Stochastic Lagrangian approach for wind farm simulation

Mireille Bossy, Aurore Dupré, Philippe Drobinski, Laurent Violeau, Christian Briard

Chapter 1
Stochastic Lagrangian approach for wind farm simulation

Mireille Bossy, Aurore Dupré, Philippe Drobinski, Laurent Violeau and Christian Briard

Abstract We present a stochastic Lagrangian approach for atmospheric boundary layer simulation. Based on a turbulent-fluid-particle model, a stochastic Lagrangian particle approach could be an advantageous alternative for some applications, in particular in the context of down-scaling simulation and wind farm simulation. This paper presents two recent advances in this direction, first the analysis of an optimal rate of convergence result for the particle approximation method that grounds the space discretisation of the Lagrangian model, and second a preliminary illustration of our methodology based on the simulation of a Zephyr ENR wind farm of six turbines.

1.1 Introduction

The stakes of the simulation of wind farm production are growing with the development of renewable energies. The various time scales involved (from wind potential evaluation, to short-term production forecast), the mix of constraints on existing sites or on new projects are all issues where simulations can play an important role.

Although some CFD models and wind turbine models and softwares are already well established in this sector of activity (see eg. Sørensen [20], Niayifar and Porté-Agel [14], and the references cited therein), the problem of how to enriching and refining a wind simulation (from a meteorological forecast, or from a larger scale

M. Bossy, L. Violeau
Université Côte dAzur, Inria, France, e-mail: firstname.lastname@inria.fr

A. Dupré and P. Drobinski
LMD/IPSL, École polytechnique, Université Paris Saclay, ENS,PSL Research University, Sorbonne Universités, UPMC Univ Paris 06, CNRS, Palaiseau France e-mail: firstname.lastname@lmd.polytechnique.fr

C. Briard
Zephyr ENR, France

information, eventual combined with measurements) remains a widely open problem, regarding the production estimation of a site, wind turbine by wind turbine. Among various existing approaches for wind-farm simulation we can distinguish

- wind extrapolation methods, and parametrization of wake effect for real-time simulation response,
- interaction fluid and structure models for wake computations, with often laminar flow hypothesis and rather simple terrain description,
- LES methods for turbulent flows, including mills contribution forces related to actuator disc modeling.

The turbulent nature of the atmospheric boundary layer contributes to the uncertainty of the wind energy estimation, and have to be taken into account in the modeling approach when assessing the power production. The purpose of the Stochastic Downscaling Model (SDM) presented here is to propose a new method to downscale the wind simulation from a mesoscale meteorological solver in the Atmospheric Boundary Layer (ABL), particularly near the ground where the effect of turbulence is predominant. The SDM main modeling features reside in the choice of a fully Lagrangian paradigm of turbulent flow modeling. This is allowed by Stochastic Lagrangian fluid-particle Modeling (SLM) approach. Such Probability Density Function (PDF) methods are computationally inexpensive when we need to refine the space scale, as particles methods are free of numerical constraints (such as CFL constraint).

SDM developpment is a collaborative long term work (see [1, 2, 4] for more detailed presentation). The development of such a model and numerical approach is an interdisciplinary task involving disciplines such as stochastic analysis and numerical analysis for the design and optimal use of a Lagrangian particle solver, physics of the Atmospheric Boundary Layer (ABL) for the calibration and validation of SDM equations and boundary conditions, and engineering for the Lagrangian adaptation of actuator disk model for the turbine wake effect.

This paper presents two recent advances in these directions: Section 1.2 is dedicated to the convergence rate analysis of the stochastic particle algorithm. We analyse the convergence rate with numericals experiments and check the adequacy with the theoretical optimal rate of convergence result obtained in [6] for the particle approximation method that grounds SDM numerical algorithm. Section 1.3 presents some SDM preliminary simulation of the energy production of the “Parc de Bonneval” wind farm operated by Zephyr ENR. With the initial and boundary conditions generated from the MERRA reanalysis, we evaluate SDM result against measurements collected at the wind farm.

1.2 Stochastic Lagrangian Models

Lagrangian approaches for turbulent flow are already well established for turbulent subgrid-scale modeling that refers to the representation of the small-scales of the

flow that cannot be adequately resolved solely on a computational mesh. In the context of atmospheric flow, the so-called Lagrangian Particle Dispersion Models (LPDM) are widely used for the analysis of air pollutants dispersion (see e.g. Stohl [21] and the references therein). Such method adopts perspective of a 'air parcel' by tracking a number of fictitious particles (with position X_t) released into a flow field:

$$dX_t = \overline{U}(t,X_t)dt + u(t)dt \tag{1.1}$$

where $u(t)$ is a random fluctuation of the mean velocity $\overline{U}$, given for example by a LES computation. The velocity fluctuation is modeled with stochastic differential equation (SDE) of various degrees of complexity according to the involved representations, but generally starting from the simplest Langevin model

$$du(t) = -\frac{u(t)}{T}dt + \sqrt{C_0\varepsilon(t,X_t)}dW_t \tag{1.2}$$

where the stochastic (or fast) part of the motion is described by the 3-dimensional Brownian motion W, amplified with the turbulent pseudo dissipation of the flow ε. Stochastic description of particles in turbulent flows are also well established in the case of disperse two-phase flows and may concern many other applications (see e.g. Minier [11]).

The SDM methodology also makes use of the air parcel view point but it replaces the given coarse mean velocity in the particle velocity dynamics (1.1) by a statistical mean velocity $\langle \mathrm{U} \rangle$ computed by solving locally a Lagrangian probability density function (PDF) model. This approach relies on the so-called fluid particle approach developed in the seminal work of S. Pope ([15], see also [12] and the references therein). In this approach, a fluid-particle, or virtual fluid parcel with position, instantaneous velocity and temperature state $(\mathrm{X}_t, \mathrm{U}_t, \theta_t)$ is described as the solution of a stochastic differential equation generically of the form

$$\begin{aligned} d\mathrm{X}_t =& \mathrm{U}_t dt, \\ d\mathrm{U}_t =& -\frac{1}{\rho}\nabla_x \langle \mathscr{P} \rangle (t,\mathrm{X}_t)dt - G(t,X_t)\left(\mathrm{U}_t - \langle \mathrm{U} \rangle (t,\mathrm{X}_t)\right)dt \\ &+ F_t dt + \sqrt{C(t,\mathrm{X}_t)\varepsilon(t,\mathrm{X}_t)}dW_t, \\ d\theta_t =& D_1(t,\mathrm{X}_t,\theta_t)dt + D_2(t,\mathrm{X}_t,\theta_t)d\widetilde{W}_t. \end{aligned} \tag{1.3}$$

$(W,\widetilde{W})$ is a $(3d \times 1d)$-Brownian motion, and we refer to Section 1.3.1 for the SDM version of (1.3) in the ABL. In SDM, the Eulerian statistical average of the velocity, classically represented by the bracket operator $\langle \mathrm{U} \rangle$ in RANS approaches or $\widetilde{\mathrm{U}}$ in LES approaches, is recovered as the probabilistic conditional expectation[1] of the particle velocity U_t, knowing that its position X_t is x, under the probability $\mathbb{P}$ of the model (1.3), provided with expectation symbol $\mathbb{E}$:

$$\langle \mathrm{U} \rangle (t,\mathrm{x}) := \mathbb{E}\left[\mathrm{U}_t | \mathrm{X}_t = \mathrm{x}\right], \tag{1.4}$$

[1] We consider here only the case of constant mass density flow, for the sake of clarity.

More generally, for any integrable function f, we set

$$\langle f(\mathrm{U},\theta)\rangle(t,\mathrm{x}) := \mathbb{E}\left[f(\mathrm{U}_t,\theta_t)|\mathrm{X}_t=\mathrm{x}\right]. \tag{1.5}$$

Equivalently, in term of PDF modeling approach (see [18] for further details) the Eulerian density ρ_{Euler} at time t and at a given x is recovered as the Lagrangian conditional density ρ, knowing the event $\{\mathrm{X}_t=\mathrm{x}\}$:

$$\langle f(U,\Theta)\rangle(\mathrm{x},t) = \int_{\mathbb{R}^3\times\mathbb{R}} f(u,\theta)\rho_{\text{Eu}}(t,\mathrm{x};u,\theta)dud\theta = \frac{\int_{\mathbb{R}^3\times\mathbb{R}} f(u,\theta)\rho(t,x,u,\theta)dud\theta}{\rho(t,x)}$$

where $\rho(t,\cdot,\cdot,\cdot)$ is the probability density law of the random variable $(\mathrm{X}_t,\mathrm{U}_t,\theta_t)$, and $\rho(t,x)=\int_{\mathbb{R}^3\times\mathbb{R}}\rho(t,x,u,\theta)d\theta dx$. Thus, the coefficients of the stochastic equation (1.3) are (function of, or derivatives of) statistical averages $\langle u^{(i)}\rangle$, $\langle u^{(i)}u^{(j)}\rangle$, defined as in (1.5). Here and in the sequel we denote by $\mathrm{U}_t=(u_t^{(1)},u_t^{(2)},u_t^{(3)})$.

