Mean wind flow reconstruction of a high-rise building based on variational data assimilation using sparse pressure measurements

Mohamed Yacine Ben Ali, Gilles Tissot, Sylvain Aguinaga, Dominique Heitz, Etienne Mémin

Mean wind flow reconstruction of a high-rise building based on variational data assimilation using sparse pressure measurements

M.Y. Ben Ali[a,b,c,*], G.Tissot[a], S. Aguinaga[b], D.Heitz[c], E. Mémin[a]

[a]*Inria/IRMAR, Fluminance team, Campus universitaire de Beaulieu, F-35042 Rennes, France*
[b]*Centre Scientifique et Technique du Bâtiment (CSTB), 11 rue Henri Picherit, F-44300 Nantes, France*
[c]*INRAE, UR OPAALE, F-35044 Rennes Cedex, France*

Abstract

Data assimilation consists in combining a model and measurements in order to estimate the state of the system. In wind engineering, the interest to take advantage of both sources of information is extremely high. As a matter of fact, on the one hand, measurements in wind tunnel experiments are in practice extremely sparse and, on the other hand, stationary numerical models based on Reynolds averaged Navier-Stokes (RANS) fail to predict the time-averaged wake accurately. Variational data assimilation (VDA) techniques are defined as high-dimensional optimisation problems, the solution of which is determined through an adjoint method. In practical applications, where only wall pressure measurements on the building are available, an accurate mean flow reconstruction is very challenging, especially in regions far away from any measurements. In the present paper, spatially distributed forces are considered as a control parameter on the momentum equation or the turbulence closure equations. These forcings are interpreted as corrections brought to the turbulence closure. Consequently, the problem corresponds to a high-dimensional optimisation problem for a spatially distributed variable with sparse pressure measurements. Some guidelines are brought here to perform accurate and physically consistent reconstructions for wind engineering applications. In order to deal as closely as possible with a practical application, this study focuses on the realizable revision of the $k-\epsilon$ RANS model applied on a high-rise building wake flow. In particular, some regularisation strategies as well as efficient techniques for control parameter selection and identification are provided.

1. Introduction

Because of their potential long-time exposure to strong atmospheric winds, high-rise buildings may exhibit problematic tearing affecting security [1, 2, 3]. To infer how wind flow affects these structures, some authors [4, 3] have used physical models and numerical simulations for a long time. In order to predict accurately the pressure distribution on the building, the surrounding flow field has to be well represented, but the development of computationally afordable models that fits experimental informations in realistic 3D configurations is still today challenging.

Wind tunnel experiments are conducted on reduced-scale models at the design stage, leading to reliable experimental load predictions thanks to years of progress on flow measurement techniques. High-Frequency Pressure Integration (HFPI) [5] is one of the most common techniques employed for wind load prediction. This sparse wall pressure measurement may lead to misinterpretations on the local pressure field when complex geometries are considered. Moreover, although techniques such as Particle Image Velocimetry (PIV) [6] allow in principle to obtain planar or even full 3D data [7] velocity fields, the cost is considerably higher.

More recently, made possible by the substantial progresses in computational fluid dynamics, numerical simulations of atmospheric boundary layer flow over these large structures have been carried out. Both steady simulations embedding turbulence models within the Reynolds averaged Navier–Stokes (RANS) and large eddy simulations (LES) were investigated to give an insight into the time-averaged flow profile. Several studies were conducted to compare their relative performances and assess their applicability to the prediction of flow around buildings. [1, 8, 9, 10, 11, 12, 13]. In most of these studies, flow-field around cubic-buildings were analysed, and deficiencies of the eddy viscosity modelling in the $k - \epsilon$ model [14] were put forward (e.g., by Murakami et al. [1]). These include the stagnation point anomaly [15, 16] with overestimation of turbulent kinetic energy near the frontal corner and the resulting underestimation of the size of separated boundary layers. Another reported anomaly is the underestimation of turbulent kinetic energy in the wake resulting in an overestimation of the size of

*Corresponding author:
Email address: mohamed-yacine.ben-ali@inria.fr (M.Y. Ben Ali)

the recirculation zone. These accuracy issues strongly hinder the predictive abilities of the model to reproduce wall-pressure measurements. Although, LES leads to intrinsically superior performances [11, 12, 17], they are computationally expensive. Moreover, they carry some difficulties, such as the definition of time-dependent inlet and wall boundary conditions [18, 19, 20].

Measurements and numerical simulations provide complementary pieces of information. The aim of the data-model coupling strategy, which constitutes the core of the present work, is to move beyond the respective limitations of data and models by taking advantage of all available information. This strategy will allow us to provide an effective solution to address turbulence modelling errors at a reasonable cost, informed by data obtained from realistic experimental procedures. The concept of data-model coupling, or more commonly referred to as data-assimilation (DA), comes from estimation theory, and was first been applied to numerical weather prediction [21].

Two large classes of DA exist. Statistical techniques based on Bayesian inference leads to sequential strategies such as the Kalman filter. It has been used to estimate optimal flow parameters from data affected by a high uncertainty level [22, 23]. It is opposed to variational methods [24, 25], where the estimation problem is written as an optimisation problem. The latter is solved through gradient descent optimization techniques based on the adjoint of the linear tangent dynamics operator or an ensemble approximation of it. These methodologies have been carried out for direct numerical simulations (DNS) or LES models in [26, 27, 28, 22, 29, 30].

In this line of thought, formal uncertainty quantification (UQ) techniques have been employed to address the modelling errors underlying the closure constants of RANS systems in probabilistic terms [31, 32, 33, 34, 35, 36]. In a recent work by Shirzadi et al. [37], global coefficients of the standard $k-\epsilon$ model were adapted for unstable atmospheric boundary layer (ABL) flow around high-rise buildings using Monte–Carlo optimisation techniques. More recently, in Ben Ali et al. [38] the global coefficients of the realizable revision of the $k-\epsilon$ model were also calibrated for wind-loads prediction on a high-rise building of ratio $H/D = 4.9$ (same configuration as the present paper). In this latter work, the authors used wind tunnel pressure measurements in a variational data assimilation (VDA) framework. The VDA is then applied for a detailed analysis of the considered turbulence model. Local sensitivity to the global turbulence coefficients is discussed, and a calibration process was conducted to investigate the limitations of the models through hold/relax closure hypothesis scenarios. It was re-

ported that the VDA technique is indeed able to provide insights on the model variability through calibration of model closure coefficients. It was observed that variabilities of the solution were strongly constrained and did not allow the flow to reach an effective region of the state space for accurate estimations of both the wind-loads and the recirculation zone, highlighting, in consequence, the rigidity of the standard closure model. To overcome such restrictions, the authors have considered a so-called *weak constraint* by defining an additive forcing term to allow the solution to deviate from the default model. Better flow reconstructions have been obtained, but no detailed studies have been performed to determine the best way to relax the dynamical constraint for flow reconstruction. Many control parameters and formulations of the optimisation problem can be considered. The present study complements the work ofBen Ali et al. [38]to explore in details the possible control parameters choices to estimate at best the flow surrounding the building.

In the last few years, studies dealing mainly with fundamental and industrial oriented flow configurations, e.g., airfoils, infinite cylinder, backward step, have been considered for data assimilation (DA) flow reconstruction. A hierarchy of procedures has been proposed. In [39, 40], laminar steady Navier–Stokes equations corrected by a direct forcing on the momentum equations have been used to assimilate synthetic particle image velocimetry (PIV) data. For turbulent flows, DA with RANS turbulence models have been explored [41, 42, 43, 44]. Franceschini and Sipp [44] have partly addressed closure errors to the transport equation of eddy-viscosity in the Spalart-Allmaras model [45]. In other works [41, 42, 43], data (either synthetic or experimental) were assimilated, and turbulence model errors were addressed in the turbulent energy production as it constitutes one common issue for the prediction of a variety of flow configurations. In there studies, reconstructions of several benchmark flow configurations, ranging from one-dimensional channel flows to 2D airfoil flows, were performed under the framework of both ensemble Kalman [46] and VDA. It should be noted that all these studies were investigated in bi-dimensional flow configurations, in which turbulence is often generated at a unique integral scale. To our knowledge, three-dimensional cases of complex flow interactions, such as building-lower atmospheric boundary layer interaction, are still largely unexplored with VDA.

Ben Ali et al. [38] focused on exploiting the VDA framework to extract physical information on the flow and identify deficiencies in the turbulence

closure modelling of the RANS. In the present study, the objective is to perform the best flow estimation as possible with a distributed control parameter based on a RANS model and on wall-pressure measurements for wind-engineering applications. The main contributions are :

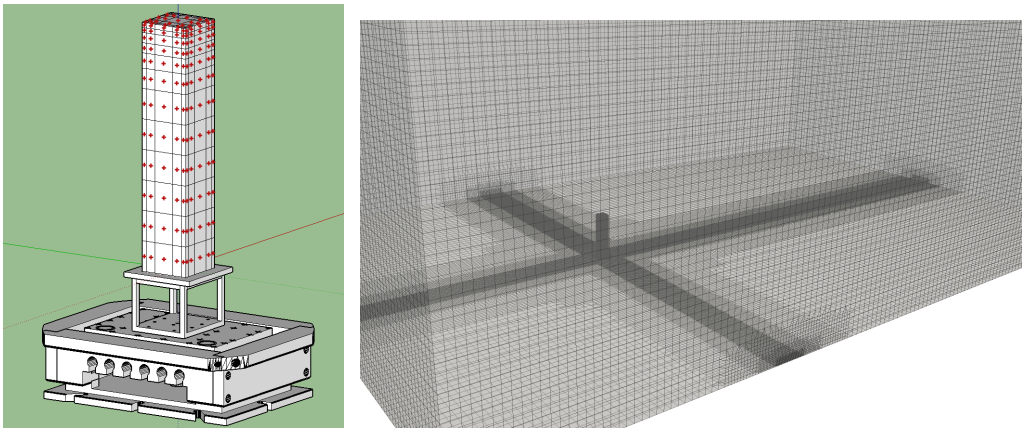
- Applying the VDA on a realistic wind engineering estimation problem combining wall-pressure measurements with a routinely used industrial three-dimensional RANS code;
- Determining the best choice of distributed control parameter to avoid overfitting and obtain accurate reconstructions;
- Proposing a Sobolev gradient descent direction and comparing it with penalty techniques to regularise the solution and drastically accelerate the convergence.

To these ends, a large set of assimilations with various choices are performed and compared to provide guidelines for VDA estimations in wind engineering.

This paper is organized as follows. In section 2, the flow configuration is presented. In section 3, the methodology to derive the VDA approach is introduced. A modified set of RANS equations is presented, motivations for the choices of parameter and regularization solution are given. In section 4, the numerical setup is presented. In section 5.1, the data-assimilation results of the various parametrization strategies are shown. Finally, to obtain accurate reconstructions, a modified model that combines previous parametrizations is investigated in section 5.2.

2. Flow configuration

The present work was conducted on an isolated high-rise building with a square section and an aspect ratio of $H/D \simeq 4.9$ for which $H = 147$ m is the height and $D = 30$ m the width at full scale. Measurements were produced from experiments held in CSTB (Nantes, France) as part of the thesis of Sheng et al. [47]. In his work, two types of experiments were carried out, using the HFPI for wall-pressure and PIV plans for the near wake flow. It should be recall that only the wall-pressure was considered in this work as input data for the data-assimilation procedure. PIV plans from the same experiment were used only for an external validation purpose. For illustration, figure 1 shows the 3D view of the case study from both the experiments and our computational setup.



(a) (b)

Figure 1: Flow configuration (a) the scaled model of the building with positions of pressure measurements (red dots), (b) the full scale building in the computational domain.

3. Methodology

This section briefly presents the variational data-assimilation framework, which was developed in [38].

With the variational approach [24, 25], the estimation problem is expressed in terms of the minimization of an objective functional, which measures the mismatch between the model predictions and the data, under the constraint that the optimal state obeys to the model dynamics. In the context of time-averaged models, the data assimilation problem can be formulated as

$$\min_{\alpha} \quad \mathcal{J}(\alpha, \mathbf{X}(\alpha), \mathbf{Y}) \tag{1a}$$

$$\text{subject to} \quad M_i(\alpha, \mathbf{X}(\alpha)) = 0 \qquad i = 1, \ldots, N, \tag{1b}$$

where $\mathcal{J}()$ is the objective functional that quantifies the discrepancy between the assimilated data and the model predictions, with $\mathbf{Y}$ referring to the measurements and $\mathbf{X}$ to the flow variables. This objective is then minimized under the constraint of N flow governing equations M_i.

