19th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
3-6 June 2019, Bruges, Belgium

DATA ASSIMILATION AT LOCAL SCALE TO IMPROVE CFD SIMULATIONS OF DISPERSION AROUND INDUSTRIAL SITES AND IN URBAN NEIGHBOURHOODS

Cécile L. Defforge[1], *Bertrand Carissimo*[1], *Marc Bocquet*[1], *Raphaël Bresson*[1], *and Patrick Armand*[2]

[1]CEREA, Joint laboratory École des Ponts ParisTech and EDF R&D, Université Paris-Est, Marne-la-Vallée, France
[2]CEA, DAM, DIF, F-91297 Arpajon, France

Abstract: Precise wind fields simulated by CFD models are used for many environmental and safety micro-meteorological applications, such as dispersion modelling or wind potential assessment. Atmospheric simulations at local scale are largely determined by boundary conditions, which are provided, for instance, by meso-scale models (e.g., WRF). In order to improve the accuracy of the boundary conditions (BC), especially in the lowest levels more perturbed by high resolution topography, data assimilation methods might be used to take available observations into account. Data assimilation methods have been generally developed for larger scale meteorology and initial conditions. Among the existing methods, the iterative ensemble Kalman smoother (IEnKS) has been chosen as it is independent of the atmospheric model and it is able to handle non-linear operators. The IEnKS has been adapted to local scale atmospheric simulations by taking BCs into account. This adapted version has previously been tested on a simple shallow-water model in 1D. In the present study, we analyse the performances of the IEnKS in 3D with the CFD model *Code_Saturne* using both twin experiments and field observations over a realistic, very complex topography. We propose a method to determine the first estimate of the control vector, which corresponds to the BCs, and to construct the associated background error covariance matrix, from the statistical analysis of three years of WRF simulations. The IEnKS is proved to greatly reduce the error and the uncertainty on the BCs and thus on the simulated wind field over the small-scale domain. The IEnKS is also tested in urban conditions with observations provided by the Mock Urban Setting Test field campaign. This study case allows to evaluate the possibility to assimilate either wind observations (speed and direction) or pollutant concentration values. We present here the first results obtained in this urban configuration.

Key words: *Data assimilation, local scale simulation, boundary conditions, CFD model, iterative ensemble Kalman smoother, air quality modelling, MUST.*

INTRODUCTION

Many environmental and safety micro-meteorological applications, such as dispersion modelling or wind potential assessment require the accurate estimation of wind fields at local scale. These wind fields are generally simulated with CFD models, such as *Code_Saturne* (Archambeau et al., 2004), especially over complex terrain (e.g., Blocken, 2014) and in urban area (e.g., Vardoulakis et al., 2003). The counterpart of CFD models is the computational cost and the sensitivity to input data, and especially boundary conditions (BC) (Yang et al., 2009). In addition to numerical tools, urban areas and prospective sites for wind en-