1.2.1 Numerical analysis of SLM: particle approximation

Solution of nonlinear SDEs with coefficients that can be written as expectations of the unknowns, can be obtained (under appropriated regularity hypothesis) as the mean field limit of a linear system of N-interacting particles, as N tends to infinity. Such particle approximation principle is at the basis of the SDM numerical method. (see e.g. Bossy [3] for a introductory review). We detail this principle in the simplified equation

$$\begin{aligned}\mathrm{X}_t &= \mathrm{X}_0 + \int_0^t \mathrm{U}_s ds\\ \mathrm{U}_t &= \mathrm{U}_0 + \int_0^t \mathbb{E}[b(\mathrm{U}_s)|X_s]ds + \sigma W_t,\end{aligned} \tag{1.6}$$

preferably to the complex model (1.3), but the resulting algorithm is similar. The wellposedness of (1.6) is establised in [5]. Particle approximation for the solution of (1.6) uses kernel regression estimator for the conditional expectation function $x\mapsto\mathbb{E}[b(\mathrm{U}_t)|\mathrm{X}_t=x]$. Such estimators use local averaging estimates on a N-particle set $(X_t^i,U_t^i,i=1,\dots,N,t\in[0,T])$. Typically:

$$\mathbb{E}[b(\mathrm{U}_t)|\mathrm{X}_t=x] \quad \text{is approximated by} \quad \sum_{i=1}^N \mathscr{W}_{N,i}(x)b(U_t^i) \tag{1.7}$$

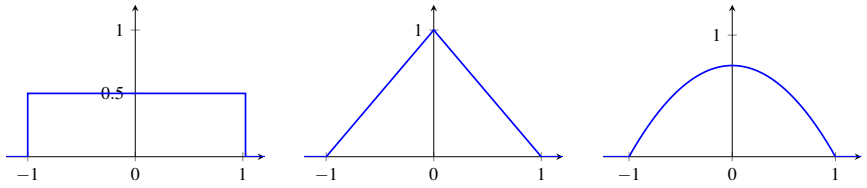
where well known propositions for the weights $\mathscr{W}_{N,i}(x)$ are Nadaraya-Watson kernel estimator

$$\mathscr{W}_{N,i}(x) = \frac{K_\varepsilon(x-X^i)}{\sum_{j=1}^N K_\varepsilon(x-X^j)}, \tag{1.8}$$

for a well chosen kernel $K_\varepsilon(x) = K(\frac{x}{\varepsilon})$, and partitioning (or mesh) estimator

$$\mathscr{W}_{N,i}(x) = \frac{\mathbb{1}_{\{X^i \in \mathscr{B}_{M,j}\}}}{\sum_{k=1}^{N} \mathbb{1}_{\{X^k \in \mathscr{B}_{M,j}\}}}, \quad \text{for } x \in \mathscr{B}_{M,j} \tag{1.9}$$

given a M-partition $\mathscr{P}_M = \{\mathscr{B}_{M,1}, \mathscr{B}_{M,2}, \ldots, \mathscr{B}_{M,M}\}$ (or a mesh) of the space domain. It is worth to notice that the algorithm complexity of a particle system based on kernel estimator is up to $\mathscr{O}(N^2)$ whereas the partitioning estimator version is up to $\mathscr{O}(N)$ (see also Section 1.2.3). We retained this last solution for SDM together with some refinement of Particle-in-cell (PIC) technics (see further details in [2]). With both M-partitioning or kernel regression estimators, the mean averages are computed with the help of two approximation parameters: N the number of particles to simulate and ε the characteristic size of the partition or the characteristic size of the support of the kernel K when it is applied on particles.



(a) Naive kernel: $K(x) = \frac{1}{2}\mathbb{1}_{\{\|x\|\le 1\}}$

(b) Triangular kernel: $K(x) = (1-\|x\|)\mathbb{1}_{\{\|x\|\le 1\}}$

(c) Epanechnikov kernel: $K(x) = (1-\|x\|^2)\mathbb{1}_{\{\|x\|\le 1\}}$

Fig. 1.1: Some examples of normalized kernel functions K.

In [6], Bossy and Violeau prove the theoretical optimal rate of convergence for the particle approximation of the solution of (1.6) that gives the relationship between the two parameters N and ε in order to achieve the optimal reduction of bias. This is the first result of this kind and to make the difficulty of the mathematical analysis more affordable, the boundary conditions are simplified by assuming them periodic. In a Lagrangian setting, periodic condition means that the particles stay in a torus domain fixed equal to $\mathscr{D} = [0,1]^d$ without loss of generality, and the model in (1.6) becomes:

$$\begin{aligned} \mathrm{X}_t &= \Big[\mathrm{X}_0 + \int_0^t \mathrm{U}_s\, ds\Big] \mod 1 \\ \mathrm{U}_t &= \mathrm{U}_0 + \int_0^t B[\mathrm{X}_s;\rho_s]\, ds + W_t, \quad \text{and } \rho_t \text{ is the density law of } (\mathrm{X}_t, \mathrm{U}_t), \end{aligned} \tag{1.10}$$

where, for the need of the approximation analysis we have rewrite $\mathbb{E}[b(\mathrm{U}_t)|\mathrm{X}_t]$ with $B[\mathrm{X}_t;\rho_t]$ for $(x,\gamma) \mapsto B[x;\gamma]$ defined for all probability density γ by

$$B[x;\gamma] = \frac{\int_{\mathbb{R}^d} b(v)\gamma(x,v)\, dv}{\int_{\mathbb{R}^d} \gamma(x,y) dv} \mathbb{1}_{\{\int_{\mathbb{R}^d} \gamma(x,y) dv > 0\}}.$$

The welposedness of (1.10) is a straightforward extension of the welposedness of (1.6). The associated particle approximation system $((X^{i,N},U^{i,N}),N\geq 1)$ is defined as the solution to

$$\begin{aligned} X_t^{i,N} &= \left[X_0^i + \int_0^t U_s^{i,N} ds\right] \mod 1, \\ U_t^{i,N} &= U_0^i + \int_0^t B_\varepsilon[X_s^{i,N};\bar{\mu}_s^{N,\varepsilon}]\,ds + W_t^i, \\ \bar{\mu}_t^{N,\varepsilon} &= \frac{1}{N}\sum_{j=1}^N \delta_{\{(X_t^{j,N},U_t^{j,N})\}} \text{ is the particles empirical measure} \end{aligned} \tag{1.11}$$

where the kernel regression version B_ε of B is defined for all density γ by

$$B_\varepsilon[x;\gamma] := \frac{\int_{[0,1]^d\times\mathbb{R}^d} b(v)K_\varepsilon(x-y)\,\gamma(y,v)dy\,dv}{\int_{[0,1]^d\times\mathbb{R}^d} K_\varepsilon(x-y)\,\gamma(y,v)dy\,dv}\mathbb{1}_{\{\int_{\mathbb{R}^d}\gamma(x,y)dv>0\}},$$

and $B_\varepsilon[x;\bar{\mu}_t^{N,\varepsilon}]$ is the dynamical version of the approximation in (1.7),(1.8). The $(W_t^i, t\leq T, 1\leq i\leq N$ are independent Brownian motions valued in $\mathbb{R}^d$, and independent from the initial variables $(X_0^i,U_0^i,1\leq i\leq N)$, independent, identically distributed with initial law ρ_0. The nonlinear equation (1.10) is thus replaced with the linear interacting system (1.11) (of dimension $2dN$) that it is possible to simulate with the help of a time-discretisation Euler scheme (see (1.17)). This algorithm is at the basis of the so-called Stochastic Lagrangian numerical algorithm (see e.g. Pope [17] and for the SDM method [2]).

The theoretical convergence analysis

A pertinent criterion for the evaluation of the algorithm (1.11) is the mean error on the conditional expectation used in the drift computation all along the time loop:

$$\mathbb{E}\left|\mathbb{E}[f(U_t)|X_t=x] - \frac{\sum_{j=1}^N f(U_t^{j,N})K_\varepsilon(x-X_t^{j,N})}{\sum_{j=1}^N K_\varepsilon(x-X_t^{j,N})}\right| \tag{1.12}$$

for any f measurable bounded on $\mathbb{R}^d$, at any time $T>0$, at any point x. We reduce the evaluation across the space domain by considering this error (1.12) seen by the particles (or weighted by the particles position distribution ρ_T):

$$\text{Error}_{L^1_{\rho_T}(\mathscr{D};L^1(\Omega))} := \int_{\mathscr{D}} \mathbb{E}\left|\mathbb{E}[f(U_T)|X_T=x] - \frac{\sum_{j=1}^N f(U_T^{j,N})K_\varepsilon(x-X_T^{j,N})}{\sum_{j=1}^N K_\varepsilon(x-X_T^{j,N})}\right|\rho_T(x)dx \tag{1.13}$$

and we denote by

$$x \mapsto F_\varepsilon[x;\bar{\mu}_T^{\varepsilon,N}] := \frac{\sum_{j=1}^N f(U_T^{j,N})K_\varepsilon(x-X_T^{j,N})}{\sum_{j=1}^N K_\varepsilon(x-X_T^{j,N})} \tag{1.14}$$

the corresponding kernel approximation function, where $\bar{\mu}_T^{\varepsilon,N}$ is the empirical measure of particles as in (1.11).

