Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Two-particle models for the estimation of the mean and standard deviation of concentrations in coastal waters

Daria Spivakovskaya[1], Arnold Heemink,[1] John Schoenmakers[2]

submitted: 6th January 2006

[1] Delft University of Technology,
Department of Mathematical Physics,
Delft Institute of Applied Mathematics (DIAM),
Mekelweg 4, 2628 CD Delft, The Netherlands,
E-Mail: d.spivakovskaya@ewi.tudelft.nl,
A.W.Heemink@ewi.tudelft.nl

[2] Weierstrass Institute for Applied Analysis and Stochastics,
Stochastic Algorithms and Nonparametric Statistics,
Mohrenstraße 39, D – 10117 Berlin, Germany,
E-Mail: schoenma@wias-berlin.de

No. 1088
Berlin 2006

2000 *Mathematics Subject Classification.* 62G07, 60H10, 65C05.

Key words and phrases. Advection-diffusion equation, Lagrangian models, two-particle model.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In this paper we study the mean and standard deviation of concentrations using random walk models. Two-particle models that takes into account the space correlation of the turbulence are introduced and some properties of the distribution of the particle concentration are studied. In order to reduce the CPU time of the calculation a new estimator based on reverse time diffusion is applied. This estimator has been introduced recently by [18]. Some numerical aspects of the implementation are discussed for relative simple test problems and finally a realistic application to predict the spreading of the pollutant in the Dutch coastal zone is described.

1 Introduction

Ship accidents have caused recently a number of serious ecological catastrophes in coastal zones. In order to reduce the possible damage for the environment it is very important to accurately predict the possible concentrations of the pollutant. In many situations it is enough to know only the ensemble mean concentration, especially for long term prediction problems. For these kind of problems one can adopt an Eulerian approach and study the transport of the pollutant with the help of a numerical approximation of the advection-diffusion equation. However, serious numerical difficulties may occur with positiveness and mass conservation, especially if the initial concentration is a delta-like function ([27], [31], [9]).

Another approach is random walk modeling ([11]). In this approach the advection-diffusion equation is interpreted as a Fokker-Planck equation and as a result it is possible to derive an Îto stochastic differential equation for the behavior the individual particles of the pollutant. By numerical integration of this stochastic differential equation, the behavior of many different particles can be simulated and the transport of the pollutant can be described ([11], [7], [23], [29]). One of the advantages of random particle models is that it is a natural way to study not only the mean ensemble concentration, but also higher order moments of the concentration. For instance, the standard deviation of the concentration is connected with the statistics of the trajectories of pairs of particles. The idea of using two-particle simulation to obtain the standard deviation of the concentration was first formulated by [5]. More recently a number of papers in which Lagrangian models have been applied to investigate the standard deviation of the concentration have been published ([4], [13], [15], [30]).

Using particle models the concentration distribution of the particles is equivalent with the probability density function of the particles. Thus the simulation of the positions of many particles is equivalent to find the probability density function. This Monte Carlo approach is conceptual very easy and can be applied for many different types of (highly nonlinear) problems. However, Monte Carlo techniques do consume a large CPU time, while the accuracy of the estimates improves very slowly for larger sample size. Especially in case we only want to determine the particle density at a few critical locations, most of the realizations of the particle tracks do hardly contribute to the final results. The efficiency of Monte Carlo methods can be improved by variance reduction, [14], [19], [22], [6], [17]. The most efficient methods of variance reduction require the analytical solution of the backward Kolmogorov equation. In many applications, however, a numerical solution is required. For high dimensional systems this may become very time consuming [22].

Recently [18] introduced the concept of reverse time diffusion. The classical Monte Carlo estimator is based on forward realizations of the original stochastic model. [18] derived a reverse system from the original model and showed that the classical Monte Carlo estimator can also be based on realizations of this reverse system. For many problems it is more efficient to use realizations of the reverse system instead of the original forward model. The most efficient implementation is, however, obtained if the forward realizations and the reverse system realizations are combined. This is called the forward reverse estimator.

In this paper the multiple particle model is formulated and the forward reverse estimator is applied for the estimation of the mean ensemble concentration and the standard deviation of the concentration of the pollutant at a number of given critical locations. In section 2 we describe the particle model that is used for the simulation of the pollutant. The multiple particle models is introduced in section 2.2. In section 3 we consider the different estimators for the probability density function. In section 4 we discuss some properties of the distribution of the concentration for a test problem. Finally in section 5 we describe an application of the two-particle model and the forward reverse estimator to calculate the mean and standard deviation of the concentration of a pollutant in the Dutch coastal waters.

2 Random walk models for modeling dispersion in shallow water

2.1 One-particle model

The ensemble mean concentration of the pollutant released at time t at the position x can be obtained from the advection-diffusion equation ([10], [25])

$$
\begin{aligned}
&\frac{\partial C}{\partial s} + \frac{\partial (u_1 C)}{\partial y_1} + \frac{\partial (u_2 C)}{\partial y_2} + \frac{\partial (u_3 C)}{\partial y_3} = \\
&\frac{\partial}{\partial y_1}\left(D_1 \frac{\partial C}{\partial y_1}\right) + \frac{\partial}{\partial y_2}\left(D_2 \frac{\partial C}{\partial y_2}\right) + \frac{\partial}{\partial y_3}\left(D_3 \frac{\partial C}{\partial y_3}\right) \\
&C(t, \boldsymbol{y}) = \delta(\boldsymbol{x} - \boldsymbol{y})
\end{aligned}
\tag{1}
$$

where $C(s, \boldsymbol{y})$ is the mean ensemble concentration in point $\boldsymbol{y}$ and at time s, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$, $\boldsymbol{u} = (u_1, u_2, u_3)^T$ is the velocity vector, D_1, D_2 and D_3 represent the dispersion in the y_1, y_2 and y_3 directions respectively. We can also use the vertically-integrated advection-diffusion equation ([25])

$$
\begin{aligned}
&\frac{\partial (CH)}{\partial s} + \frac{\partial (u_1 HC)}{\partial y_1} + \frac{\partial (u_2 HC)}{\partial y_2} = \\
&\frac{\partial}{\partial y_1}\left(HD \frac{\partial C}{\partial y_1}\right) + \frac{\partial}{\partial y_2}\left(HD \frac{\partial C}{\partial y_2}\right), \\
&C(t, \boldsymbol{y}) = \delta(\boldsymbol{x} - \boldsymbol{y}),
\end{aligned}
\tag{2}
$$

where H is water depth, D is now the horizontal dispersion coefficient, and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$. One way to solve the equation (2) is to use particle models ([11]). By introducing a new function $p = HC$, the equation (2) can be rewritten in the form:

$$
\begin{aligned}
&\frac{\partial p}{\partial s} = -\frac{\partial}{\partial y_1}\left(\left(u_1 + \tfrac{\partial D}{\partial y_1} + \tfrac{D}{H}\tfrac{\partial H}{\partial y_1}\right) p\right) - \\
&\frac{\partial}{\partial y_2}\left(\left(u_2 + \tfrac{\partial D}{\partial y_2} + \tfrac{D}{H}\tfrac{\partial H}{\partial y_2}\right) p\right) + \frac{\partial^2 (Dp)}{\partial (y_1)^2} + \frac{\partial^2 (Dp)}{\partial (y_2)^2}, \\
&p(t, \boldsymbol{x}, t, \boldsymbol{y}) = \delta(\boldsymbol{x} - \boldsymbol{y}).
\end{aligned}
\tag{3}
$$

The equation (3) can be interpreted as a Fokker-Planck equation for the transition density function $p(t, \boldsymbol{x}, s, \boldsymbol{y})$ of some underlying stochastic process $\boldsymbol{X}(s)$. This stochastic process is the solution of the following system of Îto stochastic differential equations

$$
\begin{aligned}
&d\boldsymbol{X}(s) = \left(\boldsymbol{u} + \frac{\partial D}{\partial \boldsymbol{y}} + \frac{D}{H}\frac{\partial H}{\partial \boldsymbol{y}}\right) ds + \sqrt{2D}\boldsymbol{I}_2 d\boldsymbol{B}(s), \\
&\boldsymbol{X}(t) = \boldsymbol{x},
\end{aligned}
\tag{4}
$$

where $\boldsymbol{X} = (X_1, X_2)^T$ is 2-dimensional stochastic process, $\boldsymbol{B} = (B_1, B_2)^T$ is 2-dimensional Brownian motion, $\frac{\partial}{\partial \boldsymbol{y}} = \left(\frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}\right)^T$ is the spatial gradient and $\boldsymbol{I}_2$ is 2×2 identity matrix. As a result estimating the ensemble mean concentration of the pollutant now becomes the problem of estimating the probability density function $p(t, \boldsymbol{x}, s, \boldsymbol{y})$ of the stochastic process $\boldsymbol{X}(s)$.

2.2 Multiple particle model

In the previous section we have introduced one-particle models that do not take into account the spatial correlation of the particle behavior and therefore allows only for estimating the ensemble mean concentration. However, the turbulent behavior of particles is correlated in space. As a result the actual concentration at certain locations may be much higher or lower than the ensemble mean concentration. For instance, the ensemble mean concentration may be an average of a large number of zeros (realizations when the cloud of pollutant do not reach the location) and a few very large values. This type of averaging may be meaningless, because the few high concentrations may kill the organisms in a certain area and the large number of zeros can not bring them to life again.

So if we want to describe the dispersion process in more detailed we must use multiple particle models ([2]). In a K-particle model, K particles are released at the same time and their behavior is correlated with each other. The correlation between any two-particles is assumed to depend only on the distance between them. Over large distances particles behave almost independent, but when the distance is close to zero the behavior of the particles is highly correlated. A problem is how to separate the particles when the distance is close to zero. According to [30] we should include the effect of the molecular diffusion. In one-particle models the effect of molecular diffusion is negligible compared with the turbulent diffusion, but in multiple particle models the molecular diffusion plays a significant role. Moreover, if the particles coincide, they can only separate again by the molecular diffusion. However, once the particles have separated up to a certain distance d any further separation is caused mainly by the turbulent diffusion. A discussion about this distance d can be found in [30]. As a result, we propose to write the equation (4) in the form

$$\begin{aligned} d\boldsymbol{X}(s) &= \left(\boldsymbol{u} + \frac{\partial D}{\partial \boldsymbol{y}} + \frac{D}{H}\frac{\partial H}{\partial \boldsymbol{y}}\right) ds + \\ &\sqrt{2D} I_2 \left(\sqrt{1-\beta^2} d\boldsymbol{B}(s) + \beta d\boldsymbol{W}(s)\right), \\ \boldsymbol{X}(t) &= \boldsymbol{x}, \end{aligned} \tag{5}$$

where $0 \leq \beta \leq 1$ is constant, $\boldsymbol{B}$ is a standard Brownian motion as in the one-particle model, and $\boldsymbol{W}$ is a 2-dimensional correlated Brownian motion independent of $\boldsymbol{B}$. In (5), $\sqrt{1-\beta^2}\boldsymbol{B}(s)$ represents the diffusion caused by the molecular diffusion and the small scale turbulence and $\beta \boldsymbol{W}(s)$ represents large scale turbulence.