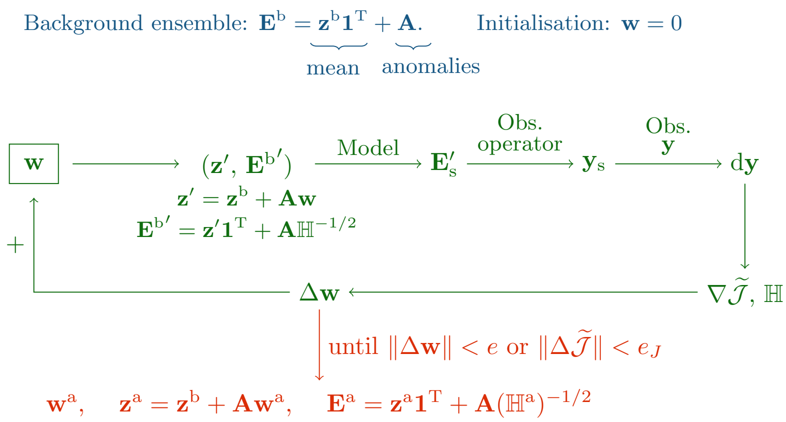

Figure 1: One analysis cycle of the IEnKS. The matrix of anomalies $\mathbf{A}$ corresponds to the departure from the background $\mathbf{z}^{\mathrm{b}}$, for each member of the background ensemble $\mathbf{E}^{\mathrm{b}}$. The goal is to find the weight vector ($\mathbf{w}$) that defines the best linear combination of the ensemble members. To do so, a cost function $\widetilde{\mathcal{J}}$ is iteratively minimised: for each value of $\mathbf{w}$, a new ensemble of BCs $\mathbf{E}^{\mathrm{b}\prime}$, centred on $\mathbf{z}'$, is generated using the *transform* method. The model and the observation operator are applied to this ensemble, which gives an ensemble of simulated observations with mean $\mathbf{y}_{\mathrm{s}}$ that can be compared to the observations $\mathbf{y}$. The increment $\mathrm{d}\mathbf{y} = \mathbf{y} - \mathbf{y}_s$ is used in the estimation of the gradient ($\nabla\widetilde{\mathcal{J}}$) and Hessian ($\mathbb{H}$) of the cost function. The weight $\mathbf{w}$ is thus updated following Gauss-Newton algorithm until the convergence criterion is reached, defined either on the increment $\Delta\mathbf{w}$ or on the cost function. At the end of the analysis cycle, we obtain the best estimate of the control vector $\mathbf{z}^{\mathrm{a}}$ and the analysis ensemble $\mathbf{E}^{\mathrm{a}}$, with a spread related to the uncertainty associated with the analysis.

ergy installation are generally equipped with some meteorological instruments, which provide observations inside the wind parc or the urban canopy.

In order to combine CFD model and in situ observations, data assimilation (DA) methods adapted to local scale simulations must be used. Up to now, the DA methods have generally been developed for larger scale simulations and deal with initial conditions (e.g., Kalnay, 2003; Asch et al., 2016). Only a few studies have used DA methods to correct the BCs for small scale simulations (e.g., Mons et al., 2017; Sousa et al., 2018). Among existing DA methods, the most recent and efficient ones are variational ensemble methods, and especially the iterative ensemble Kalman smoother (IEnKS) (Bocquet and Sakov, 2014). This method has the advantage to be independent of the dynamical model and to handle non-linear analyses. The IEnKS can be adapted to local scale atmospheric simulations by taking BCs into account. It has been previously proven to converge and tested on a simple shallow-water model in 1D (Defforge et al., 2019b). The goal of this study is to assess the ability of the adapted IEnKS to improve 3D wind simulations at local scale by assimilating observations in the lower layer of the atmosphere.

In the present study, we consider two cases with the atmospheric module of the open-source CFD model *Code_Saturne*: one corresponds to a wind resource assessment study and the other to the pollution dispersion in urban area. In both cases, real observations are available: part of them are assimilated whereas the remaining observations are used for validation.

METHODS

The iterative ensemble Kalman smoother (IEnKS)

The goal of DA is to optimally combine the information provided by the first estimate of the BCs – or *background* – and by the available observations within the domain. The iterative ensemble Kalman smoother is an ensemble variational method based on the iterative minimisation of a cost function (Bocquet and Sakov, 2014). This cost function measures both the misfit between the control vector and its first estimate (the background) and the distance between the observations and the projection of the control vector in the observation space, by the model and the observation operator. The background error is represented by an ensemble of N members and the objective is to find the best linear combination of the ensemble members. This means defining and minimising the cost function in the sub-space spanned by this ensemble (Fig. 1). Similarly, the ensemble obtained at the end of the analysis cycle provides an error estimate of the analysis, which is another advantage of the IEnKS compared to other DA methods. In the present study, it is used