Theorem 1. *(see Bossy Violeau [6].) Assume the following:*

(i) f and b are smooth and bounded functions with bounded derivatives
(ii) the kernel K is positive and bounded, with compact support in $\{x; \|x\| \leq 1\}$.
(iii) The initial law has a density ρ_0 that we assume smooth and to be bounded below by a constant $\zeta > 0$.

Then for any $1 < p < 1 + \frac{1}{1+3d}$ and $c > 0$, there exists a constant C such that for all $\varepsilon > 0$ and $N > 1$ satisfying $(\varepsilon^{(d+2)} N^{\frac{1}{p}})^{-1} \leq c$, we have for all $1 \leq i \leq N$

$$Error_{L^1_{\rho_T}(\mathscr{D};L^1(\Omega))} \leq C\Big(\varepsilon + \frac{1}{\varepsilon^{(d+1)}N} + \frac{1}{\varepsilon^{(d+1)p}N} + \frac{1}{(\varepsilon^d N)^{\frac{1}{p}}} + \frac{1}{\varepsilon^{\frac{dp}{2}}\sqrt{N}}\Big). \tag{1.15}$$

The optimal rate of convergence is achieved for $N = \varepsilon^{-(d+2)p}$ and

$$Error_{L^1_{\rho_T}(\mathscr{D};L^1(\Omega))} \leq CN^{-\frac{1}{(d+2)p}}. \tag{1.16}$$

Notice that p can be chosen almost equal to one. The global error given in (1.15) is a combination of several sources of approximations. First, the $\mathscr{O}(\varepsilon)$ term corresponds to the smoothing error for F. The $\mathscr{O}(\varepsilon^{-\frac{dp}{2}}\sqrt{N}^{-1})$ term is the Monte Carlo variance contribution to the error, next $\mathscr{O}((\varepsilon^d N)^{-\frac{1}{p}})$ is the error due to the replacement of the law ρ_t by the empirical measure $\bar{\mu}_t^{N,\varepsilon}$. There is also the approximation due to the replacement of the position of the exact process as the location where the conditioned expectation is computed by the position of a numerical particle. This is a part of the statistical error, (the use of the Nadaraya Watson estimator to compute the expectation) in $\mathscr{O}(\varepsilon + \frac{1}{\varepsilon^{d+1}N} + \frac{1}{\varepsilon^{(d+1)p}N})$.

1.2.2 Empirical numerical analysis

We analyze the convergence of the algorithm on numerical experiments in order to verify that the claimed convergence rate in Theorem 1 is optimal. For both computational time reason and clarity of the graph presentation, we limit our experiments to $d = 2$. (the wind farm simulation presented in Section 1.3.1 is for $d = 3$).

Numerical experiments imply to proceed with an Euler scheme. We decompose the time interval $[0,T]$ into M time steps of length $\Delta t := \frac{T}{M}$ and we introduce the time discretization of the interacting particle process:

$$\begin{cases} X_t^{i,N,\Delta t} = \left[X_0^i + \displaystyle\int_0^t U_{\eta(s)}^{i,N,\Delta t} ds \right] \mod 1, \\ U_t^{i,N,\Delta t} = U_0^i + \displaystyle\int_0^t B_\varepsilon[X_{\eta(s)}^{i,N,\Delta t}; \bar{\mu}_{\eta(s)}^{N,\varepsilon,\Delta t}]\, ds + W_t^i, \quad \bar{\mu}_t^{N,\varepsilon,\Delta t} = \dfrac{1}{N}\displaystyle\sum_{j=1}^N \delta_{(X_t^{j,N,\Delta t}, U_t^{j,N,\Delta t})}, \end{cases} \tag{1.17}$$

for all $1 \le i \le N$ and $t \in [0,T]$ where $\eta(t) := \Delta t \left\lfloor \frac{t}{\Delta t} \right\rfloor$ is the Δt-step time function. For all time step $k\Delta t$, $0 \le k \le M$, each random variable $(X_{(k+1)\Delta t}^{i,N,\Delta t}, U_{(k+1)\Delta t}^{i,N,\Delta t})$ is computed from the values of all the variables $(X_{k\Delta t}^{j,N,\Delta t}, U_{k\Delta t}^{j,N,\Delta t})$, $1 \le j \le N$. It is clear that this kernel estimate based algorithm has a complexity of order $\mathscr{O}(M)\,\mathscr{O}\left(N^2\right)$. The major drawback is that the computation of the drift at a point x requires a loop over all the N particles, even if they do not contribute to the final result. For this reason, we preferably use the alternative particle-mesh algorithm for SDM.

The test case description

We choose a potential P that models a pressure force as a non trivial but static external force given by

$$P(x,y) = \frac{1}{2\pi}\cos(2\pi x)\sin(2\pi y) - \frac{1}{2}x, \text{ for all } (x,y) \text{ in } \mathscr{D}.$$

The conditional drift B is a mean reverting term such that

$$B[x,u;\gamma] = \frac{\int_{\mathbb{R}^d}(v-2u)\gamma(x,v)\,dv}{\int_{\mathbb{R}^d}\gamma(x,v)\,dv} \text{ for all } (x,u) \text{ in } \mathscr{D}\times\mathbb{R}^2 \text{ and all } \gamma \text{ in } \mathscr{P}(\mathscr{D}\times\mathbb{R}^2).$$

Note that in this case, for all (x,u) in $\mathscr{D}\times\mathbb{R}^d$,

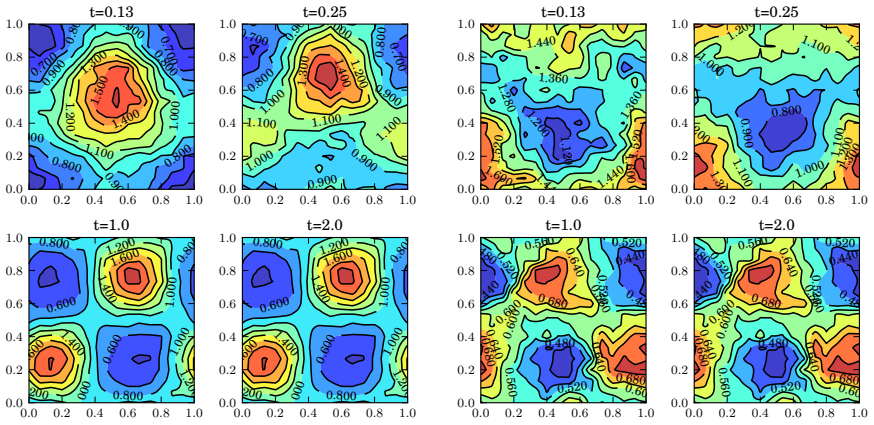
$$B[x,u;\rho_t] = \mathbb{E}[\mathrm{U}_t|\mathrm{X}_t = x] - 2u$$

and we solve for $t \le T = 2$,

$$\begin{cases} \mathrm{X}_t = \left[\mathrm{X}_0 + \displaystyle\int_0^t \mathrm{U}_s ds \right] \mod 1 \\ U_t = \mathrm{U}_0 - \displaystyle\int_0^t \nabla P(\mathrm{X}_s)\, ds + \int_0^t B[\mathrm{X}_s, \mathrm{U}_s; \rho_s]\, ds + W_t, \quad \rho_t \text{ is the density of } (\mathrm{X}_t, \mathrm{U}_t). \end{cases} \tag{1.18}$$

The initial distribution ρ_0 of $(\mathrm{X}_0, \mathrm{U}_0)$ is such that X_0 has a Gaussian distribution on $\mathbb{T}^d$ with variance σ^2 (i.e. $X_0 = \sigma Z \mod 1$, $\sigma^2 = 0.3$) and U_0 is a centered Gaussian random variable independent from X_0, with variance $v^2 = 1$. On Figure 1.2, we represent the time evolution of the *particles mass* density $\rho_t(x) = \int_{\mathbb{R}^2}\rho_t(x,u)du$ of the process X_t distributed in the torus (plot (a)), as well as the turbulent kinetic $tke(t,x) = \frac{1}{2}\mathbb{E}[(U_t - \mathbb{E}[U_t|X_t = x])^2|X_t = x]$ (plot (b)). We can observe that the density is clearly non uniform in space, and we expect this should put some stress on the

estimation of the mean fields in low density areas. Moreover, although starting from



(a) Density (b) Turbulent kinetic energy

Fig. 1.2: Evolution of the density and TKE for (X_t, U_t), $[N = 10^5, \varepsilon = 16^{-1}]$.

Gaussian initially, the density quickly converges to a stationary state and this allows to fix the final time to $T = 2$ for all the error analysis simulations, with $M = 128$ time steps. The kernel regression is performed with the Epanechnikov kernel (see Figure 1.1-(c)) and $\varepsilon = \frac{1}{16}$.