Now we are ready to introduce the K-particle model. The behavior of K particles can be described by the following $2K$-dimensional systems of the stochastic differential equations

$$\begin{aligned} d\boldsymbol{X}^{[i]}(s) &= \left(\boldsymbol{u}^{[i]} + \frac{\partial D}{\partial \boldsymbol{y}^{[i]}} + \frac{D}{H^{[i]}}\frac{\partial H^{[i]}}{\partial \boldsymbol{y}^{[i]}}\right) ds + \\ &\sqrt{2D} \boldsymbol{I}_2 \left(\sqrt{1-\beta^2} d\boldsymbol{B}^{[i]}(s) + \beta d\boldsymbol{W}^{[i]}(s)\right), \\ \boldsymbol{X}^{[i]}(t) &= \boldsymbol{x}. \end{aligned} \tag{6}$$

With the superscript $[i]$, $i = 1, \dots, K$ we indicate the functions that depend on the ith particle $\boldsymbol{X}^{[i]}$

$$\begin{aligned} \boldsymbol{u}^{[i]} &= \boldsymbol{u}(s, \boldsymbol{X}^{[i]}) \\ H^{[i]} &= H(\boldsymbol{X}^{[i]}) \end{aligned}$$

The Brownian motion processes $\boldsymbol{B}^{[i]}(s)$ are mutually independent, while any pair processes $\boldsymbol{W}^{[i]}$ and $\boldsymbol{W}^{[j]}$, $(i \neq j)$ is correlated with covariance matrix

$$E\left(\boldsymbol{W}^{[i]}(ds)\boldsymbol{W}^{[j]}(ds)^T\right) = f(r_{ij})\boldsymbol{I}_2 ds, \tag{7}$$

where $f(r)$ is a covariance function that depends on the distance r between particles $\boldsymbol{X}^{[i]}$ and $\boldsymbol{X}^{[j]}$,

$$r_{ij} = \|\boldsymbol{X}^{[i]} - \boldsymbol{X}^{[j]}\|$$

The covariance function is related to the spectrum of the turbulence. We assume that the correlation function f satisfies several conditions. First of all, we assume that this function is sufficiently smooth, at least its second derivative is continuous and bounded. We further assume that

1. $f(0) = 1$
2. $f(r) \downarrow 0$, $r \to \infty$
3. The matrix $[f(\|\boldsymbol{y}^i - \boldsymbol{y}^j\|)]_{1 \leq i \leq q, 1 \leq j \leq q}$ is positive definite for any choice of points $\boldsymbol{y}^1, \dots, \boldsymbol{y}^q$ in $\mathbb{R}^2$

For example, we can use the following function ([8])

$$f(r) = \exp(-\alpha r^2), \quad \alpha > 0 \tag{8}$$

2.3 The standard deviation of the concentration

Multiple particle models may be used to find the distribution of the concentration. However this approach is very time consuming. Two-particle models can be used to find mean and standard deviation of this concentration distribution. The behavior of a pair of particles is described by equation (6) for $K = 2$. The probability density function $p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}^{[1]}, \boldsymbol{y}^{[2]})$ (or the joint probability function of the stochastic processes $\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$) gives us the information about the standard deviation of the concentration.

Let us consider the neighborhood of the point $\boldsymbol{y}$

$$O_\epsilon(\boldsymbol{y}) = \{(\tilde{y}_1, \tilde{y}_2) : |\tilde{y}_1 - y_1| < \epsilon, |\tilde{y}_2 - y_2| < \epsilon\}.$$

The probability that two particles $\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$ will occur in $O_\epsilon(\boldsymbol{y})$ is given by $P(t, \boldsymbol{x}, \boldsymbol{x}, T, O_\epsilon(\boldsymbol{y}), O_\epsilon(\boldsymbol{y}))$. The value

$$\begin{aligned} &\frac{1}{\mu^2(O_\epsilon(\boldsymbol{y}))} P(t, \boldsymbol{x}, \boldsymbol{x}, T, O_\epsilon(\boldsymbol{y}), O_\epsilon(\boldsymbol{y})) = \\ &\frac{1}{\epsilon^4} P(t, \boldsymbol{x}, \boldsymbol{x}, T, O_\epsilon(\boldsymbol{y}), O_\epsilon(\boldsymbol{y})) \end{aligned} \tag{9}$$

is the concentration of the pairs of particles $(\boldsymbol{X}^{[1]}, \boldsymbol{X}^{[2]})$ in $O_\epsilon(\boldsymbol{y})$. Here $\mu(O_\epsilon(\boldsymbol{y}))$ is the area of the neighborhood $O_\epsilon(\boldsymbol{y})$. When the particles ($\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$) are independent this concentration can be found as

$$\begin{aligned}&\left(\frac{1}{\mu(O_\epsilon(\boldsymbol{y}))} P(t, \boldsymbol{x}, T, O_\epsilon(\boldsymbol{y}))\right)^2 = \\ &\frac{1}{\epsilon^4} P^2(t, \boldsymbol{x}, T, O_\epsilon(\boldsymbol{y})).\end{aligned} \tag{10}$$

The difference between the values (9) and (10)

$$\begin{aligned}&\overline{c^2}(T, O_\epsilon(\boldsymbol{y})) = \\ &\frac{1}{\epsilon^4} \left| P(t, \boldsymbol{x}, \boldsymbol{x}, T, O_\epsilon(\boldsymbol{y}), O_\epsilon(\boldsymbol{y})) - P^2(t, \boldsymbol{x}, T, O_\epsilon(\boldsymbol{y}))\right|\end{aligned} \tag{11}$$

gives us the information about the fluctuation of the concentration in $O_\epsilon(\boldsymbol{y})$. Proceed to the limit when $\epsilon \to 0$ we will receive the value $\overline{c^2}(T, \boldsymbol{y})$ at the point $\boldsymbol{y}$

$$\overline{c^2}(T, \boldsymbol{y}) = |p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y}) - p^2(t, \boldsymbol{x}, T, \boldsymbol{y})| \tag{12}$$

In [5] it is shown that the fluctuation defined by (12) leads to a slightly lower concentration fluctuation than the usual definition and can be used as a measure of the fluctuation amplitude. Taking into account that we use the depth averaged model and that the concentration of the pollutant is connected with the density probability function as $C(T, \boldsymbol{y}) = p(t, \boldsymbol{x}, Y, \boldsymbol{y})/H(\boldsymbol{y})$ we define the standard deviation $Dev(T, \boldsymbol{y})$ at the point $\boldsymbol{y}$ at time T with the help of the following equation

$$\begin{aligned}&Dev(\boldsymbol{y}, T) = \frac{1}{H(\boldsymbol{y})} \sqrt{\overline{c^2}(T, \boldsymbol{y})} = \\ &\sqrt{\left| \frac{p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})}{H^2(\boldsymbol{y})} - C^2(\boldsymbol{y}, T)\right|}\end{aligned} \tag{13}$$

In section (4.3) we study some properties of the concentration distribution for a relatively simple example.

3 Density estimators for the stochastic differential equations

3.1 Numerical integrating of the stochastic differential equations

Let us consider the stochastic differential equation defined in the Ito sense:

$$\begin{aligned}&d\boldsymbol{X}(s) = \boldsymbol{a}(s, \boldsymbol{X}) ds + \boldsymbol{\sigma}(s, \boldsymbol{X}) d\boldsymbol{B}(s), \quad t \le s \\ &\boldsymbol{X}(t) = \boldsymbol{x}\end{aligned} \tag{14}$$

where $\boldsymbol{X} = (X_1, \ldots, X_d)^T$. $\boldsymbol{a} = (a_1, \ldots, a_d)^T$ are d-dimensional vectors, $\boldsymbol{B} = (B_1, \ldots, B_m)^T$ ($m \geq d$ is a m-dimensional Wiener process and $\boldsymbol{\sigma} = \{\sigma_{ij}\}$ is a $d \times m$ matrix. We assume that the $d \times d$ matrix $\boldsymbol{b} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T$ is of full rank for every $(s, \boldsymbol{x})$, $s \in [t, T]$, $\boldsymbol{x} \in \mathbb{R}^d$. We also assume that the functions $a_i(s, \boldsymbol{x})$ and $\sigma_{ij}(s, \boldsymbol{x})$ and their first derivatives are continuous and bounded. This implies existence and uniqueness of the solution $\boldsymbol{X}_{t,\boldsymbol{x}}(s) \in \mathbb{R}^d$, $\boldsymbol{X}_{t,\boldsymbol{x}}(t) = \boldsymbol{x}$ of (14), smoothness of the transition density $p(t, \boldsymbol{x}, s, \boldsymbol{y})$ of the Markov process $\boldsymbol{X}$ and existence of all the moments $p(\cdot, \cdot, \cdot, \boldsymbol{y})$ ([1], [21], [12]). In general equation (14) can not be solved analytically, so we must use a numerical method to find the approximation $\overline{\boldsymbol{X}}(T)$ of the realization of the stochastic process $\boldsymbol{X}(T)$ at time T ($t < s < T$). One of the more frequently used numerical methods is the Euler method ([14], [19])

$$\begin{aligned} &\overline{\boldsymbol{X}}^h_{k+1} = \overline{\boldsymbol{X}}^h_k + \boldsymbol{a}(s, \overline{\boldsymbol{X}}^h_k)h + \boldsymbol{\sigma}(s, \overline{\boldsymbol{X}}^h_k)\Delta \boldsymbol{B}_k, \\ &\overline{\boldsymbol{X}}^h_0 = \boldsymbol{x}, \end{aligned} \tag{15}$$

where $k = 0, 1, \ldots, L-1$, $\overline{\boldsymbol{X}}^h_k = \overline{\boldsymbol{X}}^h_{t,\boldsymbol{x}}(s_k)$ is the numerical approximation of the position $\boldsymbol{X}^h_{t,\boldsymbol{x}}(s_k)$, $s_k = t+kh$, $h = (T-t)/L$ is time step of the numerical integration, $\Delta \boldsymbol{B}_k$ are mutually independent Gaussian variables with zero mean and covariance matrix $h\boldsymbol{I}_m$ ($\boldsymbol{I}_m$ is the $m \times m$ identity matrix. In many cases we do not need to know the realizations of the stochastic process $\boldsymbol{X}$, but the value of the functional

$$U = Eg(\boldsymbol{X}(s)). \tag{16}$$

To find this functional one can use the extrapolation method which is based on the Euler method. First one can apply the Euler approximation $\overline{\boldsymbol{X}}^h$ generated by (15) with time step h to simulate the functional $\overline{U}^h = Eg(\overline{\boldsymbol{X}}^h(s))$. One can then simulate the functional $\overline{U}^{2h} = Eg(\overline{\boldsymbol{X}}^{2h}(s))$ for the double time step $2h$, and finally the two results are combined to yield the approximation of (16)

$$\overline{U} = 2\overline{U}^h - \overline{U}^{2h}.$$

This method was proposed by [28].

3.2 The transition density estimator based on forward diffusion

In this chapter we consider the transition density estimator that is based on the forward system (14).