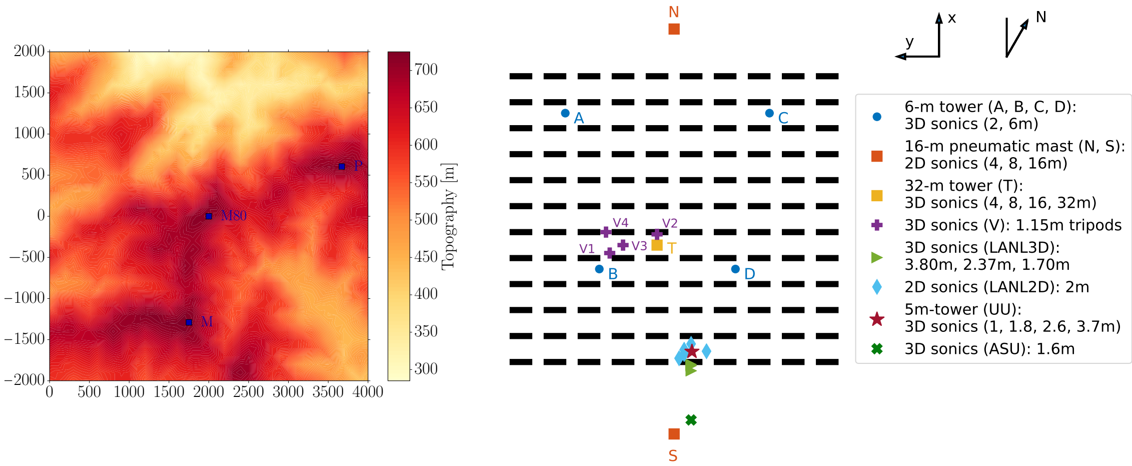

Figure 2: (left) Topography for the domain used in the first experiment and location of the three meteorological masts (blue squares). **(right)** Representation of the *Code_Saturne* domain containing the MUST array with the location of the available meteorological instruments.

with *Code_Saturne* to improve the accuracy of 3D stationary simulations through the correction of the constant BCs.

First case: Wind resource assessment

We consider a domain which extends over $4\text{km} \times 4\text{km}$ horizontally and up to 2030m vertically, above a very complex topography (Fig. 2, left). A field campaign has been performed within this domain between August and February 2007. The observations are available every 10 minutes at three meteorological masts (M80, M, and P) shown with the blue squares in Figure 2 (left). We assimilate the observations of u and v components of the wind provided by the two masts M and P at 30, 39, and 49 m above the ground (classical cup anemometers and vanes usually available for wind farms). We compare the simulation results to the observations provided by the mast M80 at 10, 25, 45, and 78 m (sonic anemometers).

The control vector corresponds to the BCs on wind (u and v), given for 20 profiles distributed around the domain, each of them being defined in 21 vertical levels. The control vector thus contains 840 variables.

WRF simulations have been performed in the same region during three years. These simulations are clustered in 50 classes according to the wind speed, the wind direction, and the departure of the WRF results from the sonic observations. Only one CFD simulation is performed per class, for the most representative date and time. For each representative situation, the background BCs are first estimated from WRF results for this time. The IEnKS is used for each of these 50 situations to correct the BCs, more strongly in the lower levels, by assimilating the available observations at the given date and time.

Second case: Atmospheric dispersion modelling

In this second case we use the observations provided by the nearly full-scale Mock Urban Setting Test (MUST) conducted in septembre 2001 at the U.S. Army Dugway Proving Ground (DPG) Horizontal Grid test site (Biltoft, 2001). Containers were aligned on a 12 by 10 grid over a 200m square area and each container was 12.2m long, 2.42m wide, and 2.54m high. The trials consisted of 15-minutes release of a tracer gas (propylene) from positions either within or immediately outside the containers array, at different heights between 0.15m and 5.2m. Numerous instruments measured both meteorological variables (wind, turbulence, etc.) and gas concentration during the trials (Fig. 2, right).

We focus here on the trial conducted on the 25$^{\text{th}}$ of September at 18h29, which corresponds to neutral stability conditions. In this case, the BCs for the *Code_Saturne* simulation is given by one profile of velocity (u and v) and turbulent kinetic energy (k). The profile is located in the center of the southern border and defined in 22 vertical levels. The control vector thus includes 66 variables. To estimate the background BCs, we use observations of wind speed and turbulence provided by meteorological instruments a few hundred meters outside of the containers array.

In this experiment, we assimilate the observations provided by 7 sonic anemometers located on the 6-m

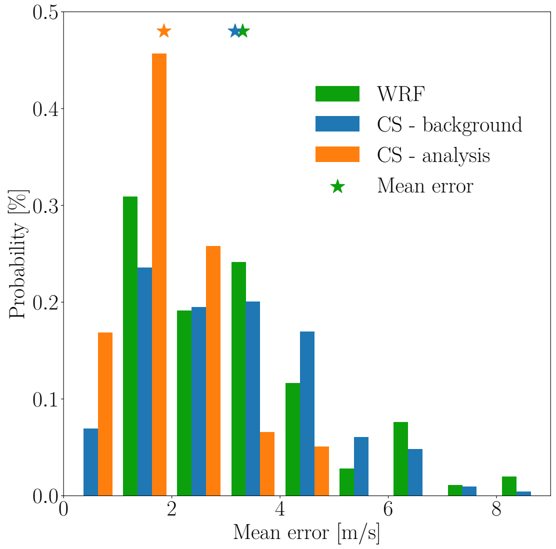

Figure 3: Histograms of the errors between sonic observations, provided by the mast M80, and the wind simulated by WRF (green), by *Code_Saturne* with BCs provided by WRF (background, blue), and by *Code_Saturne* with BCs corrected by the IEnKS (analysis, orange). The mean error made by WRF (green), *Code_Saturne* before (blue) and after (orange) DA are also shown (stars).