Expected L^1 error of the kernel method

We focus our attention on the *expected* L^1 *error* defined in (1.13). In order to estimate this quantity, we need to proceed with some approximations on the integral. In the following, we write $\pi^{\Delta x}(g)$ for the spline-interpolated function g on a grid with mesh size Δx. The reference numerical solution for $\mathbb{E}[f(U_T)|X_T = x]$ is approximated by the splined mean fields defined by

$$\overline{F_{\underline{\varepsilon}}\big[x;\bar{\mu}_T^{\underline{\varepsilon},\bar{N}}\big]}^{\Delta x} := \pi^{\Delta x}(\overline{F_{\underline{\varepsilon}}\big[\cdot;\bar{\mu}_T^{\underline{\varepsilon},\bar{N}}\big]})(x) \tag{1.19}$$

for a large number of particles $\bar{N}$ and a sufficiently small window parameter $\underline{\varepsilon}$. The numerical approximation is also splined to ease the integration step:

$$F_{\varepsilon}^{\Delta x}[x;\bar{\mu}_T^{\varepsilon,N}] := \pi^{\Delta x}\big(F_{\varepsilon}[\cdot;\bar{\mu}_T^{\varepsilon,N}](x)\big). \tag{1.20}$$

The reference mass density $\rho_T(x)$ is also estimated by using the Monte Carlo mean of kernel density estimation:

$$\bar{\rho}_T(x) := \frac{1}{N_{\text{mc}}} \sum_{k=1}^{N_{\text{mc}}} \frac{1}{\bar{N}} \sum_{j=1}^{\bar{N}} K_{\underline{\varepsilon}}(x - X_T^{j,\underline{\varepsilon},\bar{N}}(\omega_k)), \quad \text{and} \quad \bar{\rho}^{\Delta x}(x) := \pi^{\Delta x}(\bar{\rho}) \tag{1.21}$$

where the ω_k represent N_{mc} independent realizations of the simulation. The computation of the integral of splined functions can be carried out very precisely over regular grids with the help of numerical libraries. All that remains is to evaluate the expected splined L^1 error by means of a Monte Carlo simulation:

$$\text{Error}_{L^1_{\rho_T}(\mathscr{D};L^1(\Omega))} \sim \frac{1}{N_{mc}} \sum_{k=1}^{N_{mc}} \int_{\mathscr{D}} \left| \overline{F_{\underline{\varepsilon}}[x;\bar{\mu}^{\underline{\varepsilon},\bar{N}}]}^{\Delta x} - F_{\varepsilon}^{\Delta x}[x;\bar{\mu}^{\varepsilon,N}(\omega_k)] \right| \bar{\rho}^{\Delta x}(x)\, dx.$$

In Figure 1.3, we plot the expected L^1 error calculated as above as a function of the

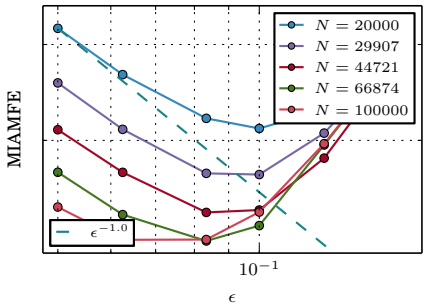


Fig. 1.3: L^1 error as a function of ε for different number of particles N

window parameter ε for different total number of particles N. We clearly recognize the typical bias-variance trade-off behavior. We can also notice that the asymptotic slope of the error when ε tends to zero is very close to -1 for a log-log scale. We expect the error to behave like $\mathscr{O}(\varepsilon + \frac{C}{\varepsilon^3 N} + \frac{C}{\varepsilon\sqrt{N}})$. Then, it seems reasonable to infer that the term of order $\mathscr{O}\left(\frac{1}{\varepsilon\sqrt{N}}\right)$ related to the variance of the stochastic integral dominates the L^1 error. Recall, however, that our theoretical analysis of the error is valid under the constraint $\frac{1}{\varepsilon^{d+2}N^{1/p}} \leq c$, for some positive constant c, so we cannot rigorously extend the bound to an asymptotic analysis when ε decreases to zero. Finally, we can observe that the slope of L^1 is bounded by one when N is sufficiently large and ε becomes large. This is in complete agreement with the bounds in Theorem 1 although this figure does not explain the relative contribution of the smoothing error and the kernel estimation error in the total L^1 error.

We can also consider the expected L^1 error as a function of $\frac{N}{\varepsilon^d}$, as in Figure 1.4. Note that $\frac{N}{\varepsilon^d}$ loosely represents the number of particles in interaction with a given particle (for compact support kernel functions), and is often referred to as the number of particle "per cell" (denoted N_{pc}), especially in the case of partitioning

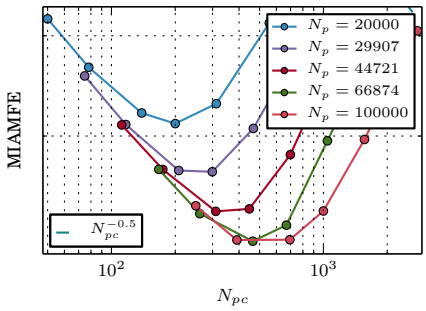


Fig. 1.4: L^1 error as a function of ε for different densities of particles $\frac{N}{\varepsilon^d}$

estimates. This figure illustrates again the concept of bias-variance trade-off and its relation with the number of particle per cell. N_p is the total number of particles in this figure and the following. For a given small number of particle per cell (compared to the optimal number of particle per cell), we can notice that the L^1 error is almost independent of the absolute value of ε. This clearly shows that the variance is directly related to the number of particles used in the computation of the regression estimate. On the contrary, when the number of particle per cell becomes large and the bias dominate, the L^1 error becomes smaller with ε. The convergence of

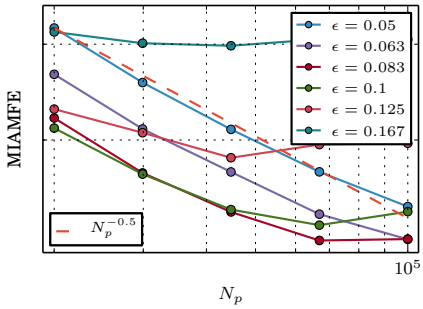


Fig. 1.5: L^1 error as a function of the total number of particles, for different value of ε

the error with respect to the number of particles N can be observed in Figure 1.5. When ε is sufficiently small, we notice as expected a convergence of order $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$, related to the reduction of the variance component of the error. On the other hand, when ε is large, increasing the number of particle does not reduce the error as the bias dominates.

In the typical bias-variance trade-off, one may be interested to find the optimal value of ε that minimizes the expected L^1 error for a given number of particles. From the simulations we ran for different couples (ε, N) of parameters, we plot the surface of the error in Figure 1.6 (left). We can then plot the curve of optimal ε as a function of the number of particles which is very close to $\frac{1}{\varepsilon^3}$ (for $d = 2$). This result is in-line with what we expected from Theorem 1 where the optimal value of window size is given by $N^{-\frac{1}{d+2}}$. Moreover, if we plot the error associated with the optimal couple of parameters as a function of ε, we can observe the optimal experimental rate of convergence of the algorithm. In Theorem 1, we have an error of order $\mathscr{O}\left(N^{-\frac{1}{d+2}}\right)$. In Figure 1.6 (right), we observe a rate of order close to $-\frac{1}{4}$ to $-\frac{1}{3}$.

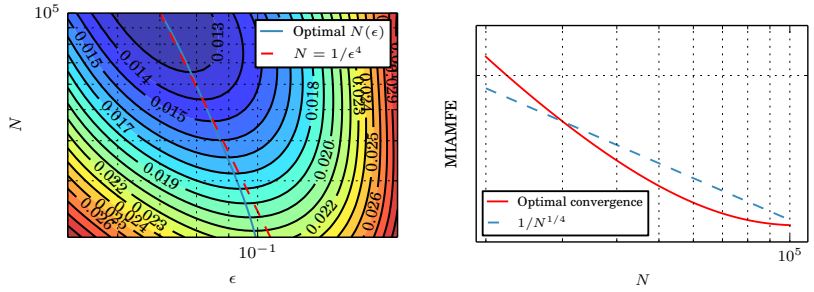


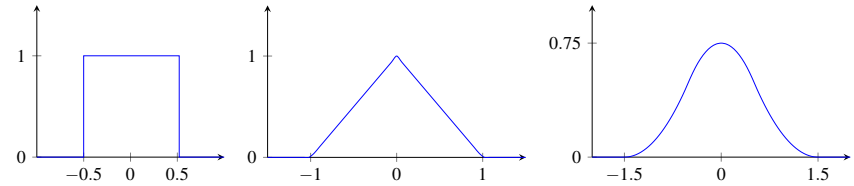
Fig. 1.6: L^1 error as a function ε and N. and optimal rate of convergence for the L^1 error

1.2.3 Particle in mesh method

We end this section with some experiments on the particle-mesh version of the algorithm.