Suppose it is required to calculate the transition density function $p(t, \boldsymbol{x}, T, \boldsymbol{y})$ of the stochastic process $\boldsymbol{X}(s)$. This function can be found from the Fokker-Planck (or forward Kolmogorov) equation

$$\begin{aligned} &\frac{\partial p}{\partial s} = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial y_i \partial y_j}(b_{ij}p) - \sum_{i=1}^{d} \frac{\partial}{\partial y_i}(a_i p), \\ &p(t, \boldsymbol{x}, t, \boldsymbol{y}) = \delta(\boldsymbol{x} - \boldsymbol{y}) \end{aligned} \tag{17}$$

or from the backward Kolmogorov equation

$$\frac{\partial p}{\partial t} = -\frac{1}{2}\sum_{i,j=1}^{d} b_{ij}\frac{\partial^2 p}{\partial x_i \partial x_j} - \sum_{i=1}^{d} a_i \frac{\partial p}{\partial x_i}, \qquad (18)$$
$$p(T, \boldsymbol{x}, T, \boldsymbol{y}) = \delta(\boldsymbol{x} - \boldsymbol{y})$$

Another approach to estimate the density function $p(t, \boldsymbol{x}, T, \boldsymbol{y})$ is to use standard methods for non-parametric statistics ([24], [32]). For instance, a kernel density estimator with kernel function K and bandwidth is given by

$$\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y}) = \frac{1}{N\delta^d}\sum_{n=1}^{N} K\left(\frac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}(T) - \boldsymbol{y}}{\delta}\right), \qquad (19)$$

where $\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}(T)$ are samples from the stochastic process $\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}$ at a given point T.

The estimation loss $|\widehat{p} - p|$ of the estimator (19) can be split into an error due to numerical approximation $\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}$ of the stochastic process $\boldsymbol{X}_{t,\boldsymbol{x}}$ and an error due to the kernel estimator. The loss of the first kind can be reduced by selecting a higher order scheme of numerical integration and choosing a small step of integration. The second error depends on the choice of the sample size, the kernel function $K(\cdot)$ and bandwidth δ. Popular choices of kernel functions are the Gaussian function

$$K(\boldsymbol{x}) = (2\pi)^{-d/2} exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) \qquad (20)$$

and the Epanechnikov symmetric multivariate kernel

$$K(\boldsymbol{x}) = \frac{1}{2}\nu_d^{-1}(d+2)(1 - \boldsymbol{x}^T\boldsymbol{x})1_{\boldsymbol{x}^T\boldsymbol{x}\leq 1} \qquad (21)$$

where $\nu_d = 2\pi^{d/2}/\{d\Gamma(d/2)\}$ is the volume of the unit d-dimensional sphere. The most important loss, however, is caused by the choice of the bandwidth δ. Even an optimal choice of the bandwidth leads to quite poor estimation quality, in particular for large dimensional d. It is well known ([24], [32]) that if the underlying density is two times continuously differentiable then the optimal bandwidth is of order $\mathcal{O}(N^{-1/(4+d)})$ and the accuracy of the estimator has the order $\mathcal{O}(N^{-2/(4+d)})$. For $d > 2$ this would require a huge sample size for providing a reasonable accuracy of estimation. The simulation can be improved by using the general form of the kernel function

$$\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y}) = \frac{|\boldsymbol{H}|^{-\frac{1}{2}}}{N}\sum_{n} K\left(\boldsymbol{H}^{-1/2}(\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}(T) - \boldsymbol{y})\right), \qquad (22)$$

where $\boldsymbol{H}$ is a symmetric positive definite $d \times d$ matrix called the bandwidth matrix. This matrix can be chosen in the form

$$\boldsymbol{H} = \delta^2 \boldsymbol{S},$$

where $\boldsymbol{S}$ is a sample matrix. This approach allows to take into account the correlation between the components of the stochastic process $\boldsymbol{X}_{t,\boldsymbol{x}}(s)$. However, the order of accuracy remains the same and it makes high dimensional models inefficient for real life application.

3.3 Representations relying on reverse diffusion

In the previous section we described the estimator based on forward system. Another approach is based on the so called reverse time diffusion and has been introduced by [29]. The main idea of the reverse time approach is that the forward Kolmogorov equation (17) can be considered as a backward Kolmogorov equation (18) connected with the stochastic process $(\boldsymbol{Y}, \mathcal{Y})$ in $\mathbb{R}^{d+1}$, hence in a space of one dimension higher.

We first introduce a reverse time variable $\tilde{s} = T + t - s$, $t \leq \tilde{s} \leq T$ and define the functions

$$\begin{aligned} \tilde{a}_i(\tilde{s}, \tilde{\boldsymbol{y}}) &= a_i(T + t - \tilde{s}, \tilde{\boldsymbol{y}}), \\ \tilde{b}_{ij}(\tilde{s}, \tilde{\boldsymbol{y}}) &= b_{ij}(T + t - \tilde{s}, \tilde{\boldsymbol{y}}). \end{aligned} \tag{23}$$

For a stochastic process $\boldsymbol{Y}_{t,\boldsymbol{y}}(\tilde{s}) \in \mathbb{R}^d$ and a scalar process $\mathcal{Y}_{t,\boldsymbol{y}}(\tilde{s})$ we then consider the system of stochastic differential equations

$$\begin{aligned} &d\boldsymbol{Y} = \boldsymbol{\alpha}(\tilde{s}, \boldsymbol{Y})d\tilde{s} + \tilde{\boldsymbol{\sigma}}(\tilde{s}, \boldsymbol{Y})d\tilde{\boldsymbol{B}}(\tilde{s}), \\ &d\mathcal{Y} = c(\tilde{s}, \boldsymbol{Y})\mathcal{Y}d\boldsymbol{s}, \\ &\boldsymbol{Y}(t) = \boldsymbol{y}, \quad \mathcal{Y}(t) = 1 \end{aligned} \tag{24}$$

with $\tilde{\boldsymbol{B}}$ being an m-dimensional standard Wiener process and

$$\begin{aligned} \alpha_i(\tilde{s}, \tilde{\boldsymbol{y}}) &= \sum_{j=1}^{d} \frac{\partial \tilde{b}_{ij}}{\partial \tilde{y}_j} - \tilde{a}_i, \\ c(\tilde{s}, \tilde{\boldsymbol{y}}) &= \frac{1}{2} \sum_{i,1=1}^{d} \frac{\partial^2 \tilde{b}_{ij}}{\partial \tilde{y}_i \partial \tilde{y}_j} - \sum_{i=1}^{d} \frac{\partial \tilde{a}_i}{\partial \tilde{y}_i} \end{aligned} \tag{25}$$

It can be proved (for example, [18], [16]) that the transition density function $p(t, \boldsymbol{x}, T, \boldsymbol{y})$ has the following probabilistic representation

$$p(t, \boldsymbol{x}, T, \boldsymbol{y}) = E\mathcal{Y}_{t,\boldsymbol{y}}(T).$$

As a result, it is possible to construct the density estimator based only on the reverse time system (24)

$$\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y}) = \frac{1}{M\delta^d} \sum_{m=1}^{M} K\left(\frac{\boldsymbol{x} - \overline{\boldsymbol{Y}}_{t,\boldsymbol{y}}^{(m)}}{\delta}\right) \overline{\mathcal{Y}}_{t,\boldsymbol{y}}^{(m)}(T), \tag{26}$$

where $(\overline{\boldsymbol{Y}}_{t,\boldsymbol{y}}^{(m)}, \overline{\mathcal{Y}}_{t,\boldsymbol{y}}^{(m)})$, $m = 1, \ldots, M$ is an independent identically distributed (i.i.d.) sample of numerical solutions of system (24)

3.4 The forward-reverse density estimator

In this section we discuss a new estimator based on both the forward system (14) and the reverse time system (24). This estimator is called the forward-reverse estimator and has been recently introduced by [18]. By taking advantage of the forward

Table 1: Accuracy of the density estimators

Estimator	bandwidth δ	Accuracy
FE/RE	$\mathcal{O}(N^{-\frac{1}{4+d}})$	$\mathcal{O}(N^{-\frac{2}{4+d}})$
FRE, $d < 4$	$\mathcal{O}(N^{-\frac{1}{d}} \log^{\frac{1}{d}} N)$	$\mathcal{O}(N^{-\frac{1}{2}})$
FRE, $d \geq 4$	$\mathcal{O}(N^{-\frac{2}{4+d}})$	$\mathcal{O}(N^{-\frac{4}{4+d}})$

and reverse-time systems via the Chapman-Kolmogorov equation with respect to an intermediate time t^* one can obtain the following estimator

$$\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y}) = \frac{1}{MN\delta^d} \sum_{n=1}^{N} \sum_{m=1}^{M} K \left(\frac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}(t^*) - \overline{\boldsymbol{Y}}_{t^*,\boldsymbol{y}}^{(m)}(T)}{\delta} \right) \overline{\mathcal{Y}}_{t^*,\boldsymbol{y}}^{(m)}(T). \tag{27}$$

In this equation t^* is an arbitrary point from the time interval $[t, T]$. If the $t^* = T$ the forward-reverse estimator collapses to the pure forward estimator (19) and if we take $t^* = t$ we will obtain the reverse estimator (26). In table (1) the accuracy of the forward estimator (FE), the reverse estimator (RE) and forward-reverse estimator (FRE) are shown. Here, we assume that $N = M$ in equation (27) and that we use a second order kernel $K(\cdot)$.

4 Test problem

In this section first we study the pure dispersion process with zero velocity vector $\boldsymbol{u} = 0$, constant water depth $H = 1$m and constant dispersion coefficient $D = 5\text{kg/m}^3$. The domain of interest (see figure 4.1) is scaled with the grid size DX. The constant $\beta = \sqrt{0.9}$ (see the equation(5)) .

4.1 100-particle model

Figure (2) shows an example of a simulation of the dispersion process for a 100-particle model after 1, 3 and 7 days (figures (a), (c), (e)) with the correlation function (8), $\alpha = 5/(DX * DX)$ and a one-particle model with 100 realizations (figures (b), (d), (f)). If we repeat the simulation many times and then average, the center of the cloud of particles will be in the origin for both the multiple and the one-particle models. However, from figures 2(a) and 2(c) it is seen that the center of the individual clouds can diverge from the origin. If we wait long enough the center of each cloud will tend to return the origin (figure 2(e)).

Suppose we want to know the distribution of the concentration in the square $\Omega_0 = [-0.2DX, 0.2DX] \times [-0.2DX, 0.2DX]$ (see figure 4.1) after 2.5 days. We release 100 particles in the origin $\boldsymbol{x} = 0$ and fix the number of particles that occur in this

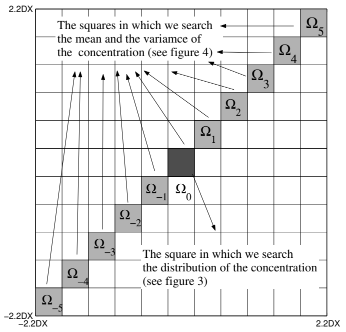

Figure 1: The domain of test problem

square after 2.5 days. Figures 3(a), (b) and (c) show the concentration distribution for different values of parameter α ($\alpha = 1/(DX * DX)$, $\alpha = 5/(DX * DX)$ and $\alpha = 10/(DX * DX)$). In figure 3 (d) all these distributions are shown together and can be compared with each other. It is clear that by increasing the parameter α, the number of observations with ßeroconcentration decreases and therefore the standard deviation of the concentration goes down (see figure 4(b)).