3.1. Objective functional

Following [38], the objective functional is constructed as follows. First, the experimental pressure is scaled consistently with the pressure of the

model using the pressure coefficient $C_p = \frac{P-P_\infty}{\frac{1}{2}\rho U_{\text{ref}}^2}$, with the reference velocity U_{ref} (resp. reference pressure P_∞) being the far upstream velocity (resp. pressure) at $H_{ref} = \frac{2}{3}H$. The specific functional reads as follows

$$\mathcal{J}(P,\alpha) = \frac{1}{2}||\rho U_{\text{ref}}^2 \Delta C_p^w||_{R^{-1}}^2 + \|\nabla\alpha\|_{B^{-1}}^2, \tag{2}$$

where $\Delta C_p^w = C_{p_{obs}}^w - \mathcal{I}(C_p^w)$ with the subscript w standing for wall-pressure and *obs* for measurements. The interpolation operator $\mathcal{I}$ maps the wall-pressure coefficient of the numerical model to the pressure coefficient at the sensor position. The covariance matrix, R, between measurements allows to take into account experimental uncertainties and potential correlations between measurements. The norm $\|\|_R^{-1}$ is induced by the inner product $(\cdot,\cdot)_{R^{-1}} = (\cdot, R^{-1}\cdot)$, where $(\cdot,\cdot)$ stands for the Euclidean inner product. Similarly, $\|\|_B^{-1}$ is induced by the weigted $L^2(\Omega)$ inner product $(\cdot,\cdot)_{B^{-1}} = (\cdot, B^{-1}\cdot)_{L^2(\Omega)}$.

It can outline here the strongly sparse nature of the partial pressure observations making the estimation problem difficult. To cope with this difficulty, an L^2 regularization term is added to penalize the gradient magnitude of the spatially distributed control parameter α. The covariance matrix B penalises allows to adjust the degree of regularity to be prescribed. A too-small penalty leads to irregular solutions, while strong penalties deteriorate reconstruction performances by imposing a too-smooth solution. In our case, spatial homogeneity for this constraint is assumed, and the confidence matrix B is considered diagonal and uniform. Its inverse is straightforwardly expressed as

$$B_{ii}^{-1} = \zeta_i \left(\max \left| \frac{\partial \mathcal{J}}{\partial \alpha}^0 \right| \right).$$

in which ζ_i is a positive free parameter and subscript i stands for the vector and matrix component. The other term is a scaling to ensure dimensional homogeneity.

3.2. Constraint definition

This section is dedicated to the specification of the dynamical model $\mathbf{M}$. The Reynolds averaged Navier-Stokes equations are considered, with the realizable version of the $k-\epsilon$ turbulence model [48]:

$$\frac{\partial(\rho U_j U_i)}{\partial x_j} + \frac{\partial P}{\partial x_i} - \frac{\partial}{\partial x_j}\left[\mu_{eff}\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right)\right] = F_i, \tag{3a}$$

$$\frac{\partial U_j}{\partial x_j} = 0, \tag{3b}$$

$$\frac{\partial \rho U_j k}{\partial x_j} - \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right] - \mu_t\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right)\frac{\partial U_i}{\partial x_j} + \rho\epsilon = 0, \tag{3c}$$

$$\frac{\partial \rho U_j \epsilon}{\partial x_j} - \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_\epsilon}\right)\frac{\partial \epsilon}{\partial x_i}\right] - C_1(S,k,\epsilon)S\epsilon + C_2\frac{\epsilon^2}{k + \sqrt{\mu\epsilon}} = F_\epsilon. \tag{3d}$$

where U, ρ and P denote the average velocity, the density and the pressure, respectively. The turbulent kinetic energy and the energy dissipation density are denoted k and ϵ. Here, $\mu_{eff} = (\mu_t + \mu)$ stands for an effective viscosity gathering both the molecular and the isotropic eddy viscosity, with

$$\mu_t = C_\mu \rho \frac{k^2}{\epsilon}. \tag{4}$$

Details on the closure constants are provided in Appendix A. It is worth noting that, through the Boussinesq approximation, the isotropic component $\frac{2}{3}\rho k$ of the Reynolds stress is absorbed into P. Therefore, in the computation of the pressure coefficient required for the objective functional (2), the isotropic part must be subtracted to obtain $C_p = \frac{P - P_\infty - \frac{2}{3}\rho k}{1/2\rho U_{\text{ref}}^2}$. Note the addition of two forcing terms $\mathbf{F}$ and F_ϵ, respectively to the RANS equations and to the turbulence dissipation transport equation. Those forcing terms constitute the control parameter α as detailed in section 3.3. Recalling that the corrections brought by the forcing term F_ϵ solely correspond to the previous work of [38].

3.3. Parameter choice

In the construction of the Reynolds averaged Navier-Stokes equations, some assumptions are performed at different levels. The first one pertains to the Boussinesq hypothesis and consists in modelling the Reynolds stress tensor in the momentum equation by a diffusive term weighted by a scalar eddy viscosity. The value of the eddy viscosity is determined by an algebraic equation involving statistical quantities related to the turbulence. At a second level, transport equations of these turbulent quantities are constructed.

Some semi-empirical terms acting as sources and sinks in these transport equations allow closing the equations.

In the present work, for each layer of assumptions, a possible parametrization is proposed as given in the system (3d). Herein, some detailed motivations are provided for the parameter choice.

Parametrising the momentum equation. At the first level, one aim to bring a direct correction to the mean deformation field. In theory, the exact modeling of the deformation should properly include the mechanisms of energy transfer between the mean field towards the turbulence scales and vice versa. However, the Boussinesq hypothesis leads to a unidirectional energy transfer. From this point of view, a volume force term added to the mean momentum equations can be proposed as a control parameter enabling us to relax the Boussinesq hypothesis and bring non-diffusive mechanisms. Without a priori knowledge of its form, it is straightforward to consider a vector force

$$\mathbf{F}_1 = \mathbf{f_u}$$

which is later-on referred to as a *raw* form. As this raw form lacks of structuration, a more refined shape, yet with a simple expression, of this force is also investigated. Here, it is aimed to ensure some consistency between the deformation correction prescribed by this force and the fact that the Reynolds stress induces it. Corrections can be restricted to regions where the turbulence appears to be the most active and where the turbulence model is likely to be prone to strong errors. Shear layers in separated flows are challenging for turbulence models due to strong flow inhomogeneity and anisotropy, which is poorly represented by eddy diffusion models. We consider here addressing only corrections due to the flow inhomogeneity, that is, regions with a strong gradient of turbulent kinetic energy. Thus, a simple model of more refined forcing term for the direct deformation correction may be defined as

$$\mathbf{F}_2 = \mathbf{f_u} \left| \nabla k \right| .$$

This choice of constraint enables to bring some structure to the parameter field in a simple way. It can be noticed that this pre-factor is consistent in terms of physical dimension with the divergence of the Reynolds stress tensor. Given these two forms, two corrective models are then investigated.

Parametrising the turbulence closure. In contrast to a direct forcing on the mean deformation field, corrections of the mean flow can still be embedded under the Boussinesq hypothesis. Parameters are then included in the

turbulence model equations rather than at the level of the momentum equations, and their effects on the mean flow are restricted to a purely diffusive mechanism. Under the realizable $k-\epsilon$ model framework, two possible corrective models can be designed: on the transport equation of kinetic energy k or on the equation of the dissipation rate ϵ.

In this work, only the correction of the transport of ϵ is considered since it is the equation where the closure terms are present. This choice keeps as much as possible the physical structure imposed by the RANS modelling.

As for its form, it can be assumed either that F_ϵ is an additive term without any particular structure as with $\mathbf{F}_1$, or that it is an explicit function of ϵ. Consistently with [38], authors propose to pre-multiply the forcing term by ϵ in order to reject unphysical corrections at locations where turbulence is weak.

The following simple definition is considered

$$F_\epsilon = -f_\epsilon \epsilon,$$

in order to improve the robustness of the DA procedure.

3.4. Descent direction

The minimization is performed by a gradient-based algorithm which requires the evaluation of the objective functional gradient w.r.t the control variables α. To that end, the adjoint formulation [24, 49] is considered. In fact, the problem (1) is equivalent to the problem of determining the optimal state $\mathbf{X}$ and the set of parameters α in addition to an adjoint state $\mathbf{X}^*$ of a Lagrangian functional $\mathcal{L}(\mathbf{X}, \mathbf{X}^*, \alpha)$ [50] . It is defined by augmenting the functional $\mathcal{J}$ with the constraint $\mathbf{M}$ weighted by the adjoint state:

$$\mathcal{L}(\mathbf{X}, \mathbf{X}^*, \alpha) = \mathcal{J}(\mathbf{X}, \alpha) + \int_\Omega (\mathbf{X}^*)^T \mathbf{M}(\mathbf{X}, \alpha)\, \mathrm{d}\Omega. \tag{5}$$

Then, by considering an infinitesimal perturbation of its arguments, setting the first variation of $\mathcal{L}$ equal to zero leads to an optimality system, which provides the optimal solution.

In [38], a differentiation consistent both with the realizable $k-\epsilon$ transport equations and with the near-wall boundary layer law has been performed. This has resulted in the definition of a continuous adjoint RANS model, together with consistent associated boundary conditions of the RANS tangent linear operator.

Moreover, a gradient projection onto the Sobolev space [51] is also considered for descent direction as an alternative regularization solution to penalization. As in the work of [38], the descent direction with Sobolev gradient is determined by solving the following equation

$$\frac{\partial \mathcal{L}}{\partial \alpha}^{H_1} = \left(\frac{1}{1 + l_{sob}^2} (\mathbf{I} - l_{sob}^2 \Delta) \right)^{-1} \frac{\partial \mathcal{L}}{\partial \alpha}, \tag{6}$$

in which the characteristic length scale l_{sob}, i.e. the filtering size, is chosen to smooth non-physical small scales and Δ stands for the Laplacian operator, while $\mathbf{I}$ stands for identity. Note that, unlike with the penalization approach which requires an estimate of a non-dimensional confidence factor ζ with no direct physical significance, the free parameter l_{sob} has the dimensions of a length. This parameter can thus been chosen *a priori* based on a fraction of a relevant characteristic scales of the domain, such as the building width.

It is worth noting that this Sobolev gradient step can be presented as a standalone regularization technique without the addition of any penalization term on the control parameter. In this work, both techniques are considered individually. For conciseness, it is chosen to discuss their impact on the performances of the data-model coupling procedure when considering the additive forcing $\mathbf{F_1}$.

4. Computational settings

Complying with the common standard [52], and consistently with[38], numerical computations are dimensioned with a full scale building.

