An adjoint approach for the analysis of RANS closure using pressure measurements on a high-rise building

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

An adjoint approach for the analysis of RANS closure using pressure measurements on a high-rise building

M.Y. Ben Ali[1,2]†, G. Tissot[2], D. Heitz[3], S. Aguinaga[1] and E. Mémin[2]

[1]Centre Scientifique et Technique du Bâtiment (CSTB), 11 rue Henri Picherit, F-44300 Nantes, France

[2]INRIA, Fluminance team, Campus universitaire de Beaulieu, F-35042 Rennes, France

[3]INRAE, UR OPAALE, F-35044 Rennes Cedex, France

(Received xx; revised xx; accepted xx)

In this study we investigate the closure of a common turbulence model for Reynolds averaged Navier-Stokes (RANS) in the framework of variational data assimilation prediction of wind flows around big structures. This study considers practical experimental settings where only sparse wall-pressure data, measured in a wind tunnel, are available. The evaluation of a cost functional gradient is efficiently carried out with the exact continuous adjoint of the RANS model. Particular attention is given to the derivation of the adjoint turbulence model and the adjoint wall law. Given the dual description of the dynamics, composed of the RANS model and its adjoint, some methodological settings that enable diagnosis of the turbulence closure are explored here. They range from adjoint maps analysis to global constants calibration and finally consider the adjunction of a distributed parameter. Numerical results on a high-rise building reveal a high reconstruction ability of the adjoint method. A good agreement in wind load and wake extension was obtained. As with sparse observations, the sensitivity field is generally not very regular for distributed parameters, a projection onto a space of more regular functions belonging to the Sobolev space (H^1) is also proposed to strengthen the efficiency of the method. This has been shown to lead to a very efficient data assimilation procedure as it provides an efficient descent direction as well as a useful regularisation mechanism. Beyond providing an efficient data-driven reconstruction technique, the proposed adjoint methodology enables an in-depth analysis of the turbulence closure and finally improves it significantly.

Key words:

MSC Codes *(Optional)* Please enter your MSC Codes here

† Email address for correspondence: mohamed-yacine.ben-ali@inria.fr

1. Introduction

During there lifetime, buildings are continuously exposed to wind coming from all directions. Particularly, due to their extended exposed surface, high-rise buildings and big structures undergo extremely strong lateral aerodynamic forces. As a consequence, large lateral deflections or, even more, some problematic tearing affecting security may be observed. Yet, by having a prior understanding of the airflow surrounding these big structures, wind loads can be predicted and such issues avoided. To understand how theses turbulent flows affect the structures, physical models along with numerical simulations are usually deployed. On the one hand, for more than a century, experiments with scaled models of buildings have been carried out in wind tunnels (Cermak & Koloseus 1954; Jensen 1958). After years of advancement on measurement techniques together with an increased knowledge on wind dynamics, tunnel experiments have proven their reliability for loads prediction issues (Surry 1999; Cochran & Derickson 2011). High-Frequency Force Balance method (HFFB) (Tschanz & Davenport 1983) and High-Frequency Pressure Integration (HFPI) (Irwin & Kochanski 1995) are two examples of techniques employed for such effort measurements. Despite many improvements brought to deal with turbulent flows, such techniques provide only partial information of the complex wind-structure interactions involved. For instance, when the structure has a complex geometrical shape, the very sparse nature of the cladding pressure measurements brought by HFPI techniques may lead to a misrepresentation of the local pressure field.

On the other hand, more recently, thanks to the significant progress of computational capabilities, computational fluid dynamics (CFD) techniques have proved their value to give a complete representation of these flows, enabling a better understanding of the relation between the flow structures and the wind loads. However, since an accurate description of such turbulent flows requires a fine enough resolution, this technique may rapidly become impractical due to the large computational resources required. To go beyond this computational limitation, turbulence model closures associated with the Reynolds averaged Navier-Stokes (RANS) simulation were widely adopted to give an insight into the time-averaged flow profile. With such models, turbulence characteristics can be reasonably represented at lower computational costs. Over the years, motivated by the available computational wind engineering guidelines (Tominaga *et al.* 2008; EN 2005), several established turbulence models have been deeply investigated (Cochran & Derickson 2011; Meroney & Derickson 2014). While RANS simulations offer good qualitative results that are physically relevant, due to their inherent assumptions built from accumulated knowledge on real turbulent flows, close inspection on wind loads reveals typical failures in their prediction. For instance, early studies by Murakami (1990, 1997) compared the standard $k-\epsilon$ model (Launder & Sharma 1974) analysis with unsteady large-eddy simulations (LES) and wind tunnel experiments. They revealed the model's poor accuracy resulting in an over-production of turbulent kinetic energy in the flow impingement region. Various revisions of the model (*e.g.*, RNG by Yakhot *et al.* (1992), realizable by Shih *et al.* (1995), MMK by Tsuchiya *et al.* (1997)) have provided results close to measures obtained in wind tunnels. Yet, recent studies have shown that such models still fail to reproduce an accurate recirculation region behind the building (Tominaga & Stathopoulos 2010, 2017; Yoshie *et al.* 2007).
This accuracy issue may strongly hinder the model predictive skill when compared to real-world measurements. One way to correct this deficiency consists in devising methods allowing to couple turbulence modelling with measurements.

Indeed, during the last decades, a wide variety of coupling techniques has been increasingly considered in fluid mechanics applications. Such techniques, commonly termed as data assimilation (DA), have been employed to estimate an optimal flow state provided by a

given dynamical model such that it remains close enough to observations. So far, two different classes of DA techniques have been applied to that end. On the one hand, Bayesian techniques, often referred to as sequential DA techniques, have been used to estimate optimal flow parameters from data affected by a high uncertainty level (Meldi & Poux 2017; Mons *et al.* 2016). On the other hand, optimal control techniques, like variational DA or ensemble-variational DA, have been proposed for direct and large eddy numerical description models (Mons *et al.* 2016; Gronskis *et al.* 2013; Yang *et al.* 2015; Mons *et al.* 2017; Li *et al.* 2020; Chandramouli *et al.* 2020). In this latter kind of approaches, DA is formulated as a constrained optimisation problem (Bryson 1975). A cost functional, reflecting the discrepancies between some (incomplete) measurements of the flow variables and a numerical representation of the flow dynamics, is minimised using a gradient-based descent method. In such an optimisation problem, the functional gradient's evaluation is efficiently carried out through the dynamical model's adjoint instead of a costly finite difference approach (Errico 1997; Plessix 2006). At this point, these DA methods were often used to reconstruct initial conditions and/or boundary conditions for nonstationary flow simulation issues (such as large eddy simulations). In the last few years, mean flow reconstruction problems have also been considered with data assimilation techniques. In some studies (Foures *et al.* 2014; Symon *et al.* 2017), built from variational DA techniques, the authors considered laminar steady Navier-Stokes equations corrected by an unknown volume-force to directly model the turbulence effects. These studies showed that, in laminar or transitional flows, such models perform well to assimilate synthetic particle image velocimetry (PIV) data. Other DA studies at high Reynolds number were performed with RANS turbulence models (Li *et al.* 2017; Singh & Duraisamy 2016; Franceschini *et al.* 2020). In these works, mean flow DA approaches exploited the turbulence models' structure, which results from a trade-off between asymptotic theories on turbulence mixing and empirical tuning to fit experimental data. This was expressed through a calibration process of the closure constants or of a corrective source term added to the turbulence model. Experimental knowledge plays here a crucial role. Such studies dealt mainly with fundamental and industrial oriented flow configurations in which turbulence is often generated at a unique integral scale. However, to the authors' knowledge, for flow configurations involving complex flow interactions as in the case of an atmospheric boundary layer around a bluff-body, turbulence closure structure analysis using DA techniques are still largely unexplored. Nevertheless, it is noteworthy that formal uncertainty quantification (UQ) techniques have been employed to interpret these closure models in probabilistic terms (Etling *et al.* 1985; Duynkerke 1988; Tavoularis & Karnik 1989; Edeling *et al.* 2014; Margheri *et al.* 2014). For example, in a recent work by Shirzadi *et al.* (2017), global coefficients of the standard $k-\epsilon$ model were adapted for unstable atmospheric boundary layer (ABL) flow around high-rise buildings using a forward UQ technique (*e.g.*, Monte Carlo simulations).

In the present work, we propose investigating one of the most common turbulence closure models for RANS modelling in a variational data assimilation procedure framework. A continuous adjoint approach is considered and then discretised using a 3D finite volume scheme. Without loss of generality, we will, first, use this methodology to investigate the sensitivity fields of the global closure coefficients of the high Reynolds realizable revision of the $k-\epsilon$ model (Shih *et al.* 1995). Their physical interpretation will enable us to point out limits on such closure models' applicability for data-model coupling purposes, particularly for wind flows around buildings. Contrary to previous data-model studies in which velocity measurements were considered on significant parts of the flow domain, we rely on sparse pressure data measured on the building surface. This difference is far from being cosmetic as it leads so far to a much more practical experimental setting for large-scale volumetric measurements. Besides, we point out some difficulties faced in the literature in coupling RANS modeled 3D flows with only parietal experimental measurements. This

framework will discuss limitations and improvements in estimating wind loads and mean velocities surrounding a high-rise building. To address the limitations of common turbulence modelling more efficiently, we will then relax the model rigidity by considering a distributed additive control parameter in the turbulent dissipation transport equation, where the closure is performed. Beyond the fact that it provides a better agreement with real flow data due to the richer control parameter space and avoids overfitting to the data thanks to the prior information brought by the RANS model structure, the optimal control parameter enables us to identify features that are missing in the initial RANS closure hypotheses. To that end, a modified dissipation rate equation and its adjoint equation are introduced. A physical interpretation of the reconstructed field will then be addressed to point out the limits of models' closure applicability for data-model coupling purposes. In the optimisation procedure, the adjoint sensitivity field and the associated cost functional gradient is generally irregular for distributed control parameters due to the lack of specific treatment. This lack of regularity often even hinders a proper estimation of the sensitivity map and requires the adjunction of regularisation terms whose calibration is not straightforward. In contrast to this conventional approach, a projection onto a space of regular functions: the Sobolev space $H^1(\Omega)$, is proposed to regularise the descent direction. As will be shown, this leads to a very efficient data assimilation procedure.

The paper is organised as follows. We first describe the adjoint-based turbulence model and wall-pressure measurements coupling for flow reconstruction around a high-rise building. The next section improves the adjoint-based turbulence models' sensitivity analysis tool and proposes a corrective turbulence model. Then, the case studied is described. The models' sensitivities are discussed, and their performances for flow reconstruction from wall-pressure data are presented. Finally, a summary and further outlook are given.

2. Development for an adjoint-based diagnostics

In this section, we set up the variational data-model coupling framework, based on optimal control techniques. A particular attention is given to the analytical derivation of the adjoint model of one of the most common turbulence models, the realizable revision of $k-\epsilon$, coupled with near wall closure.

2.1. *Variational approach*

A generic variational data-coupling problem can be formally described by the following optimisation problem :

$$\begin{aligned} \min_{\alpha} \quad & \mathcal{J}(\alpha, \mathbf{X}(\alpha), \mathbf{Y}_{\text{obs}}) \\ \text{subject to} \quad & M_i(\alpha, \mathbf{X}(\alpha)) = 0 \qquad i = 1, \dots, N \end{aligned} \tag{2.1}$$

where $\mathcal{J}()$ is the cost function that quantifies the misfit between observations and the model, *i.e.* here measurements and CFD solution respectively, penalised by an *a priori* statistical knowledge of these discrepancies in the form of a covariance matrix. Here we refer to the flow measurements $\mathbf{Y}_{\text{obs}}$, and the predicted flow $\mathbf{X}$. N stands for the number of independent variables necessary for a full description of the flow. The minimisation of this function is constrained by the set of flow governing equations M_i. Such problem may be solved using a gradient-based algorithm. It consists in iteratively evaluating the cost functional and its sensitivity derivatives in order to find the minimum by successive updates of the control variables α. The sketch of the procedure is given in algorithm 1.

