

Ergodic Estimations of Upscaled Coefficients for Diffusion in Random Velocity Fields

Nicolae Suciu and Călin Vamoş

Abstract Upscaled coefficients for diffusion in ergodic velocity fields are derived by summing up correlations of increments of the position process, or equivalently of the Lagrangian velocity. Ergodic estimations of the correlations are obtained from time averages over finite paths sampled on a single trajectory of the process and a space average with respect to the initial positions of the paths. The first term in this path decomposition of the diffusion coefficients corresponds to Markovian diffusive behavior and is the only contribution for processes with independent increments. The next terms describe memory effects on diffusion coefficients until they level off to the value of the upscaled coefficients. Since the convergence with respect to the path length is rather fast and no repeated Monte Carlo simulations are required, this method speeds up the computation of the upscaled coefficients over methods based on long-time limit and ensemble averages by four orders of magnitude.

1 Introduction

Direct Monte Carlo estimations of diffusion coefficients by averaging over ensembles of realizations of the process and taking the large time limit often constitute a numerical challenge in simulation studies. Owing to the ergodicity of the process, the numerical burden can be reduced to a great extent.

For ergodic transport processes, the ensemble average of the observables can be estimated by the arithmetic mean of the observables resulting from repeated simulations of diffusion, done for the same realization of velocity field and for point-like

Nicolae Suciu

Mathematics Department, Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen, Germany

url: <http://www1.am.uni-erlangen.de/~suciu/>

Călin Vamoş

Tiberiu Popoviciu Institute of Numerical Analysis, Romanian Academy, Cluj-Napoca, Romania

url: <http://www.ictp.acad.ro/vamos/>

sources with different locations uniformly distributed over large enough spatial domains [8]. Even though promising results can be obtained in this way, this “ergodic simulations method” depends on the quality of the numerically generated processes: worse are the ergodic properties of the latter, greater is the number of simulations for different initial positions required to achieve the desired accuracy. For instance, results of two-dimensional ergodic simulations of diffusion in random velocity fields presented in [8], obtained by averaging over a moderate number of 121 initial positions, indicated the approach of ergodic estimates to the corresponding ensemble averages but the accuracy of the upscaled diffusion coefficients was not yet satisfactory. To increase the accuracy, more initial positions should be considered, which would increase the computational costs and render the ergodic simulations less competitive with respect to the direct Monte Carlo approach.

The “path decomposition method” proposed in Sect. 2 provides ergodic estimates of diffusion coefficients by sums of correlations of increments on paths of increasing but finite lengths on a single trajectory of the diffusion process. To increase the accuracy, the path correlations are further averaged over a large number of paths. In the case of diffusion in random velocity fields, the upscaled diffusion coefficients can be explicitly written in terms of correlations of the Lagrangian velocity, sampled on the same trajectory (Sect. 3). Summing up autoregressive processes of order 1, diffusion processes with memory and exactly computable diffusion coefficients are constructed and are used to test the path decomposition method (Sect. 4). Finally, in Sect. 5, the new approach is applied to a problem of diffusion in random velocity fields which occurs in modeling groundwater contamination. Some conclusions are drawn in Sect. 6.

2 Path Decomposition of Diffusion Coefficients

Let $X = \{X_t, t \geq 0\}$ be a stochastic process of mean zero starting from $X_0 = 0$. If after a transient time X behaves as a normal diffusion, the diffusion coefficient is related to the expectation $E\{X_t^2\}$ by the Einstein formula [1, 15]

$$D = \lim_{t \rightarrow \infty} \frac{1}{2t} E\{X_t^2\}. \quad (1)$$

Dividing a finite time interval $[0, t]$ in S subintervals of equal length τ , $t = S\tau$, the position X_t and its square X_t^2 can be expressed in terms of random variables $\delta X_s = X_{s\tau} - X_{(s-1)\tau}$,

$$X_t = \sum_{s=1}^S \delta X_s, \quad X_t^2 = \sum_{s=1}^S (\delta X_s)^2 + 2 \sum_{r=1}^{S-1} \sum_{s=1}^{S-r} \delta X_s \delta X_{s+r}.$$