The principle of the Particle-Mesh methods is to aggregate the N scattered data points $(X^i, f(U^i))$, for $1 \leq i \leq N$ onto a regular mesh covering the simulation domain $\mathscr{D}$, thus reducing the size of the data set to the number of nodes in the mesh. The mean field is evaluated from the mesh charges at each particle position using standard regression techniques as in (1.7),(1.9). If we design the charge assignment and the force interpolation operation such that they can be performed in constant time for each particle, the Particle-Mesh algorithm has a $\mathscr{O}(M)\,\mathscr{O}(N)$ complexity, i.e. it has linear complexity with respect to the total number of particles. This is a tremendous improvements over the previous kernel regression method, and the speed-up is not only theoretical but is actually achieved in practical simulations.

The drawback of this approach is that it introduces new sources of numerical errors, and unlike classical particle computer simulations, increasing the number of nodes in the mesh does not necessarily reduce the error if the total number of particles is left unchanged. Moreover, refining the mesh increases the computational cost, so it is particularly important to be able to reduce the errors for a given mesh size in order to achieve the best compromise between quality and computational cost. In this regard, we will consider three charge assignment and interpolation functions that are designed to be optimal according to smoothness and spatial localization of errors criteria: the Nearest Grid Point (NGP), the Cloud in Cell (CIC), and the triangular Shaped Cloud (TSC) (see Figure 1.7 for details).



(a) Nearest Grid Point: $K_h(x) := \mathbb{1}_{\{\|x\| \le \frac{h}{2}\}}$

(b) Cloud in Cell: $K_h(x) := (1 - \|x\|)\mathbb{1}_{\{\|x\| \le h\}}$

(c) Triangular Shaped Cloud: $K_h(x) := (\frac{3}{4} - \frac{x^2}{h^2})\mathbb{1}_{\{\|x\| \le \frac{h}{2}\}} + \frac{1}{2}(\frac{3}{2} - \frac{|x|}{h})^2\mathbb{1}_{\{\frac{h}{2} \le \|x\| \le \frac{3h}{2}\}}$

Fig. 1.7: Charge assignment functions (from left to right: NGP, CIC, TSC)

Charge assignment

Consider a mesh of cell size h (also called window size). Let x_i be the position of the i-th node. Then the charge c_i and the charge density d_i assigned at node i are defined by

$$c_i := \frac{1}{N}\sum_{j=1}^{N} f(U^j)K_h(x_i - X^j) \quad d_i := \frac{1}{N}\sum_{j=1}^{N} K_h(x_i - X^j),$$

where K is a charge assignment function. By definition of c_i and d_i the ratio $\frac{c_i}{d_i}$ is simply the kernel regression estimate at the node point x_i as in (1.7):

$$\mathbb{E}[b(\mathrm{U}_t)|\mathrm{X}_t = x] \sim \frac{c_i}{d_i} = \frac{\frac{1}{N}\sum_{j=1}^{N} f(U^j)K_h(x_i - X^j)}{\frac{1}{N}\sum_{j=1}^{N} K_h(x_i - X^j)}.$$

The computation of the mesh charge values can be performed efficiently in $\mathcal{O}(N)$ with an outer loop on the particles and the use of a mesh localization procedure that makes it possible to loop only on the nodes charged by a given particle.

Of course, it is important that the localization of the particle in the mesh and the computation of the list of nodes charged by the particle be performed in constant time. In practice, the lists of neighbor cells are computed once and for all (in linear time) at the beginning of the procedure to speed up the execution of the algorithm.

In Figure 1.8, we measured the influence of the regularity order of the charge assignment function K_h. Aside from the smoothing aspect of the obtained velocity field, we can observe a gap between the error produced by the partitioning estimates (corresponding to NGP assignment charge) and the higher order CIC or TSC functions, and CIC appears to be a good compromise between the error level and the ease of implementation.

1.3 Wind farm simulation experiment with SDM

Our SDM model has been evaluated against measurements collected at a wind farm located in Bonneval, a small town 100 km Southwest of Paris, France (48.20°N and 1.42°E). The wind farm is operated by Zephyr ENR, a private company managing 5 other wind farms. The Bonneval wind farm, called "Parc de Bonneval", has been implemented in 2006 and is composed of six Vestas V80-2 MW turbines. In order to evaluate the simulations with the data collected at "Parc de Bonneval", wind turbines have been numerically integrated in SDM based on an actuator disk model. This model allows the simulation of the dynamical effect of the presence of wind turbines, in the form of trailing wakes, as well as the computation of the wind energy production.

1.3.1 SDM for atmospheric boundary layer simulation

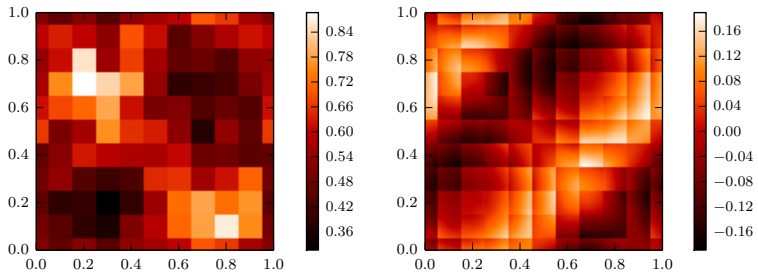
We run SDM for winter day of December 22th 2016, with the equation (1.3) configured for the case of the neutral atmosphere hypothesis. Here and in the sequel we denote by

$$\mathrm{U}_t = (u_t^{(1)}, u_t^{(2)}, u_t^{(3)}) = (u_t, v_t, w_t)$$

the velocity components (with numbering or with letters, depending on how it is convenient in the equations), and for the components of the instantaneous turbulent velocity:

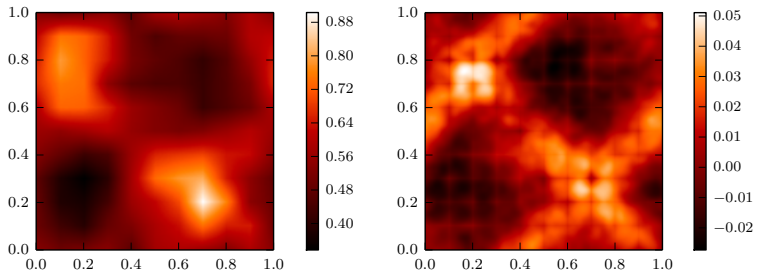
$$\mathrm{U}_t - \langle \mathrm{U} \rangle(t, \mathrm{X}_t) = (u_t'^{(1)}, u_t'^{(2)}, u_t'^{(3)}) = (u_t', v_t', w_t').$$

We start from the General Langenvin model introduced by Pope [18]:



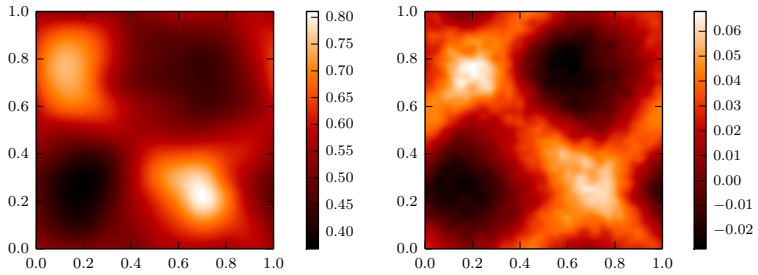
(a) Velocity norm for the NGP scheme

(b) Velocity error for the NGP scheme



(c) Velocity norm for the CIC scheme

(d) Velocity error for the CIC scheme



(e) Velocity norm for the TSC scheme

(f) Velocity error for the TSC scheme

Fig. 1.8: Velocity norm and average error for the NGP, CiC and TSC schemes

$$\begin{cases} dX_t = \mathrm{U}_t dt, \quad \text{with } \mathrm{U}_t = (u_t^{(i)}, i = 1,2,3) \text{ and } \mathbf{u}_i'(t) = u_t^{(i)} - \langle u_t^{(i)} \rangle \\ du_t^{(i)} = -\partial_{x_i} \langle \mathscr{P} \rangle (t, X_t) dt + \left(\sum_j G_{ij} \left(u_t^{(j)} - \langle u^{(j)} \rangle \right) \right) (t, X_t) dt + \sqrt{C_0 \varepsilon(t, X_t)} dB_t^{(i)} \end{cases} \tag{1.22}$$

and we follow [22] in order to incorporate the effects of the wall blocking of normal velocity component and pressure reflection from the surface in our stand-alone PDF method, by adapting the Durbin elliptic relaxation method, [8]: we set

$$G_{ij} = -\gamma_{ij} - \frac{1}{2}\frac{\varepsilon}{tke}\delta_{ij}, \qquad C_0\varepsilon = \frac{2}{3}(\gamma_{ij})\langle u_i' u_j'\rangle$$
$$\gamma_{ij} = (1-\alpha tke)\gamma_{ij}^{\text{wall}} + \alpha tke\ \gamma_{ij}^{\text{homogeneous}}$$

The elliptic blending coefficient α solves near the ground:

$$L^2\nabla^2\alpha - \alpha = -\frac{1}{tke}$$

where L is a length scale defined as a maximum of the turbulent scale and the scale connected with dissipative eddies.

$$-\gamma_{ij}^{\text{homogeneous}} = -\frac{1}{2}(C_R - 1)\frac{\varepsilon}{tke}\delta_{ij} + C_2\frac{\partial\langle u^{(i)}\rangle}{\partial x_j}, \qquad \text{and } -\gamma_{ij}^{\text{wall}} = -7.5\frac{\varepsilon}{tke}n_i n_j$$

where n is the wall-normal unit vector. The coefficients C_0 and C_2 have to satisfy some realizability constraints (see [9], [16]).