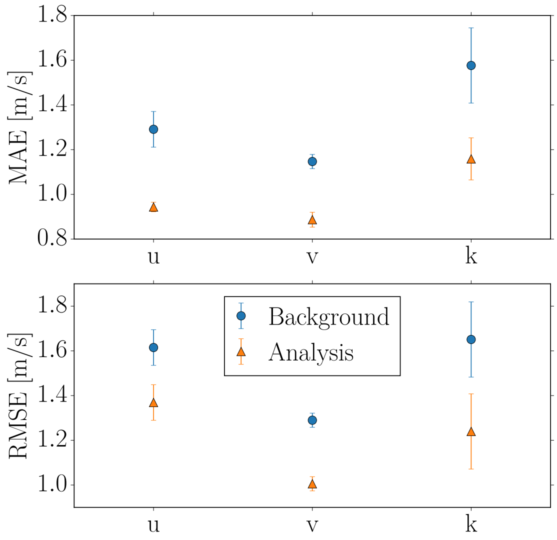

Figure 4: Mean absolute error (MAE, above) and root mean square error (RMSE) of the simulated wind components (u and v) and turbulent kinetic energy (k), before (blue) and after (orange) the analysis cycle of the IEnKS. The errorbars show the standard deviation of the background and analysis ensembles for the MAE and the RMSE.

tower A (2m and 6m), the tripod V3 (1.15m), and the 5-m tower (1m, 1.8m, 2.6m, and 3.7m) in the DA process. All the other observations are used for validation.

Estimation of the background ensemble

Because the cost function is minimised in the ensemble space, the definition of the background ensemble is of primary importance. This ensemble represents the background error covariance matrix (**B**) and it is equivalent to estimate this covariance matrix and then select the leading modes, i.e. the eigenvectors associated with the largest eigenvalues. To estimate **B**, we decompose the covariances ($C_{i,j}$) as the product of standard deviations and correlation coefficients: $C_{i,j} = \text{corr}_{ij}\sigma_i\sigma_j$, where corr_{ij} is the correlation coefficient between the i^{th} and the j^{th} variables and σ_i is the standard deviation of the i^{th} variable. Both the correlation coefficients and the standard deviations are estimated from statistical analyses of the climatology. In the first case, we use the three years of WRF simulations at the location of the *Code_Saturne* BCs to estimate the correlations and the variances of the wind components. In the second case, we use the observations above the canopy provided by the MUST campaign during all the trials. Once the background error covariance matrix is estimated, we construct the ensemble from the $N-1$ leading modes to which is added a N^{th} member, necessary to ensure that the ensemble mean is equal to the background.

RESULTS

First case: Wind resource assessment

For each representative situation, we perform an analysis cycle of the IEnKS to assimilate the 30 observations of u and v provided by the masts M and P. The wind fields simulated with the analysis BCs thus obtained are compared with the observations from the sonic anemometers of the third mast (M80). The results show that the IEnKS helps reduce the error for most of the situations, except in some cases for which the background error is already small. For all the situations, the IEnKS converges in 4 to 6 iterations. We also show that the use of the CFD model does not reduce much the error if the BCs are imprecise. However, after a cycle of the IEnKS with 5 members, the distribution of the errors on the simulated wind field at the location of the mast M80 is largely shifted toward smaller values and the mean error is nearly divided by 2 (Fig. 3).