In order to properly define the cost function, one may proceed as follows. The only available experimental inputs are wall-pressure measurements. Ideally, the discrepancy between the

Algorithm 1 $\min_{\alpha} \mathcal{J}(\alpha, \mathbf{X}(\alpha), \mathbf{Y}_{\text{obs}})$

Initialisation: $\alpha^m = \alpha_{\text{b}}$ and m=0
repeat
  Solve $\mathbf{M}(\mathbf{X}^m, \alpha^m) = 0$
  Compute sensitivity $\frac{\mathrm{D}\mathcal{J}}{\mathrm{D}\alpha}(\mathbf{X}^m, \alpha^m)$
  update $\alpha \rightarrow \alpha^{m+1}$; $m = m + 1$
until $\left\|\mathcal{J}^m - \mathcal{J}^{m-1}\right\| < \varepsilon$

experimental pressure measure and the model wall-pressure can be expressed as $\Delta P^w = P^w_{\text{obs}} - \mathcal{H}(P^w)$ where $\mathcal{H}(.)$ restricts the pressure at the measurement positions (the subscript w stands for wall). However, due to measurement errors, this difference must be weighted by their associated uncertainties. Having no access to the real pressure values, these uncertainties have to be estimated. By assuming a normal distribution around the measured value, this can be introduced by means of an empirical covariance matrix. Concerning the projector $\mathcal{H}(.)$, in this work, we have considered interpolation by a Gaussian kernel of half size D, i.e. the building diameter, from the computational grid to the position of the measurements to ensure consistency between estimated observation and pressure measurements. So far, it should be pointed out that under the assumption of incompressible flow, the pressure solved numerically is only defined up to a constant. Thus, the experimental and numerical pressures can best be collated by comparing their respective pressure coefficient $C_p = \frac{P - P_\infty}{1/2\rho U^2_{\text{ref}}}$ with P_∞ denoting the reference static pressure. We note the difference between numerical and experimental value by $\Delta C_p = C^w_{p_{obs}} - C^w_p$.

To ensure that the set of parameters α remains in a realistic set of values, we define a physically likely range for each component α_i. This can be formalised by a penalisation term on the cost functional, leading to

$$\mathcal{J}(P, \alpha) = ||\frac{1}{2}\rho U^2_{\text{ref}} \Delta C^w_p||^2_{R^{-1}} + ||\alpha - \alpha_b||^2_{B^{-1}}, \tag{2.2}$$

where R is the covariance matrix defined from measurement's uncertainties, B denotes the covariance matrix associated to the parameter validity range, and U_{ref} stands for a reference velocity. Without loss of generality, a diagonal measurement covariance with a constant standard deviation $\sqrt{R_{ii}} = 1$ is used. This uniformity represents an equal degree of confidence for each measurement, and the diagonal structure ensues from an assumption of spatially uncorrelated errors, which can be assumed for sufficiently distant measurements. Matrix B corresponds to *a prior* knowledge on the range of values of the parameters. In practice, it is worth noting that the role of B matrix is twofold. In the one hand, it imposes a realistic interval in which the parameters can be optimised; while, on the other hand, it ensures a consistent scaling between inhomogeneous terms.

2.2. *A RANS model*

The incompressible airflow surrounding the building can be fully described by its velocity $\mathbf{u}$ and pressure p . This unsteady state, solution of the Navier-Stokes equations, can be further decomposed in terms of its mean, $(\mathbf{U}, P)$, that will be resolved and a modeled fluctuation $(\mathbf{u}', p')$. By applying time averaging to the Navier-Stokes equations, one can obtain the partial differential equations (PDEs) of the RANS equations in a conservative form, whose solution provides the mean wind flow :

$$\frac{\partial(\rho U_j U_i)}{\partial x_j} = -\frac{\partial P}{\partial x_i} + \mu \frac{\partial}{\partial x_j}\left[\frac{\partial U_i}{\partial x_j}\right] - \frac{\partial}{\partial x_j}\left(\rho \overline{u'_i u'_j}\right), \tag{2.3}$$

$$\frac{\partial U_j}{\partial x_j} = 0.$$

Due to the non-linear term, the averaging procedure leads to a second-order moment $\overline{u_i' u_j'}$, called the Reynolds stress. Since all the unsteadiness and turbulence effects of the wind flow are contained in this term, without *a priori* specification of this term, the above system is not closed and cannot be solved. The immediate solution for the closure is to include additional transport equations to predict the turbulence second-order statistics. Relying on the Boussinesq analogy between large-scale dissipation and molecular friction, Reynolds stresses are commonly modeled using a turbulent diffusion-like term, so-called eddy viscosity model. Several models have been proposed to relate quantities describing turbulent fluctuations. A common practice associates the turbulent kinetic energy k, representing the isotropic part of the exact Reynolds stress, with the turbulence length scale l. Due to their simple structures, these models often require empirical closure functions or constants which are established and determined from experimental knowledge, trying to ensure their widest possible applications. For instance, the steady realizable k$-\epsilon$ turbulence model (Shih *et al.* 1995), in which $\epsilon \sim \frac{k3/2}{l}$ models the turbulence dissipation rate at the viscous scale, is often adopted for wind flow expertise around real-world buildings. The Reynolds stress is linearly linked to the mean shear stress by an eddy viscosity as follows:

$$-\rho\overline{u_i u_j} = \mu_t \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} \right) - \frac{2}{3}\delta_{ij}\rho k, \tag{2.4}$$

where μ_t stands for the isotropic (*i.e.* assuming that length and time scales of turbulence are smaller than those of the mean flow with no preferential direction) eddy viscosity coefficient. Its value is calculated using the relation

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

The coefficient C_μ, following the work of Shih *et al.* (1995), is a non uniform constant defined by

$$C_\mu = \frac{1}{A_0 + A_s U_s \frac{k}{\epsilon}}, \tag{2.5}$$

where A_s and U_s are functions of the mean strain and rotation rates and A_0 is a closure tuning coefficient. Substituting the Reynolds stress model (2.4) in the mean momentum equation (2.3) yields to

$$\frac{\partial(\rho U_j U_i)}{\partial x_j} = -\frac{\partial}{\partial x_i}\left(P + \frac{2}{3}\rho k\right) + \frac{\partial}{\partial x_j}\left[\mu_{eff}\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right)\right], \tag{2.6}$$

where $\mu_{eff} = (\mu_t + \mu)$ stands for an effective viscosity. It can be noted that the isotropic component $\frac{2}{3}\rho k$ is absorbed in a modified mean pressure and only the anisotropic part plays an effective role in transporting momentum. It is worth noting that anisotropy here arises only from the mean flow strain and does not depend on the turbulent fluctuations. Moreover, in the computation of the pressure coefficient (required for the observation error in the cost functional (2.2)), we subtract the isotropic part to obtain $C_p = \frac{P - P_\infty - \frac{2}{3}\rho k}{1/2\rho U_{\text{ref}}^2}$. With regards to the turbulence closure model, the transport of mean turbulent kinetic energy, k, is described by

$$\frac{\partial \rho U_j k}{\partial x_j} = \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right] + \mu_t\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right)\frac{\partial U_i}{\partial x_j} - \rho\epsilon, \tag{2.7}$$

where σ_k is a closure constant that enables the scalar mixing of k to be affected by other mechanisms than eddy viscosity. The hypothesis, associated to the value $\sigma_k = 1$, in which the eddy diffusion affects in the same way the momentum and the turbulent kinetic energy k, is analysed in section 5.2.

The turbulent dissipation rate transport is described by the model proposed by Shih *et al.* (1995)

$$\frac{\partial \rho U_j \epsilon}{\partial x_j} = \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_\epsilon}\right)\frac{\partial \epsilon}{\partial x_i}\right] + C_1(S,k,\epsilon)S\epsilon - C_2\frac{\epsilon^2}{k+\sqrt{\mu\epsilon}}, \tag{2.8}$$

where $S = 2(S_{ij}S_{ij})^{\frac{1}{2}}$ is the magnitude of mean strain rate where $S_{ij} = \frac{1}{2}\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right)$ and $C_1 = \max(0.43, \eta/(5+\eta))$ where $\eta = Sk/\epsilon$ is the normalised strain rate. The constants σ_ϵ and C_2 are closure coefficients that need to be calibrated. Assuming that the generation of ϵ is linked to its redistribution everywhere, as it is was established in an inertial layer, leads to a relationship between these coefficients (details in section 5.2.2). This closure hypothesis is analysed as well in section 5.2. To properly close this set of equations (since the model is valid only for high Reynolds regimes), an asymptotic behaviour has to be imposed near the wall where this assumption is no longer valid in the viscous and buffer layers. Indeed, for large scale configurations (*i.e.*, ABL, high-rise buildings), the first grid centre closest to the wall usually falls at the high end of the logarithmic layer. Thus, with a finite volume scheme, in the domain Ω_c, covered by the first grid cell closest to the wall centered on $|_c$ and with a boundary face centered on $|_f$ (see figure 1), these relations hold

$$U^+|_c = \frac{1}{\kappa} f(y^+|_c), \tag{2.9}$$

$$\frac{\partial k}{\partial x_j} n_j|_f = \frac{\partial \epsilon}{\partial x_j} n_j|_f = 0, \tag{2.10}$$

$$P_k|_c = \epsilon|_c = \frac{u_\tau^3|_c}{\kappa y|_c} \tag{2.11}$$

$$y^+|_c = \frac{\rho u_\tau|_c y|_c}{\mu}, \tag{2.12}$$

$$\tau_{wall}|_c = \rho u_\tau^2|_c = \mu_{eff}\frac{U_i t_i|_c}{y|_c}, \tag{2.13}$$

$$U^+|_c = \frac{U_i t_i|_c}{u_\tau|_c}, \tag{2.14}$$

where U^+, y^+ are the dimensionless wall unit tangential velocity component and distance from wall, respectively, n_i and $t_i = 1 - n_i$ are the projections of normal and tangential unit vectors onto the boundary face in the orthonormal frame (x, y, z). The log function f is an empirical function parametrised by constants which depend on the wall type (such as smooth or rough); and $\kappa = 0.41$ is the Von Karman constant. As viscous effects is being neglected at the center of the first cell where $y^+ >> O(1)$, the friction velocity u_τ is scaled by the square root of the fluctuations, following the empirical expression $u_\tau = C_\mu^{1/4} k^{1/2}$ with $C_\mu = 0.09$. Let us note that (2.13) extends the constant shear stress to the wall as we assume linear behaviour (in second order finite volume scheme FVM) inside the cell Ω_c. In practice, this extension is of great importance as it enables the enforcement of the logarithmic law while providing numerical stability for FVM scheme. For instance, in evaluating the production term $P_k|_c$ in (2.11), the friction velocity to the power three is often split into two contributions: turbulence and the mean flow shears.

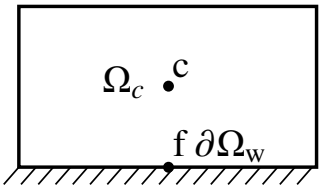


Figure 1: First grid cell close to the wall.