We take benefit of the Lagrangian methodology to easily introduce complex terrain description when we specify the wall-boundary condition by a particle reflection at the wall (see [7, 13] for more details)

$$w_{\text{in}} = -w_{\text{out}}, \quad u_{\text{in}} = u_{\text{out}} - 2\frac{\langle u'w'\rangle}{\langle w'^2\rangle}w_{\text{out}}, \qquad v_{\text{in}} = v_{\text{out}} - 2\frac{\langle v'w'\rangle}{\langle w'^2\rangle}w_{\text{out}}.$$

$\langle u'w'\rangle$ and $\langle v'w'\rangle$ are computed according to the *log law of the wall* with

$$u_*(t,(x,y)) = \kappa\frac{\sqrt{\langle u\rangle^2 + \langle v\rangle^2}}{\log(z_c/z_0)}(t,(x,y))$$

Lagrangian actuator disk model

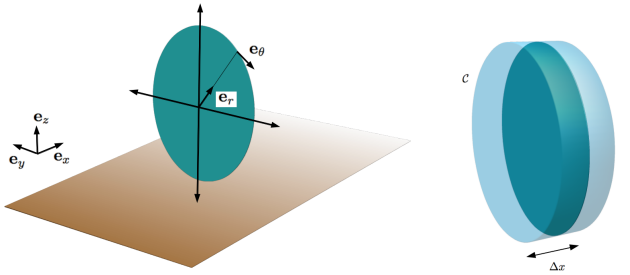
SDM method allows also to introduce local interaction with objects, in particular porous objects like actuator models for mills.

The SDM approach could be used with various actuator disck modelling options (see [4] and the references therein). In the Actuator Disc approach, each mill is represented as an immersed surface which concentrates all forces exerted by the mill on the flow. In the SDM context, the presence of wind mills is taken into account thanks to additional force f that represents the body forces that the blades exert on the flow in the stochastic differential equations that govern the movement of the particles. To this end, Equation (1.22), which governs the time evolution of the velocity $\mathrm{U}_t = (u_t, v_t, w_t)$ of a particle, is modified as follows:

$$
\begin{aligned}
d\mathrm{U}_t = &\left(-\frac{1}{\rho}\nabla_x\langle\mathscr{P}\rangle(t,\mathrm{X}_t)\right)dt + f(t,\mathrm{X}_t,\mathrm{U}_t)\,dt \\
&- G(t,\mathrm{X}_t)\Big(\mathrm{U}_t - \langle\mathrm{U}\rangle(t,\mathrm{X}_t)\Big)dt + C(t,\mathrm{X}_t)dW_t
\end{aligned}
\tag{1.23}
$$

where the term $f(t,\mathrm{X}_t,\mathrm{U}_t)$ represents the body forces of the turbine seen by the particle. We refer to [] for a detailed discussion on the turbine force terms implementation in the Lagrangian context (including nacelle and mast force terms).

For "Parc de Bonneval" wind farm simulation presented here, we choose a rather basic non rotating uniformly loaded actuator disc model. Such model can be easily parametrized with the characteristic data of thrust coefficient C_T and power coefficient C_p, provided by the turbine manufacturer, and varying with the dynamics of the inflow wind seen by each of the six Vestas V80 turbines.



(a) Local coordinates (b) The cylinder $\mathscr{C}$

Fig. 1.9: (a) The local reference frame at the actuator disc of the turbine, using cylindrical coordinates; (b) The cylinder $\mathscr{C}$ that extends the actuator disc. Mill forces are applied to particles that lie inside.

In order to derive the force f, it is convenient to define local reference frame of cylindrical coordinates centered at the hub of the turbine, with basis vectors $\mathbf{e}_x$, $\mathbf{e}_r$ and $\mathbf{e}_\theta$ as shown in Figure 1.9a. Assuming that the flow moves along the positive direction of the x axis, and that the turbine's main axis is aligned with the x axis, so that it faces the wind directly, the total thrust force exerted by the turbine is formally given by (see e.g. [])

$$F_x = -\frac{1}{2}\rho A C_T U_\infty^2 \mathbf{e}_x,$$

where $\mathbf{U}_\infty$ is the unperturbed velocity far upstream from the turbine's location, U_∞ is its norm, A is the surface area of the turbine's disc, ρ is the density of air, and C_T is a dimensionless, flow dependent parameter called the *thrust coefficient*. As in Réthoré et al. [19], the local velocity magnitude U_D is used instead of U_∞ and the thrust force expression becomes

$$F_x = -\frac{1}{2}\rho A C_T U_D^2 \mathbf{e}_x \quad \text{with} \quad U_D(t) = \mathbb{E}[\mathrm{U}_t^2 | \mathrm{X}_t \in D]. \tag{1.24}$$

In order to adapt this thrust force model to particles, the disc is extended to a cylinder $\mathscr{C}$ of length Δx and mass $\rho A \Delta x$ (see Figure 1.9b). The force per unit mass inside region $\mathscr{C}$ is then given by:

$$f_x = -\frac{1}{\Delta x} C_T U_D^2 \mathbb{1}_{\{x \in \mathscr{C}\}} \mathbf{e}_x. \tag{1.25}$$

The available power is computed following the same idea:

$$P = \frac{1}{2}\rho A C_p U_D^3.$$

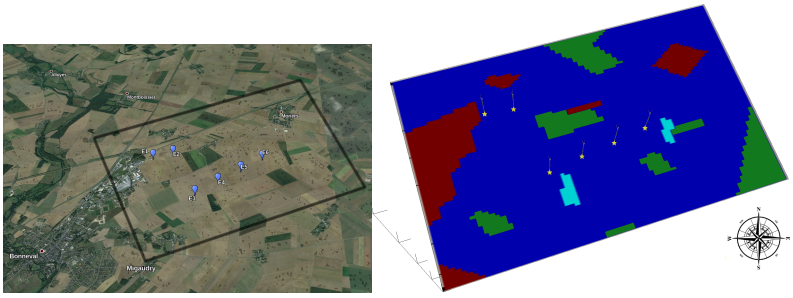
Numerical setup

The modeled domain is flat with a variable roughness length inferred from Google-Earth and lookup tables of roughness lengths for typical types of land-use. Four different roughness lengths have been used with respect to the land-use pattern shown in the Figure 1.10. The roughness length varies between 0.01 and 0.4 m. The characteristics of the numerical domain of the simulation and of the turbines are summarized in Table 1.1.

The initial and boundary conditions are generated from the MERRA reanalysis with a 3-hourly time sampling [10]. All fields are provided on the same 5/8 degree longitude by 1/2 degree latitude grid. The data used as initial and boundary conditions are those of the closest grid point located at 25 km Southwest of "Parc de Bonneval" (48°N and 1.25°E). The vertical mesh has 72 pressure levels but only the first three levels from the surface up to 970 hPa (about 400 m) are used. The pressure level coordinates are converted into altitude coordinates using the surface pressure from the MERRA reanalysis. The wind components are then interpolated onto the refined grid of SDM. The time step of the SDM simulation is 5 s. The profiles extracted from the MERRA reanalysis at the closest grid point are therefore interpolated linearly in time with a 5 s time sampling.

Case study description

The simulated case study corresponds to the 22^{th} December 2016. Figure 1.11 displays the time evolution of the measured wind direction, wind speed and wind energy production at the 6 turbines. The wind speed and direction are measured directly at "Parc de Bonneval" by anemometers located on the hub of each turbine. The wind energy production is also provided directly from the generator. The sampling period is 1 s but the data are averaged over a 10-min period. Those time series are used to evaluate SDM model performance.



(a) (b)

Fig. 1.10: (a) Aerial view of the "Parc de Bonneval" from Google-Earth; (b) Aerial view of the simulated wind farm. The pattern define the roughness length. Blue part represents farmland (0.04 m), red are small town (0.4 m), green are uncut grass (0.01 m), cyan are small forest (0.15 m). Yellow stars represent the turbines.

Simulation parameters	
Domain size x	3000 m
Domain size y	4787 m
Domain size z	408 m
75 cells in x	$\Delta x = 40$ m
75 cells in y	$\Delta y = 63.83$ m
85 cells in z	$\Delta z = 4.8$ m
Particles per cell	80
Final time is 24 h	Time step is 5 s

(a) Configuration of the simulation

Mill configuration	
Hub height	100 m
Radius	40 m
Nacelle radius	4 m
Rotational speed	1.75 rad s^{-1}

(b) Parameters of the mill

Table 1.1: Main parameters of the simulation.