To estimate the impact of DA in a context of wind potential assessment, we assume that a unique wind turbine of $6\mathrm{MW}$ is installed at the location of the mast M80. For each representative situation, we consider the wind speed, at $78\ \mathrm{m}$ above the ground, given by: the sonic anemometer, WRF simulations, *Code_Saturne* simulations with the BCs provided by WRF (referred to as CS^{b}), and *Code_Saturne* simulations with the BCs corrected by DA (referred to as CS^{a}). The wind resource is obtained as an average of the 50 power values, weighted according to the size of the classes, and compared to the one computed from the measurements. The wind resource estimated with WRF results is underestimated by 30%, with CS^{b} it is underestimated by 42%, and with CS^{a} by less than 10%. Moreover, the uncertainty on the wind resource estimation is reduced from 4.7% with CS^{b} to 1.3% with CS^{a}. Consequently, the use of the IEnKS to correct the BCs of the 50 representative situations allow to largely reduce the error of the wind potential assessment as well as the uncertainty on this estimate.

Second case: Atmospheric dispersion modelling
The IEnKS is used with 5 members to assimilate the 14 observations of u and v in order to update the values of wind and turbulence which define the BCs. The IEnKS converges in 4 iterations here. The wind fields simulated with *Code_Saturne* when the background and analysis BCs are prescribed are compared with all the observations that are not assimilated. The mean absolute error (MAE) and the root mean square error (RMSE) are computed for u, v, and k over the 14 observations available at this date within the urban canopy ($< 2\mathrm{m}$ above the ground) (Fig. 4). In this case again, the IEnKS helps reduce the error on the simulated values and to reduce the uncertainty. Further investigations will analyse the impact of data assimilation on pollutant dispersion. Finally, we intend to assess the ability of the method to correct the boundary conditions of wind and turbulence by assimilating observations of pollutant concentration.

CONCLUSION

In this study the IEnKS adapted to local scale atmospheric simulations is used to assimilate observations near the ground. The method is tested in two different cases with field observations. In both cases the IEnKS is proved to improve the correctness and the accuracy of the boundary conditions and thus of the simulated wind field in operationally affordable conditions.

ACKNOWLEDGEMENT

The CEREA is a member of the Pierre-Simon Laplace Institute (IPSL).

REFERENCES

Archambeau, F., N. Méchitoua, and M. Sakiz, 2004: Code Saturne: A Finite Volume Code for the computation of turbulent incompressible flows - Industrial Applications. *Int. J. on Finite Vol.*, **1**.

Asch, M., M. Bocquet, and M. Nodet, 2016: *Data assimilation: methods, algorithms, and applications*. Society for Industrial and Applied Mathematics.

Biltoft, C. A., 2001: Customer report for mock urban setting test. *DPG Document Number 8-CO-160-000-052. Prep. for Def. Threat. Reduct. Agency*.

Blocken, B., 2014: 50 years of Computational Wind Engineering: Past, present and future. *J. Wind. Eng. Ind. Aerodyn.*, **129**, 69–102.

Bocquet, M., and P. Sakov, 2014: An iterative ensemble Kalman smoother. *Quart. J. Royal Meteor. Soc.*, **140**, 1521–1535.

Defforge, C. L., B. Carissimo, M. Bocquet, P. Armand, and R. Bresson, 2019a: Data assimilation at local scale to improve CFD simulations of atmospheric dispersion: application to 1D shallow-water equations and method comparisons. *Int. J. Environ. Pollut.*.

Defforge, C. L., B. Carissimo, M. Bocquet, R. Bresson, and P. Armand, 2019b: Improving CFD atmospheric simulations at local scale for wind potential estimate using the iterative ensemble Kalman smoother. *J. Wind. Eng. Ind. Aerodyn.*, **In revisio**.

Kalnay, E., 2003: *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press, 341 pp.

Mons, V., L. Margheri, J.-C. Chassaing, and P. Sagaut, 2017: Data assimilation-based reconstruction of urban pollutant release characteristics. *J. Wind. Eng. Ind. Aerodyn.*, **169**, 232–250.

Sousa, J., C. García-Sánchez, and C. Gorlé, 2018: Improving urban flow predictions through data assimilation. *Build. Environ.*, **132**, 282–290.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Springer Netherlands, 670 pp.

Vardoulakis, S., B. E. Fisher, K. Pericleous, and N. Gonzalez-Flesca, 2003: Modelling air quality in street canyons: a review. *Atmospheric Environ.*, **37**, 155–182.

Yang, Y., M. Gu, S. Chen, and X. Jin, 2009: New inflow boundary conditions for modelling the neutral equilibrium atmospheric boundary layer in computational wind engineering. *J. Wind. Eng. Ind. Aerodyn.*, **97**, 88–95.