2.3. *A continuous adjoint RANS model*

The optimisation problem (2.1) can be solved by augmenting the cost function with the constraint, *i.e.* the RANS model. This is done through Lagrange multipliers also called adjoint variables. The resulting unconstrained optimisation problem can be written in a compact form as

$$\mathcal{L}(\mathbf{X}, \mathbf{X}^*, \alpha) = \mathcal{J}(P, \alpha) + \langle \mathbf{X}^*, \mathbf{M}(\mathbf{X}, \alpha) \rangle_\Omega, \tag{2.15}$$

where $\langle ., . \rangle_\Omega$ stands for the spatial L^2 inner-product in the flow domain Ω. The mean flow state $\mathbf{X}$ is defined by the set $(\mathbf{U}, P, k, \epsilon, \mu_t)$. As for the adjoint state $\mathbf{X}^*$, we define it as $(\mathbf{U}^*, P^*, k^*, \epsilon^*, \mu_t^*)$. The term $\mathbf{U}^*$ stands for the adjoint velocity, P^* is the adjoint pressure field, k^* is the adjoint turbulent kinetic energy, ϵ^* the adjoint kinetic dissipation rate and μ_t^* the adjoint eddy-viscosity.

Solving the optimisation problem implies to find the set of parameters, the state vector and the adjoint state such that the derivatives of $\mathcal{L}$ with respect to all variables vanish. To this end, based on the application of the Green-Gauss theorem and the use of integrations by parts – see the Appendix and the work of (Othmer 2008)– the adjoint system reads as follows:

$$-\rho\frac{\partial U_j U_i^*}{\partial x_j} - \rho U_j \frac{\partial U_j^*}{\partial x_i} - \frac{\partial}{\partial x_j}\left[\mu_{eff}\left(\frac{\partial U_i^*}{\partial x_j} + \frac{\partial U_j^*}{\partial x_i}\right)\right] + \frac{\partial P^*}{\partial x_i} = D_{U^*,i} \tag{2.16}$$

$$\frac{\partial U_j^*}{\partial x_j} = 0 \tag{2.17}$$

$$-\frac{\partial \rho U_j k^*}{\partial x_j} - \frac{\partial}{\partial x_i}\left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial k^*}{\partial x_j}\right] - \rho\frac{2}{3}\frac{\partial U_i}{\partial x_i}k^* = D_{k^*} \tag{2.18}$$

$$-\frac{\partial \rho U_j \epsilon^*}{\partial x_j} - \frac{\partial}{\partial x_i}\left[\left(\mu + \frac{\mu_t}{\sigma_\epsilon}\right)\frac{\partial \epsilon^*}{\partial x_j}\right] - \rho\frac{\partial P_\epsilon}{\partial \epsilon}\epsilon^* + \rho\frac{\partial s_\epsilon}{\partial \epsilon}\epsilon^* = D_{\epsilon^*} \tag{2.19}$$

$$\frac{\partial P_k}{\partial \mu_t}k^* - \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} -\right)\frac{\partial U_j^*}{\partial x_i} - \frac{1}{\sigma_k}\frac{\partial k}{\partial x_i}\frac{\partial k^*}{\partial x_i} - \frac{1}{\sigma_\epsilon}\frac{\partial \epsilon}{\partial x_i}\frac{\partial \epsilon^*}{\partial x_i} = \mu_t^*. \tag{2.20}$$

The right-hand sides are expressed as

$$\begin{aligned}
D_{U^*,i} &= k\frac{\partial k^*}{\partial x_i} + \epsilon\frac{\partial \epsilon^*}{\partial x_i} + \frac{2}{3}\frac{\partial k k^*}{\partial x_i} - \frac{\partial P_\epsilon}{\partial U_i}\epsilon^* - \frac{\partial P_k}{\partial U_i}k^* - \frac{\partial}{\partial x_i}\left(\frac{\partial \mu_t}{\partial U_i}\mu_t^*\right) \\
D_{k^*} &= \rho\frac{\partial P_\epsilon}{\partial k}\epsilon^* - \rho\frac{\partial s_\epsilon}{\partial k}\epsilon^* - \frac{\partial \mu_t}{\partial k}\mu_t^* \\
D_{\epsilon^*} &= -\frac{\partial \mu_t}{\partial \epsilon}\mu_t^*,
\end{aligned}$$

where P_k and P_ϵ stand for the production terms of turbulent energy (second term in the

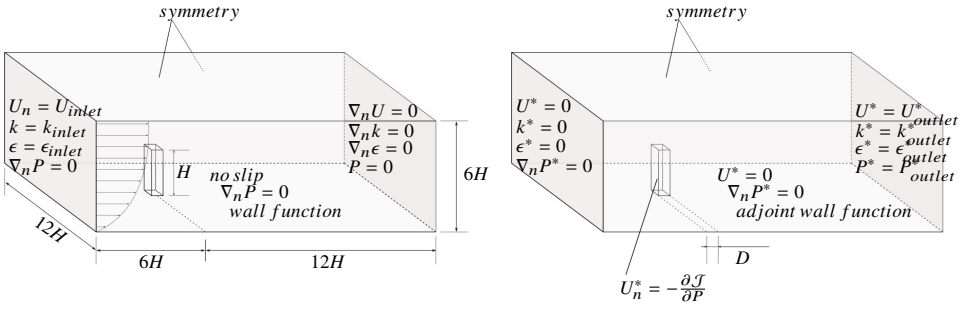


Figure 2: Settings and boundary conditions of flow solution (a) and adjoint solution (b).

RHS of (2.7)) and for the turbulence dissipation rate (second term in the RHS of (2.8)), respectively. As for s_ϵ, it denotes the modeled sink of turbulence dissipation rate (third term in the RHS of (2.8)). To solve this adjoint system, adjoint boundary conditions have to be derived consistently with the boundary conditions of the direct problem. The next section is dedicated to this crucial point.

2.4. *Adjoint boundary conditions*

The treatment of the adjoint boundary conditions is a central piece in adjoint methods in order to obtain consistency in the gradient computation. In our case with the transport equations of turbulent quantities, some treatments are not standard, particularly at the adjoint wall law. Moreover, some specific treatments are performed at the discrete level of the finite volume formulation. In this section, we propose to recall the general procedure to obtain adjoint boundary conditions, and then to detail the conditions to enforce at each boundary. Boundary conditions for the flow and adjoint fields are summarised in figure 2.

Derivation of (2.15) leads, in addition to system (2.16)-(2.19), to a system of boundary terms. Directional derivative with respect to P, leads to

$$\left[U_i^* n_i \delta P\right]_{\delta\Omega} = -\frac{\partial J}{\partial P}\delta P, \tag{2.21}$$

where the boundary integral is defined as $[\cdot]_{\partial\Omega} = \int_{\partial\Omega}(\cdot)\,\mathrm{d}\partial\Omega$. By deriving with respect to U_i, we obtain

$$\begin{aligned}
&[P^*\delta U_i n_i]_{\delta\Omega} - \left[\rho\left((U_i^* n_i)(U_i n_i) + U_j U_j^* n_i\right)\delta U_i n_i\right]_{\delta\Omega} - \left[\mu_{eff}(U_i^* n_i + U_j^* n_j)\frac{\partial \delta U_i}{\partial x_j}\right]_{\delta\Omega} \\
&- \left[\delta U_i \mu_{eff}\left(\frac{\partial U_i^* n_i}{\partial x_j}n_i + \frac{\partial U_j^* n_j}{\partial x_i}n_i\right)\right]_{\delta\Omega} \\
&- \left[\delta U_i\left(\frac{5}{3}\rho k k^* n_i + \left(\frac{\partial P_k}{\partial U_j}n_i\right)k^* + \rho\epsilon\epsilon^* n_i + \left(\frac{\partial P_\epsilon}{\partial U_j}n_i\right)\epsilon^* + \left(\frac{\partial \mu_t}{\partial U_j}n_i\right)\mu_t^*\right)\right]_{\delta\Omega}.
\end{aligned} \tag{2.22}$$

Derivative with respect to k, leads to

$$\left[-\frac{2}{3}\rho U_i^* n_i \delta k\right]_{\delta\Omega} + \left[\rho U_i n_i k^* \delta k\right]_{\delta\Omega} - \left[\left(\mu + \frac{\mu_t}{\sigma_k}\right) k^* \frac{\partial \delta k}{\partial x_j} n_j\right]_{\delta\Omega} + \left[\frac{\partial\left(\mu + \frac{\mu_t}{\sigma_k}\right) k^*}{\partial x_j} n_j \delta k\right]_{\delta\Omega}, \tag{2.23}$$

and with respect to ϵ to

$$\left[\rho U_i n_i \epsilon^* \delta\epsilon\right]_{\delta\Omega} - \left[\left(\mu + \frac{\mu_t}{\sigma_k}\right) \epsilon^* \frac{\partial \delta\epsilon}{\partial x_j} n_j\right]_{\delta\Omega} + \left[\frac{\partial\left(\mu + \frac{\mu_t}{\sigma_k}\right) \epsilon^*}{\partial x_j} n_j \delta\epsilon\right]_{\delta\Omega}. \tag{2.24}$$

Finally by deriving with respect to μ_t, we obtain

$$\left[\delta\mu_t \left(\mu_t^* - \left(\frac{\partial U_j}{\partial x_i} n_j + \frac{\partial U_i}{\partial x_j} n_j\right) U_i^* - \frac{\rho}{\sigma_k} k^* \frac{\partial k}{\partial x_j} n_j - \frac{\rho}{\sigma_\epsilon} \epsilon^* \frac{\partial \epsilon}{\partial x_j} n_j\right)\right]_{\delta\Omega}. \tag{2.25}$$

Then at each boarder, variations of the direct boundary conditions are injected in the system (2.21)–(2.25), to obtain the corresponding adjoint boundary conditions.

At the inlet, the direct boundary conditions lead to

$$\delta U_i t_i = 0 \quad ; \quad \delta U_i n_i = 0 \quad ; \quad \delta k = 0 \quad ; \quad \delta\epsilon = 0. \tag{2.26}$$

Substituting this in (2.21)–(2.24), the system reduces to

$$U_i^* t_i = 0 \quad ; \quad U_i^* n_i = 0 \quad ; \quad k^* = 0 \quad ; \quad \epsilon^* = 0. \tag{2.27}$$

This quite standard result at the continuous level is not straightforward to implement in the finite volume formulation. No condition is imposed on P^* and the inlet boundary condition for the adjoint pressure is left arbitrary. But, in accordance with (Zymaris *et al.* 2010; Othmer 2008) and identically as the numerical treatment of the direct inlet pressure P, zero Neumann condition on P^* is imposed to ensure numerical stability. To obtain the other boundary conditions, the same approach is employed.

At the outlet, the pressure value is prescribed while the other flow variables have their normal gradient imposed, leading to

$$P^* n_i = (U_i^* n_i)(U_i n_i) + (U_j^* U_j) n_i + 2 * \mu_{eff} \frac{\partial (U_j^* n_j)}{\partial x_i} n_i + \frac{5}{3}\rho k k^* n_i + \rho\epsilon\epsilon^* n_i + \left(\frac{\partial \mu_t}{\partial U_j} n_j\right) n_i \mu_t^*. \tag{2.28}$$

This provides a constraint on the boundary condition to determine the adjoint pressure at the outlet. In equation (2.28), the adjoint pressure at the next iteration is determined explicitly by evaluating U^* at the previous iteration. Projecting then the fluxes on the outlet tangent plane, we obtain:

$$\mu_{eff} \frac{\partial (U_j^* t_j)}{\partial x_i} n_i + (U_i^* t_i)(U_i n_i) = -\left(\frac{\partial \mu_t}{\partial U_i} n_i\right) t_i \mu_t^* - \left(\frac{\partial P_k}{\partial U_j} n_j\right) t_i k^* - \left(\frac{\partial P_\epsilon}{\partial U_j} n_j\right) t_i \epsilon^*. \tag{2.29}$$

This equation provides a boundary condition for the tangential component of the adjoint velocity.