4.1. Inflow wind profile: a neutral ABL

The atmospheric flow is modeled as a horizontally-homogeneous turbulent boundary layer (HHTBL) [53, 54]. This consists in considering constant properties in the streamwise and spanwise directions. Thus, only variations along the vertical axis are considered. To enforce the inlet wind flow, profiles for U, k and ϵ are defined as

$$U_{in} = \frac{u_\tau^{ABL} \ln(z + z_0)}{\kappa z_0}, \qquad k_{in} = \frac{(u_\tau^{ABL})^2}{\sqrt{C_\mu}} \qquad \text{and} \qquad \epsilon_{in} = \frac{(u_\tau^{ABL})^3}{\kappa (z + z_0)}, \tag{7}$$

where $C_\mu = 0.09$ is chosen as for a standard $k - \epsilon$ model and u_τ^{ABL} is the friction velocity associated with the constant shear stress

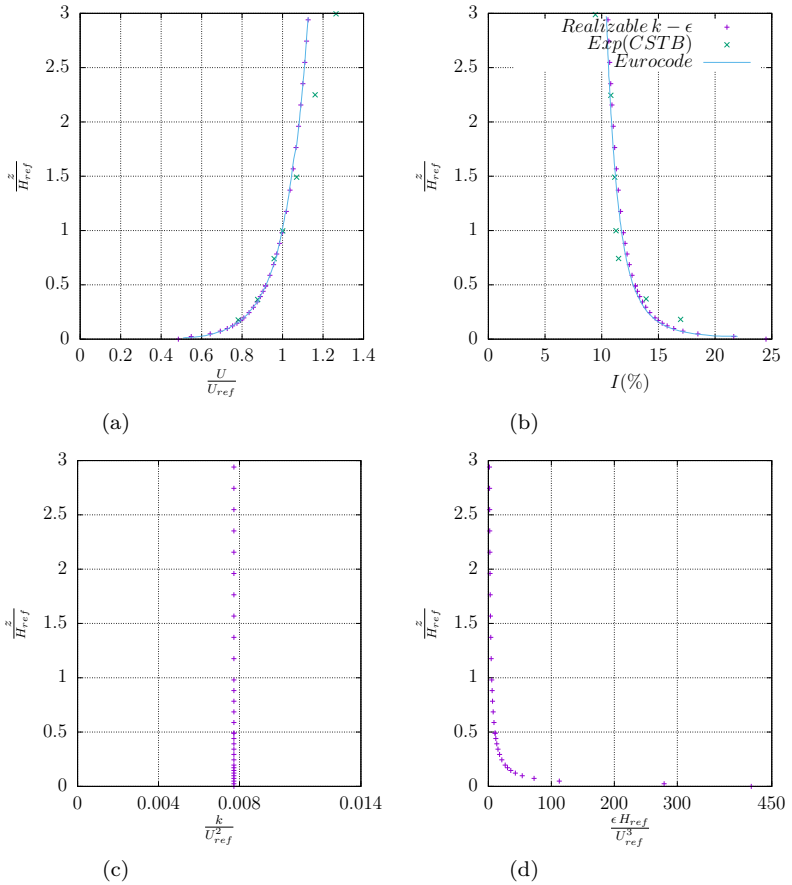


Figure 2: Profiles of the neutral atmospheric boundary layer: (a), mean wind velocity; (b), turbulent intensity $I = \frac{\sqrt{k}}{U_{\text{ref}}}$; (c), turbulent kinetic energy and (d), turbulence dissipation rate.

$$u_\tau^{ABL} = \frac{\kappa U_{ref}}{\ln(\frac{H_{ref}+z_0}{z_0})}. \tag{8}$$

The roughness height z_0 is set to 0.02, which was chosen as an intermediate between class I and class II roughness [52]. The inflow velocity profiles are shown in figure 2.

4.2. Computational domain

Based on recommendation by [16], in this work, the dimension of computational domain is fixed as $18H \times 12H \times 6H$ (length $\times$ width $\times$ height). This sufficiently large domain helps to avoid the influence of the domain

boundaries on the flow near the building and the wall-pressure. Meanwhile, the upstream and downstream distances were set as $\sim 6H$ and $\sim 12H$, respectively. The blockage ratio is 0.28 % less than the threshold of 3% [16].

Hexagonal structured grid was adopted to generate the background empty fetch, after which local mesh refinement approach was employed to densify the grid around the building model, as shown in figure 1b. The minimum distance of grid centroid to the ground was set to $0.006 \times H$ ($\sim$ one meter) while it reaches down to $0.0008H$ for the centroid of the cell adjacent to the building walls. The whole domain corresponds to approximately 3.5 millions cells.

4.3. Numerical method

The open source library OpenFOAM (5th version of the OpenFOAM foundation) [55] was used to implement the CFD and adjoint governing equations. The library utilizes a second order finite volume discretization approach and a fully implicit first order method for time integration. The variables are defined at the centers of each control volume (CV). A prediction-correction procedure is used for the pressure-velocity coupling using the SIMPLE-type methods with Rhie and Chow [56] interpolation. With regard to the discretization schemes, the second order upwind (or linear-upwind) scheme [57] is used for the advection terms as it is shown to be one of the most efficient and accurate scheme for Reynolds Averaged Navier–Stokes (RANS) simulations on bluff body flows [58].

4.4. VDA settings

Regarding the minimization procedure, a steepest descent algorithm is used with an adaptive step. A maximum step size is fixed for each of the parameters based on a prior sensitivity validation test, while a minimum pre-factor for the step size of 10^{-4} is considered as an optimization convergence criteria for all parametrisations.

Regularisation of the force parameter $\mathbf{F}_1$ is proceeded separately either by the penalisation or by the prescription of the descent direction with Sobolev gradient. The penalty-free parameter ranges in $\zeta_i = 0.1, 0.25, 0.5$. Concerning the filtering length l_{sob}, as suggested in previous works [38, 59], a filtering length scale equivalent to 10% of the building width seems to give a fair compromise to filter unphysical small scales.

For the other parametrisation cases, only the Sobolev gradient is considered, with the same filtering length scale.

In the further investigation of combined parametrisations, penalty and Sobolev gradient will be employed together to obtain an efficient regularisation and capture physical flow features.

5. Results

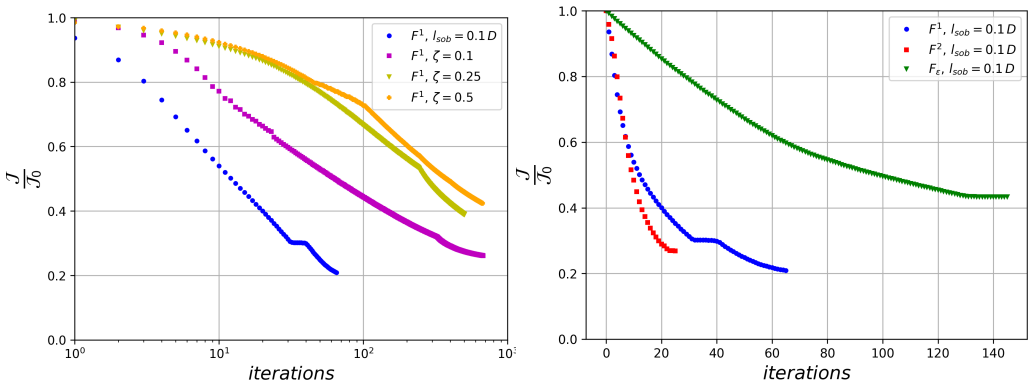
This section is dedicated to the comparison of the different choices of control parameter and regularisation techniques introduced in our data assimilation problem. The impact of their individual contribution on the wall-pressure estimation and the wake flow reconstruction will be first explored. Among all calculations presented in this paper, the case of control parameter f_ϵ with Sobolev gradient regularisation is common with the study [38]. In a second step, an efficient flow reconstruction by combining the advantages identified in the previous techniques is proposed.

5.1. Choice of control parameter and regularisation

5.1.1. Cost reduction and convergence rate

Evaluation of the regularization. Reduction of the (normalized) cost functional in the case of the forcing $\mathbf{F}_1$ on the momentum equation is compared in figure 3a between Sobolev gradient and the penalty constraint. With Sobolev gradient, the objective function has reached its lowest value ($\mathcal{J}/\mathcal{J}_0 \simeq 0.2$). It should be noted nevertheless that with a penalty parameter set to $\zeta = 0.1$, a very close level of reduction $\mathcal{J}/\mathcal{J}_0 \simeq 0.26$ has been reached, but after approximatively 5 times more iterations ($n_{it} \sim 500$ against 100 with Sobolev gradient). The faster minimisation and the lower objective demonstrate the Sobolev gradient ability to improve significantly the gradient-based algorithm. This is consistent with the observation in [38] using the alternative control parameter f_ϵ. This improvement is at the price of a low additional cost corresponding to the resolution of a Helmholtz like equation, yet, of the same order as an additional iteration of the prediction-correction loop. The filtering length l_{sob} can be chosen *a priori*, by taking 10% of the diameter of the building as characteristic length scale. It seems to constitute a solid baseline. Indeed, a higher fraction leads to an overly-diffused parameter, while lower values did not provide enough regularization.

Note however that close to the optimal solution, the penalty provides a better numerical behaviour, showing less oscillations around the minimum.



(a) (b)

Figure 3: Cost function reduction with different parametrizations; (a) an added force F^1 with Sobolev gradient projection $l_{sob} = 0.1D$ against penalisation $\zeta = 0.1, 0.25, 0.5$, (b) comparison between all parameters with only Sobolev gradient projection.

Evaluation of the parameter choice. A comparison of the cost reduction with the different choices of parametrization is shown in figure 3b. It can be seen that using a structured force on the momentum equation $\mathbf{F}_2$ allows reaching an optimal objective at a level very close to the raw force $\mathbf{F}_1$ ($\mathcal{J}/\mathcal{J}_0 \simeq 0.27$), yet with a twice faster descent. This faster convergence suggests that adding a physical-based structure to the correction force seems to lead to a more efficient descent direction. Meanwhile, forcing with f_ϵ on the transport equation of ϵ leads to a lower cost reduction with more iterations. This is an expected result since it relies on a model that is further constrained, as highlighted in [38]. Recalling that the cost functional is representative of the building wall pressure. Due to potential overfitting effects, it is hard to conclude just in light of this result. Sections 5.1.3 and 5.1.4 explore velocity field reconstructions in the wake to obtain more prescriptive conclusions.

5.1.2. Wall pressure

Results of the pressure loads at and around measurement locations are compared with the experimental data [47] and the non-assimilated model in figure 4.

In the back facade, where $C_p < 0$, reconstructed wall-pressure with the raw force $\mathbf{F}_1$ shows perfect fitting to data, while the pre-multiplied force $\mathbf{F}_2$ and the forcing f_ϵ on the transport of ϵ lead to an intermediate estimation between model and data. Observing the windward and side facades and the roof top, it can be noticed that the forcing f_ϵ associated

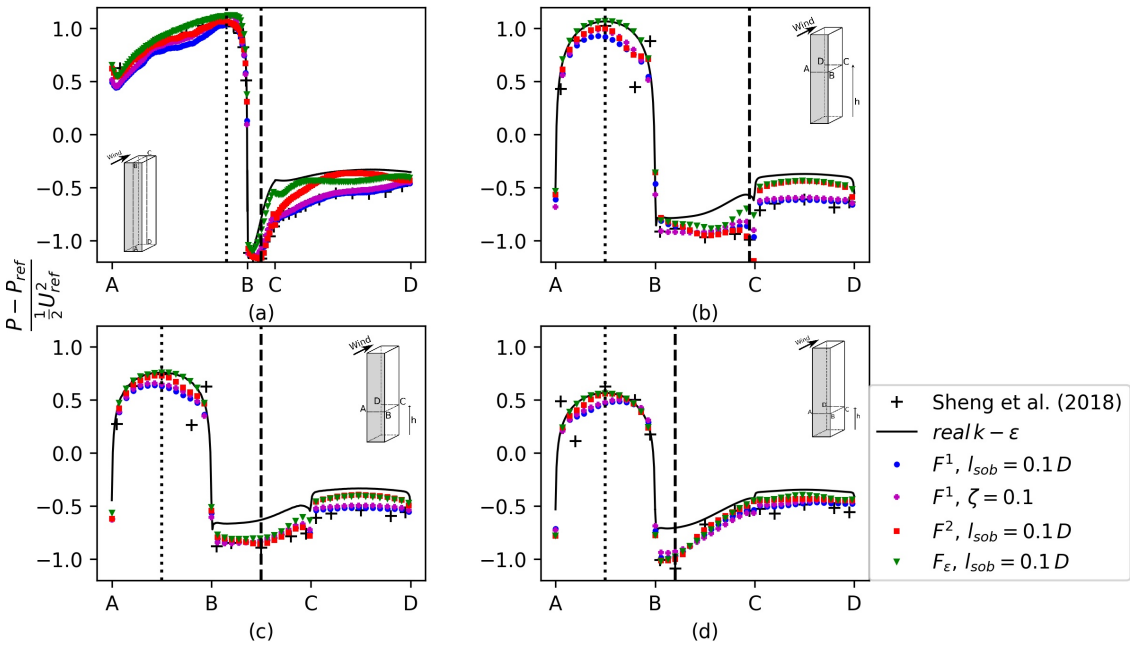


Figure 4: Pressure coefficient profiles along building facades with different parametrizations; (a) along the symmetry plane and along horizontal transverse plane at three heights, (b) $z/D = 3.33$, (c) $z/D = 0.9$, (d) $z/D = 0.63$.

with the most structured model leads to the closest results to the non-assimilated model, while the less structured model by a raw forcing $\mathbf{F}_1$ fits better the measurements. This is consistent with the results of section 5.1.1. A structured forcing on the momentum equation $\mathbf{F}_2$ leads to intermediate results.

If the objective of data assimilation is to perform wall pressure interpolation between pressure taps enhanced by a RANS model, then the raw forcing on the momentum equation is well suited. The present results show that more structure forcings give more confidence in the model. In the next sections, a detailed comparison on the ability of these strategies is given in terms of velocity field reconstruction.