This day is characterized by a strong wind blowing until 5:00 local time (LT). Between 5:00 and 16:00 LT, the wind speed weakens from 10 m s^{-1} to 2 m s^{-1}. It increases again up to 6 m s^{-1} and decreases down to less than 2 m s^{-1} in 2 hours. The wind shifts progressively from the South to the North. According to the position of the turbines in Figure 1.10 a wind direction around 230° lines up the turbines 3 to 6 and a direction around 250° lines up the turbines 1 and 2. Indeed we can see the wake effect in Figure 1.11 between 7:00 and 9:00 LT. This phenomenon decreases the production downstream by 50%.

Results

Figure 1.12 displays the time evolution of the simulated wind direction, wind speed and wind energy production at the 6 turbines. It can be directly compared to Figure 1.11. The time variability is well reproduced with a slightly increasing wind

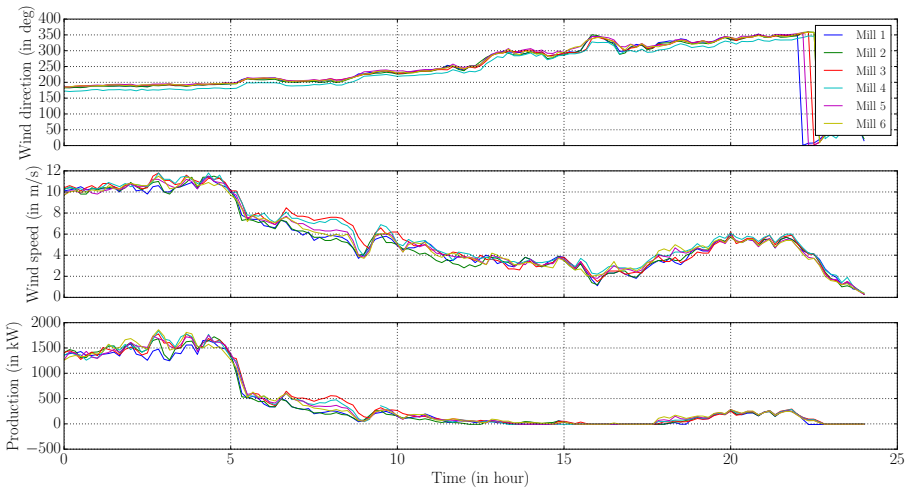


Fig. 1.11: Time evolution of "Parc de Bonneval" measurments during the 22^{th} December 2016

speed between 0:00 and 3:30 LT and a constant wind direction. The wind speed increases between 8 and 9.2 m s^{-1}. The simulated wind speed is slightly weaker than the measured wind speed which remains constant and equal to 10 m s^{-1} over this period of time. Such underestimation is caused by the initial and boundary conditions from MERRA reanalysis which provide a weaker wind speed at the hub height. The wind direction is also slightly biased by about 10°. The simulated wind speed then decreases at a similar rate than the observed wind speed. The short increase of the wind speed followed by a fast decrease between 15:00 and 23:00 LT is underestimated in the simulation as the wind speed peaks at about 3.4 m s^{-1} in the simulation versus 6 m s^{-1} in the measurements. The bias in wind direction disappears after 8:00 LT. Finally the high frequency variablilty is much too smooth in the simulation as the small domain size produces a strong control by the lateral boundary conditions in the domain.

However, the wake effect is well reproduced in the simulation between 7:00 and 12:00 LT. The magnitude is underestimated but the sheltering effect by the forefront turbines is clearly visible. The difference of wind speed between the forefront turbines and those located downstream is about 1-1.5 m s^{-1} in the simulation against 2 m s^{-1} in the measurements. Figure 1.13 displays a zoom between 6:00 and 13:00 LT of the measured and simulated wind direction, wind speed and wind energy production. In detail, the measured wind speed and energy production displays a continuously decrease between the forefront turbines and the most downstream turbines. At "Parc de Bonneval" we can distinguish two groups of wind turbines. The forefront turbine 3 with turbines 4, 5 and 6 downstream in the wake between 6:30 and 9:00 LT and forefront turbine 1 with turbine 2 downstream in the wake be-

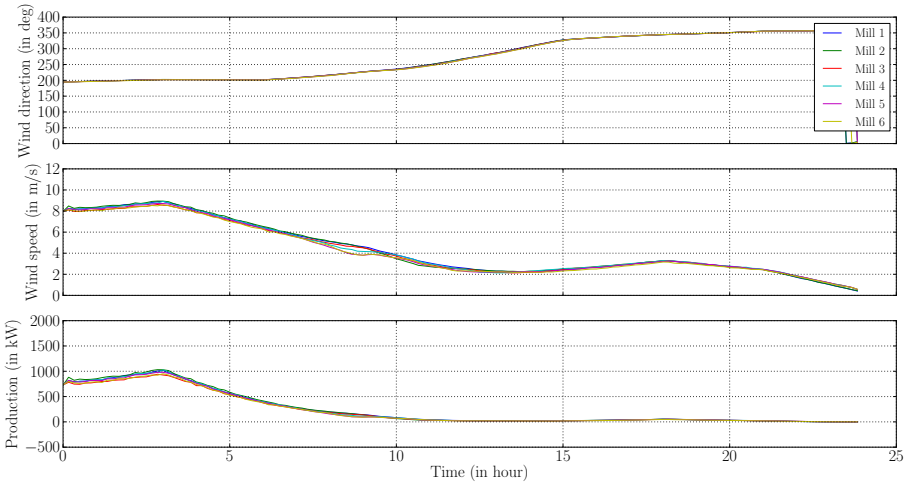


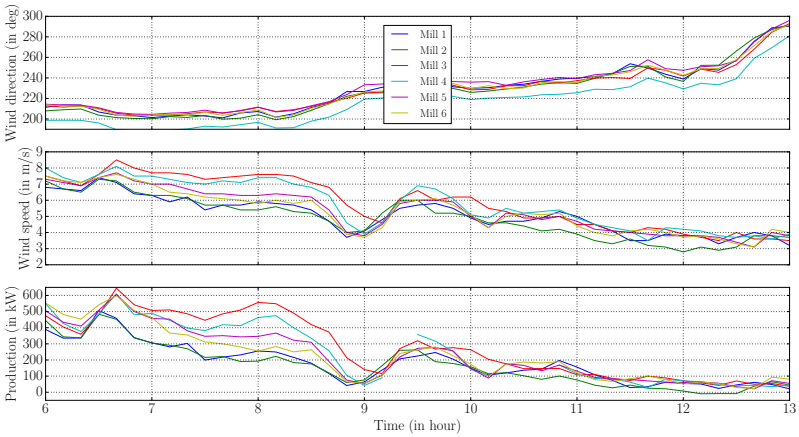
Fig. 1.12: Time evolution of SDM results for the 22^{th} December 2016

tween 10:00 and 12:00 LT. The simulation displays a similar behavior with however significant differences. Between 6:30 and 9:00 LT, wind speed and energy production at turbines 1 and 2 are similar to wind speed and energy production simulated at turbine 3, and turbines 4, 5 and 6 are in the wake of turbine 3 as observed. Between 10:00 and 12:00 LT, the simulated wind speed and energy production varies as observed at the locations of the wind turbines with however a weaker difference between the forefront and the trailing wind turbines.

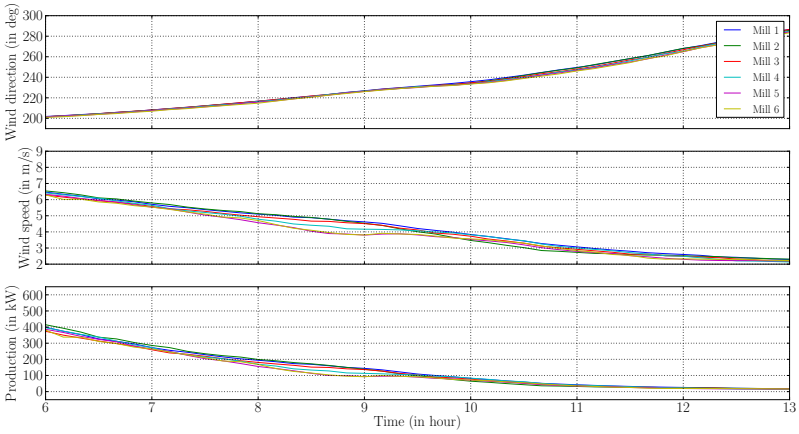
Figure 1.14 shows surface views of the simulated turbulent kinetic energy at the hub height (100 m) at different times (0:20, 8:00 and 11:00 LT). At this altitude the only source of turbulence is the presence of the turbines. Figure 1.14a displays the turbulent kinetic energy pattern 20 minutes after the beginning of the simulation at 00:20 LT. At this moment the turbines are not lined up and they all produce the same energy. Figure 1.14b is similar as Figure 1.14a at 8:00 LT. At this time, the wind direction is around 220°. Consequently, the turbines 3, 4, 5 and 6 are lined-up. Figure 1.14b displays the sheletring effect by the forefront wind turbine and the turbulence generated in its wake. At 11:00 LT (Figure 1.14c), the wind veers so that turbine 1 creates a wake which reaches turbine 2.