Defining the (time average) correlations

$$\rho(r) = \frac{1}{S-r} \sum_{s=1}^{S-r} \delta X_s \delta X_{s+r}$$

one obtains

$$X_t^2 = S\rho(0) + 2 \sum_{r=1}^{S-1} (S-r)\rho(r). \quad (2)$$

The expectation of (2) is a discrete form of Taylor's formula (see e.g. [3]). If X is an ergodic process, the diffusion coefficient can be estimated as $D = \frac{1}{2S\tau} X_t^2$, without taking the expectation of X_t^2 . For some lattice gas systems and molecular dynamics simulations the hydrodynamic long-time limit in (1) can be indeed replaced by $\tau \rightarrow 0$ (fixed t) and expectations can be approximated by time averages [15]. The space average of (2) over N different paths of length S sampled on a single trajectory of the process X further improves the estimates. Replacing the correlations in (2) by their space averages $\varrho = \frac{1}{N} \sum_{n=1}^N \varrho_n$ and using (1) one obtains the following estimation of the diffusion coefficient for a given length S of the paths:

$$D = \frac{1}{2\tau} \rho(0) + \frac{1}{\tau} \sum_{r=1}^{S-1} \left(1 - \frac{r}{S}\right) \rho(r). \quad (3)$$

For a Markovian diffusive behavior, as for instance random walk or Wiener process, the diffusion coefficients is solely defined by the first term of (3). The correlations $\rho(r)$, with $r > 0$, describe the transient regime of diffusion processes with memory. One expects that if the process takes place in a random environment, (3) also estimates the up-scaled diffusion coefficient provided the environment has suitable ergodic properties.

3 Upscaled Coefficients for Diffusion in Random Fields

The advection-dispersion model for transport in heterogeneous media such as turbulent atmosphere, plasmas, or aquifers [7] is a vector valued process described by the Itô equation $d\mathbf{Y}(t) = \mathbf{V}(\mathbf{Y}(t))dt + d\mathbf{W}(t)$, where $\mathbf{V}$ is a realization of a random velocity field and $\mathbf{W}$ a Wiener process of mean zero and variance $2D_0 t$. The diagonal ll components of the upscaled diffusion coefficient are given by the long time limit of $D_{ll}^* = \frac{1}{2t} \Sigma_{ll}$, where $\Sigma_{ll} = \langle E\{X_l^2\} \rangle$ is the average over the ensemble of velocity realizations of the expected squared displacement from the mean plume center of mass, $X_l = Y_l - \langle E\{Y_l\} \rangle$ [5]. If the Eulerian velocity field $\mathbf{V}$ is statistically homogeneous and has suitable smoothness properties which ensure the existence of pathwise unique solutions of the Itô equation, then the Lagrangian velocity field $\mathbf{V}(\mathbf{Y})$ is also statistically homogeneous [6]. Under these conditions, from the Itô-Euler scheme $\delta X_{l,s} = u_l(\mathbf{Y}_{s-1})\tau + \delta W_{l,s}$, where $u_l = V_l - \langle E\{V_l\} \rangle$ is the velocity fluctuation, one obtains

$$\langle E\{\delta X_{l,s} \delta X_{l,s+r}\} \rangle = \delta_{r,0} 2D_0 \tau + \tau^2 \rho_{u,ll}(r),$$

where

$$\rho_{u,ll}(r) = \langle E\{u_l(Y_{l,s-1})u_l(Y_{l,s-1+r})\} \rangle.$$

If the Eulerian velocity field $\mathbf{V}$ has a finite correlation range then it is ergodic [14]. Assuming that the Lagrangian velocity field inherits the ergodicity property, one expects that the velocity correlation above can be approximated by an average over N paths of length S on a single trajectory of the process $\mathbf{X}(t)$,

$$\rho_{u,ll}(r) = \frac{1}{N} \sum_{n=1}^N \frac{1}{S-r} \sum_{s=1}^{S-r} u_l(Y_{l,n+s-1})u_l(Y_{l,n+s-1+r})$$

and from (3) one obtains

$$D_{ll}^* = D_0 + \frac{\tau}{2} \rho_{u,ll}(0) + \tau \sum_{r=1}^{S-1} \left(1 - \frac{r}{S}\right) \rho_{u,ll}(r). \quad (4)$$

