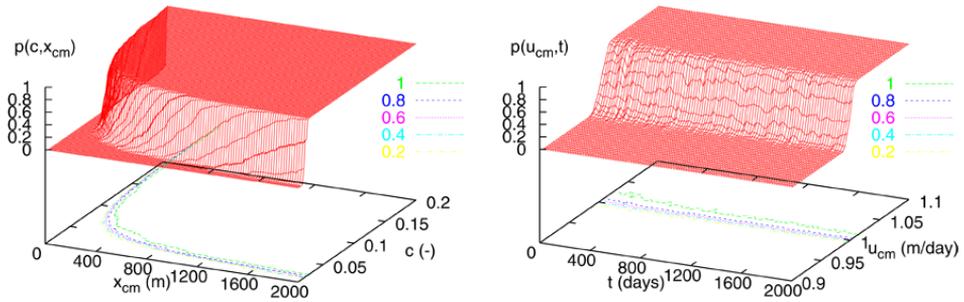
Figure 7: Probability distributions of the concentration estimated along the longitudinal component of the center of mass $c(x_{cm})$ (left panel) and of the longitudinal component of the center of mass velocity as function of time $u_{cm}(t)$ (right panel), for $\sigma^2 = 0.1$.

References

[1] P. J. Colucci, F. A. Jaberi, P. Givi. Filtered density function for large eddy simulation of turbulent reacting flows. *Phys. Fluids* 10:499-515, 1998.

[2] G. Dagan. *Flow and Transport in Porous Formations.* Springer, Berlin, 1989.

[3] J. Eberhard, N. Suciu, C. Vamoş. On the self-averaging of dispersion for transport in quasi-periodic random media. *J. Phys. A: Math. Theor.* 40:597-610, doi: 10.1088/1751-8113/40/4/002, 2007.

[4] A. Fannjiang, T. Komorowski. Turbulent diffusion in Markovian flows, *Ann. Appl. Probab.* 9:591-610, 1999.

[5] P. E. Kloeden, E. Platen. *Numerical Solutions of Stochastic Differential Equations.* Springer, Berlin, 1999.

[6] T. Komorowski, G. Papanicolaou. Motion in a Gaussian incompressible flow, *Ann. Appl. Probab.* 7:229-264, 1997.

[7] F. A. Radu, N. Suciu, J. Hoffmann, A. Vogel, O. Kolditz, C.-H. Park, S. Attinger. Accuracy of numerical simulations of contaminant transport in heterogeneous aquifers: a comparative study. *Adv. Water Resour.*, doi:10.1016/j.advwatres.2010.09.012, 2010 (in press).

[8] K. Sabelfeld. *Monte Carlo methods in boundary value problems.* Springer, Berlin, 1991.

[9] N. Suciu, C. Vamoş, J. Vanderborght, H. Hardelauf, H. Vereecken. Numerical investigations on ergodicity of solute transport in heterogeneous aquifers. *Water Resour. Res.* 42:W04409, doi:10.1029/2005WR004546, 2006.

[10] N. Suciu, C. Vamoş, J. Eberhard. Evaluation of the first-order approximations for transport in heterogeneous media. *Water Resour. Res.* 42:W11504, doi:10.1029/2005WR004714, 2006.

[11] N. Suciu, C. Vamoş, I. Turcu, C.V.L. Pop, and L. I. Ciortea. Global random walk modeling of transport in complex systems. *Comput. Visual. Sci.*, 12:77-85, doi:10.1007/s00791-007-0077-6, 2007.

[12] N. Suciu, C. Vamoş, H. Vereecken, K. Sabelfeld, P. Knabner. Memory effects induced by dependence on initial conditions and ergodicity of transport in heterogeneous media. *Water Resour. Res.* 44:W08501, doi:10.1029/2007WR006740, 2008.

[13] N. Suciu, P. Knabner. Comment on 'Spatial moments analysis of kinetically sorbing solutes in aquifer with bimodal permeability distribution' by M. Massabo, A. Bellin, and A. J. Valocchi. *Water Resour. Res.* 45:W05601, doi:10.1029/2008WR007498, 2009.

[14] N. Suciu, C. Vamos, F. A. Radu, H. Vereecken, P. Knabner. Persistent memory of diffusing particles. *Phys. Rev. E* 80:061134, doi:10.1103/PhysRevE.80.061134, 2009.

[15] N. Suciu. Spatially inhomogeneous transition probabilities as memory effects for diffusion in statistically homogeneous random velocity fields. *Phys. Rev. E* 81:056301, doi:10.1103/PhysRevE.81.056301, 2010.

[16] C. Vamoş, N. Suciu, H. Vereecken. Generalized random walk algorithm for the numerical modeling of complex diffusion processes. *J. Comput. Phys.* 186:527-544, doi:10.1016/S0021-9991(03)00073-1, 2003.