To summarize the performance of the simulation against the measurements, Table 1.2 displays skill scores: the Normalized Root Mean Square Error (NRMSE) and the MAE (Mean Absolute Error) defined by

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widehat{y_i}-y_i)^2}}{y_{max}-y_{min}}, \qquad \text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\widehat{y_i}-y_i|. \tag{1.26}$$



(a) Observed wake effect

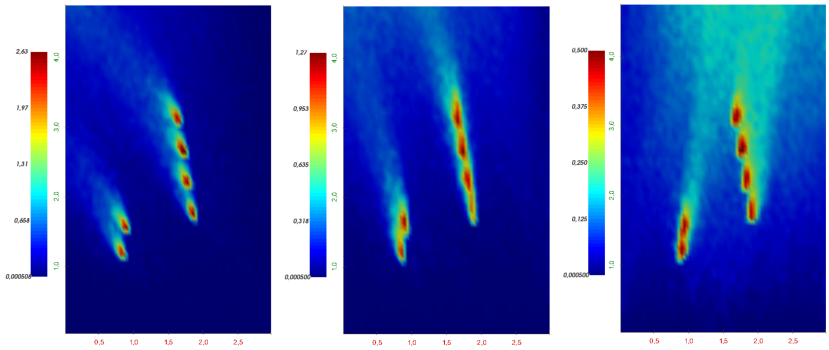


(b) Simulated wake effect

Fig. 1.13: Zoom between 6:00 and 12:00 LT

N is the number of measurements. It is equal to 145 (one measurement every 10 minutes from the 22^{th} December 2016 00:00 LT to the 23^{th} December 2016 00:00 LT). We dispose of the same number of simulated data saved at the same time. y is the measured wind speed and $\widehat{y}$ is the simulated wind speed.

Table 1.2 shows a systematic bias of 1.5 m s^{-1} between the simulation and the measurements which is in part due to the initial and lateral boundary conditions from MERRA reanalysis. The NRMSE range varies between about 14.5 to 17% which can be attributed to the too smooth high frequency variability in the simulation in part attributed to the strong control by the lateral boundary conditions.



(a) at 00:20 LT (b) at 8:00 LT (c) at 11:00 LT

Fig. 1.14: Surface view at hub height (100m) at different times. The three panels show the turbulent kinetic energy.

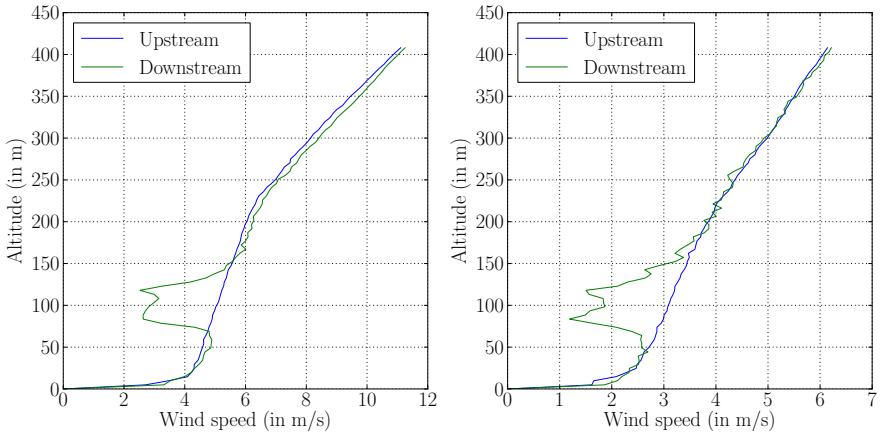
	NRMSE (in %)	MAE (in m/s)
Turbine 1	14.57	1.369
Turbine 2	14.56	1.334
Turbine 3	15.88	1.578
Turbine 4	16.83	1.681
Turbine 5	14.92	1.455
Turbine 6	14.71	1.425

Table 1.2: Indicator of the deviation between the simulated wind $\widehat{y}$ and the observed wind y over the six turbines

Figure 1.15 shows four vertical profiles at different times and locations. At 8:00 LT (Figure 1.15a), the profile displaying a continuously increasing wind speed corresponds to a profile upstream of the turbines. The profile displaying a strong wind speed decreased between 60 and 150 m height is extracting downstream turbine 6. At 11:00 LT (Figure 1.15b) the continuously increasing profile is taken at the same location than in Figure 1.15a. The other profile is extracting downstream turbine 2. In both case, the presence of the turbines decreases the wind speed from 2 m s^{-1} maximum at 80 m and 120 m height (just under and above the hub).

References

1. F. Bernardin, M. Bossy, C. Chauvin, P. Drobinski, A. Rousseau, and T. Salameh. Stochastic Downscaling Methods : Application to Wind Refinement. *Stoch. Environ. Res Risk. Assess*, 23(6):851–859, 2009.



(a) at 8:00 LT (b) at 11:00 LT

Fig. 1.15: Vertical profils taken at different time and place. (a) is taken when the turbines 3, 4, 5 and 6 are lined up; (b) is taken when the turbines 1 and 2 are lined up.

2. F. Bernardin, M. Bossy, C. Chauvin, J.-F. Jabir, and A. Rousseau. Stochastic Lagrangian Method for Downscaling Problems in Computational Fluid Dynamics. *ESAIM: M2AN*, 44(5):885–920, 2010.
3. M. Bossy. Some stochastic particle methods for nonlinear parabolic PDEs. In *GRIP—Research Group on Particle Interactions*, volume 15 of *ESAIM Proc.*, pages 18–57. EDP Sci., Les Ulis, 2005.
4. M. Bossy, J. Espina, J. Morice, C. Paris, and A. Rousseau. Modeling the wind circulation around mills with a Lagrangian stochastic approach. *SMAI-Journal of Computational Mathematics*, 2:177–214, 2016.
5. M. Bossy, J.-F. Jabir, and D. Talay. On conditional McKean Lagrangian stochastic models. *Probab. Theory Relat. Fields*, 151:319–351, 2011.
6. M. Bossy and Violeau L. Optimal rate of convergence of particle approximation for conditional McKean-Vlasov kinetic processes. *arxiv preprint*, (–), 2018.
7. T. D. Dreeben and S. B. Pope. Probability density function and Reynolds stress modeling of near wall turbulent flows. *Physics of Fluids (1994-present)*, 9(1):154–163, 1997.
8. P. A. Durbin. A reynolds stress model for near-wall turbulence. *Journal of Fluid Mechanics*, 249:465498, 1993.
9. P.-A. Durbin and C.-G. Speziale. Realizability of second-moment closure via stochastic analysis. *J. Fluid Mech.*, 280:395–407, 1994.
10. W. McCarty, L Coy, R. Gelano, A. Huang, Merkova D., Smith E. B., M. Sienkiewicz, and K. Wargan. Merra-2 Input Observations: Summary and Assessment. *NASA Technical Report Series on Global Modeling and Data Assimilation*, 46, 2016.
11. J.-P. Minier. Statistical descriptions of polydisperse turbulent two-phase flows. *Physics Reports*, 665(Supplement C):1 – 122, 2016. Statistical descriptions of polydisperse turbulent two-phase flows.
12. J.-P. Minier, S. Chibbaro, and S. B. Pope. Guidelines for the formulation of Lagrangian stochastic models for particle simulations of single-phase and dispersed two-phase turbulent flows. *Physics of Fluids*, 26(11):113303, 2014.

13. J.-P. Minier and J. Pozorski. Wall-boundary conditions in probability density function methods and application to a turbulent channel flow. *Physics of Fluids*, 11(9):2632–2644, 1999.
14. A. Niayifar and F. Porté-Agel. Analytical modeling of wind farms: A new approach for power prediction. *Energies*, 9(9, 741), 2016.
15. S. B. Pope. Lagrangian PDF methods for turbulent flows. In *Annual review of fluid mechanics, Vol. 26*, pages 23–63. Annual Reviews, Palo Alto, CA, 1994.
16. S. B. Pope. Lagrangian pdf methods for turbulent flows. *Annu. Rev. Fluid Mech.*, 26:23–63, 1994.
17. S. B. Pope. Particle method for turbulent flows: Integration of stochastic model equations. *Journal of Computational Physics*, 117(2):332 – 349, 1995.
18. S. B. Pope. *Turbulent flows*. Cambridge University Press, Cambridge, 2000.
19. P.E. Réthoré, N.N. Sørensen, A. Bechmann, and F. Zahle. Study of the atmospheric wake turbulence of a cfd actuator disc model. In *In Proceedings of European Wind Energy Conference*, Marseille, France, 2009. 16-19 March.
20. J.N. Sørensen. Aerodynamic aspects of wind energy conversion. *Annual Review of Fluid Mechanics*, 43(1):427–448, 2011.
21. A. Stohl. Computation, accuracy and applications of trajectoriesa review and bibliography. *Atmospheric Environment*, 32(6):947 – 966, 1998.
22. M. Waclawczyk, J. Pozorski, and J.-P. Minier. Probability density function computation of turbulent flows with a new near-wall model. *Physics of Fluids*, 16(5):1410–1422, 2004.