5.1.3. Wake centreline profiles

Results of the streamwise velocity on the wake centreline are here confronted to data from PIV experiments [47] and the uncorrected model in figure 5. It should be recalled that these data were not used in the assimilation problem. Hereafter, an up-wash (resp. down-wash) denotes the ascending (resp. descending) motions in the wake velocity field. In order to span both up-wash and down-wash motions, profiles are compared at four height levels ($z/D = 0.63, 1.26$ for up-wash and $z/D = 2.5, 3.33$ for down-wash). Moreover, at each height, a different recirculation length is identified, by the intersection of the velocity profile with the axis $U/U_{ref} = 0$. As a first remark, all corrective models show some reduction at all heights. Yet, depending on the chosen parameter and the considered wake region, different reduction levels are achieved. With a parameter f_ϵ, data assimilation leads to the shortest recirculation among all parameters and regardless of the considered region, with a very good agreement with PIV data. For both momentum forcing strategies $\mathbf{F}_1$ and $\mathbf{F}_2$, model abilities for retrieving the wake flow extension vary with the height and the forcing type. The most noticeable relative reductions are observed closer to the ground for the up-wash flow, with a very close prediction for both force shapes. However, in the down-wash flow region, a raw form leads to better wake reduction, while the structured forcing $\mathbf{F}_2$ leads to poor results.

At this stage, forcing on f_ϵ leads to good wake velocity field reconstruction, but lower performances for pressure loads predictions and wall pressure interpolation. Meanwhile, a direct forcing on the momentum equation improves pressure prediction, but suspiciously by an unphysical mechanism, since it does not reduce the wake consistently with PIV measurements.

A second key aspect in the velocity profiles is the maximum reverse

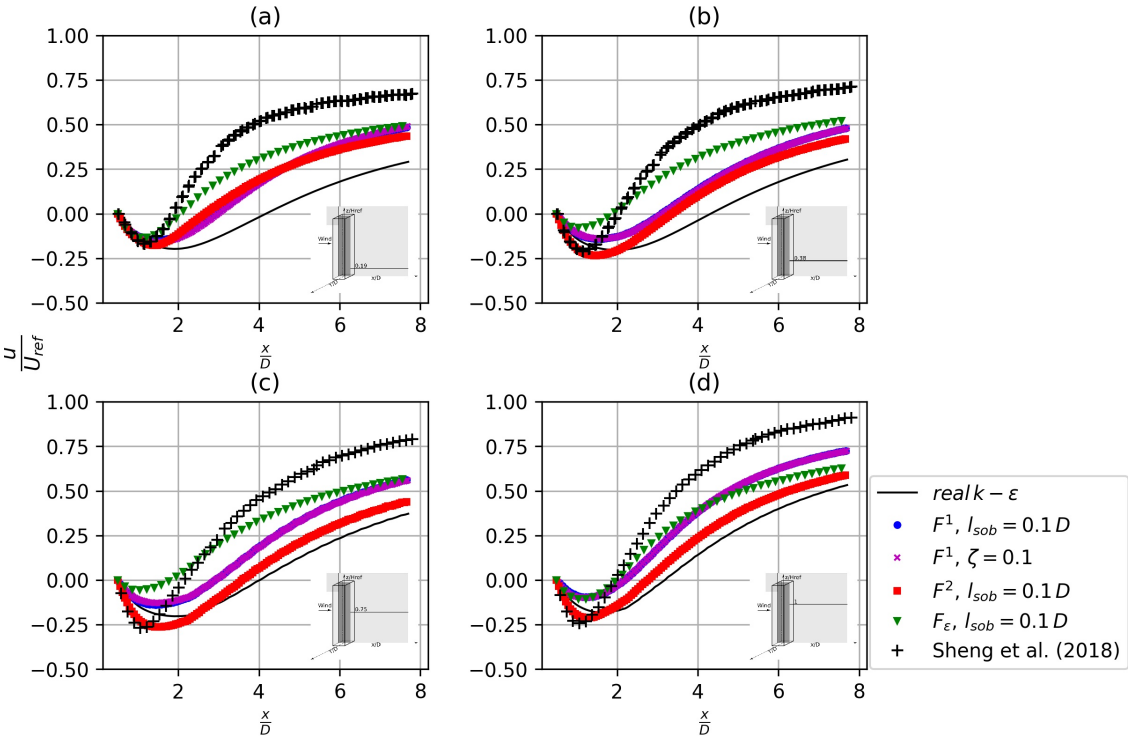


Figure 5: Streamwise velocity comparison along centreline at heights, (a) $z/D = 0.63$, (b) 1.25, (c) 2.5, (d) 3.33.

velocity, happening approximatively around $x/D = 1.5$ in the PIV measurements. Strong reverse flow is associated with intense depression and is thus indicative of the ability of the estimated wake to match the leeward pressure. The mismatch of the minimum velocity between PIV and model is especially strong for large z/D. Based on this criteria, forcing on f_ϵ and the direct raw forcing $\mathbf{F}_1$ shows bad performances, but surprisingly, the structured forcing on the momentum equation $\mathbf{F}_2$ leads to accurate maximum reversed velocity in the wake. This is corroborated by the leeward pressure predictions in figure 4a), suggesting that the reasonable pressure predictions close to the high end (near point C) are obtained due to a good velocity prediction in the wake close to the building. It can be noted that in the case of raw forcing, regularisation effects are not visible in the wake centreline velocity profiles.

5.1.4. Reconstructing flow features

For a better insight into the reconstructed flow fields, figure 6 shows time-averaged sectional streamlines at symmetry plane. Note that sectional streamlines are computed by in-plane velocity components. In order to discuss the flow topology, two specific points are to be observed. The first one is the saddle point observed at the frontier of the recirculation region, resulting from the interaction between the descending velocities (down-wash) and ascending velocities (up-wash) (for the uncorrected RANS model it is around $x/D \simeq 4$,$z/D \simeq 2$). The location of this point is strongly related to the recirculation lengths discussed in section 5.1.3. The second point is the focal point around ($x/D \simeq 1.5$,$z/D \simeq 4.5$) reminiscent of a large spanwise vortex structure, connecting the two symmetric recirculation vortices visible in the horizontal plans in figures 7 and 8. This focal point is a sign of a local depression.

In the PIV symmetry plan, compared to the non-corrected models, the saddle point is located more upstream and slightly higher in z. Consistently with the results of section 5.1.3, the optimisations pull the saddle point upstream, but not sufficiently, and lower in z instead of being raised in the case of the structure-free forcing $\mathbf{F}_1$. With the structured forcing $\mathbf{F}_2$, the z position of the saddle point is correct, but the recirculation length is too large. Forcing on the turbulence closure equation through f_ϵ leads to a drastic reduction of the recirculation region, and an accurate prediction of the saddle point position. As pointed out in [38], this striking result is mitigated by the fact that this saddle point has been pulled slightly too far upstream. Horizontal cross sections in figures 7 and 8 complement the view

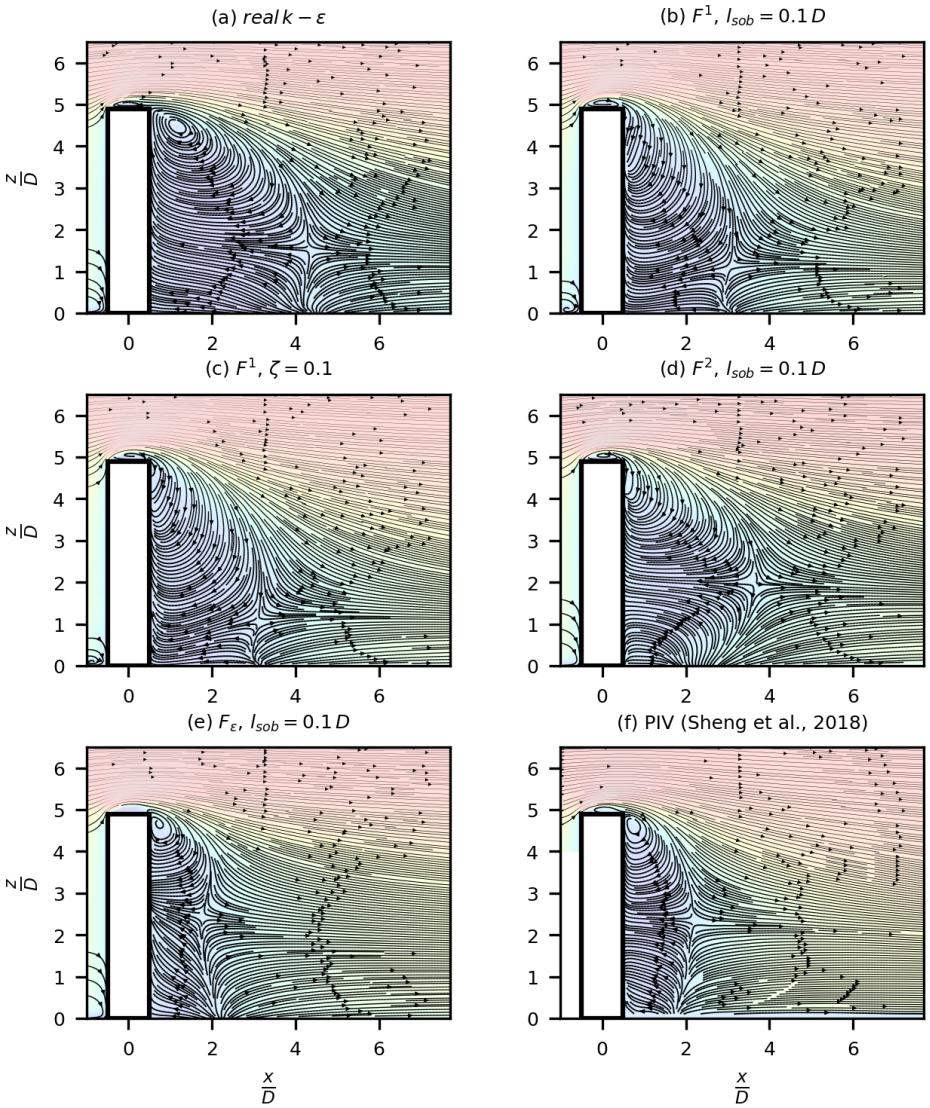


Figure 6: Flow topology (2D) comparison with the different proposed parametrization at symmetry plane.

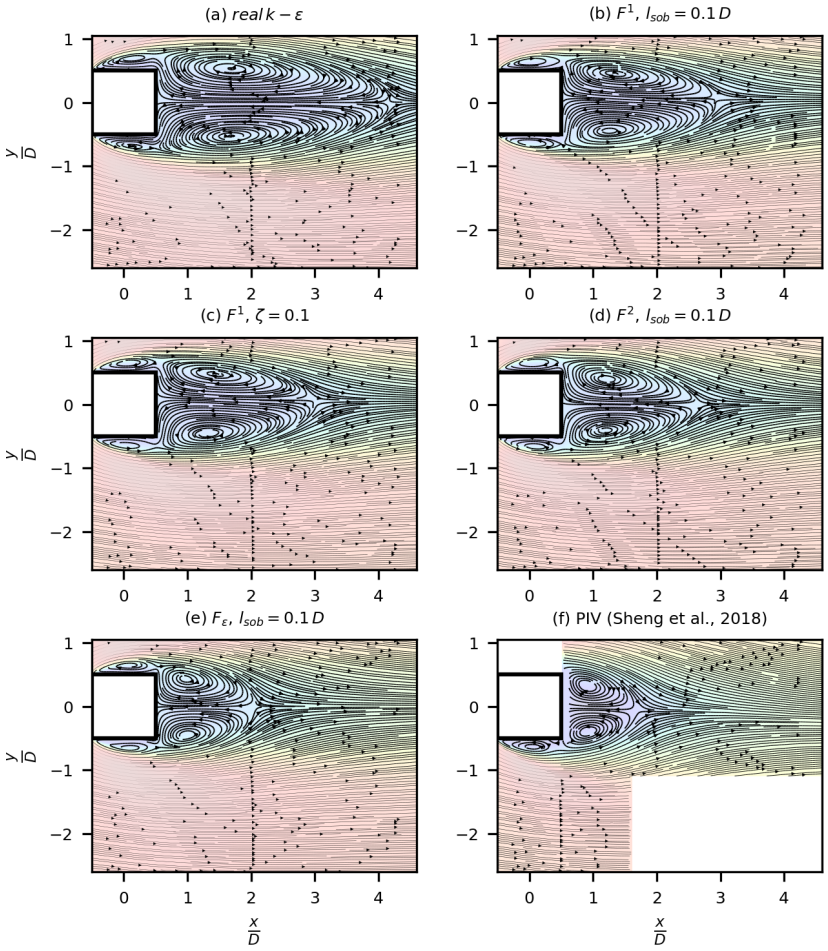


Figure 7: Flow topology (2D) comparison with the different proposed parametrization on horizontal plane at normalized height $z/D = 0.63$.

of the 3D flow structure.