In Sect. 5 below it is shown that the heuristic formula (4) yields quite good estimates of upscaled coefficients, which, in turn, supplies a numerical indication for the ergodicity of the Lagrangian velocity field.

The path correlations decomposition (4) of the diffusion coefficients resembles a discrete Green-Kubo formula [1] and has been used in [9] to calibrate parameters for massive parallel simulations of transport in saturated aquifers [4, 5].

4 Exactly Computable Diffusion Coefficients

The trajectory of a diffusion process can be simulated numerically by summing up realizations of a Gaussian white noise, $Z = \{Z_n, n = 0, 1, 2, \dots\}$, of mean zero and constant variance σ_0^2 . This is a particular case of the more general algorithm for autoregressive processes of order 1 (AR(1))

$$X_n = \phi X_{n-1} + Z_n. \quad (5)$$

AR(1) processes are often used to generate synthetic time series with given statistical properties [10, 13]. For $\phi = 1$, (5) defines a discrete time diffusion process, for $\phi = 0$ it reduces to the white noise, for $0 < \phi < 1$ one obtains a correlated AR(1) process, and for $-1 < \phi < 0$ one obtains an anticorrelated AR(1) [12]. For $|\phi| < 1$ the infinite AR(1) processes, $\{X_n, n = 0, \pm 1, \pm 2, \dots\}$, are stationary and, as follows from (5), their expectation is zero, $E\{X_n\} = 0$, and the variance and the autocovariance are given by the relations

$$\sigma^2 = \frac{\sigma_0^2}{1-\phi}, \quad \gamma(r) = E\{X_n X_{n+r}\} = \sigma^2 \phi^r. \quad (6)$$

Since in numerical simulations AR(1) processes always have a finite length, there is a transient regime until the process reaches a stationary state. If the first term X_1 in (5) is a realization of a random variable whose variance equals the variance σ^2 of the theoretical AR(1) of infinite length, then the transient regime is completely eliminated and one obtains an AR(1) process of finite length with the variance and autocovariance (6) of the infinite AR(1) process [12].

In the following we show that AR(1) processes can be used to generate diffusion processes with memory. This can be achieved by replacing in the algorithm for the generation of the discrete time diffusion processes the Gaussian noise by the AR(1) noise $\{X_n\}$ of mean zero and constant variance, defined by the relation (5) with $|\phi| < 1$. The process $\{Y_n\}$ starting from $Y_1 = 0$ is then generated by

$$Y_n = Y_{n-1} + X_n = \sum_{s=1}^n X_s. \quad (7)$$

From (7) it follows that $E\{Y_n\} = 0$, and using (2) and the stationarity of the AR(1) noise $\{X_n\}$ one finds that the variance is given by a Taylor formula,

$$\begin{aligned} \sigma_{Y_n}^2 &= E\{Y_n^2\} = n\sigma^2 + 2 \sum_{r=1}^{n-1} \sum_{s=1}^{n-r} E\{X_s X_{s+r}\} = n\sigma^2 + 2 \sum_{r=1}^{n-1} \sum_{s=1}^{n-r} \gamma(r) \\ &= n\sigma^2 + 2 \sum_{r=1}^{n-1} (n-r) \gamma(r). \end{aligned}$$