It is worth noting that, instead, an alternative choice could be made by imposing $P^* = 0$ and determining the adjoint velocity by solving equations (2.28) and (2.29). Previous works, for a different turbulence model (Zymaris *et al.* 2010) or for frozen turbulence assumption (Othmer 2008), showed that both implementations yields to identical sensitivities.

Derivation w.r.t. k, ϵ and μ_t leads, respectively, to

$$\rho k^* U_i n_i + \rho D_k \frac{\partial k^*}{\partial x_i} n_i = \frac{\partial \mu_t}{\partial k} \mu_t^*, \tag{2.30}$$

$$\rho \epsilon^* U_i n_i + \rho D_\epsilon \frac{\partial \epsilon^*}{\partial x_i} n_i = \frac{\partial \mu_t}{\partial \epsilon} \mu_t^*, \tag{2.31}$$

$$\left(\frac{\partial U_j}{\partial x_i} n_j + \frac{\partial U_i}{\partial x_j} n_j \right) U_i^* = \mu_t^*. \tag{2.32}$$

Therefore, for known outlet direct and adjoint velocities, the adjoint eddy viscosity is updated through equation (2.32). Then, conditions (2.30) and (2.31) can be imposed to solve k^* and ϵ^* respectively.

As for the side and top free-stream boundaries under symmetry condition, we assume a zero flux of all flow variables,

$$\frac{\partial P}{\partial x_i} n_i = \frac{\partial (U_j t_j)}{\partial x_i} n_i = \frac{\partial k}{\partial x_i} n_i = \frac{\partial \epsilon}{\partial x_i} n_i = 0, \tag{2.33}$$

and zero normal velocity,

$$U_i n_i = 0. \tag{2.34}$$

We thus obtain the following boundary conditions for the adjoint variables:

$$U_i^* n_i = 0, \quad \frac{\partial (U_i^* t_i)}{\partial x_j} n_j = 0, \quad \frac{\partial k^*}{\partial x_i} n_i = 0, \quad \frac{\partial \epsilon^*}{\partial x_i} n_i = 0, \quad \mu_t^* = 0. \tag{2.35}$$

This shows that symmetric boundary conditions are conserved with the adjoint model.

The wall boundaries are split into two parts, namely $\partial\Omega_{\text{tower}}$ for the part where data are provided, *i.e.* the tower, and $\partial\Omega_{\text{gr}}$ for the walls at the ground modelling the surrounding environment, for which there is no data. Based on equation (2.21) we have

$$U_i^* n_i = \frac{\partial \mathcal{J}}{\partial P} \qquad \text{at} \qquad \partial\Omega_{\text{tower}}, \tag{2.36}$$

$$U_i^* n_i = 0 \qquad \text{at} \qquad \partial\Omega_{\text{gr}}, \tag{2.37}$$

$$U_i^* t_i = 0 \qquad \text{at} \qquad \partial\Omega_{\text{tower}} \cup \partial\Omega_{\text{gr}}. \tag{2.38}$$

Therefore exactly in the same way as for the inlet, the no-slip condition on the velocity, associated with a zero Neumann condition on the mean pressure, implies a homogeneous Dirichlet boundary condition for the adjoint velocity and a zero Neumann condition for the adjoint pressure at the ground walls $\partial\Omega_{\text{gr}}$. Let us note that due to the wall-pressure measurements, the Dirichlet condition on the adjoint variable U^* is inhomogeneous on the normal component along $\partial\Omega_{\text{tower}}$ at the sensor positions.

Considering the adjoint turbulence variables (k^*, ϵ^*, μ_t^*), it is important to consider the expression of the wall-law in order to derive their boundary conditions. From the equality (2.13), the log law is imposed by re-evaluating the wall viscous fluxes through a prescribed value of the eddy-viscosity. From the relations

$$U^+ = \frac{1}{\kappa} \ln(E y^+) \qquad \text{at} \qquad \partial\Omega_{\text{tower}},$$

and

$$U^+ = \frac{1}{\kappa} \ln\left(\frac{y + z_0}{z_0}\right) \qquad \text{at} \qquad \partial\Omega_{\text{gr}},$$

the eddy viscosity is then defined such that

$$\mu_t = \mu\left(\frac{y^+\kappa}{\ln(Ey^+)} - 1\right) \qquad \text{at} \qquad \partial\Omega_{\text{tower}}, \tag{2.39}$$

$$\mu_t = \mu\left(\frac{y^+\kappa}{\ln(\frac{y+z_0}{z_0})} - 1\right) \qquad \text{at} \qquad \partial\Omega_{\text{gr}}, \tag{2.40}$$

where $E = 9.8$ is the roughness parameter for smooth walls (Versteeg & Malalasekera 2007) and $z_0 = 0.02\,\text{m}$ is the roughness length which is relevant for an ABL flow scale. Recalling from figure 1 that $|_c$ and $|_f$ denote respectively the first cell centre and the boundary face values, on both walls $\mu_t|_f$ is seen to be dependent on $k|_c$ solely. Furthermore, by imposing the inertial balance (2.11) we observe that $\epsilon|_c$ is an explicit function of $k|_c$. As a consequence, we conclude that the logarithmic wall-closure can actually be defined uniquely through the knowledge of the turbulent kinetic energy value at the first cell-center only. We propose thus to reconsider the derivation of these terms inside the domain included in the cells adjacent to the wall. Replacing $\epsilon|_c$ by its function of $k|_c$ (equation (2.11)), we are led to consider a zero Neumann boundary condition, which is fully consistent with the homogeneity assumption near the wall. This yields to the following modification to the sources terms at the wall adjacent cell

$$D_{k^*} = \rho\frac{\partial P_\epsilon}{\partial k}\epsilon^* - \rho\frac{\partial s_\epsilon}{\partial k}\epsilon^* + \rho\left(\frac{\partial P_k}{\partial k}|_c - \frac{\partial \epsilon}{\partial k}|_c\right)k^*|_c \qquad \text{and} \qquad D_{\epsilon^*} = 0 \qquad \text{in} \qquad \Omega_c, \tag{2.41}$$

with P_ϵ and s_ϵ, the dissipation production and sink terms, being modified accordingly. At the continuous level, the third term in the RHS vanishes assuming turbulence homogeneity. However, at the discrete level, as discussed in section 2.2, it is kept for numerical consistency.

Furthermore, as we impose a homogeneous Neumann boundary condition for k in equation (2.10), it can be observed that this leads to the same set of conditions as in the outlet (equations (2.30), (2.31) and (2.32)). Moreover, with the no slip condition and the set of conditions for the adjoint velocity ((2.36), (2.37), (2.38)), the wall boundary conditions for the adjoint dissipation rate and eddy viscosity reads as

$$\frac{\partial \epsilon^*}{\partial x_i}n_i = 0 \qquad \text{at} \qquad \partial\Omega_{\text{tower}} \cup \partial\Omega_{\text{gr}}, \tag{2.42}$$

$$\mu_t^* = 2\left(\frac{\partial(U_i n_i)}{\partial x_j}n_j\right)U_i^* n_i \qquad \text{at} \qquad \partial\Omega_{\text{tower}}, \tag{2.43}$$

$$\mu_t^* = 0 \qquad \text{at} \qquad \partial\Omega_{\text{gr}}. \tag{2.44}$$

Concerning the boundary condition on k^*, we propose to consider (2.30) at $|_c$ while we impose zero Neumann condition at $|_f$. This leads to the following boundary condition for k^*:

$$\frac{\partial k^*}{\partial x_j}n_j|_f = 0 \qquad \text{at} \qquad \partial\Omega_w \tag{2.45}$$

and

$$k^*|_c = \frac{\partial \mu_t}{\partial k}|_f \frac{\mu_t^*|_f}{\rho U_i n_i} \qquad \text{in} \qquad \Omega_c. \tag{2.46}$$

Some numerical tests have shown that this alternative solution led to identical results on the initial sensitivity field than the non-homogeneous Neumann boundary condition. Hence, with this treatment, wall conditions for the adjoint system are now fully consistent with the initial RANS model and leads us thus to a consistent minimisation procedure.

Now that we have at our disposal a dual description of the dynamics composed of a RANS direct model and the adjoint of its tangent linear representation, we explore three methodological settings for an in-depth diagnosis of the turbulence closure. The first tool at hand consists simply to inspect the adjoint state maps. The second one consists in optimising global constants parameters of the turbulence model for reducing the observation error; and to relax/enforce constraints on these parameters to test physical hypotheses. The last one, goes one step further and considers the adjunction of distributed unknowns which enables to identify a missing term in the equation where the turbulence closure is performed, here, in the transport equation for the energy dissipation rate. In order to conduct an efficient structural inspection, distributed parameter is sought in the Sobolev space and further estimated through a data-assimilation procedure. These three settings are described in section 3 and applied then to a high-rise building case study described section 4. The numerical results on this case study for the three different sensitivity analyses are presented in section 5.

3. Adjoint-based diagnostic tool for turbulence models

In the previous section we detailed the construction of a continuous adjoint model (together with its consistent boundary conditions) of the tangent linear operator of a RANS model for very high Reynolds flow, associated with large integral/body length scales. In the current section, we present methodological tools derived from this adjoint operator. Beyond providing a data-driven flow reconstruction, it enables us an in-depth analysis of the turbulence closure.

3.1. *Adjoint state as a basis for sensitivity analysis*

While in general, adjoint variables are usually considered as a purely mathematical object, they do have physical meaning which have been explored in several works (Hall & Cacuci 1983; Giles & Pierce 2000; Gunzburger 2003). Although we know that RANS models (at least with Boussinesq eddy viscosity hypothesis for its closure) does not allow an accurate representation of complex flows, they nevertheless provide some usefull global insights on the flow. With this in mind, the reconstructed adjoint state enables to highlight a misrepresentation of the turbulent flow by the RANS model, hence, pointing where it is possible to optimise the RANS model parameters to optimally reduce the difference between the CFD state and a given experimental dataset. Moreover, from an optimisation perspective, one may interpret them as a steepest descent direction of an objective cost function, for a particular control parameter. For instance, the adjoint velocity can be seen as the influence of an arbitrary forcing $\mathbf{f}_u$ acting on the mean momentum equation. This can be shown when deriving an optimality condition by a perturbation of this force,

$$\begin{aligned} \mathbf{M_u}(\mathbf{X},\alpha) = \mathbf{f}_u \qquad &\rightarrow \qquad \frac{\partial \mathcal{J}}{\partial \mathbf{f}_u} = \mathbf{U}^*. \\ \mathit{Momentum\ equation} \qquad &\rightarrow \qquad \mathit{Optimality\ condition}. \end{aligned}$$

Then, as we perform a first update of this forcing by a gradient-descent minimisation algorithm, one obtain

$$\mathbf{f}_u^{it_1} = (\mathbf{f}_u^{it_0} = 0) - \lambda U^*,$$

while λ is a positive non-dimensional marching step factor. Through dimensional analysis, one may scale $\mathbf{U}^*$, which has the dimension of an acceleration, as $U^* \sim \frac{U_{\text{ref}}^2}{H_{\text{ref}}}$. The adjoint velocity features are thus immediately representing a missing term, that can be interpreted as a correction of the Reynolds stress. Examining the adjoint turbulence variables of a $k-\epsilon$ model, namely k^*, ϵ^* and μ_t^* , similar interpretations can be drawn as we consider arbitrary

forcing such as

$$
\begin{aligned}
M_k(\mathbf{X},\alpha) = f_k \quad &\rightarrow \quad \frac{\partial \mathcal{J}}{\partial f_k} = k^*, \\
M_\epsilon(\mathbf{X},\alpha) = f_\epsilon \quad &\rightarrow \quad \frac{\partial \mathcal{J}}{\partial f_\epsilon} = \epsilon^*, \\
M_{\mu_t}(\mathbf{X},\alpha) = f_{\mu_t} \quad &\rightarrow \quad \frac{\partial \mathcal{J}}{\partial f_{\mu_t}} = \mu_t^*.
\end{aligned}
$$

The adjoint variables shed some light on the flow regions that are sensitive to an eventual correction of the turbulence model. As this will be shown in the case study of section 5, this interpretation is very helpfull to analyze the incorporation of additional variables to define efficient data-driven turbulence closure. These forcings can indeed be associated to any of modeled terms meant to adress a particular turbulence modelling error (e.g. energy production, backscaterring, redistribution or dissipation). The sensitivity of the associated parametric shapes and hyper-parameter can be efficiently obtained and inspected through the adjoint operator.