Concerning the focal point in figure 6, complete disappearance is noted by the raw direct forcing $\mathbf{F}_1$; whereas it is moved upstream in the vicinity of the leeward facade in the case of structured force $\mathbf{F}_2$. Yet, with the proposed closure parameter f_ϵ, this critical point is retrieved.

By taking these observations globally, it could be argued that the recirculation region is globally reduced. This is, however, not achieved in an entirely satisfactory way with a raw force.

The results suggest that the pressure reduction in the leeward facade is performed in the optimisation by approaching the focal point very close

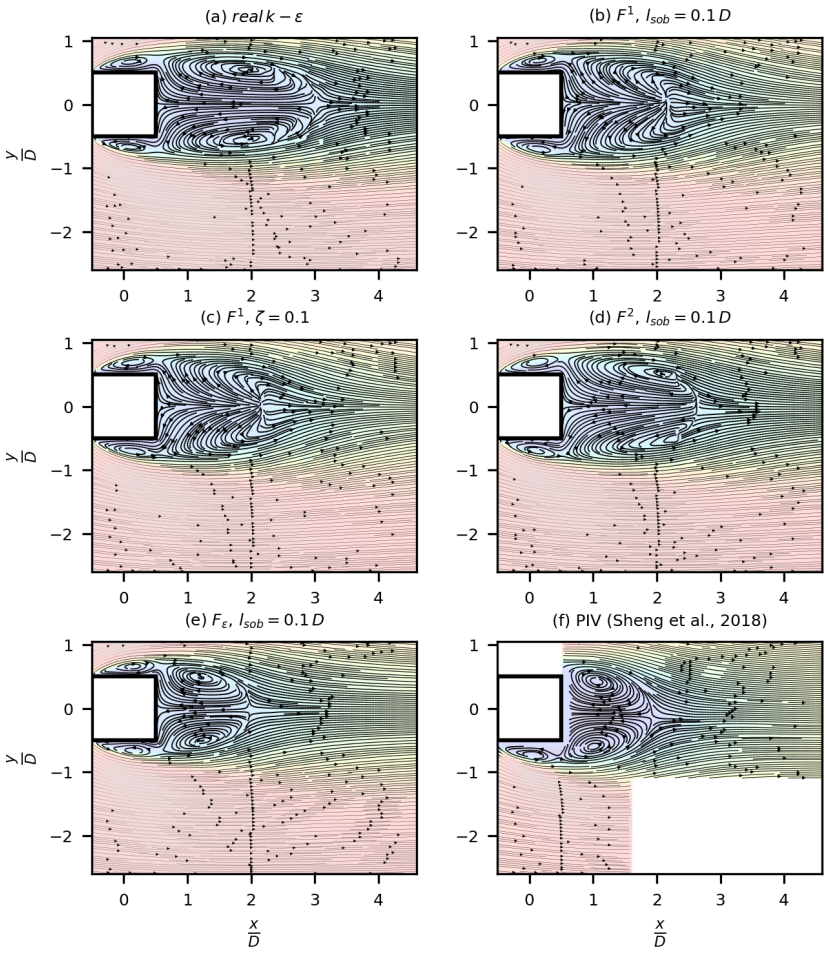


Figure 8: Flow topology (2D) comparison with the different proposed parametrization on horizontal plane at normalized height $z/D = 3.3$.

to the wall instead of significantly reducing the overall recirculation region. This was not the case with f_ϵ as a control parameter, as the two-dimensional vortices at both elevations ($z/D = 0.63, 3.33$) are still captured. Note nonetheless that as one span downward, far enough from the high-end, such that descending motions are negligible (say at $z/D = 0.63$), all models tend to preserve the physical features (see figure 7).

Clearly, each parametrization has its pro and cons when applied individually. Direct forcing, either structured or not, provides a superior ability for wall-pressure interpolation, while embedding parameters in the turbulence closure equation does preserve better the physical features of the wake velocity field. To take advantage of both strategies, it is proposed next to investigate the coupling of these two kinds of corrections. The choice of the control parameter shape plays an important role here. Hence, performing an efficient hybridisation is far from being an easy task as it requires proper scaling between the two corrections.

5.2. Exploiting complementary effects for reconstruction

5.2.1. Control parameters

In order to refine flow reconstructions, this section considers simultaneously two kinds of control parameters; one forcing vector field acting on the momentum equations and a scalar forcing field acting on the transport equation of ϵ. In order to control better the rigidity of the model, a structured form of the forcing on the momentum equation is chosen. The present choice is motivated based on the numerical tests performed in the previous sections to act on two different physical mechanisms.

More precisely, the choice of the control parameter is motivated based on the following observations. It has been shown in section 5.1 that forcing directly on the momentum equation tends to lead to an overfitting of the data. As a matter of fact, the wall pressure is well represented, but the flow field is unphysical. A way to provide a better structure to this forcing is to premultiply it by $|\nabla k|$. This still leads to a good reduction of the wall pressure discrepancy, with performances in terms of the cost functional decrease comparable to the direct forcing. Regarding the flow field, despite a moderate reduction of the wake extension, the vertical position of the saddle point is better recovered.

On the other hand, forcing on the transport of ϵ leads to impressive reduction of the recirculation length. this can be interpreted by the fact that it acts on the turbulent mixing, which is a key driver to simulate correctly the wake shape. The disadvantage of this very structured model is that it

constraints severly the data assimilation process to reach in its exploration a sufficiently large domain of physical solution in the state space[38].

With the objective of performing the best flow reconstruction as possible, since they apparently act on two separate mechanisms, a combination of both strategies is investigated in the following. The model proposed in that aim reads as

$$\frac{\partial(\rho U_j U_i)}{\partial x_j} + \frac{\partial}{\partial x_i}\left(P + \frac{2}{3}\rho k\right) - \frac{\partial}{\partial x_j}\left[\mu_{eff}\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right)\right] = -F_2 \tag{9a}$$

$$\frac{\partial U_i}{\partial x_i} = 0 \tag{9b}$$

$$M_k = 0 \tag{9c}$$

$$M_\epsilon = -f_\epsilon \epsilon \tag{9d}$$

where M_k and M_ϵ corresponds to the tubulent kinetic energy and turbulent dissipation evolutions of the base model.

5.2.2. Quality assessment for flow reconstructions

Validation metrics are defined here to evaluate the quality of the reconstruction. These metrics are proposed to compare the interpolation quality, i.e., the ability to reconstruct the wall pressure on the building based on sparse measurements, and the extrapolation/generalisation quality, i.e., the velocity field reconstruction. For the former purpose, the objective function $\mathcal{J}$ is simply considered.

The wake reconstruction assessment is not straightforward. Consistently with the discussion of the velocity profiles in section 5.1.3, quantities representative respectively of the recirculation length and the maximum reverse flow velocity are defined. At a height z_j, $\mathcal{M}_2^j$ measures the recirculation length relative error, defined as

$$\mathcal{M}_2^j = \left|\Delta L_x^j / \Delta L_x^{j,0}\right|, \tag{10}$$

with ΔL_x^j the error in recirculation length of the optimised model compared with PIV, and $\Delta L_x^{j,0}$ the error of recirculation length of the non-assimilated RANS model compared with PIV.

Similarly, the relative error of maximum reverse velocity at height z^j is defined by

$$\mathcal{M}_3^j = \left|\Delta U_{x_{\min}}^j / \Delta U_{x_{\min}}^{j,0}\right|, \tag{11}$$

	$L_x^\star$	$U_{x_{\min}}^\star$	$\mathcal{M}_2(\%)$	$\mathcal{M}_3(\%)$
Sheng et al. [47]	1.86, 2.13, 1.93	-0.170, -0.470, -0.421	-	-
realizable $k-\epsilon$	4.15, 4.01, 3.18	-0.197, -0.203, -0.181	100, 100, 100	100, 100, 100
F_1, $l_{sob} = 0.1\,D$	2.92, 2.89, 2.16	-0.146, -0.138, -0.095	46.5, 40.3, 18.2	86.8, 124, 136
$F_2 = \|\nabla k\| f_u$, $l_{sob} = 0.1\,D$	2.64, 3.56, 2.75	-0.177, -0.264, -0.211	33.9, 76.0, 65.6	26.7, 77.1, 87.5
F_2, f_ϵ, $l_{sob} = 0.1\,D$	2.43, 2.40, 2.33	-0.158, -0.120, -0.158	25.1, 14.5, 32.0	43.2, 131, 110
F_2, f_ϵ, $l_{sob} = 0.1\,D$, $(\zeta_{f_u}, \zeta_{f_\epsilon}) = (0.1, 15)$	2.29, 2.42, 2.36	-0.115, -0.105, -0.144	18.9, 15.3, 34.4	201, 137, 116
F_2, f_ϵ, $l_{sob} = 0.1\,D$, $(\zeta_{f_u}, \zeta_{f_\epsilon}) = (0.25, 15)$	2.18, 2.32, 2.30	-0.099, -0.091, -0.135	14.5, 10.0, 29.7	260, 142, 120
F_2, f_ϵ, $l_{sob} = 0.1\,D$, $(\zeta_{f_u}, \zeta_{f_\epsilon}) = (0.5, 15)$	2.22, 2.17, 2.25	-0.104, -0.080, -0.120	15.7, 2.20, 25.0	241, 147, 126
f_ϵ, $l_{sob} = 0.1\,D$	2.06, 1.77, 1.98	-0.135, -0.055, -0.106	8.69, 18.9, 3.94	129, 156, 131

Table 1: Normalized Recirculation length $L_x^\star$ and peak streamwise velocity $U_x^\star$ inside the recirculating flow at two height $z/D = (0.63, 2.5, 3.33)$ on the symmetry plane ($y/D = 0$) with different parametrization. Variables are normalized by D and U_{ref}, respectively.

where $\Delta U^j_{x_{\min}}$ stands for the error on the wake's peak velocity of the optimised model compared with PIV, and $\Delta U^{j,0}_{x_{\min}}$ for the error on the maximum reverse velocity of the non-assimilated RANS model compared with PIV.

As seen in section 5.1.3, it turns out that the two criteria $\mathcal{M}_2$ and $\mathcal{M}_3$ follow an opposite trend by varying the model rigidity, from an overfitted to an over-constrained situation. A good estimation of the wake velocity field consists of performing a trade-off between these quantities for all heights z^j. It is then proposed to define the new quantity

$$A^2(\mathcal{M}_2, \mathcal{M}_3) = \frac{(1 - \frac{1}{n_z}\sum_j^{n_z} \mathcal{M}_2^j)}{\frac{1}{n_z}\sum_j^{n_z} \mathcal{M}_3^j} \tag{12}$$

where n_z denotes the number of considered heights. Here it should be pointed out that a large value of $A^2(\mathcal{M}_2, \mathcal{M}_3)$ is the footprint of a good estimation.

Yet, the choice of metric is not unique. The present choice has the advantage of being more sensitive than a standard $L^2(\Omega)$ norm, by being oriented toward a worst-case scenario in identifying maximum values. The average over several peak values confers robustness to the metric.

The first two metrics are given in table 1, measured at three heights, namely $z/D = 0.33, 2.5, 3.33$ in the symmetry plane.

Averaged metrics are collected in table 2. Tables summarize the current (hybrid) parametrization results with different regularization levels: the individual, direct forcing form, the structured force alone, and the corrected closure on ϵ budget alone.