According to Einstein's relation (1), the diffusion coefficient is

$$D = \lim_{n \rightarrow \infty} \frac{\sigma_{Y_n}^2}{2n\tau} = \frac{\sigma^2}{2\tau} + \frac{1}{\tau} \lim_{n \rightarrow \infty} \sum_{r=1}^{n-1} \left(1 - \frac{r}{n}\right) \gamma(r) = \frac{\sigma^2}{2\tau} + \frac{1}{\tau} \lim_{n \rightarrow \infty} \sum_{r=1}^{n-1} \gamma(r),$$

and can be effectively computed, for $|\phi| < 1$, by using the explicit form of the autocovariance (6),

$$D = \frac{\sigma^2}{2\tau} + \frac{\sigma^2}{\tau} \lim_{n \rightarrow \infty} \phi \frac{1 - \phi^{n-1}}{1 - \phi} = \frac{\sigma^2}{2\tau} + \frac{\sigma^2}{\tau} \frac{\phi}{1 - \phi}.$$

Denoting by $D_0 = \sigma^2/(2\tau)$ the diffusion coefficient of the process generated by a white noise of variance σ^2 , the diffusion coefficient of the process generated by a general AR(1) noise with $|\phi| < 1$ takes the explicit form

$$D = D_0 \left(1 + 2 \frac{\phi}{1 - \phi}\right). \quad (8)$$

Since the finite limit (8) exists, the process (7) generated by a AR(1) noise is a diffusion process.

To test the path decomposition method against exact results given by (8), we consider a variance $\sigma^2 = 1 \text{ m}^2$ and a time step $\tau = 0.5 \text{ s}$, so that the diffusion

coefficient for the case $\phi = 0$ takes on the value $D_0 = \sigma^2/(2\tau) = 1 \text{ m}^2/\text{s}$. Using $N = 50,000$ paths with fixed length $S = 100$ we computed the correlations $\rho(r)$ entering the path decomposition (3) for diffusion processes generated by AR(1) noises with different values of the parameter ϕ .

Figure 1 shows the progress of partial sums in (3), $D(r)$, for increasing r . One can see that the path decomposition (3) converges, after a transient regime, to the exact coefficients (8), represented by straight lines in left panel of Figure 1. The right panel Figure 1 shows that the super-diffusive transient regime is expanded as ϕ tends to 1 when, according to (8), D tends to infinity.

The computation time, of about 4 seconds in all cases presented in Figure 1, is still of the order of the CPU time required by direct estimations of diffusion coefficients with the Einstein's relation (2). However, the path decomposition method demonstrates its advantages when applied to diffusion processes in random velocity fields, as shown by the example presented in the next section.

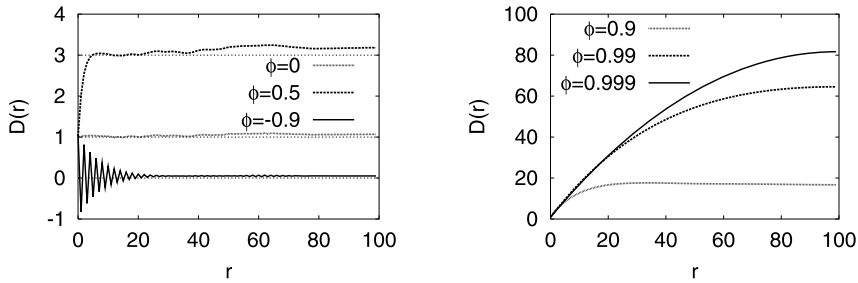


Fig. 1 Diffusion coefficients for processes generated by AR(1) noises with $\phi = 0, 0.5, -0.9$ (left), and with $\phi = 0.9, 0.99$, and 0.999 (right).

5 Estimated Upscaled Coefficients for Diffusion in Random Velocity Fields

We consider a two dimensional process of diffusion in random velocity fields, used to model a typical situation of contaminant transport in groundwater systems, consisting of a superposition of a Wiener process with diffusion coefficient $D_0 = 0.01 \text{ m}^2/\text{day}$ and a random velocity field of mean $U = 1 \text{ m/day}$, exponentially correlated, with finite correlation range $\lambda = 1 \text{ m}$ and variance of the order of 0.01 m/day [4, 5, 6]. The diffusion processes were simulated by simultaneously tracking 10^{10} computational particles in every velocity realization with the “global random walk”

algorithm [11]. The simulations and the Monte Carlo estimations of the upscaled diffusion coefficients are described in detail in [4].