3.2. *Adjoint diagnosis on global closure coefficients*

Given the adjoint dynamics, the sensitivity of any parameter can be obtained from the optimality condition (6.3). Free constants of the RANS model can be finely tuned knowing their sensitivities. As we will show for a particular turbulence model (realizable $k-\epsilon$), the quality of the associated numerical reconstructions appears to be quite restricted, for a range of physically acceptable values of these parameters. The reason of the inefficiency of the calibration procedure is interpreted below in terms of a too strong model "rigidity". This hindering facts will be later illustrated when we will perform the sensitivity analysis and data assimilation on the high-rise building case.

3.2.1. *A sensitivity field*

Considering the vector $\alpha = (A_0, C_2, \sigma_k, \sigma_\epsilon)$ of closure parameters, the optimality condition (6.3) is obtained by differentiating the Lagrangian (2.15) in the directions $\delta\alpha = (\delta A_0, \delta C_2, \delta\sigma_k, \delta\sigma_\epsilon)$:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial A_0}\delta A_0 = \left\langle \frac{\partial M_{\mu_t}}{\partial A_0}\delta A_0, \mu_t^* \right\rangle_\Omega &= \langle -\mu_t C_\mu \delta A_0, \mu_t^* \rangle_\Omega = \left\langle -\mu_t C_\mu, \mu_t^*, \right\rangle_\Omega \delta A_0 \\
\frac{\partial \mathcal{L}}{\partial C_2}\delta C_2 &= \left\langle \frac{\epsilon^2}{k+\sqrt{\mu\epsilon}}, \epsilon^* \right\rangle_\Omega \delta C_2 \\
\frac{\partial \mathcal{L}}{\partial \sigma_k}\delta\sigma_k &= \left\langle -\frac{\partial}{\partial x_j}\left[\frac{\mu_t}{\sigma_k^2}\frac{\partial k}{\partial x_i}\right], k^* \right\rangle_\Omega \delta\sigma_k \\
\frac{\partial \mathcal{L}}{\partial \sigma_\epsilon}\delta\sigma_\epsilon &= \left\langle -\frac{\partial}{\partial x_j}\left[\frac{\mu_t}{\sigma_\epsilon^2}\frac{\partial \epsilon}{\partial x_i}\right], \epsilon^* \right\rangle_\Omega \delta\sigma_\epsilon.
\end{aligned}
\tag{3.1}
$$

So far, this does not include any explicit dependency of the cost on the set of parameters, such as a penalisation term. It can be observed that the optimality conditions reduce drastically the high dimensional dependency of the model to the lower dimensional parameter space. This reduction, performed via the inner product $\langle ., . \rangle_\Omega$, does reflect the global compromising character of the closure coefficients. This results in a rather rigid situation when seeking data-model fitting. However, this rigidity can be understood as a strong confidence in the model

structure. This strongly constrains the solutions but enables, on the other hand, to assimilate very sparse measurements. The examination of the spatially varying adjoint variables to diagnose the parameter sensitivity provides useful piece of information even though we deal with a rigid parametric model for the reconstruction. This type of analysis will be exploited in our case study.

3.2.2. *Penalty range*

To ensure realistic numerical solutions, relevant physical range for each of the closure coefficients is defined. These constraints are introduced via penalty terms on the control parameters through an error covariance matrix, in cost function (2.2). Values of the parameters outside of a range defined by the standard deviation are hence strongly penalised. These standard deviations are in practice fixed from experiments on prototypical configurations of boundary layer or decaying turbulence. This may become questionable in regions where the fluid and building interact and near flow separations associated with strong shears. As a matter of fact, the assumptions underlying the concept of eddy viscosity starts to be less reasonable in these regions (Pope 2001). Thus, as a compromise, the range limits are fixed from experiments as intervals $\Delta\alpha_i$ centered around the background a priori value. The covariance matrix is finally expressed as follows

$$B_{ii}^{-1} = \zeta_i \left(\frac{|\frac{\partial \mathcal{J}_0}{\partial \alpha}|}{\Delta\alpha_i} \right).$$

As mentioned earlier, this covariance has two roles: first, to impose the trusted (or recommended) ranges $\Delta\alpha_i$ and secondly, to ensure a dimensional homogeneity of the cost function through the norm of the sensitivity derivative given at the first minimisation iteration $|\frac{\partial \mathcal{J}_0}{\partial \alpha}|$. The importance of control variables with high values of sensitivity derivatives is strengthen in the objective function in comparison to less sensitive parameters. The parameters ζ_i are dimensionless free parameters allowing to give more or less global *a priori* confidence on each parameter. In practice these parameters can be fixed from a priori considerations.

3.3. *Adjoint diagnosis on spatially distributed closure: correction in the dissipation transport equation*

In contrast to the previous section, we consider now the adjoint system as a basis for the inspection of the model misrepresentations through a distributed parameter. Instead of correcting coefficients of the model, we consider here the adjunction of a force to one of the model equations. We consider a corrective forcing term at the level of the dissipation transport equation, where the closure take place. Then, with the aim to further investigate such a closure through a data-assimilation procedure, a specific optimisation in the $H^1(\Omega)$ Sobolev space is proposed in order to provide a regularisation procedure that guaranties an efficient descent direction as well as an implicit spatial smoothing of the forcing. With this regularisation, a significant improvement of the results will be shown in our study case.

3.3.1. *A corrective model*

As it was noted in the previous section, due to the model rigidity arising from the drastically small parameter space, the flow is not free to visit a sufficiently large domain of the state space that is too far from the basic RANS model. To overcome such restrictions, one may straightforwardly consider a set of closure parameters with a higher dimension. Still maintaining the validity of the Boussinesq approximation, a strategy consists in enriching the turbulence model structure. We choose to add a forcing term in the transport equation of ϵ (equation (2.8)) in order to correct what we may call structural errors, *i.e.* error arising

from the choice of the turbulence model equation. We could have considered as well a control parameter defined directly as the forcing of equation (2.8), however, we preferred to introduce some dependency of this forcing to the state variable. We choose a forcing term of the form $-f_\epsilon\epsilon$, where f_ϵ is the control parameter. The sign convention is chosen so as the added forcing corresponds to a sink of dissipation. The objective of the pre-multiplication by ϵ is to prevent unphysical dissipation corrections at locations where there is no turbulence and to focus specifically on relevant regions such as the shear layers and the wake. Numerical tests presented in section 5.3.1, demonstrate that this term behaves indeed much better than a direct forcing term. With this additional forcing, the dissipation transport equation becomes

$$\frac{\partial \rho U_j \epsilon}{\partial x_j} - \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_\epsilon}\right)\frac{\partial \epsilon}{\partial x_i}\right] - C_1(S,k,\epsilon)S\epsilon + C_2\frac{\epsilon^2}{k+\sqrt{\mu\epsilon}} = -f_\epsilon\epsilon. \tag{3.2}$$

This model remains close to the RANS structure as to avoid overfitting effects in the context of severe differences between the state space and measurements.

3.3.2. *Sensitivity field*

The fact that the added distributed parameter depends on the state variable ϵ, requires some modifications of the adjoint equations. In compact form, the adjoint equation on ϵ^* reads now as

$$M_{\epsilon^*} = -f_\epsilon\epsilon^*, \tag{3.3}$$

where M_{ϵ^*} contains all the adjoint terms derived from equation (2.19). Regarding the adjoint boundary conditions, since no face flux are involved through the additive term, no changes have to be made. The optimality condition associated with the control parameter f_ϵ is obtained by considering the directional derivative

$$\left\langle \frac{\partial \mathcal{L}}{\partial f_\epsilon}, \delta f_\epsilon \right\rangle_\Omega = \langle \epsilon\delta f_\epsilon, \epsilon^* \rangle_\Omega,$$

leading straightforwardly to express the Lagrangian sensitivity to f_ϵ as

$$\frac{\partial \mathcal{L}}{\partial f_\epsilon} = \epsilon\epsilon^*. \tag{3.4}$$

3.3.3. *Descent direction*

With very sparse partial observations and the consideration of spatially distributed control parameters, the risks of obtaining local minima or unphysical flow reconstructions is much stronger. The control parameter can be any function of $L^2(\Omega)$, which allows highly irregular functions. Regularisation is a classical way to reduce the number of local minima eventually associated to unphysical solutions. To that purpose, penalty of the spatial gradients of the control parameter is often considered (Franceschini *et al.* 2020). Such regularisations introduces a smoothing penalty parameter on which the solution strongly depends and whose value is in general non-trivial to choose. In the following, we consider as an alternative a Sobolev gradient regularisation (Protas *et al.* 2004; Tissot *et al.* 2020). It consists to define the control parameter in the Sobolev space $H^1(\Omega)$ which is more regular than $L^2(\Omega)$. With this approach the functional is still defined in its basic form as

$$\mathcal{J}(P) = \frac{1}{2}\left\|\rho U_{\text{ref}}^2 \Delta C_p^w\right\|_{R^{-1}}^2.$$