Usually one is only interested in the ensemble mean and the standard deviation of the concentration. The diffusion process in this test case is symmetric in all directions (it means that the concentration at some point depends only on the distance between this point and the origin). Therefore, we study the mean and the standard deviation of the concentration not in the whole domain, but only in the squares Ω_i, $i = -5, \dots, 5$. On figure 4 it is shown that the standard deviation decreases, when the value of α increases.

We are also interested how the distribution of the concentration changes with time. According to the theory (see, for example, [10],[20]) the difference between correlated and independent simulations of the particle spreading becomes less if time increases. In the next experiment we fix the parameter $\alpha = 10/(DX * DX)$, but we change the size of the central area Ω_0 in such way, that in average 5 particles from 100 will occur in this area. We consider 4 time steps

- T=2.5 days
  $\Omega_0 = [-0.52DX, 0.52DX] \times [-0.52DX, 0.52DX]$
- T=5 days
  $\Omega_0 = [-0.74DX, 0.74DX] \times [-0.74DX, 0.74DX]$
- T=7.5 days
  $\Omega_0 = [-0.9DX, 0.9DX] \times [-0.9DX, 0.9DX]$
- T=10 days
  $\Omega_0 = [-1.04DX, 1.04DX] \times [-1.04DX, 1.04DX]$