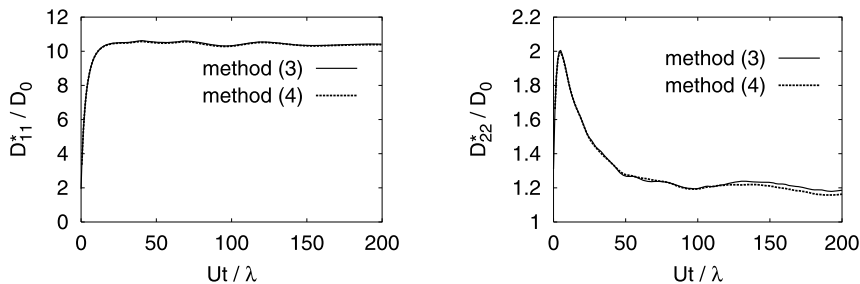


Fig. 2 Longitudinal (left) and transverse (right) upscaled diffusion coefficients estimated by the methods (3) and (4).

To estimate upscaled diffusion coefficients we considered a time step $\tau = 0.5$ day and $N = 2,000,000$ paths of fixed length $S = 400$, on a single trajectory of the process, which required a CPU time of about 65 minutes. Figure 2 shows the partial sums in (3) and (4) as functions of dimensionless time Ut/λ . The longitudinal direction is that of the mean velocity. The upscaled coefficients computed by the method (4) are practically identical with those given by (3). This result supplies a numerical indication that the Lagrangian velocity inherits the ergodic properties of the Eulerian velocity field.

For comparison with the Monte Carlo estimations presented in [4, 6], we used a reduced number of paths, $N = 200,000$, and we estimated mean values and standard deviations of the upscaled coefficients by averaging over 100 independent estimations (3) performed for different realizations of the velocity field. Figure 3 shows that the path decomposition method yields estimations of the upscaled coefficients close to the Monte Carlo results in the limits of one standard deviation.

The Monte Carlo convergence with the number of velocity realizations R of a statistical estimate f can be assessed with an estimation error defined, for given increment ΔR of the number of realizations, by

$$E(f(R)) = |f(R + \Delta R) - f(R)|. \quad (9)$$

Figure 4 shows results on the convergence of the mean coefficients $\langle D_{ll}^* \rangle$ estimated by the Monte Carlo simulations presented in Figure 3. For longitudinal quantities ($l = 1$) the errors (9) were computed at dimensionless times of 1 and 20 (where the ensemble averages were most sensitive to R) while for transverse ones ($l = 2$) the dimensionless times were 1 and 10. The number of realizations was gradually increased by $\Delta R = 32$ in both cases. For $R \geq 1000$ realizations E falls under the desired accuracy threshold of $D_0/2$. The errors (9) for the path decomposition method

(3), computed at the dimensionless time of 200 with an increment of the number of realizations of $\Delta R = 5$, are presented in Figure 5. The results show that the oscillations of both mean values $\langle D_{ll}^* \rangle$ and standard deviations $SD(D_{ll}^*)$, $l = 1, 2$ are already smaller than D_0 for about 10 realizations.