Provided the optimality condition (3.4) and for an arbitrary functions ψ and ϕ in $H^1(\Omega)$, the Sobolev gradient is defined such that

$$\left\langle \frac{\partial \mathcal{L}}{\partial f_\epsilon}, \psi \right\rangle = \left\langle \frac{\partial \mathcal{L}}{\partial f_\epsilon}^{H^1}, \psi \right\rangle_{H^1}, \tag{3.5}$$

with the inner product definition

$$\langle \phi, \psi \rangle_{H^1} = \int_\Omega \phi\psi + l_{sob}^2 (\nabla\phi \cdot \nabla\psi) d\Omega,$$

in which l_{sob}^2 is a free parameter. Through integration by part of the second term of the inner product (involving the function at gradients), the equality (3.5) leads to the new optimality condition

$$\frac{\partial \mathcal{L}}{\partial f_\epsilon}^{H^1} = \left(\frac{1}{1 + l_{sob}^2} (\mathbb{I} - l_{sob}^2 \nabla^2) \right)^{-1} \frac{\partial \mathcal{L}}{\partial f_\epsilon}, \tag{3.6}$$

in which ∇^2 stands for the Laplacian operator. Equation (3.6) is a projection of the sensitivity in $L^2(\Omega)$ onto the Sobolev space. With this approach, the sensitivity field is consequently regularised through the solution of an Helmoltz equation. Since Matrix inversion is not an option in such large system, the Poisson equation (3.6) is here solved through an iterative technique expressed within the same finite volume scheme as for the direct RANS equations. It is worth to mention that this type of formulation offers two main advantages compared to classical regularisation terms. In the one hand, as opposed to the global penalty coefficient introduced in those latter, the free parameter involved in the projection approach is indeed a physical quantity. As a matter of fact, it is easy to see through dimensional analysis of (3.6) that l_{sob} has the dimensions of a length. More precisely this parameter can be seen as a filtering length scale below which the sensitivity field is smoothed. It provides us a way to introduce a characteristic length scale relevant with the flow (*e.g.*, the building width for instance). In the other hand, the Sobolev gradient does ensure a descent direction. Indeed, applying a Taylor expansion of the cost function around an initial guess f_ϵ in the direction $\delta f_{eps} = -\frac{\partial \mathcal{L}}{\partial f_\epsilon}^{H^1}$ can be expressed as follows

$$\mathcal{J}(f_\epsilon + h\delta f_\epsilon) = \mathcal{J}(f_\epsilon) + h \left\langle \frac{\partial \mathcal{L}}{\partial f_\epsilon}, \delta f_\epsilon \right\rangle + \mathcal{O}(h^2).$$

Substituting the second term in the RHS by using the equality (3.5) yields to

$$\mathcal{J}(f_\epsilon + h\delta f_\epsilon) = \mathcal{J}(f_\epsilon) - h \left\| \frac{\partial \mathcal{L}}{\partial f_\epsilon}^{H^1} \right\|_{H^1}^2 + \mathcal{O}(h^2),$$

in which we define the norm $\|a\|_{H^1}^2 = \langle a, a \rangle_{H^1}$. Thus, for a small enough perturbation $h\delta f_\epsilon$, we have $\mathcal{J}(f_\epsilon + h\delta f_\epsilon) < \mathcal{J}(f_\epsilon)$. Now, injecting this optimality condition into a steepest descent algorithm, an update of the forcing at an iteration n reads:

$$f_\epsilon^{n+1} = f_\epsilon^n - \lambda \left. \frac{\partial \mathcal{L}}{\partial f_\epsilon}^{H^1} \right|_n, \tag{3.7}$$

in which the step size is constrained by $\lambda = \beta \frac{1}{\max(\frac{\partial \mathcal{L}}{\partial f_\epsilon}^{H^1})}$ where $\beta = 2 \cdot 10^{-2}$ is chosen based on the sensitivity validation. In the next section, we present some numerical results obtained

on a realistic case study in terms of the turbulence model parameters estimation and in terms of their sensitivity analysis, with the objective of analysing the closure hypotheses of a given RANS model using the data assimilation framework.

4. Case study

In this section, we first describe the wind tunnel experiments. Then, we present the numerical setup.

4.1. *Description of the wind tunnel experiment*

Experimental data were provided by the CSTB (Nantes, France) from the work of Sheng *et al.* (2018). Measurements were performed in the atmospheric wind tunnel (NSA) with a test section of 20 m long, 4 m wide and 2 m high. Upstream of the isolated building, roughness elements and turbulence generator were set to reproduce the wind profile perceived by the full scale building. The floor of the wind tunnel is equipped with a turntable that enables the flow incidence to vary from 0 to 360^o. In the present paper, only one wind direction is considered. In these experiments, the building was modeled with a wall-mounted prism of square cross-section with the dimensions: 10 cm × 10 cm × 49 cm which corresponds to a tower of height $H = 147$ m and a width $D = 30$ m at full scale. To perform measurements, two tower models were built. The first model was made of Plexiglas which allows for optical access and, thus, to use particle image velocimetry (PIV). The second model was equipped with 265 pressure taps to measure the unsteady pressure distribution on the modeled building.

4.2. *Numerical setup*

The open source library OpenFOAM (Jasak 1996) was used to implement the CFD and adjoint governing equations. The library utilizes a second order finite volume discretisation approach (Moukalled *et al.* 2016) and a fully implicit first order method for time integration. A prediction-correction procedure is used for the pressure-velocity coupling based on the Rhie-Chow interpolation (Rhie & Chow 1983). Correction for mesh non orthogonality was applied for the Poisson solver. A full scale building is modeled and a neutral atmospheric boundary model was used (Richards & Hoxey 1993) to enforce the inlet wind profiles. Profiles for U, k and ϵ are defined as

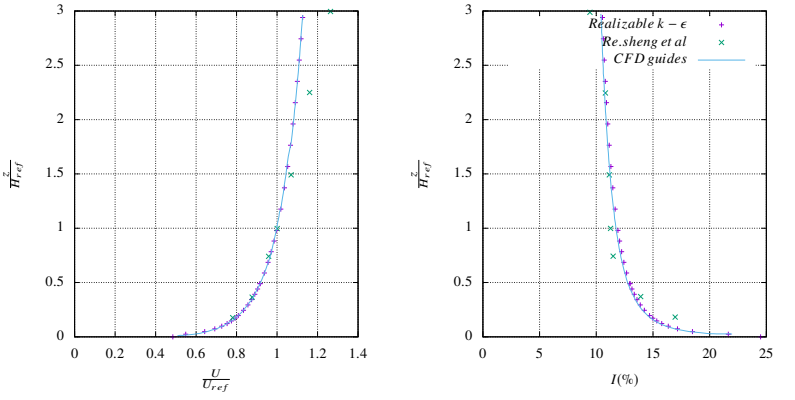
$$U_{in} = \frac{u_\tau^{ABL} \ln\left(\frac{z+z_0}{z_0}\right)}{\kappa}, \qquad k_{in} = \left(\frac{u_\tau^{ABL}}{C_\mu}\right)^2 \qquad and \qquad \epsilon_{in} = \frac{(u_\tau^{ABL})^3}{\kappa(z+z_0)}, \tag{4.1}$$

where u_τ^{ABL} is the friction velocity associated with the constant shear stress along the ABL width

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

in which U_{ref} and $H_{ref} = \frac{2}{3}H$ are, respectively, reference velocity and height chosen to match with the experimental profiles (and thus the eurocode (EN 2005)) (see figure 3). These profiles are consistent with the wall treatment as we prescribes eddy viscosity's ground-value by (2.40), such that $u_\tau = u_\tau^{ABL}$. As for the roughness height z_0, it was set to 0.02, as an intermediate between the roughness class I and class II (EN 2005).

The size of the computational domain was fixed to ensure that the blockage effects are inferior to 3% (Tominaga *et al.* 2008; EN 2005). Grid refinement was chosen to ensure a good representation of the wind gradient at the inlet. Unstructured grid was then adopted



(a) (b)

Figure 3: Profile of the neutral atmospheric boundary layer: (a), mean wind velocity; (b), turbulent intensity I as defined in (EN 2005).

with the minimum distance of the centroid of the cell adjacent to the building walls set to $0.001\,H$. This grid refinement reached approximately 3.5 million cells.

The adjoint differential equations were discretised using the same CFD library as for the direct equations. As for the direct simulation, the adjoint pressure and velocity were iteratively solved using a prediction-correction procedure. The discretisation schemes used for the flow equations were maintained. Moreover, we note that the derivation of the non-linear terms leads to an explicit dependency of the adjoint solution on the direct flow solutions that prevents a parallel computation of the two solvers.

5. Results

In this section, we validate the proposed data assimilation scheme for global and distributed turbulence model parameters. An adjoint state analysis is conducted to obtain the sensitivities to both model control closure parameters, global and distributed coefficients. We assess the limits of global closure optimization performances and exhibit the ability of the proposed distributed closure method not only to reconstruct wall-pressure-driven wake flow accurately but also to enable turbulence closure analysis.

5.1. *Adjoint state analysis*

The normalised adjoint fields (by their maximum in-plane values), shown in figures 4, 5 and 6, highlight the areas of interest in terms of turbulence modelling on two horizontal plans (at normalised height $z/H_{ref} = \{0.19, 1\}$) and on the symmetry plan (at $y/D = 0$). These areas correspond to regions, whose state is observable by the sensors, and where the turbulence closure model fails to reproduce the physical behaviour of the flow; this corresponds to the recirculation regions behind and at the top of the building (as seen in the centered streamwise vertical plans on figure 6), the area of the vortex shedding due to flow separation (seen in the horizontal plans on figure 4 and 5) and the flow impingement region of the building. Based on these adjoint fields, the cost functional's sensitivity to any control parameter (be distributed or not) can be obtained through its associated optimality condition.

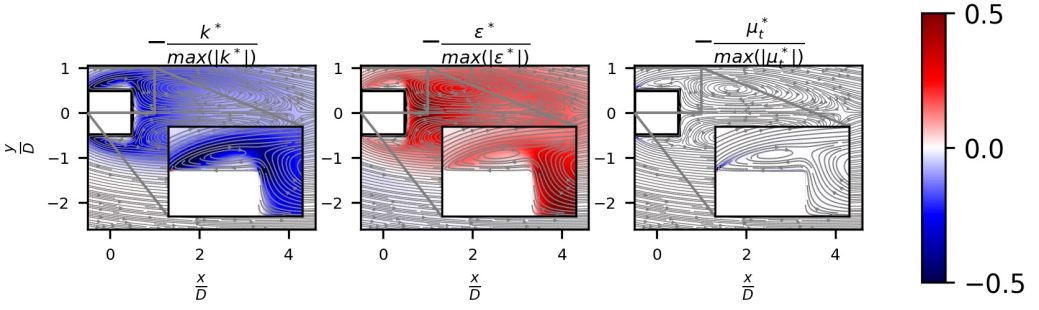


Figure 4: Adjoint turbulence variables at horizontal plan with normalised height $z/H_{ref} = 0.19$:(a), ϵ^*; (b), k^*; (c), μ_t^*. Variables are normalised by their in-plan peak values.

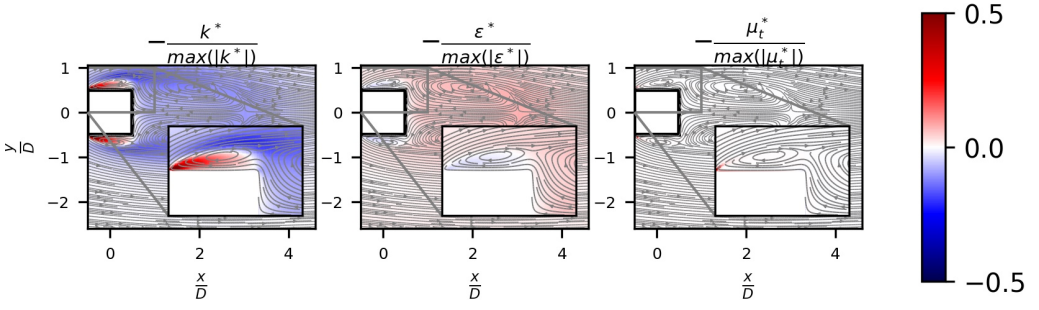


Figure 5: Adjoint turbulence variables at horizontal plan with normalised height $z/H_{ref} = 1$:(a), ϵ^*; (b), k^*; (c), μ_t^*. Variables are normalised by their in-plan peak values.