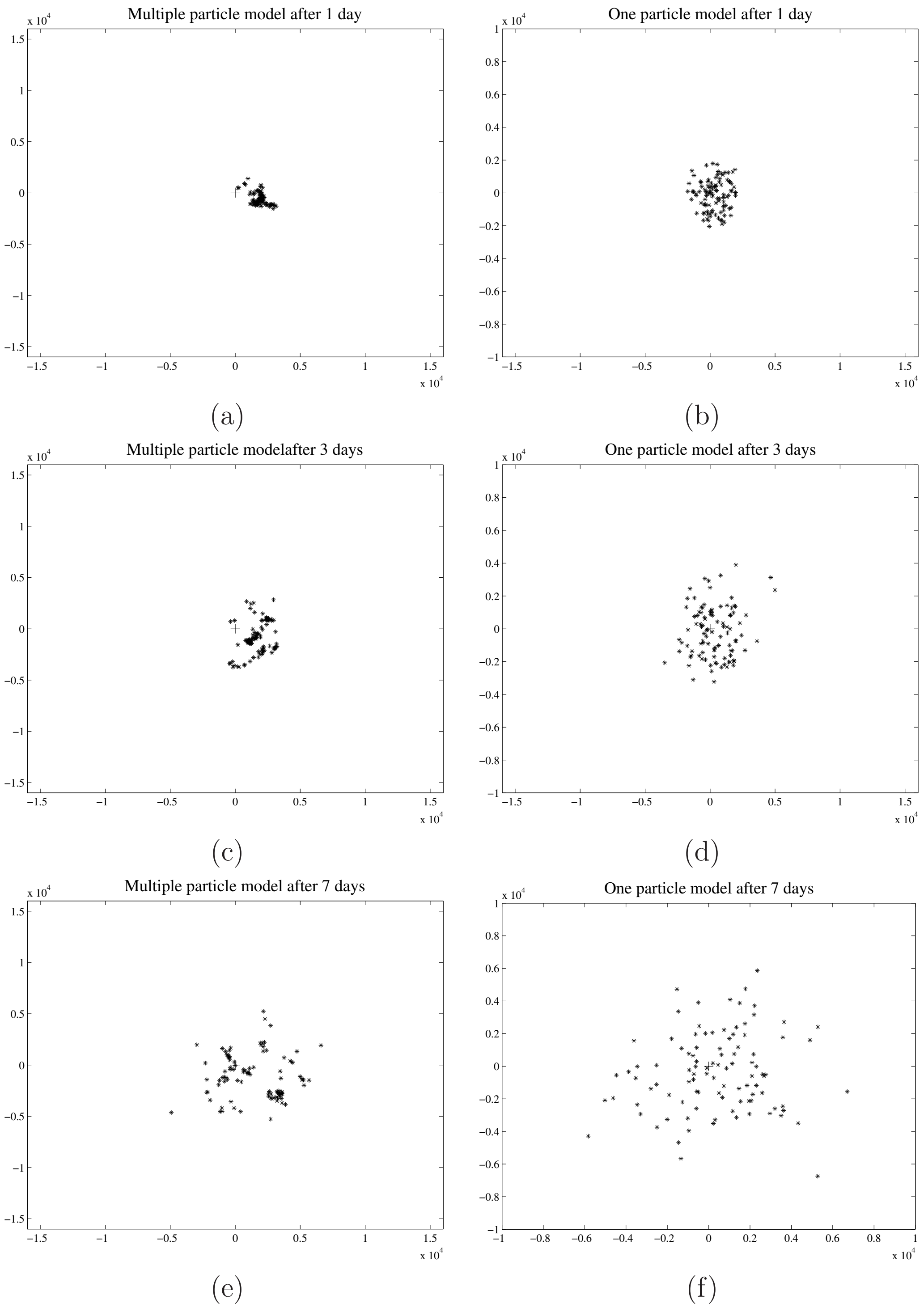

Figure 2: The multiple particle model and one-particle model

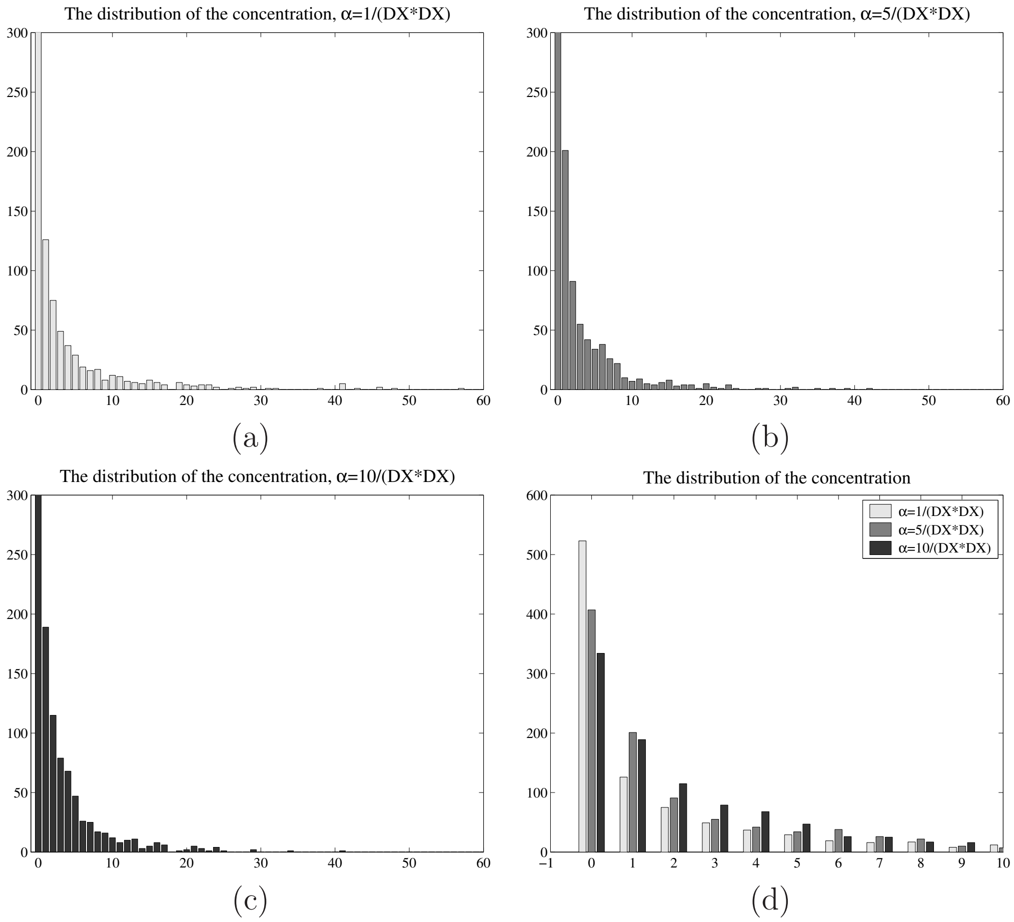

Figure 3: The distribution of the concentration

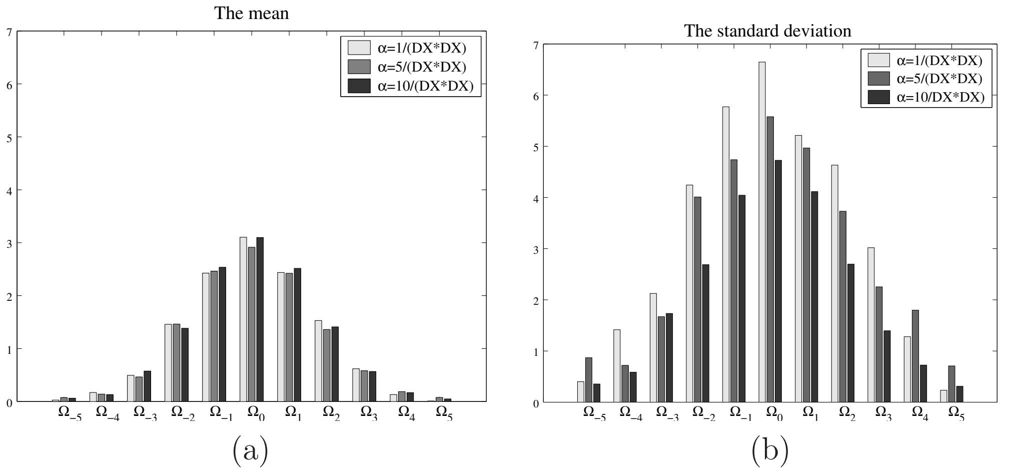

Figure 4: The mean and the standard deviation

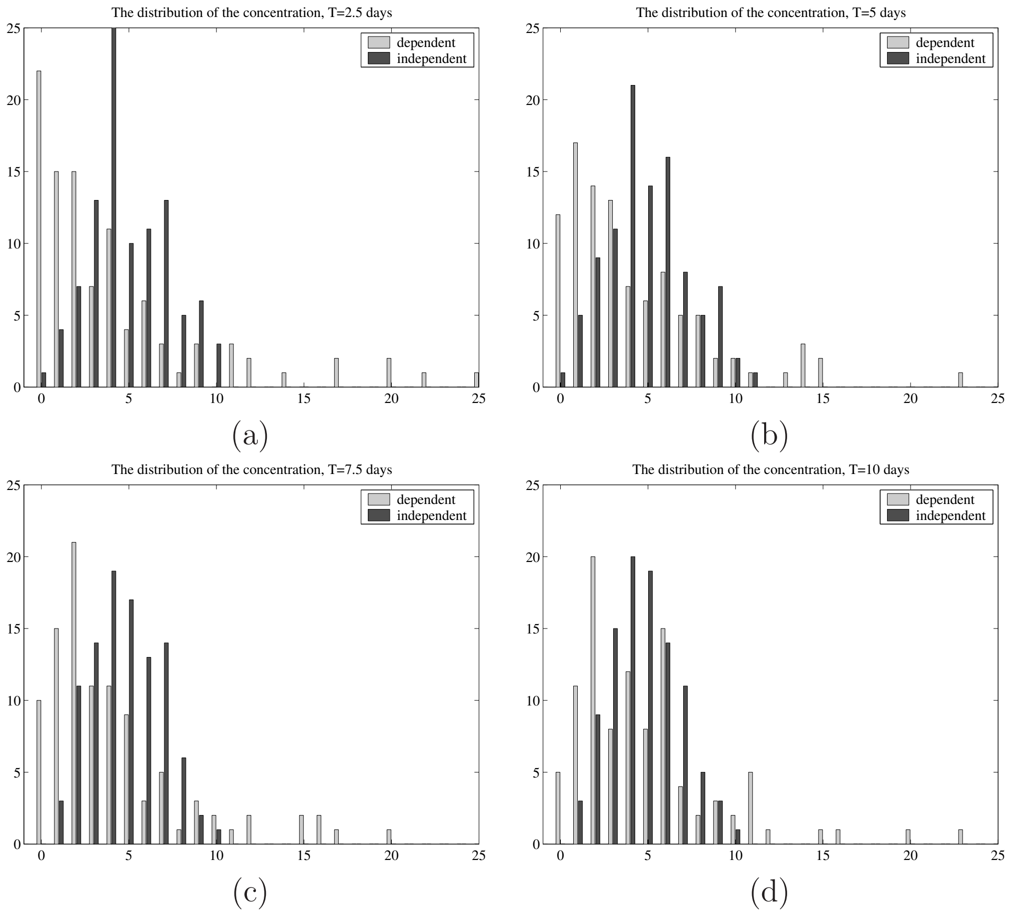

Figure 5: The distribution of the concentration with different time steps

and for these times we compare the distribution of the concentration for correlated and independent simulations (see figure 5). We have repeated each experiment 100 times. All distributions in figure 5 have the same mean, but, in case of multiple particle model, this mean is the result of many 'zero' concentrations and few very high concentrations. However, from the figure 5 it can be seen that the number of the 'zero' concentrations goes down with the time and the difference between multiple particle model and one-particle model disappear.

4.2 One- and two-particle models

In practice one does not need to start the multiple particle simulation to find the ensemble mean concentration and the concentration deviation. It is enough to consider only one- or two-particle models. We consider in this section a one-particle model of the form

$$\begin{aligned} d\boldsymbol{X}(s) &= \sqrt{2D} I_2 d\boldsymbol{B}(s), \\ \boldsymbol{X}(t) &= \boldsymbol{x}, \end{aligned} \tag{28}$$

where the diffusion coefficient D is assumed to be constant. The equation (28) can be solved exactly and the density function $p(t, \boldsymbol{x}, T, \boldsymbol{y})$ of the stochastic process $\boldsymbol{X}(T)$ is a normal distribution with mean $\boldsymbol{x}$ and variance $2D(T-t)$

$$p(t, \boldsymbol{x}, T, \boldsymbol{y}) = \frac{1}{4\pi D(T-t)} exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{x}\|^2}{4D(T-t)}\right). \tag{29}$$

We will consider the averaged concentration in the domain Ω_0 defined as

$$\begin{aligned} C(\Omega_0, T) &= \frac{1}{\mu(\Omega_0)} \int_{\boldsymbol{y}\in\Omega_0} C(\boldsymbol{y}, T) d\boldsymbol{y} = \\ &\frac{1}{\mu(\Omega_0)} \int_{\boldsymbol{y}\in\Omega_0} \frac{p(t, \boldsymbol{x}, T, \boldsymbol{y})}{H(\boldsymbol{y})} d\boldsymbol{y} = \\ &\frac{1}{\mu(\Omega_0)} \int_{\boldsymbol{y}\in\Omega_0} p(t, \boldsymbol{x}, T, \boldsymbol{y}) d\boldsymbol{y} = p(t, \boldsymbol{x}, T, \Omega_0), \end{aligned} \tag{30}$$

where $\mu(\Omega_0)$ is the area of Ω_0 and the water depth $H(\boldsymbol{y}) \equiv 1$ is taken to be constant in our test problems. The exact value of the integral in the equation (30) can be found from standard tables. We can use the forward estimator to find $p(t, \boldsymbol{x}, T, \Omega_0)$ in the following way,

$$\widehat{p}(t, \boldsymbol{x}, T, \Omega_0) = \frac{1}{N\delta^2} \sum_{n=1}^{N} K\left(\frac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)} - \boldsymbol{\eta}^{(n)}}{\delta}\right), \tag{31}$$

where $\eta^{(n)}$, $n = 1, \ldots, N$ are independent random numbers, uniformly distributed in the square Ω_0. In a similar way we can rewrite the forward-reverse estimator (27)

$$\begin{aligned} &\widehat{p}(t, \boldsymbol{x}, T, \Omega_0) = \frac{1}{N^2\delta^2} \\ &\sum_{n,m=1}^{N} K\left(\frac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)} - \overline{\boldsymbol{Y}}_{t^*,\boldsymbol{\eta}^{(m)}}^{(m)}(T)}{\delta}\right) \overline{\mathcal{Y}}_{t^*,\boldsymbol{\eta}^{(m)}}^{(m)}(T) \end{aligned} \tag{32}$$

We consider a two-particle forward system of the form

$$\begin{aligned} d\boldsymbol{Z}(s) &= \sqrt{2D}\boldsymbol{I}_4\left(\sqrt{1-\beta^2}d\boldsymbol{B}(s)+\beta d\boldsymbol{W}(s)\right)\\ \boldsymbol{Z}(t) &= (\boldsymbol{x},\boldsymbol{x})^T \end{aligned} \tag{33}$$

where $\boldsymbol{Z}=(\boldsymbol{X}^{[1]},\boldsymbol{X}^{[2]})^T$, $\boldsymbol{B}=(\boldsymbol{B}^{[1]},\boldsymbol{B}^{[2]})^T$, $\boldsymbol{W}=(\boldsymbol{W}^{[1]},\boldsymbol{W}^{[2]})^T$. The Brownian motions $\boldsymbol{W}^{[1]}$ and $\boldsymbol{W}^{[2]}$ are assumed to be correlated with covariance matrix (7). From here on we will use the notation r instead of r_{12} to denote the distance between particles $\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$. The reverse time system corresponding to (33) has the following form

$$\begin{aligned} d\boldsymbol{R}(s) &= \begin{pmatrix} \boldsymbol{Y}^{[1]}(s)\\ \boldsymbol{Y}^{[2]}(s)\end{pmatrix} = \begin{pmatrix} 2D\dfrac{\partial f}{\partial \tilde{\boldsymbol{y}}^{[2]}}\\ 2D\dfrac{\partial f}{\partial \tilde{\boldsymbol{y}}^{[1]}}\end{pmatrix} ds+\\ &\sqrt{2D}\boldsymbol{I}_4(\sqrt{1-\beta^2}d\tilde{\boldsymbol{B}}(s)+\beta d\tilde{\boldsymbol{W}}(s))\\ d\gamma(s) &= 2D\beta^2\left(\frac{\partial^2 f}{\partial\tilde{y}_1^{[1]}\partial\tilde{y}_1^{[2]}}+\frac{\partial^2 f}{\partial\tilde{y}_2^{[1]}\partial\tilde{y}_2^{[2]}}\right)\gamma ds\\ \boldsymbol{R}(t^*) &= (\boldsymbol{y},\boldsymbol{y})^T,\ \gamma(t^*)=1 \end{aligned} \tag{34}$$

It should be noted that the introduction of the reverse system in section 3 is based on a system of uncorrelated Brownian motions. It is not difficult to see however that the reverse system has an equivalent representation with respect to correlated Brownian motions which is given by (34) for this example.

The deviation of the concentration in the area Ω_0 at time T is defined as (we use the fact that $H=1$)

$$Dev(T,\Omega_0)=\sqrt{|p(t,\boldsymbol{x},\boldsymbol{x},T,\Omega_0,\Omega_0)-p^2(t,\boldsymbol{x},T,\Omega_0)|}, \tag{35}$$

where

$$\begin{aligned} &p(t,\boldsymbol{x},\boldsymbol{x},T,\Omega_0,\Omega_0)=\\ &\frac{1}{\mu^2(\Omega_0)}\int_{\Omega_0}\int_{\Omega_0} p(t,\boldsymbol{x},\boldsymbol{x},T,\boldsymbol{y}^{[1]},\boldsymbol{y}^{[2]})d\boldsymbol{y}^{[1]}d\boldsymbol{y}^{[2]}. \end{aligned} \tag{36}$$

Similarly to the one-particle model we can use the forward and forward-reverse estimators to find the value of the joint probability function $p(t,\boldsymbol{x},\boldsymbol{x},T,\Omega_0,\Omega_0)$

$$\begin{aligned} &\widehat{p}_{FE}(t,\boldsymbol{x},\boldsymbol{x},T,\Omega_0,\Omega_0)=\\ &\frac{1}{N\delta^4}\sum_{n=1}^{N}K\left(\frac{\overline{\boldsymbol{Z}}^{(n)}_{t,\boldsymbol{x},\boldsymbol{x}}(T)-\boldsymbol{\zeta}^{(n)}}{\delta}\right) \end{aligned} \tag{37}$$

$$\begin{aligned} &\widehat{p}_{FRE}(t,\boldsymbol{x},\boldsymbol{x},T,\Omega_0,\Omega_0)=\\ &\frac{1}{N^2\delta^4}\sum_{n,m=1}^{N}\tilde{K}\left(\frac{\overline{\boldsymbol{Z}}^{(n)}_{t,\boldsymbol{x},\boldsymbol{x}}(t^*)-\overline{\boldsymbol{R}}^{(m)}_{t^*,\boldsymbol{\zeta}^{(m)}}(T)}{\delta}\right)\gamma_{t^*,\boldsymbol{\zeta}^{(m)}}(T), \end{aligned} \tag{38}$$

where $\boldsymbol{\zeta}^{(n)}\in\mathbb{R}^4$, $n=1,\ldots,N$ are the random numbers uniformly distributed in $\Omega_0\times\Omega_0$. In tables 2,3,4,5,6 the results of the one- and two-particle models are

Table 2: The results of the one-particle model, the exact value $p(t, \boldsymbol{x}, T, \boldsymbol{y}) = 0.1886$

method	N	$p(t, \boldsymbol{x}, T, \boldsymbol{y})$
FE	10^4	0.1879 ± 0.0131
FE	10^5	0.1895 ± 0.0070
FE	10^6	0.1880 ± 0.0025
FRE	10^3	0.1875 ± 0.0050
FRE	10^4	0.1889 ± 0.0016
FRE	10^5	0.1885 ± 0.0007

Table 3: The joint probability function $p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})$

method	N	$p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})$
	$\alpha = 1/(DX * DX)$	
FE	10^5	0.2405 ± 0.0259
FE	10^6	0.2468 ± 0.0163
FE	10^7	0.2520 ± 0.0066
FRE	10^3	0.2546 ± 0.0209
FRE	10^4	0.2568 ± 0.0070
FRE	10^5	0.2604 ± 0.0023
	$\alpha = 5/(DX * DX)$	
FE	10^5	0.2303 ± 0.0218
FE	10^6	0.2622 ± 0.0118
FE	10^7	0.2933 ± 0.0070
FRE	10^3	0.3138 ± 0.4193
FRE	10^4	0.3715 ± 0.1077
FRE	10^5	0.3567 ± 0.0353

shown. The time of the simulation has been 2.5 days for all experiments. Each experiment was repeated 30 times in order to find the error due to the numerical simulation. In tables 2 and 4 one can see the results of the one-particle model for the point $\boldsymbol{y} = \boldsymbol{0}$ and for the area Ω_0 (see the section 4.1). In this case we know the solution of the stochastic differential equation (28) and can compare the numerical results with the exact solution. From these tables it can be seen that the forward-reverse estimator (FRE) is at least two orders of magnitude more accurate than the pure forward estimator (FE). Also one can see that the value of the function $p(t, \boldsymbol{x}, T, \boldsymbol{y})$ does not differ too much from $p(t, \boldsymbol{x}, T, \Omega_0)$.