	$\varepsilon_{\mathcal{J}} = \mathcal{J}^{end}/\mathcal{J}^{0}(\%)$	$A^1(\mathcal{M}_2)$ (%)	$A^1(\mathcal{M}_3)$ (%)	$A^2(\mathcal{M}_2, \mathcal{M}_3)$
Sheng et al. [47]	-	-	-	-
realizable $k-\epsilon$	100	100	100	-
F_1, $l_{sob} = 0.1\,D$	20.9	35	116	0.56
$F_2 = \lvert\nabla k\rvert f_u$, $l_{sob} = 0.1\,D$	26.9	58.5	<u>63.8</u>	0.65
F_2, f_ϵ, $l_{sob} = 0.1\,D$	23.7	23.9	94.6	<u>0.80</u>
F_2, f_ϵ, $l_{sob} = 0.1\,D$, $(\zeta_{f_u}, \zeta_{f_\epsilon}) = (0.1, 15)$	<u>20.3</u>	22.8	151	0.51
F_2, f_ϵ, $l_{sob} = 0.1\,D$, $(\zeta_{f_u}, \zeta_{f_\epsilon}) = (0.25, 15)$	25.2	18.1	174	0.47
F_2, f_ϵ, $l_{sob} = 0.1\,D$, $(\zeta_{f_u}, \zeta_{f_\epsilon}) = (0.5, 15)$	29	14.3	171	0.50
f_ϵ, $l_{sob} = 0.1\,D$	43.4	<u>10.5</u>	139	0.64

Table 2: Summarizing scores for reconstruction quality.

These quantitative comparisons are accompanied by graphs of the cost function reduction, illustrated in figure 9, and wall-pressure distributions, in figure 10. In the following, these tables and figures are discussed by comparing the proposed criteria.

Objective reduction assessment. Among all tests performed, the best cost functional reduction is obtained with the coupled scaling $(\zeta_{f_{u_i}}, \zeta_{f_\epsilon}) = (0.1, 15)$. The reached performance is even better than when a direct raw forcing $\mathbf{F}^1$ without penalty (highly overfitted) was considered. In the following sections, the physical relevance of the flow structure in these cases is explored.

Employing penalty in conjunction with Sobolev gradient seems to improve the cost functional reduction. Here, the effect of this is explored as follows. Three penalization scaling factors of the momentum forcing were considered: $\zeta_{f_{u_i}} = 0.1, 0.25, 0.5$, with a given penalization to ϵ fixed at $\zeta_{f_\epsilon} = 15$. Fixing instead $\zeta_{f_{u_i}}$ and sweeping ζ_{f_ϵ}, leads to an opposite trend since it is the ratio between them which drives the overall behaviour. As expected, high values of $\zeta_{f_{u_i}}$ leads to large constraints for the optimisation, while lower values does not efficiently suppress non-physical oscillations. Moreover, it has been observed that $\zeta_{f_{u_i}}$ has a significantly higher impact on the cost functional than ζ_{f_ϵ}. This is consistent with the fact that forcing on the momentum equations allows modifying the wall-pressure easily. Nonetheless, a high impact of ζ_{f_ϵ} on the flow field should be expected, as f_ϵ affects the bulk momentum diffusion.

Wake predictions. Examining the averaged flow metrics $A^1(\mathcal{M}_2), A^1(\mathcal{M}_3)$ reveals a reversed trend between them. The larger the reduction of the recirculation length, the worst the prediction of the peak streamwise velocity. With no penalization considered, the only adjustment of the closure (by

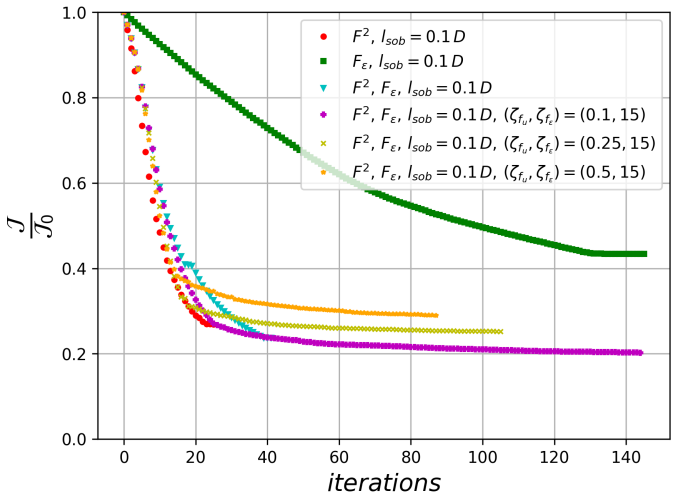


Figure 9: Cost function reduction under a correction of a structured force with a modified closure applied on ϵ budget for different penalization $(\zeta_{f_u}, \zeta_{f_\epsilon}) = ([0.1\,0.25\,0.5], 15)$. Comparison is made between raw-force (black circle), structured force (blue square), default closure, and modified closure made at the ϵ budget (yellow triangle).

f_ϵ) enables the lowest wake extension but leads to the slightest accurate predictions, falling outside the base model error. On the other hand, the only correction of the structured force leads to the least wake contraction and the best error decrease.

To measure the trade-off between recirculation length and peak velocity, the quantity $A^2(\mathcal{M}_2, \mathcal{M}_3)$ (table 2) is introduced. From table 2, it can be observed that the synchronous adjustment of the closure with the force provides the better compromise ($A^2 = 0.8$), which is better achieved with no penalization. It can also be seen that the direct forcing gives a very low compromise ($A^2 = 0.56$) as compared with the most constrained parametrization by only correcting f_ϵ ($A^2 = 0.64$). This supports the idea, that this metric is only partially representative of the quality of reproduction of flow features.

By examining the flow topology, illustrated in figures 11 and 12, it can be seen that correcting at the same time $(|\nabla k| f_u, f_\epsilon)$ leads to an intermediate solution between both individual corrections $(|\nabla k| f_u, 0)$ and $(0, f_\epsilon)$.

It can be seen that recirculation region in the symmetry plane has significantly reduced and a saddle point location in good agreement with PIV data. However, it is still noted that the model excessively pulls the focal point near the high end to the leeward facade, the mechanism employed

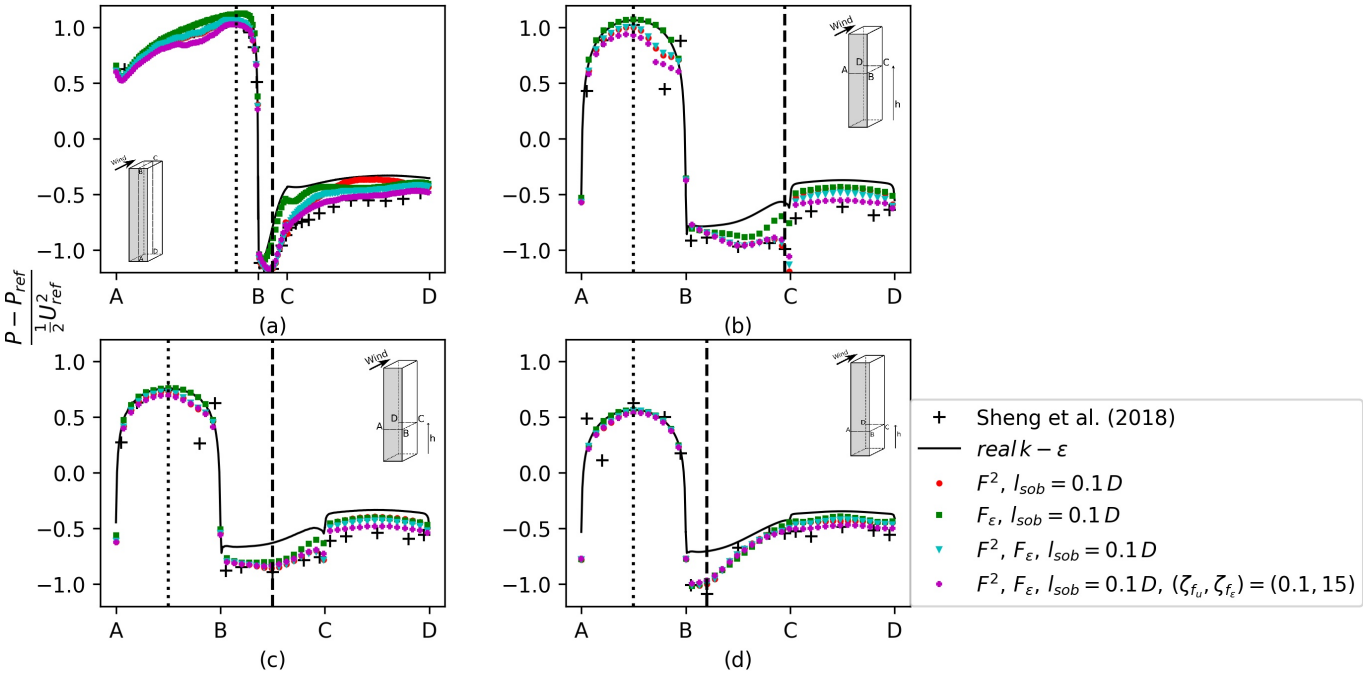


Figure 10: Pressure coefficient profiles along building facades under a correction of a structured force with a modified closure applied on ϵ budget for different penalization $(\zeta_{f_u},\zeta_{f_\epsilon}) = ([0.1\,0.25\,0.5],15)$. Comparison is made between raw-force (black circle), structured force (blue square), default closure, and modified closure made at the ϵ budget (yellow triangle).

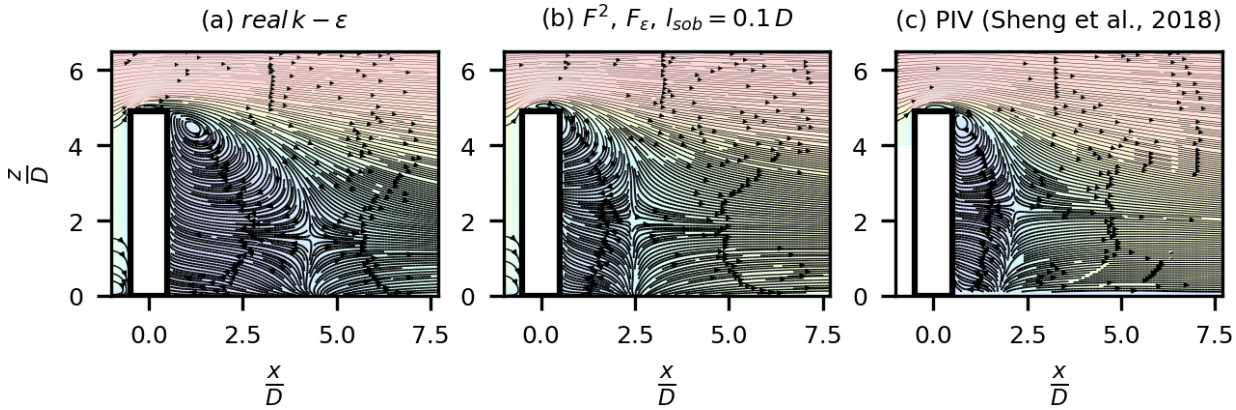


Figure 11: Flow topology (2D) with reconstruction under a correction of a structured force with a modified closure on symmetry plane.

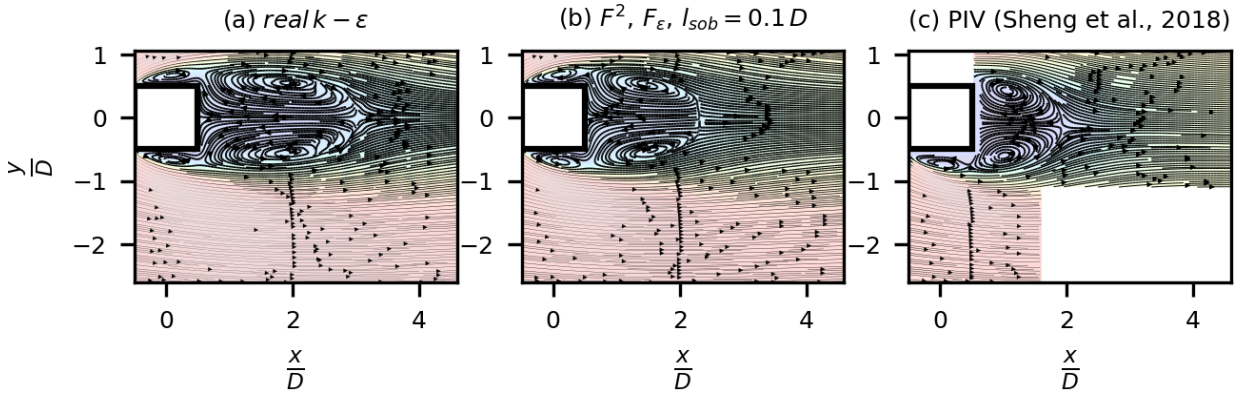


Figure 12: Flow topology (2D) with reconstruction under a correction of a structured force with a modified closure on horizontal plane at normalized height $z/D = 3.33$.

by the optimisation to fit the wall-pressure at the high end of the leeward facade. Nonetheless, this remaining overfitting artifact is very limited in space as the transverse structures are well retrieved. The reconstructed flow still recovers the lateral vortex rolls, which extend from the ground and meet at the high-end vortex on the symmetry plane, with the focal points in the horizontal planes being well captured until height $z/D = 3.33$ with the downstream location of $x/D \simeq 1.5$. This location is intermediate between the base RANS and the PIV ($x/D \simeq 1$) (see figure 12). Note that, except with the correction of f_ϵ solely, both structured/raw form of forcing with default closure does not capture them at all.