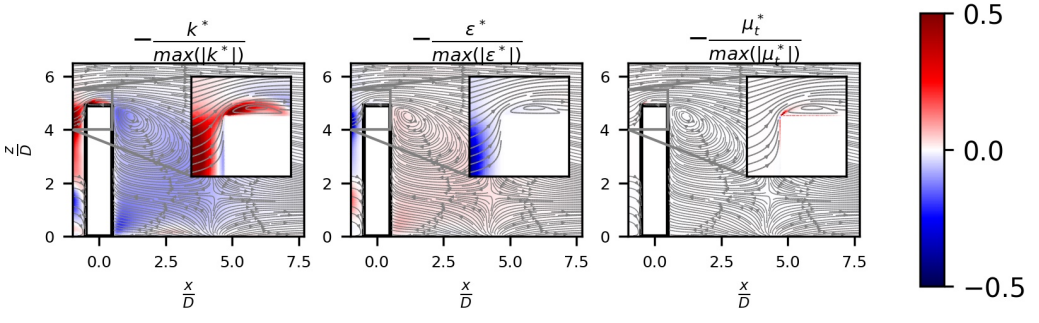


Figure 6: Adjoint turbulence variables at symmetry plan :(a), ϵ^*; (b), k^*; (c), μ_t^*. Variables are normalised by their in-plan peak values.

5.2. *Results for the global coefficients*

This section exhibits the adjoint approach's capability to provide a complete information on the cost sensitivity to the model's global coefficients. First, we analyse the sensitivity fields to highlight the spatial locations where a modification of these global coefficients could efficiently correct the model errors. Then, we discuss the results of a data assimilation procedure. The data assimilation is performed to investigate some closure hypotheses validity in the RANS modelling.

	$\frac{\partial \mathcal{J}}{\partial A_0}(\times 10^7)$	$\frac{\partial \mathcal{J}}{\partial \sigma_k}(\times 10^7)$	$\frac{\partial \mathcal{J}}{\partial \sigma_\epsilon}(\times 10^7)$	$\frac{\partial \mathcal{J}}{\partial C_2}(\times 10^7)$
AD	1.371	-2.756	0.712	-8.583
FD ($\delta\alpha_i = 10^{-1}$)	2.39	-3.11	0.683	-6.699
FD ($\delta\alpha_i = 10^{-2}$)	2.07	-3.76	0.77	-11.4
FD ($\delta\alpha_i = 10^{-3}$)	2.05	-3.04	–	-11.9

Table 1: Sensitivity derivative values computed from the proposed adjoint model (AD) and the finite difference (FD).

5.2.1. *Sensitivity analysis*

First, a validation test has to be done to check the validity of the sensitivity fields. To that end, a specific innerproduct is employed in the variational formulation by considering an integration over a volume of interest, where the sensitivities are the most important. This volume of interest, Ω_{in}, is centred around the building and has a size of $2.5H \times 4H \times 1.5H$. It is introduced in equation (2.15), and used to compute the optimality condition through the following weighted sum

$$\left.\frac{\partial \mathcal{J}}{\partial \alpha_i}\right|_{AD} = \left\langle \frac{\partial \mathcal{M}}{\partial \alpha_i}, \mathbf{X}^* \right\rangle_{\Omega_{in}} \equiv \sum_{j}^{\text{nCells}} \left(\frac{\partial \mathcal{M}}{\partial \alpha_i} . \mathbf{X}^* \right) \omega_j,$$

where ω_j denotes the volume of the j^{th} cell. The validation metric is chosen to be the error between this weighted sum and the finite difference gradient computed by the cost function variations resulting from a small variation of the global coefficients in the volume of interest,

$$\left.\frac{\partial \mathcal{J}}{\partial \alpha_i}\right|_{FD} = \frac{(\mathcal{J}(\alpha_i + \delta\alpha_i) - \mathcal{J}(\alpha_i))}{\delta\alpha_i}.$$

The sensitivity (gradient) of $\mathcal{J}$ with respect to the four closure coefficients computed using the proposed adjoint approach and the finite differences are given in table 1. The comparison shows that the adjoint-based sensitivities are very close to the finite difference values. A fair agreement is obtained for the sensitivity associated with the coefficients involved in the redistribution of k and ϵ. However, deviations appear to be notably more important for the eddy viscosity pre-factor A_0. These deviations are typical of the error levels associated to continuous adjoint methods. (Othmer 2008; Zymaris *et al.* 2010). This validates hence the procedure.

In order to explore the effect of the turbulence model's global coefficients, their associated sensitivity maps (plotted in figure 7, 8, and 9, and defined by the spatially distributed operand inside the integral in the optimality condition (3.1)) are discussed.

We can see there is high interest in optimising these coefficients at the shear layers resulting from flow separations at the leading lateral edges and on top of the building. However, there is very little sensitivity in the bulk of the recirculation wake region. Moreover, with regards to the regularity of the sensitivity fields, $\frac{\partial \mathcal{J}}{\partial \sigma_k}$ and $\frac{\partial \mathcal{J}}{\partial \sigma_\epsilon}$ (Figures 7 and 8 (a) and (b)) have the largest local variations compared to the others. In fact, this is explained by the high (second) order derivative associated with the diffusion of k and ϵ, in the optimality conditions. Now, regarding the local signs and overall values of each sensitivity field, we observe systematic change of sign over the domain. The L^2 inner product in (3.1) leads to an averaged compromise solution over the whole domain for the global coefficient values. This compromise is likely to provide a far too weak amplitude for these coefficients in key regions of the flow.

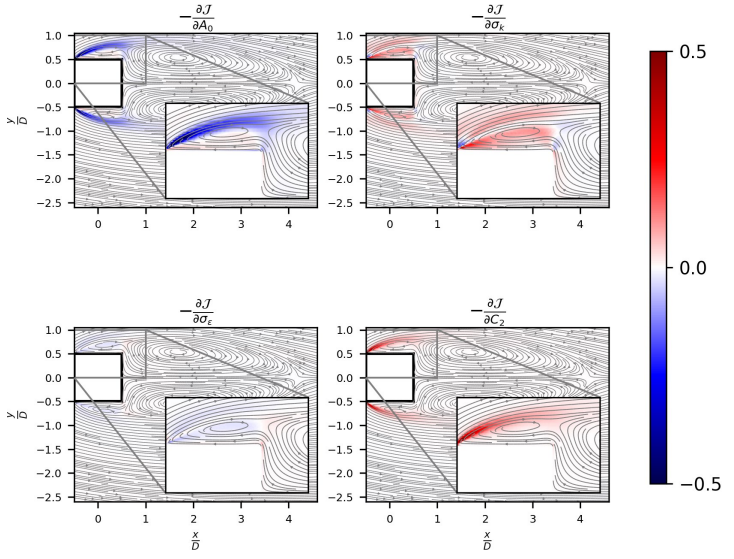


Figure 7: Closure coefficient sensitivities at horizontal plan with normalised height $z/H_{ref} = 0.19$: (a), sensitivity to A_0, (b), σ_k, σ_ϵ (c), and, (d), C_2 (d). Sensitivities are normalised by their in-plan peak values.

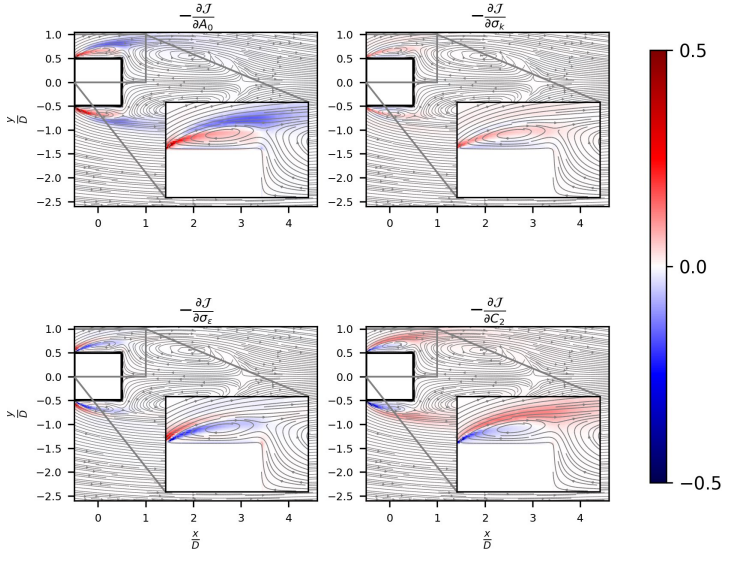


Figure 8: Closure coefficient sensitivities at horizontal plan with normalised height $z/H_{ref} = 1$: (a), sensitivity to A_0, (b), σ_k, σ_ϵ (c), and, (d), C_2 (d). Sensitivities are normalised by their in-plan peak values.

As shown in table 1, C_2 is the most sensitive coefficient. Its large value suggests that the coefficient is less subject to sign cancellations in the sensitivity maps. This is especially observed in the horizontal map at lower elevations far enough from the high-end (plot at $z/H_{ref} = 0.19$ where $H_{ref} = 2/3H = 3.3D$). With regards the other coefficients, their global sensitivity derivatives suggest, a decrease of both the eddy-viscosity pre-factor A_0, the turbulent energy diffusivity pre-factor $1/\sigma_k$ and an increase of the diffusivity pre-factor $1/\sigma_\epsilon$ of the energy dissipation. This implies, in the one hand, that the turbulence dissipation rate

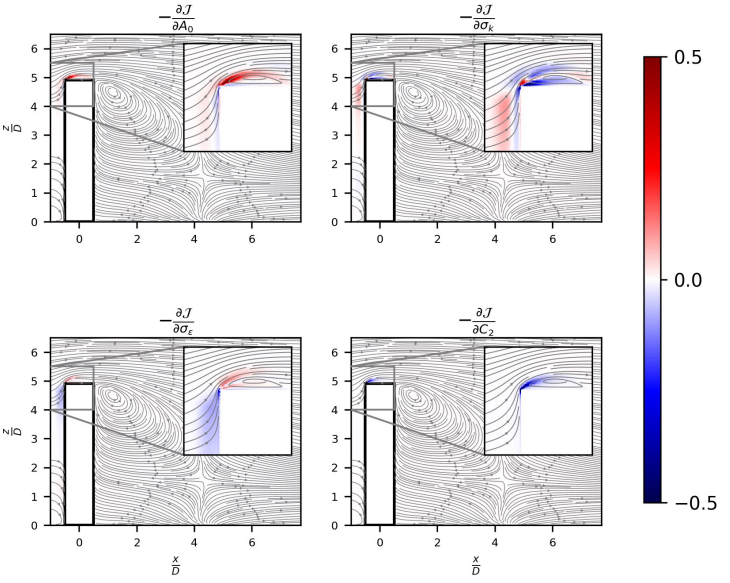


Figure 9: Closure coefficient sensitivities on the symmetry plan with normalised: (a), sensitivity to A_0, (b), σ_k, σ_ϵ (c), and, (d), C_2 (d). Sensitivities are normalised by their in-plan peak values.

is under-predicted all over the domain and that a better redistribution of the flow field can be obtained by increasing C_2. Inversely, by decreasing the eddy viscosity pre-factor A_0 (which increases C_μ), this trend shows that the original turbulence model tends globally to under-estimates the momentum mixing and thus it globally advocates higher turbulent diffusion. By looking at local sensitivities to this coefficients, these global directions are strongly driven by the high interest giving to the lateral free-stream (surrounding lateral recirculating flow) entrained toward the wake region. Thus, upon calibration, we may expect some preferential improvement of the wake extension and, inversely, a retrogression in the extension of lateral separated flow.