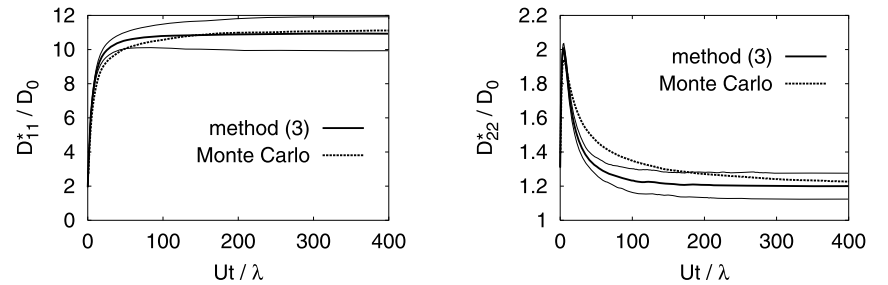


Fig. 3 Longitudinal (left) and transverse (right) upscaled diffusion coefficients estimated by method (3) and by Monte Carlo simulations. The thick full lines represent the arithmetic average $\langle D_{ll}^* \rangle$ of 100 estimations (3) and the thin full lines represent $\langle D_{ll}^* \rangle \pm SD(D_{ll}^*)$.

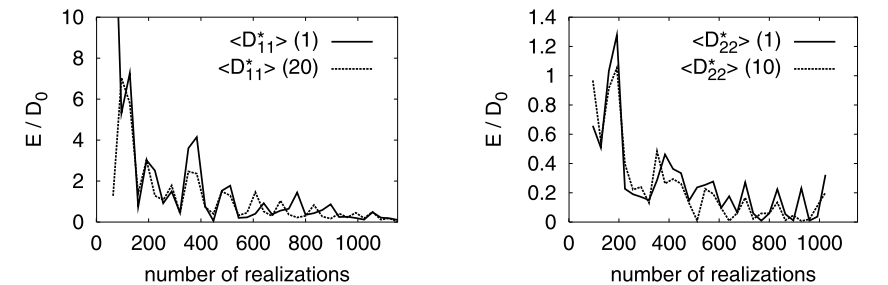


Fig. 4 Convergence of the mean longitudinal (left) and transverse (right) upscaled diffusion coefficients for the Monte Carlo simulations.

A single estimation (3) of the upscaled diffusion coefficients lasted about 6 minutes, so that the total computation time for the statistical estimations presented in Figure 3 was of about 10 CPU hours. Thus the estimates with the decomposition (3) were thousand times faster than those obtained with the Monte Carlo approach (1024 realizations and about 12 CPU hours / realization, on the same computing platform). Moreover, since quite good estimates of the upscaled coefficients can

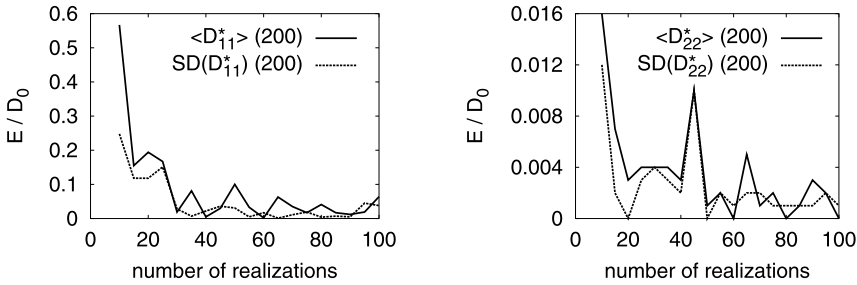


Fig. 5 Convergence of the mean and standard deviation of the longitudinal (left) and transverse (right) upscaled diffusion coefficients estimated by the path decomposition method (3).

be obtained by using a single trajectory of the process (see Figure 3), the effective speed up of the computations is of about four orders of magnitude.

6 Conclusions

The path decomposition method proposed in this paper avoids the cumbersome tasks of taking the long time hydrodynamic limit and of averaging over large statistical ensembles, like in usual Monte Carlo approaches. Instead, the diffusion coefficients are computed as contributions of correlations of the process increments on paths of finite lengths, sampled on a single trajectory.

The advantage of this method is obvious for processes with memory, consisting of diffusion in random velocity fields. As compared with the Monte Carlo approach, the speed up of the computations can be as large as four orders of magnitude. The comparison presented in Figure 3 shows that the path decomposition method yields fairly good estimations of the upscaled coefficients.

Since the method allows fast estimation of the diffusion coefficients, it can be used to calibrate more complex simulations for large scale transport problems. A promising field of applications is the Lagrangian approach for turbulent dispersion, where the Lagrangian velocity samples are readily available as solutions of Itô-type equations [2].