Therefore, by considering both the qualitative observations of the flow features and the proposed wall/flow metrics, it can be summarized that combination of a direct forcing on the momentum equation with simple structuring form, along with a forcing on the transport equation of ϵ is a promising strategy. It ensures overall a good quality to the reconstruction. It is worth noting that although the quantitative flow metric A^2 did not favors the penalization, examining the qualitative flow topology suggests a better restitution of some physical features. It was observed by the re-capturing of the high-end focal point with the couple of scaling $(\zeta_{f_{u_i}}, \zeta_{f_\epsilon}) = (0.5, 15)$, which is otherwise absent (see figure 13). Therefore, providing a penalization plays an important role in refining the flow features. Yet, an adequate couple of scaling is required to provide a good compromise between a well-interpolated wall pressure and the best retrieval of realistic flow patterns.

Incorporating the forcing in the transport equation of turbulence dis-

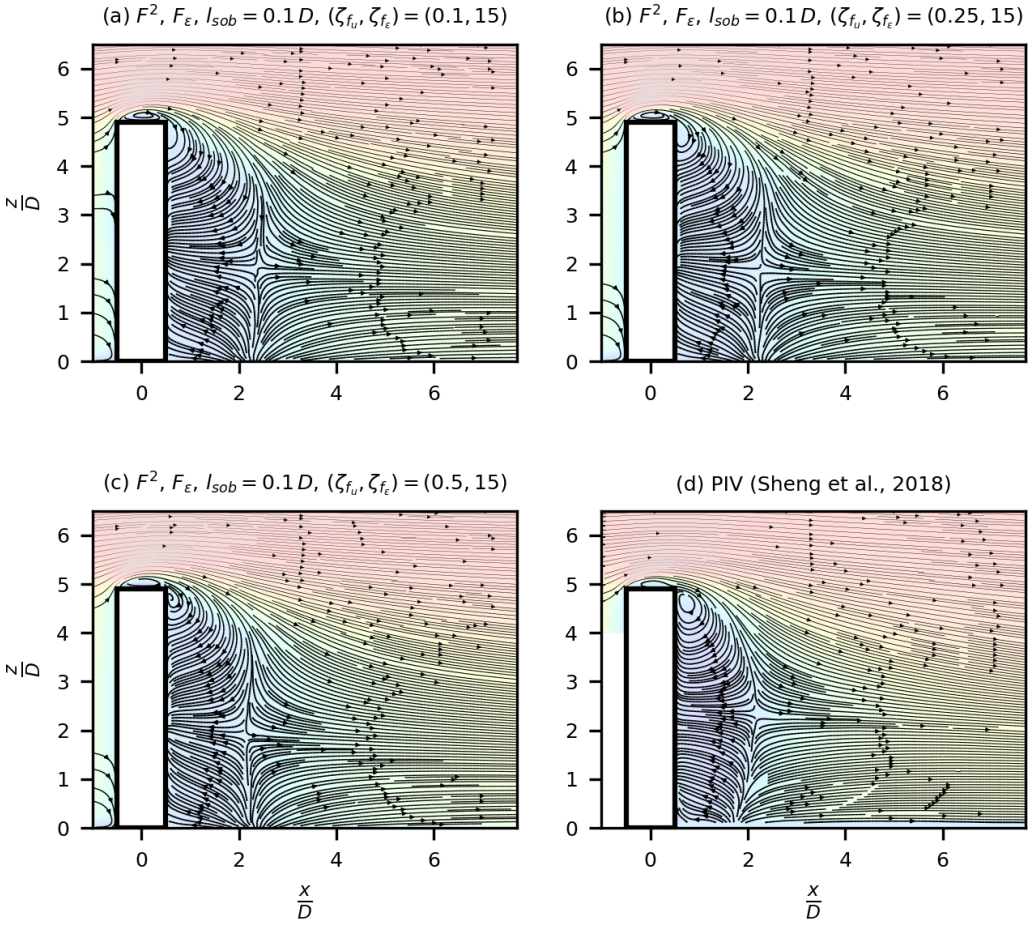


Figure 13: Flow topology (2D) on symmetry plane. Comparison with flow reconstruction,(a) under a correction of a structured force with a modified closure applied on ϵ budget with no penalization, (b) with penalization $(\zeta_{f_{u_i}}, \zeta_{f_\epsilon}) = (0.1, 15)$, (c) $(\zeta_{f_{u_i}}, \zeta_{f_\epsilon}) = (0.5, 15)$.

sipation rate reduces the overfitting. It suggests that the structure of the RANS model is relevant and that the corrections to be added lye where the closure is performed. This principle is likely generalisable for other models: the control parameters should correct the closure models which are by definition not fully known. The second main observation is that a very good improvement is reached relaxing the Boussinesq assumption by a forcing on the momentum equation. This suggests that in wake flows, the effect of large scale structures and/or strong inhomogeneities and anisotropy (as in shear layers) have a significant effect on the mean flow. Having accurate results through a forcing on the momentum equation is a hard task since an overfitting issue is faced. It is achieved thanks to the hybrid approach that provides a forcing on both equations.

6. Conclusion

The present study aimed to investigate variational data assimilation (DA) methodologies devised in previous work [38] to reconstruct 3D wind flows around structures and provide guidelines toward an efficient reconstruction. In this framework, the Reynolds Averaged Navier–Stokes equations (RANS) constitute the base mathematical model to describes the mean wind flow. In particular, the case on which this study focuses on was dedicated to the analysis of a high-rise building with a square section and an aspect ratio $H/D \sim 5$. For its improved capability to account for the energy transfer with strong strain rates, the realizable revision developed by Shih et al. [48] was also chosen as a reference turbulence closure. The inflow wind profile models the lower part of the atmospheric boundary layer under the assumption of horizontal homogeneity. The data considered for the assimilation were the averaged wall-pressure measured on a scaled model of the building in a wind tunnel experiment by Sheng et al. [47] at CSTB (Nantes, France).

Several important findings arising from this study can be highlighted:

- In 3D wake flows, despite the fact than RANS models systemically overpredict the recirculation length, wall-pressure information is a meaningful enough piece of information to recover acurately experimental wake extension;

- In weak constraint VDA, a control parameter on the transport equation within which the closure is performed avoids overfitting. Over-

679 fitting can be acceptable for wall-pressure interpolation, but reveals
680 catastrophic for wake flow reconstructions;

681 • A "hybrid" control parameter on the transport of ϵ and on the mo-
682 mentum equation leads to very accurate results due to a relaxation of
683 the Boussinesq assumption;

684 • Sobolev gradient descent direction leads to efficient regularisation and
685 to a fast convergence.

686 The present study constitutes a first step towards applying DA proce-
687 dure to real-world applications in structural wind engineering, where veloc-
688 ity field measurements of the whole 3D domain are not accessible. Note
689 that in this work, the quality of the data-assimilated flow field is quantified
690 by validation PIV data plans (but only wall-pressure measurements were
691 used for estimation). With the guidelines provided in this study, we can
692 imagine validations with more sparse and less costly measurements (sonic
693 anemometers) for estimations in operational conditions.
694 The reconstruction strategy employed in the present paper considered
695 a distributed additive forcing control parameter, acting on the momentum
696 equation and/or the transport equation of turbulent dissipations, where the
697 closure is performed.
698 In the first step, to recover the mean flow from the pressure measure-
699 ments, an unknown spatially distributed forcing was added to the momen-
700 tum equation in order to infer corrections to the Reynolds induced force.
701 Without a priori knowledge of its nature, an initial raw form was consid-
702 ered but led to unphysical oscillations. To avoid that, two regularization
703 approaches were investigated; the first one proceeded by penalizing the gra-
704 dient of the control parameter while the second was instead conducted by
705 Sobolev gradient [51, 59]. Regarding the effects of regularisation, cost re-
706 duction results with the Sobolev gradient yielded a much faster convergence,
707 lower discrepancy levels, and an excellent agreement with the wall-pressure
708 experiments in most building's wall regions. An insufficient reduction of
709 the recirculation region, associated with non-physical features, has been ob-
710 served regarding flow reconstruction ability. It has raised the point about
711 the lack of structuration in the forcing.
712 In order to remedy this situation, authors proceeded by refining the
713 shape of the force with the aim to ensure some consistency with the forc-
714 ing directly related to the Reynolds stresses. To that purpose, a pre-
715 multiplication by the gradient of the kinetic energy favours corrections

where Reynolds stresses are likely active. In terms of objective, globally, the discrepancy with wall-pressure experiments has been reduced significantly and in a faster way, allowing us to reach a state very close to the one achieved with raw forcing. Moreover, a better retrieval of the wake features was noticed, particularly close to the ground reattachment point. Unfortunately, it still leads to an insufficient reduction of the wake.

Alternatively, correction on the transport equation of ϵ was also considered. The modified closure produced a drastic improvement of the wake extension, associated with a moderate performance of wall-pressure reconstruction. This behaviour is well understood because acting near the turbulence closure enables the maintenance of the model structure and avoiding overfitting the wall pressure measurements. However, as the flow adjustment comes only from diffusive effects, it suggests some room for improvement with other forms of the (relaxing for instance the Boussinesq eddy viscosity assumption) distributed parameters to obtain accurate flow estimations.

With these results in mind, an intermediate solution was proposed. While maintaining the Boussinesq assumption and correcting the turbulence closure equation, an addition of a direct correction of the deformation with a premultiplication factor is considered. To go toward an efficient hybridization, control parameter forms have played a significant role here. The reconstruction results were then discussed, and the quality of the reconstruction was evaluated qualitatively and by the definition of some wall/flow validations metrics. Both the qualitative and quantitative observations have shown that the combination of a direct forcing with a simple structuring form and local closure adjustment is a promising strategy as it ensures overall a good quality to the reconstruction.

It is interesting here to point out the great improvement in the prediction of the tearing pressures, often the most important information needed for dimensioning, goes with the improvement in the prediction of the wake. Moreover, the conjunction of descent direction in H^1 and gradient penalization in L^2 allows us to refine the quality of the flow estimation. Besides the quality benefits on the final flow field, it was noticed that a robust DA procedure could be ensured with the addition of penalty.

This allows to consider this kind of parametrization as a potential candidate for an efficient complex flow reconstruction given the limited amount of data in hand. More so, with regards to the chosen parameters *forms*, the authors are confident that this is quite generalizable as Boussinesq assup-

tion represents one fundamental bottleneck of the RANS modelling. It is believed that this may represents an advancement over the state of the art on flow reconstruction, and it constitutes an additional step toward more challenges in wind engineering such as flow reconstructions around complex buildings or within urban cities.

Appendix A. Closure parameters in the realizable $k-\epsilon$ RANS equations

In this appendix, closure coefficients of the realizable $k-\epsilon$ RANS equations are detailed. First, σ_k (resp. σ_ϵ) is a closure constant, that allows the turbulent mixing to have a different intensity in the momentum equations than in the transport equation of kinetic energy k (resp. dissipation rate ϵ). Besides, in the transport equation of ϵ,

$$C_1 = \max\left(0.43, \frac{\eta}{(5+\eta)}\right) \tag{A.1}$$

with

$$\eta = Sk/\epsilon \tag{A.2}$$

is the normalised strain rate. The constant C_2 is a closure coefficients.