In tables 3 and 5 the results of the two-particle model are presented for different value of the parameter α. There are two aspects of the approach that we will discuss in the next two sections. Firstly, it can be seen from these tables that the FRE is much more efficient (three orders of magnitude) than the FE when the value of

Table 4: The results of the one-particle model, the exact value $p(t, \boldsymbol{x}, T, \Omega_0) = 0.1856$

method	N	$p(t, \boldsymbol{x}, T, \Omega_0)$
FE	10^4	0.1797 ± 0.0141
FE	10^5	0.1859 ± 0.0077
FE	10^6	0.1844 ± 0.0030
FRE	10^3	0.1863 ± 0.0051
FRE	10^4	0.1853 ± 0.0020
FRE	10^5	0.1857 ± 0.0006

Table 5: The joint probability function $p(t, \boldsymbol{x}, \boldsymbol{x}, T, \Omega_0, \Omega_0)$

method	N	$p(t, \boldsymbol{x}, \boldsymbol{x}, T, \Omega_0, \Omega_0)$
	$\alpha = 1/(DX * DX)$	
FE	10^5	0.1606 ± 0.0232
FE	10^6	0.1662 ± 0.0128
FE	10^7	0.1705 ± 0.0070
FRE	10^3	0.1682 ± 0.0244
FRE	10^4	0.1705 ± 0.0049
FRE	10^5	0.1725 ± 0.0017
	$\alpha = 5/(DX * DX)$	
FE	10^5	0.1239 ± 0.0157
FE	10^6	0.1360 ± 0.0092
FE	10^7	0.1367 ± 0.0067
FRE	10^3	0.1250 ± 0.1953
FRE	10^4	0.1250 ± 0.0585
FRE	10^5	0.1440 ± 0.0185

parameter α is relatively small, but for a large value of α the FRE is not so efficient. We will discuss this in more detail in section 4.4. Secondly, if one compares the joint probability function at the point $p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})$ and the average value of this function on Ω_0, it can be seen that these values differ significantly. Moreover, the results from the table 6 show that the deviation in the point $Dev(T, \boldsymbol{y})$ grows with larger values of α, while the averaged value in Ω_0 of deviation $Dev(T, \Omega_0)$ decreases. This behavior will be discussed in more detail in the next section.

4.3 The definition of the standard deviation at a point

In this section we discuss some properties of the joint probability function $p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}^{[1]}, \boldsymbol{y}^{[2]})$, in particularly, how this function depends on the choice of the correlation function $f(\rho)$. In figure 6 the correlation function (8) for $\alpha =$

Table 6: The concentration and the standard deviation of the concentration

	in the point $\boldsymbol{y}$	
α	$C(\boldsymbol{y}, T)$	$Dev(\boldsymbol{y}, T)$
$1/(DX * DX)$	0.1886	0.4737
$5/(DX * DX)$	0.1886	0.5142
	in the area Ω_0	
α	$C(\Omega_0, T)$	$Dev(\Omega_0, T)$
$1/(DX * DX)$	0.1856	0.3682
$5/(DX * DX)$	0.1856	0.3249

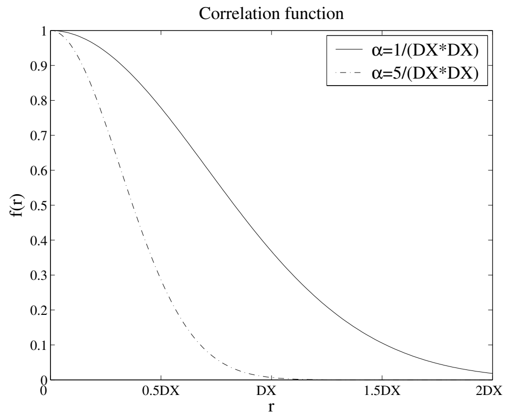

Figure 6: The correlation function

$1/(DX * DX)$ and for $\alpha = 5/(DX * DX)$ is shown. It is clear at the larger α is, the faster the correlation function goes down and as a result the particles become independent much faster. Therefore it is natural to expect that the deviation of the concentration decreases when α increases, but the results of the tables 3, 5 and 6 show the opposite. To explain this behavior we consider the joint probability function $p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}^{[1]}, \boldsymbol{y}^{[2]}$ in some neighborhood of the point $\boldsymbol{y}$. Let introduce the following function

$$\overline{C(\boldsymbol{y})C(\boldsymbol{y}; r)} = p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{u}) \mid_{\|\boldsymbol{y}-\boldsymbol{u}\|=r} \tag{39}$$

The function (39) can be interpreted as the dependence between the concentration at a point $\boldsymbol{y}$ and at a location at a distance r from this point $\boldsymbol{y}$. In figure (7) the function $\overline{C(\boldsymbol{y})C(\boldsymbol{y}; r)}$ is shown for $\alpha = 1/(DX * DX)$ and for $\alpha = 5/(DX * DX)$. The value of this function for $\alpha = 5/(DX * DX)$ at $r = 0$ is larger than for $\alpha = 1/(DX * DX)$. But this function decreases very fast and as a consequence it is not enough to study this function when $r = 0$ as it has been done before, but we need to consider some neighborhood of this point. We propose to define the

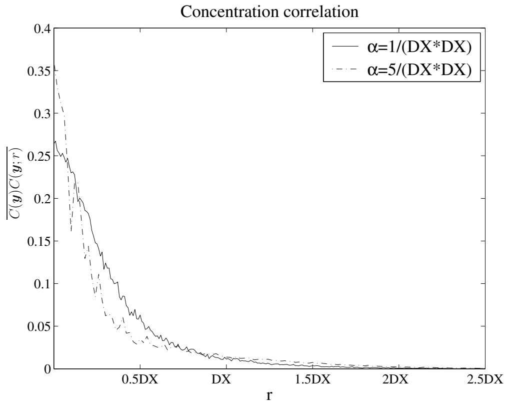

Figure 7: The correlation of the concentration

deviation at the point $\boldsymbol{y}$ at time T $\widehat{Dev}(T, \boldsymbol{y})$ as followed

$$\widehat{Dev}(T,\boldsymbol{y}) = \frac{1}{H(\boldsymbol{y})}\times \left| \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} p(t,\boldsymbol{x},\boldsymbol{x},T,\boldsymbol{u},\boldsymbol{v})g(\boldsymbol{u})g(\boldsymbol{v})d\boldsymbol{u}d\boldsymbol{v} - \left(\int\limits_{\mathbb{R}^2} p(t,\boldsymbol{x},T,\boldsymbol{u})g(\boldsymbol{u})d\boldsymbol{u}\right)^2\right|^{\frac{1}{2}} \tag{40}$$

where

$$g(\boldsymbol{u}) = \frac{1}{2\pi\sigma^2}\exp\left(\frac{\|\boldsymbol{y}-\boldsymbol{u}\|^2}{2\sigma^2}\right). \tag{41}$$

The function $g(\boldsymbol{u})$ is a density function of a 2-dimensional normal distribution with the vector of means $\boldsymbol{y}$ and and covariance matrix$\sigma^2\boldsymbol{I}_2$. We can use the forward and forward-reverse estimators (31) and (32) to find the integral

$$\widehat{p}(t,\boldsymbol{x},T,\boldsymbol{y}) = \int\limits_{\mathbb{R}^2} p(t,\boldsymbol{x},T,\boldsymbol{u})g(\boldsymbol{u})d\boldsymbol{u}$$

assuming that $\boldsymbol{\eta}^{(n)}$, $n = 1,\ldots,N$ are independent realizations of the random variable $\boldsymbol{\eta}$ with density function $g(\boldsymbol{u})$. In a similar way we can use the formulae (37) and (38) to find the integral

$$\widehat{p}(t,\boldsymbol{x},\boldsymbol{x},T,\boldsymbol{y},\boldsymbol{y}) = \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} p(t,\boldsymbol{x},\boldsymbol{x},T,\boldsymbol{u},\boldsymbol{v})g(\boldsymbol{u})g(\boldsymbol{v})d\boldsymbol{u}d\boldsymbol{v}.$$

In this case we assume that $\boldsymbol{\zeta}^{(n)}$, $n = 1,\ldots,N$ are independent realizations of the 4-dimensional random variable $\boldsymbol{\zeta}$ normally distributed with the vector of means $(\boldsymbol{y},\boldsymbol{y})^T$ and the covariance matrix $\sigma^2\boldsymbol{I}_4$. The value of the parameter σ^2 should be relatively small compared with the variance of the stochastic process $\boldsymbol{X}(s)$ at time T. For example, in our simulation $\sigma^2 = 0.04DX^2$, while the variance of $\boldsymbol{X}(T)$ is $2D(T-t) \approx 0.85DX^2$.

It can be shown easily using the equation (29) and the properties of density function of a normally distributed random variable that

$$\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y}) = \\ \frac{1}{2\pi(2D(T-t)+\sigma^2)} \exp\left(\frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{2D(T-t)+\sigma^2}\right)$$

The results for functions $\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y})$ and for $\widehat{p}(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})$ are shown in table 7. Comparing this results with the results from tables 2 and 3 it can be seen that

Table 7: The concentration and the standard deviation of the concentration

method	N	$\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y})$	$\widehat{p}(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})$	$\widehat{Dev}(T, \boldsymbol{y})$
		$\alpha = 1/(DX * DX)$		
FE	10^5	0.1801	0.1074 ± 0.0162	0.2738
FE	10^6	0.1801	0.1161 ± 0.0091	0.2892
FE	10^5	0.1801	0.1146 ± 0.0060	0.2866
FRE	10^3	0.1801	0.1142 ± 0.0154	0.2859
FRE	10^4	0.1801	0.1148 ± 0.0047	0.2870
FRE	10^5	0.1801	0.1166 ± 0.0014	0.2901
		$\alpha = 5/(DX * DX)$		
FE	10^5	0.1801	0.0822 ± 0.0126	0.2231
FE	10^6	0.1801	0.0866 ± 0.0079	0.2327
FE	10^5	0.1801	0.0847 ± 0.0050	0.2286
FRE	10^3	0.1801	0.0788 ± 0.0749	0.2153
FRE	10^4	0.1801	0.0871 ± 0.0446	0.2338
FRE	10^5	0.1801	0.0882 ± 0.0173	0.2361

the differences between the functions $p(t, \boldsymbol{x}, T, \boldsymbol{y})$ and $\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y})$ are not very big, while the functions $p(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})$ and $\widehat{p}(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})$ differ very strongly.