The decomposition of the upscaled diffusion coefficients in terms of velocity correlations is another feature, which, as illustrated by the results presented in Figure 2, is useful in investigations on relationships between ergodic properties of the Eulerian and Lagrangian velocity fields.

Acknowledgements The work presented in this paper was supported by the Deutsche Forschungsgemeinschaft grant SU 415/1-2 and the Romanian Ministry of Education and Research grant 2-CEX06-11-96.

References

1. Harayama, T., Klages, R., Gaspard, P.: Deterministic diffusion in flower shape billiards. *Phys. Rev. E.*, **66**, 02611 (2002)
2. Kurbanmuradov, O., Sabelfeld, K.: Lagrangian stochastic models for turbulent dispersion in the atmospheric boundary layer. *Boundary-Layer Meteorology*, **97**, 191-218, (2000)
3. Mandelbrot, B. B., Wallis, J. R.: Noah, Joseph, and operational hydrology. *Water Resour. Res.*, **4**(5), 909-918 (1968)
4. Suciú, N., Vamoş, C., Vanderborght, J., Hardelauf, H., Vereecken, H.: Numerical investigations on ergodicity of solute transport in heterogeneous aquifers. *Water Resour. Res.*, **42**, W04409, doi:[10.1029/2005WR004546](https://doi.org/10.1029/2005WR004546) (2006)
5. Suciú, N., Vamoş, C., Eberhard, J.: Evaluation of the first-order approximations for transport in heterogeneous media. *Water Resour. Res.*, **42**, W11504, doi:[10.1029/2005WR004714](https://doi.org/10.1029/2005WR004714) (2006)
6. Suciú, N., Vamoş, C., Vereecken, H., Sabelfeld, K., Knabner, P.: Memory effects induced by dependence on initial conditions and ergodicity of transport in heterogeneous media. *Water Resour. Res.*, **44**, W08501, doi:[10.1029/2007WR006740](https://doi.org/10.1029/2007WR006740) (2008)
7. Suciú, N., Vamoş, C., Vereecken, H., Sabelfeld, K., Knabner, P.: Itô equation model for dispersion of solutes in heterogeneous media. *Revue D'Analyse Numérique et de Théorie D'Approximation*, **37**(2), 221-238, (2008)
8. Suciú, N., Vamoş, C., Sabelfeld, K.: Ergodic simulations of diffusion in random velocity fields. In: Keller, A., Heinrich, S., and Niederreiter, H. (eds) *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pp. 659-668. Springer, Heidelberg (2008)
9. Vamoş, C.: C++ codes 'descin.cpp' and 'descre.cpp'. Unpublished manuscript (2001)
10. Vamoş, C.: Automatic algorithm for monotone trend removal. *Phys. Rev. E* **75**, 036705, doi:[10.1103/PhysRevE.75.036705](https://doi.org/10.1103/PhysRevE.75.036705) (2007)
11. Vamoş, C., Suciú, N., Vereecken, H.: Generalized random walk algorithm for the numerical modeling of complex diffusion processes. *J. Comput. Phys.*, **186**(2), 527-544 (2003)
12. Vamoş, C., Şoltuz, Ş. M., Crăciun, M.: Order 1 autoregressive processes of finite length. E-print [arXiv: 079.2963v1](https://arxiv.org/abs/079.2963v1) [physics.data-an], 19 Sep 2007 (2007)
13. Vamoş, C., Crăciun, M.: Serial correlation of detrended time series. *Phys. Rev. E* **78**, 036707 (2008)
14. Yaglom, A. M.: *Correlation Theory of Stationary and Related Random Functions, Volume I: Basic Results*, Springer, New York (1987)
15. Ying, S. C., Vattulainen, J., Merikoski, J., Hjelt, T., Als-Nissila, T.: Memory expansion for diffusion coefficients. *Phys. Rev. B*, **58**(4), 2170-2178 (1998)