Recall that the particularity of the realizable revision by Shih et al. [48] is that the coefficient C_μ is a non uniform coefficient that depends on the mean strain-rate and ensures the realizability conditions, i.e. Schwartz inequality and the non-negativity of the diagonal Reynolds stress. It is defined by

$$C_\mu = \frac{1}{A_0 + A_s U_s \frac{k}{\epsilon}}, \tag{A.3}$$

with

$$U_s = \sqrt{S_{ij}S_{ij} + \Omega_{ij}\Omega_{ij}} \tag{A.4a}$$

$$S_{ij} = \frac{1}{2}\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right) \tag{A.4b}$$

$$\Omega_{ij} = \frac{1}{2}\left(\frac{\partial U_i}{\partial x_j} - \frac{\partial U_j}{\partial x_i}\right) \tag{A.4c}$$

and

$$A_s = \sqrt{6}\cos(\phi), \quad \phi = \frac{1}{3}\arccos(\sqrt{6}W) \tag{A.5a}$$

$$W = \frac{S_{ij}S_{jk}S_{ki}}{S}, \quad S = \sqrt{2S_{ij}S_{ij}}, \tag{A.5b}$$

where S_{ij} and Ω_{ij} are, respectively, the mean strain (with its magnitude S) and rotation rates. The parameter A_0 is a tuning coefficient.

References

[1] S. Murakami, A. Mochida, Y. Hayashi, S. Sakamoto, Numerical study on velocity-pressure field and wind forces for bluff bodies by $k-\epsilon$, ASM and LES, Journal of Wind Engineering and Industrial Aerodynamics 44 (1992) 2841–2852.

[2] P. A. Irwin, Bluff body aerodynamics in wind engineering, Journal of Wind Engineering and Industrial Aerodynamics 96 (2008) 701–712.

[3] L. Cochran, R. Derickson, A physical modeler's view of computational wind engineering, Journal of Wind Engineering and Industrial Aerodynamics 99 (2011) 139–153.

[4] D. Surry, Wind loads on low-rise buildings: Past, present and future, Wind Engineering into the 21^{st} Century (1999).

[5] P. A. Irwin, W. W. Kochanski, Measurement of structural wind loads using the high frequency pressure integration method, in: Restructuring: America and Beyond, ASCE, 1995, pp. 1631–1634.

[6] R. J. Adrian, Particle-imaging techniques for experimental fluid mechanics, Annual review of fluid mechanics 23 (1991) 261–304.

[7] G. Elsinga, F. Scarano, B. Wieneke, B. W. van Oudheusden, Tomographic particle image velocimetry, Experiments in fluids 41 (2006) 933–947.

[8] S. Murakami, A. Mochida, K. Hibi, Three-dimensional numerical simulation of air flow around a cubic model by means of large eddy simulation, Journal of Wind Engineering and Industrial Aerodynamics 25 (1987) 291–305.

[9] S. Murakami, Computational wind engineering, Journal of Wind Engineering and Industrial Aerodynamics 36 (1990) 517–538.

[10] S. Murakami, Comparison of various turbulence models applied to a bluff body, in: Computational Wind Engineering 1, 1993, pp. 21–36.

[11] W. Rodi, Comparison of LES and RANS calculations of the flow around bluff bodies, Journal of Wind Engineering and Industrial Aerodynamics 69-71 (1997) 55–75.

[12] T. G. Thomas, J. J. R. Williams, Simulation of skewed turbulent flow past a surface mounted cube, J. Wind Eng. Ind. Aerodyn. (1999) 14.

[13] K. Nozawa, T. Tamura, Large eddy simulation of the flow around a low-rise building immersed in a rough-wall turbulent boundary layer, Journal of Wind Engineering and Industrial Aerodynamics 90 (2002) 1151–1162.

[14] B. E. Launder, B. I. Sharma, Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc, Letters in Heat and Mass Transfer 1 (1974) 131–137.

[15] R. I. Issa, Rise of total pressure in frictional flow, AIAA journal 33 (1995) 772–774.

[16] J. Franke, C. Hirsch, A. Jensen, H. Krüs, M. Schatzmann, P. Westbury, S. Miles, J. Wisse, N. Wright, Recommendations on the use of CFD in predicting pedestrian wind environment, in: Cost action C, volume 14, 2004.

[17] A. Sohankar, L. Davidson, C. Norberg, Large Eddy Simulation of Flow Past a Square Cylinder: Comparison of Different Subgrid Scale Models, Journal of Fluids Engineering 122 (2000) 39–47.

[18] A. Gosman, Developments in CFD for industrial and environmental applications in wind engineering, Journal of Wind Engineering and Industrial Aerodynamics 81 (1999) 21–39.
[19] G. Tabor, M. Baba-Ahmadi, Inlet conditions for large eddy simulation: A review, Computers & Fluids 39 (2010) 553–567.
[20] M. S. Thordal, J. C. Bennetsen, H. H. H. Koss, Review for practical application of CFD for the determination of wind load on high-rise buildings, Journal of Wind Engineering and Industrial Aerodynamics 186 (2019) 155–168.
[21] P. Bergthörsson, B. R. Döös, Numerical weather map analysis, Tellus 7 (1955) 329–340.
[22] V. Mons, J. C. Chassaing, T. Gomez, P. Sagaut, Reconstruction of unsteady viscous flows using data assimilation schemes, Journal of Computational Physics 316 (2016) 255–280.
[23] M. Meldi, A. Poux, A reduced order model based on Kalman filtering for sequential data assimilation of turbulent flows, Journal of Computational Physics 347 (2017) 207–234.
[24] F. X. Le Dimet, O. Talagrand, Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, Tellus A: Dynamic Meteorology and Oceanography 38 (1986) 97–110.
[25] P. Courtier, O. Talagrand, Variational assimilation of meteorological observations with the adjoint vorticity equation. ii: Numerical results, Quarterly Journal of the Royal Meteorological Society 113 (1987) 1329–1347.
[26] A. Gronskis, D. Heitz, E. Mémin, Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation, Journal of Computational Physics 242 (2013) 480–497.
[27] Y. Yang, C. Robinson, D. Heitz, E. Mémin, Enhanced ensemble-based 4DVar scheme for data assimilation, Computers & Fluids 115 (2015) 201–210.
[28] V. Mons, J. Chassaing, T. Gomez, P. Sagaut, Is isotropic turbulence decay governed by asymptotic behavior of large scales? An eddy-damped quasi-normal markovian-based data assimilation study, Physics of Fluids 26 (2014) 115105.
[29] V. Mons, L. Margheri, J. C. Chassaing, P. Sagaut, Data assimilation-based reconstruction of urban pollutant release characteristics, Journal of Wind Engineering and Industrial Aerodynamics 169 (2017) 232–250.
[30] P. Chandramouli, E. Memin, D. Heitz, 4D large scale variational data assimilation of a turbulent flow with a dynamics error model, Journal of Computational Physics (2020) 109446.
[31] D. Etling, H. W. Detering, F. Theunert, On the simulation of wind-driven currents in shallow water, Archives for meteorology, geophysics, and bioclimatology, Series A 33 (1985) 355–363.
[32] P. G. Duynkerke, Application of the $E - \epsilon$ turbulence closure model to the neutral and stable atmospheric boundary layer, Journal of the Atmospheric Sciences 45 (1988) 865–880.
[33] S. Tavoularis, U. Karnik, Further experiments on the evolution of turbulent stresses and scales in uniformly sheared turbulence, Journal of Fluid Mechanics 204 (1989) 457–478.
[34] H. Kato, S. Obayashi, Data Assimilation for Turbulent Flows, in: 16th AIAA

Non-Deterministic Approaches Conference, National Harbor, Maryland, 2014.
[35] W. N. Edeling, P. Cinnella, R. P. Dwight, H. Bijl, Bayesian estimates of parameter variability in the $k-\epsilon$ turbulence model, Journal of Computational Physics 258 (2014) 73–94.
[36] L. Margheri, M. Meldi, M. V. Salvetti, P. Sagaut, Epistemic uncertainties in rans model free coefficients, Computers & Fluids 102 (2014) 315–335.
[37] M. Shirzadi, P. A. Mirzaei, M. Naghashzadegan, Improvement of k-epsilon turbulence model for CFD simulation of atmospheric boundary layer around a high-rise building using stochastic optimization and Monte-Carlo sampling technique, Journal of Wind Engineering and Industrial Aerodynamics 171 (2017) 366–379.
[38] M. Y. Ben Ali, G. Tissot, D. Heitz, S. Aguinaga, E. Mémin, An adjoint approach for the analysis of RANS closure using pressure measurements on a high-rise building, 2020. URL: https://hal.archives-ouvertes.fr/hal-03076369.
[39] D. P. G. Foures, N. Dovetta, D. Sipp, P. J. Schmid, A data-assimilation method for Reynolds-Averaged Navier-Stokes-driven mean flow reconstruction, Journal of Fluid Mechanics 759 (2014) 404–431.
[40] S. Symon, N. Dovetta, B. J. McKeon, D. Sipp, P. J. Schmid, Data assimilation of mean velocity from 2D PIV measurements of flow over an idealized airfoil, Experiments in Fluids 58 (2017) 61.
[41] K. Duraisamy, Z. J. Zhang, A. P. Singh, New approaches in turbulence and transition modeling using data-driven techniques, in: 53rd AIAA Aerospace Sciences Meeting, Kissimmee, Florida, 2015.
[42] E. J. Parish, K. Duraisamy, A paradigm for data-driven predictive modeling using field inversion and machine learning, Journal of Computational Physics 305 (2016) 758–774.
[43] A. P. Singh, K. Duraisamy, Using field inversion to quantify functional errors in turbulence closures, Physics of Fluids 28 (2016) 045110.
[44] L. Franceschini, O. Sipp, D.and Marquet, Mean-flow data assimilation based on minimal correction of turbulence models: Application to turbulent high Reynolds number backward-facing step, Phys. Rev. Fluids 5 (2020) 094603.
[45] P. Spalart, S. Allmaras, A one-equation turbulence model for aerodynamic flows, in: 30th Aerospace Sciences Meeting and Exhibit, Reno,NV,U.S.A., 1992.
[46] M. A. Iglesias, K. J. H. Law, A. M. Stuart, Ensemble Kalman methods for inverse problems, Inverse Problems 29 (2013) 045001.
[47] R. Sheng, L. Perret, I. Calmet, F. Demouge, J. Guilhot, Wind tunnel study of wind effects on a high-rise building at a scale of 1: 300, Journal of Wind Engineering and Industrial Aerodynamics 174 (2018) 391–403.
[48] T. H. Shih, W. W. Liou, A. Shabbir, Z. Yang, J. Zhu, A new $k-\epsilon$ eddy viscosity model for high Reynolds number turbulent flows-model development and validation, Document ID: 19950005029 (Acquired Dec 28, 1995) Accession (1995).
[49] J. L. Lions, Optimal control of systems governed by partial differential equations problèmes aux limites (1971).
[50] M. D. Gunzburger, Perspectives in flow control and optimization, volume 5, 2003.
[51] B. Protas, T. R. Bewley, G. Hagen, A computational framework for the regularization of adjoint analysis in multiscale PDE systems, Journal of Computational Physics 195 (2004) 49–89.

[52] N. EN, 1-4: Eurocode 1: Actions sur les structures–partie 1-4: Actions générales–actions du vent, 2005.

[53] P. J. Richards, R. P. Hoxey, Appropriate boundary conditions for computational wind engineering models using the $k-\epsilon$ turbulence model, in: Computational Wind Engineering, 1993, pp. 145–153.

[54] P. Richards, S. Norris, Appropriate boundary conditions for computational wind engineering models revisited, Journal of Wind Engineering and Industrial Aerodynamics 99 (2011) 257–266.

[55] H. G. Weller, G. Tabor, H. Jasak, C. Fureby, A tensorial approach to computational continuum mechanics using object-oriented techniques, Computers in physics 12 (1998) 620–631.

[56] C. M. Rhie, W. L. Chow, Numerical study of the turbulent flow past an airfoil with trailing edge separation, AIAA journal 21 (1983) 1525–1532.

[57] R. F. Warming, R. M. Beam, Upwind second-order difference schemes and applications in aerodynamic flows, AIAA Journal 14 (1976) 1241–1249.

[58] E. Robertson, V. Choudhury, S. Bhushan, D. Walters, Validation of OpenFOAM numerical methods and turbulence models for incompressible bluff body flows, Computers & Fluids 123 (2015) 122–145.

[59] G. Tissot, R. Billard, G. Gabard, Optimal cavity shape design for acoustic liners using Helmholtz equation with visco-thermal losses, Journal of Computational Physics 402 (2020) 109048.