4.4 The implementation of the FRE

As it was mentioned before the accuracy of the forward-reverse estimator strongly depends an the correlation function $f(r)$, in particular on the parameter α. The difficulties arise with the weighting coefficient $\mathcal{Y}$. First of all, the differential equation for the scalar function $\mathcal{Y}(s)$ is very sensitive to the choice of the time step h of the numerical integration of the reverse time system (34). This problem is discussed in detail in [26]. As a consequence we must chose the time step h sufficiently small. Unfortunately, decreasing the time step we also increase the time of the simulation and affects the efficiency of the forward-reverse estimator. A possible solution is to compare the results of the simulations with time steps h and $2h$ (using the Talay-Tubaro method we repeat the simulation with double time step) and if the solutions of the reverse time system (34) differ very much, we do not take into consideration the results of these particles. For example, one can compare the results for time step $h = 300$s and for time step $h = 30$s ($N = 10^4$, $\alpha = 5/(DX * DX)$, $t^* = 0.5(T-t)+t$)

$dt = 300s \quad \widehat{p}(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y}) = 0.0871 \pm 0.0446$

$dt = 30s \quad \widehat{p}(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y}) = 0.08435 \pm 0.0367$

Another problem occurs when the weighting coefficient $\mathcal{Y}(T)$ varies very much for different particles. For example, when $\alpha = 5/(DX * DX)$ the values $\mathcal{Y}(T)$ can be around 300 and, of course, the contribution of this particle will dominate in the sum (27). What we can do here is to throw away the particles with "very large" coefficient. For instance, one can see the results of the forward-reverse estimator including all particles and without the particles with the coefficient $\mathcal{Y}$ of more than 250

all particles	$\widehat{p} = 0.0871 \pm 0.0446$
throw away particles	$\widehat{p} = 0.0782 \pm 0.0239$

This method may help in many situations, but when the correlation function is too steep, the forward-reverse method will fail and we have to choose the internal point t^* closer to the end point T.

5 Application

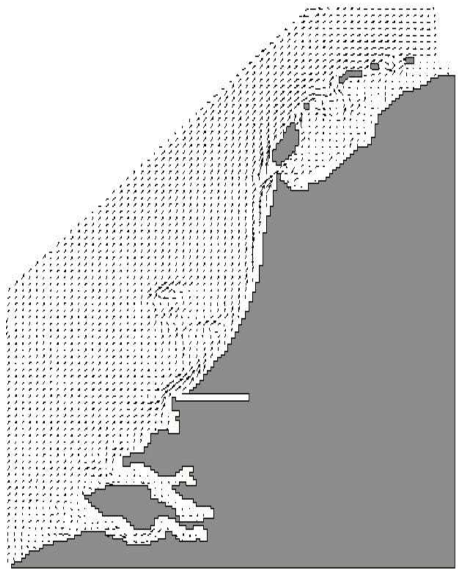

Figure 8: The tidally-averaged flow

In this section we study the mean ensemble concentration and the deviation of the concentration in some critical locations along the Dutch seaside using one- and two-particle models. Furthermore, we apply the forward-reverse estimator for both models and compare the results with the results of the classical forward estimator.

In our application we have used a tidally-averaged numerical flow model with grid size $DX = 1600m$ (see figure 8). Because we know the velocities and water depth only in the grid points of the model, we use bilinear interpolation to obtain the velocities and water depth at arbitrary locations. The dispersion coefficient is chosen constant: $D = 5\, m^2/s$. We assume that at initial time 10^4 kg of contaminant was released.

5.1 One-particle model

The reverse system associated with the one-particle model (3) has the following form (we take into account that D is constant)

$$\begin{aligned} d\boldsymbol{Y}(s) &= \left(-\boldsymbol{u} - \frac{D}{H}\frac{\partial H}{\partial \tilde{\boldsymbol{y}}}\right) ds + \sqrt{2D}\boldsymbol{I}_2 d\boldsymbol{W}(s) \\ d\mathcal{Y}(s) &= -\mathcal{Y}(s)\frac{\partial}{\partial \tilde{\boldsymbol{y}}} \cdot \left(\boldsymbol{u} + \tfrac{D}{H}\tfrac{\partial H}{\partial \tilde{\boldsymbol{y}}}\right) ds \\ \boldsymbol{Y}(t^*) &= \boldsymbol{y},\, \mathcal{Y}(t^*) = 1, \end{aligned} \tag{42}$$

To solve the systems (3) and (42) we use the Euler method (15) and then apply the Talay-Tubaro method to find the probability density function. We suppose that the contaminant is released in the location $(20, 40)$. We consider some critical locations in which we want to know the ensemble mean concentration and the standard deviation of the concentration. For two of them we compare the forward estimator and the forward-reverse estimator (see table 8). The time step of the simulation is supposed $h = 5$ min. We use the Euler scheme (15) and then applied the Talay-Tubaro method. The bandwidth δ for the estimators is chosen according table 1. In order to reduce the numerical error of the forward estimator we chose the bandwidth δ for one-particle model as $\delta = CN^{-1/6}$, where

$$C = \left(\frac{1}{N-1}\sum_{n=1}^{N} \|\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)} - \frac{1}{N}\sum_{n=1}^{N} \overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}\|^2\right)^{\frac{1}{2}}.$$

For the two-particle model we use the Fukunaga method (22). In figure 9 one can see the example of the forward and reverse time simulation during 10 days. The internal point point t^* is chosen in the middle of the time interval $[t, T]$. In practice t^* is an additional parameter which can be chosen at any point of the time interval $[t, T]$. In our application we use the method that was proposed by [3]. The concentration of the pollutant in the location $\boldsymbol{y} = (27, 47)$ after $T = 2.5$ days is:

$$C(T, \boldsymbol{y}) = \frac{p(t, \boldsymbol{x}, T, \boldsymbol{y})}{H(\boldsymbol{y})} = 2.15 * 10^{-5}\, kg/m^3$$

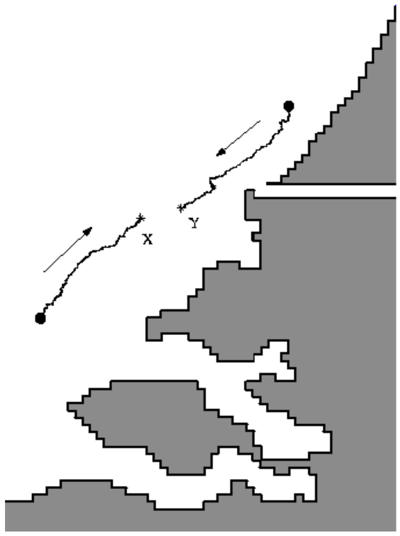

Figure 9: The example of the forward and reverse time simulation for the one-particle model

The concentration of the pollutant in the location $\boldsymbol{y} = (35, 55)$ after $T = 5$ days is equal

$$C(T, \boldsymbol{y}) = \frac{p(t, \boldsymbol{x}, T, \boldsymbol{y})}{H(\boldsymbol{y})} = 6.7 * 10^{-6}\, kg/m^3$$

5.2 Two-particle model

Now we consider the two-particle model. The forward system has the following form

$$\begin{aligned}
&\boldsymbol{Z}(s) = d\begin{pmatrix} \boldsymbol{X}^{[1]} \\ \boldsymbol{X}^{[2]} \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}^{[1]} + \frac{D}{H^{[1]}}\frac{\partial H^{[1]}}{\partial \boldsymbol{y}^{[1]}} \\ \boldsymbol{u}^{[2]} + \frac{D}{H^{[2]}}\frac{\partial H^{[2]}}{\partial \boldsymbol{y}^{[2]}} \end{pmatrix} ds + \\
&\sqrt{2D} I_4 \begin{pmatrix} \sqrt{1-\beta^2} d\boldsymbol{B}^{[1]}(s) + \beta d\boldsymbol{W}^{[1]}(s) \\ \sqrt{1-\beta^2} d\boldsymbol{B}^{[2]}(s) + \beta d\boldsymbol{W}^{[2]}(s) \end{pmatrix} \\
&\boldsymbol{Z}(t) = (\boldsymbol{x}, \boldsymbol{x})^T
\end{aligned} \tag{43}$$

The reverse time system that corresponds with the two-particle forward system (43) can be written as follows

$$\begin{aligned}
&d\boldsymbol{R} = \begin{pmatrix} d\boldsymbol{Y}^{[1]} \\ d\boldsymbol{Y}^{[2]} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{u}^{[1]} - \frac{D}{H^{[1]}}\frac{\partial H^{[1]}}{\partial \tilde{\boldsymbol{y}}^{[1]}} + 2D\frac{\partial f}{\partial \tilde{\boldsymbol{y}}^{[2]}} \\ -\boldsymbol{u}^{[2]} - \frac{D}{H^{[2]}}\frac{\partial H^{[2]}}{\partial \tilde{\boldsymbol{y}}^{[2]}} + 2D\frac{\partial f}{\partial \tilde{\boldsymbol{y}}^{[1]}} \end{pmatrix} ds + \\
&\sqrt{2D} I_4 \begin{pmatrix} \sqrt{1-\beta^2} d\boldsymbol{B}^{[1]}(s) + \beta d\boldsymbol{W}^{[1]}(s) \\ \sqrt{1-\beta^2} d\boldsymbol{B}^{[2]}(s) + \beta d\boldsymbol{W}^{[2]}(s) \end{pmatrix} \\
&d\gamma(s) = \Big(\varphi_1(s, \boldsymbol{Y}^{[1]}) + \varphi_2(s, \boldsymbol{Y}^{[2]}) + \varphi_3(s, \boldsymbol{R})\Big)\gamma(s) ds \\
&\boldsymbol{R}(t^*) = (\boldsymbol{y}, \boldsymbol{y})^T, \quad , \gamma(t^*) = 1,
\end{aligned} \tag{44}$$

Table 8: The results for the one-particle model

method	N	$\widehat{p}(t, \boldsymbol{x}, T, \boldsymbol{y})$
	$\boldsymbol{y} = (27, 47)$	$T = 2.5$ days
FE	10^4	0.1218 ± 0.0089
FE	10^5	0.1211 ± 0.0054
FE	10^6	0.1223 ± 0.0024
FRE	10^3	0.1220 ± 0.0057
FRE	10^4	0.1213 ± 0.0015
FRE	10^5	0.1216 ± 0.0006
	$\boldsymbol{y} = (35, 55)$	$T = 5$ days
FE	10^4	0.0398 ± 0.0040
FE	10^5	0.0389 ± 0.0017
FE	10^6	0.0392 ± 0.0010
FRE	10^3	0.0392 ± 0.0018
FRE	10^4	0.0393 ± 0.0007
FRE	10^5	0.0393 ± 0.0002

where

$$\varphi_i(s, \boldsymbol{Y}^{[i]}) = -\frac{\partial}{\partial \tilde{\boldsymbol{y}}^{[i]}} \cdot \left(\boldsymbol{u}^{[i]} + \frac{D}{H^{[i]}} \frac{\partial H^{[i]}}{\partial \partial \tilde{\boldsymbol{y}}^{[i]}} \right), \quad i = 1, 2$$

represents the contribution of the flow and the function

$$\varphi_3(s, \boldsymbol{R}) = 2D\beta^2 \left(\frac{\partial^2 f}{\partial \tilde{y}_1^{[1]} \partial \tilde{y}_1^{[2]}} + \frac{\partial^2 f}{\partial \tilde{y}_2^{[1]} \partial \tilde{y}_2^{[2]}} \right)$$

represents the contribution of the correlation between particles. Figure 4 shows an example of the forward and reverse time simulation for a pair of particles. The time of the simulation is 10 days. In table 9 the results for the second moment of the particle distribution (or the joint probability function of process $\boldsymbol{Z}(T)$) are shown. We compare the forward and forward-reverse estimator for two locations.

The deviation $\widehat{Dev}(T, \boldsymbol{y})$ can be found with help of the equation (40)

$\boldsymbol{y} = (27, 47) \quad \widehat{Dev}(T, \boldsymbol{y}) = 4 * 10^{-5}\, kg/m^3$
$\boldsymbol{y} = (35, 55) \quad \widehat{Dev}(T, \boldsymbol{y}) = 1.8 * 10^{-5}\, kg/m^3$

From the table 8 we can see that the mean ensemble concentration calculated by the FRE using 10^3 particles is $C(T, \boldsymbol{y}) = 6.7 \cdot 10^{-6} \pm 0.3 \cdot 10^{-6}$ kg/m^3. This deviation $\pm 0.3 \cdot 10^{-6}$ kg/m^3 is caused by statistical error of the estimation process and converges to zero with the increasing the number of particles. Contrary, the standard deviation $\widehat{Dev} = 1.8 \cdot 10^{-5}$ kg/m^3 as calculated in table 9, is the result of the spatial correlation of the turbulence and does not depend on the number of particles.

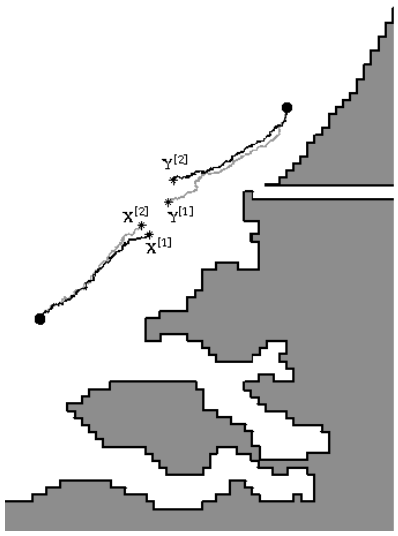

Figure 10: The example of the forward and reverse time simulation for two-particle model

The results for the concentration and the standard deviation for different locations along the Dutch coast ar(see figure 11) are shown in table 10. Here we used the FRE with number of particles $N = 10^5$. The results show that even 10 days after the calamity the concentration fluctuation are still very large and should be taken into account in order to assess the impact of the calamity.

6 Conclusion

In this paper we studied one- and two-particle models for computing the mean and standard deviation of the concentration in the Dutch coastal zone using the forward-reverse estimator. The results show that the actual concentration may become much higher than the ensemble mean concentration as computed by the traditional transport model. As a consequence, for providing an accurate prediction of the spreading of the pollutant, we need to use two-particle models and have to take into account the spatial correlation of the turbulence.

The results also show that the forward-reverse estimator is at least two orders of magnitude more efficient than the classical pure forward estimator.

References

[1] Arnold L (1974) Stochastic differential equations. Willey, New York London Sydney Toronto

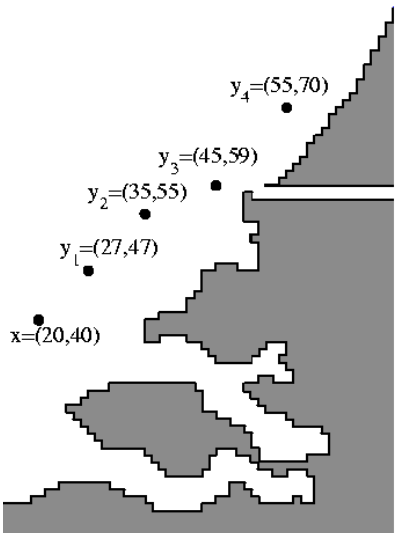

Figure 11: The locations of interest

[2] Baas AF, De (1988) Some properties of the Langevin Model for dispersion. Ph.D. Dissertation, Delft University of Technology

[3] Berg E, van der, Heemink AW, Lin HX, Schoenmakers JGM (in press) Probability density estimation in stochastic environmental models using reverse representations, Stochastic Environmental Research and Risk Assessment DOI: 10.1007/s00477-005-0022-5

[4] Borgas MS, Sawford BL (1994) A family of stochastic models for two-particle dispersion in isotropic homogeneous stationary turbulence. J. Fluid Mech. 279: 69-99

[5] Durbin PA (1980) A stochastic model of two-particle dispersion and concentration fluctuations in homogeneous turbulence. J. Fluid Mech. 100: 279-302

[6] Carlfisch R, Morokoff W (1995) Quasi-Monte Carlo integration. Journal of Computational Physics 122: 218-230

[7] Costa M, Ferreira JS (2000) Discrete particle distribution model for advection-diffusion transport. Journal of Hydraulic Engineering 126, nr. 525-532

[8] Cressie NAC (1991) Statistics for spatial data. Willey, New York Chichester Toronto Brisbane Singapore

[9] Ewing RE, Wang H (2001) A summary of numerical methods for time-dependent advection-dominated partial differential equations. Journal of Computational and Applied Mathematics 128, nr. 1-2: 423-445

Table 9: The results for two-particle model

method	N	$\widehat{p}(t, \boldsymbol{x}, \boldsymbol{x}, T, \boldsymbol{y}, \boldsymbol{y})$
	$\boldsymbol{y} = (27, 47)$	$T = 2.5$ days
FE	10^5	0.0594 ± 0.0086
FE	10^6	0.0648 ± 0.0049
FE	10^7	0.0653 ± 0.0027
FRE	10^3	0.0645 ± 0.0096
FRE	10^4	0.0662 ± 0.0034
FRE	10^5	0.0659 ± 0.0010
	$\boldsymbol{y} = (35, 55)$	$T = 5$ days
FE	10^4	0.0125 ± 0.0015
FE	10^5	0.0126 ± 0.0014
FE	10^6	0.0126 ± 0.0067
FRE	10^3	0.0119 ± 0.0076
FRE	10^4	0.0123 ± 0.0023
FRE	10^5	0.0126 ± 0.0010

Table 10: The concentration $\widehat{C}(T, \boldsymbol{y})$, kg/m^3 and the standard deviation of the concentration $\widehat{Dev}(T, \boldsymbol{y})$, kg/m^3

$\boldsymbol{y}$	T, days	t^*	$\widehat{C}(T, \boldsymbol{y})$	$\widehat{Dev}(T, \boldsymbol{y})$
$(27, 47)$	2.5	$0.5T$	$2.15 * 10^{-5}$	$4 * 10^{-5}$
$(35, 55)$	5	$0.5T$	$6.7 * 10^{-6}$	$1.8 * 10^{-5}$
$(45, 59)$	7.5	$0.6T$	$2.8 * 10^{-6}$	$7.1 * 10^{-6}$
$(55, 70)$	10	$0.7T$	$2.3 * 10^{-6}$	$6.1 * 10^{-6}$

[10] Fischer HB, List EJ, Koh RCY, Imberger J, Brooks NH (1979) Mixing in inland and coastal waters. Academic Press, New York

[11] Heemink AW (1990) Stochastic modeling of dispersion in shallow water. Stochastic Hydrol. and Hydraul. 4: 161-174

[12] Jazwinski AW (1970) Stochastic differential processes and filtering theory. Academic Press, New York

[13] Kaplan H, Dinar N (1987) A three dimensional stochastic model for concentration fluctuation statistics in isotropic homogeneous turbulence. Journal of computational physics 79: 317-335

[14] Kloeden PE, Platen E (1992) Numerical solution of stochastic differential equations. Springer, Berlin Heidelberg New York

[15] Kurbanmuradov OA, Orslag SA, Sabelfeld KK, Yeung PK (2001) Analysis of relative dispersion of two particles by Lagrangian stochastic models and DNS methods. Monte Carlo Methods and Applications 7, nr.3-4: 245-264

[16] Kurbanmuradov OA, Rannik U, Sabelfeld K, Vesala T (1999) Direct and adjoint Monte Carlo for the footprint problem. Monte Carlo Methods and Applications 5, nr. 2: 85-111

[17] Little T, Pant V (2001) A finite difference method for the valuation of variance swaps. Journal of Computational Finance 5, nr.1: 81-102

[18] Milstein GN, Schoenmakers JGM, Spokoiny V (2004) Transition density estimation for stochastic differential equations via forward-reverse representations. Bernoulli 10, nr. 2: 281-312

[19] Milstein GN, Tretyakov MV (2004) Stochastic numerics for mathematical physics. Springer, Berlin Heidelberg New York

[20] Monin AS, Yaglom AM (1979) Statistical fluid mechanics: mechanics of turbulence. MIT Press,Cambridge.

[21] Oksendal B (1985) Stochastic differential equations. Springer, Berlin Heidelberg New York

[22] Schoenmakers JGM, Heemink AW, Ponnambalam K, Kloeden PE (2002) Variance reduction for Monte Carlo simulation of stochastic environmental models. Appl. Mat. Mod. 26: 785-795

[23] Scott CF (1997) Particle tracking simulation of pollutant discharges. Journal of Environmental Engineering 123, nr. 9: 919-927

[24] Silverman BW (1986) Density estimation for statistics and data analysis. Champan and Hall, London

[25] Holley ER (1996) Diffusion and Dispersion.In: Singh VP, Hager WH (eds) Environmental hydraulics Kluwer Academic Press, Dodrecht, pp 111-151

[26] Spivakovskaya D, Heemink AW, Milstein GN, Schoenmakers JGM (2005) Simulation of the transport of particles in coastal waters using forward and reverse time diffusion, Advances in Water Resources 28: 927-938

[27] Stijn ThL, Praagman N, van Eijkeren J (1987) Positive advection schemes for environmental studies. In: Numerical methods in laminar and turbulent flow, Pineridge Press, Swansea: 1256-1267

[28] Talay D, Tubaro L (1991) Expansion of the global error for numerical schemes solving stochastic differential equations. Stochastic Ananlysis and Applications 8, nr. 4: 485-509

[29] Thomson DJ (1987) Criteria for the selection of stochastic models of particles trajectories in turbulent flow. J. Fluid Mech. 180: 529-556

[30] Thomson DJ (1990) A stochastic model for the motion of particle pairs in isotropic high-Reynolds-number turbulence, and its application to the problem of concentration variance. J. Fluid Mech. 210: 113-153

[31] Yang Y, Wilson LT, Makela ME, Marchetti MA (1998) Accuracy of numerical methods for solving the advection-diffusion equation as applied to spore and insect dispersal. Ecological Modelling 109, nr. 1:1-24

[32] Wand MP, Jones MC (1995) Kernel smoothing. Champan and Hall, London