5.2.2. *Closure hypothesis analysis through data assimilation*

In this section, the optimisation problem is solved iteratively by following Algorithm 1, and we discuss the data assimilation procedure's ability to estimate the flow state. Guided by the work of Shih *et al.* (1995), we intend here to devise some penalty ranges for the coefficients. Concerning the coefficients which are involved in the energy dissipation rate budget, referring to the work (Shih *et al.* 1995), the C_2 coefficient is actually expressed as $C_2 = \beta/\eta$, in which $\beta = \eta + 1$ is the dissipation decay rate (such as $\epsilon/\epsilon_{t_0} = (t/t_0)^{-\beta}$ where t_0 is an initial time) and η is the energy decay exponent (such as $k/k_{t_0} = (t/t_0)^{-\eta}$) that varies from 1.08 to 1.3 in decaying homogeneous turbulence experiments (Shih *et al.* 1995). Thus, a range for this coefficient can be set as $C_2 \in [1.76, 1.93]$, where the background value is 1.9.
For σ_ϵ, the inertial turbulence assumption near the wall allows to establish

$$\sigma_\epsilon = \frac{\kappa^2}{C_2\sqrt{C_\mu} - C_1} \tag{5.1}$$

were the von Kármán constant $\kappa = 0,41$, the eddy viscosity coefficient $C_\mu = 0.09$ and $C_1 = 0.43$. Assuming a quasi-linear dependency between the two constants (see figure 11), knowing the range on C_2 we obtain $\sigma_\epsilon \in [1.14, 1.71]$. To possibly relax the underlying

Scenario	σ_k	σ_ϵ	C_2	A_0	$\frac{\mathcal{J}-\mathcal{J}_0}{\mathcal{J}_0}$
Default value	1.0	1.2	1.9	4.0	
A	1.14	0.99	1.92	4.05	10.7%
B	1.0	0.74	1.96	4.08	15.7%
C	1.03	1.07	1.95	3.99	9.8%

Table 2: Summary of the optimisation results, closure coefficients and the relative decrease of cost function.

assumption of decaying turbulence for this range, two cases study will be considered for this constant. In the first scenario, it will be assumed that relation (5.1) holds beyond the inertial layer, as established by Shih *et al.* (1995). On the second scenario, this constraint is relaxed and coefficient σ_ϵ is assumed to be an independent control parameter. In that case, the closure is thus performed by the data. The second case is expected to bring more degree of freedom in the optimisation process, due to the independent adaptation of the two coefficients.

In the transport equation of k, the coefficient σ_k, which adjusts the level of turbulent energy mixing with respect to the momentum eddy diffusivity, is in commonly fixed to unity (as in any $k-\epsilon$ turbulent model). This generally assumes a quasi-equality between the scalar and the momentum mixing. Due to the lack of comparative studies in the literature between the $realizable$ model results and experiments, estimating a physical range for this coefficient is not possible. Therefore, we considered two optimisation procedures where in the first one we maintain $\sigma_k = 1$ while in the other case we relax this constraint letting σ_k evolves in the arbitrary chosen range: $\sigma_k \in [0.9, 1.1]$. Similarly, for the bounds on A_0, without any *a priori* informations on its physical range, we fixed a larger range of possible value: $A_0 \in [3.6, 4.4]$, where the background usual value is 4.0.

Based on the remarks of the previous section, the results of three data-assimilation scenarios are discussed and compared. A first straightforward approach corresponds to the optimisation of the four coefficients independently. This is referred to as scenario **A**. Then, two scenarios are considered to investigate the two closure assumptions mentioned in the previous section. First, we consider the equality between the mixing of turbulent kinetic energy and momentum, referred as scenario **B**. Secondly, keeping σ_k a free parameter, the scenario **C** consists in enforcing the inertial constraint and defining σ_ϵ using (5.1). Three criteria are considered to evaluate the agreement between the CFD results and the measurements. The first is the relative reduction of cost function $\frac{\mathcal{J}-\mathcal{J}_0}{\mathcal{J}_0}$, $\mathcal{J}_0$ being the initial cost. This depicts the improvement of the global effect of wind on the building. Next, C_p, the dimensionless pressure, is compared locally on the facades of the building. Third, to quantify the accuracy of the recovered mean flow field, the streamwise length of the recirculation region behind the building is compared to the one observed from the PIV plans.

Regarding the update of the coefficients, a steepest descent algorithm is used with an adaptive step. A maximum step size is set to 10^{-2} while a minimum step size inferior to 10^{-4} is considered as an optimisation convergence criteria. The confidence coefficients are all set to $\zeta_i = 5 \times 10^{-2}$. This low uniform values represent a relative degree of confidence on the background closure values. The variations of the closure coefficients along the optimisation iterations are shown in figure 10. The maximum reduction of the cost and the optimal coefficients values for the three considered scenarios are summarised in table 2. In terms of mismatch between CFD and experimental mean pressure, it is shown that the highest reduction can be achieved through the optimisation scenario **B**. Conversely, scenario **C** leads to the least improvement in the cost function. However, we note a faster convergence rate for

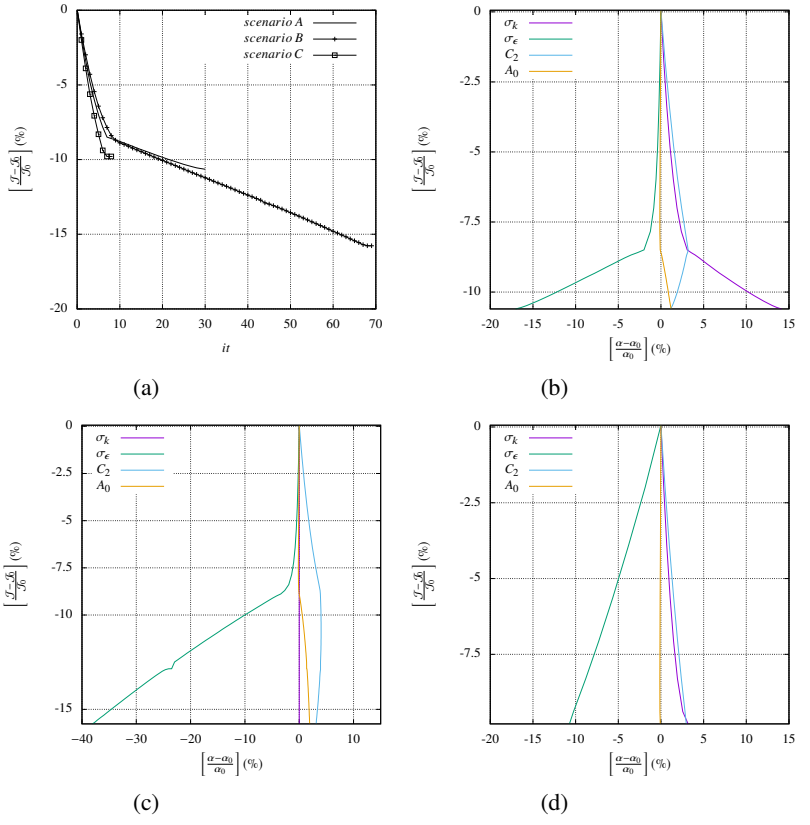


Figure 10: Cost function reduction, (a), and closure coefficient variations: (b), scenario **A**; (c), scenario **B**; (d), scenario **C**.

C, for which the optimal solution is reached 10 times faster than for **B**. Furthermore, a shift between the two regimes can be noticed in scenarios **A** and **B**. Indeed, this shifting occurs when the penalisation on the variation of C_2 becomes of the same order of magnitude than the required advancement for the cost minimisation. Whereas in **B**, σ_k is not optimised and the trend on C_2 until convergence is mainly dominated by its penalisation. In all scenarios, the optimal value of C_2 increases while it stays within 5% of the background value. Considering A_0, a minor variation is observed during optimisation in all scenarios. On the contrary, a higher variation of σ_ϵ below the recommended range is necessary to reduce the cost function. In figure 11, we show the variation of σ_ϵ with respect to C_2. In scenario **C**, a quasi-linear dependency is established through relation (5.1). However, we retrieve the two regimes in **B** and **A** where this dependency is broken.

In general, it can be concluded that a better agreement between the turbulence model (e.g *realizable* $k-\epsilon$) and wind tunnel experiments, in terms of wind load on the facades of high-rise buildings, can be achieved through optimisation of the closure coefficients. Even if it offers less degrees of freedom in the optimisation, better results are obtained when enforcing the constraint that equals turbulence mixing in the equation of transport of k to momentum mixing by the eddy viscosity (scenario **B**). This suggests that it is a physically valid hypothesis in our case study. It helps structuring the data assimilation process and leads to a robust procedure. It states that the turbulent mixing of the momentum and kinetic energy are of same nature. At the opposite, by relaxing the constraint and establishing

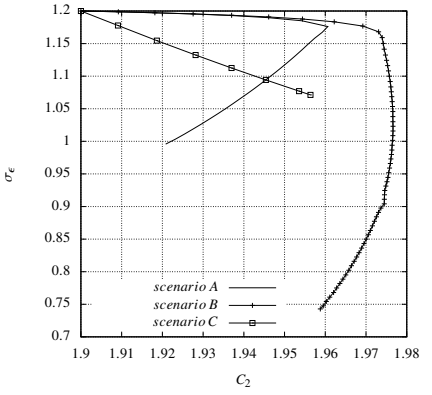


Figure 11: σ_ϵ variations with respect to C_2 for all scenarios.

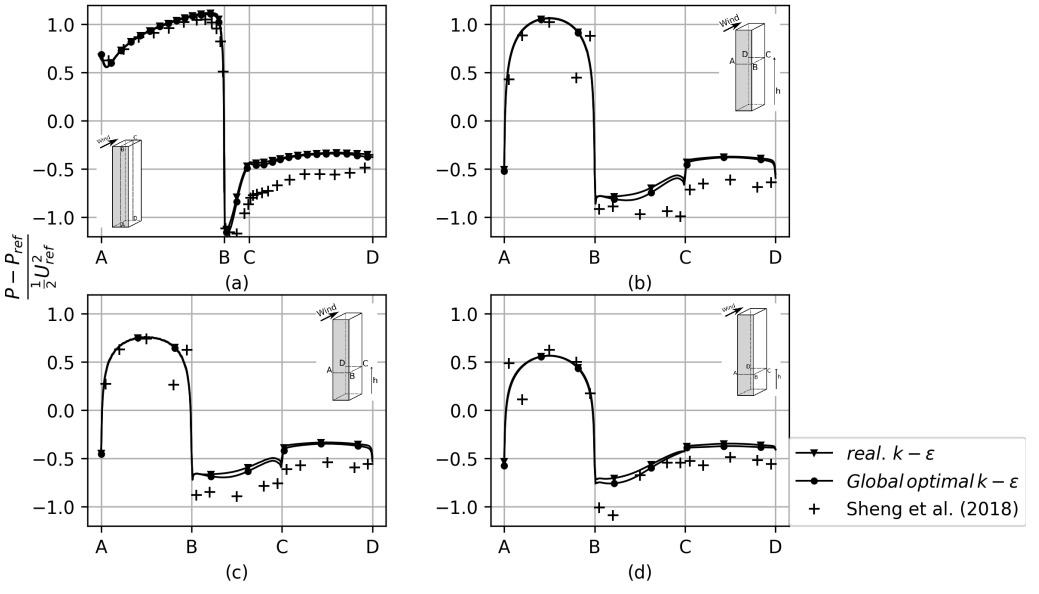


Figure 12: Comparison of pressure coefficient profiles between CFD (scenario **B**) and experimental results along building facades. Contours are token at building symmetry plan, (a) and three horizontal plans at $z/H_{ref} = 1$, (b), $z/H_{ref} = 0.27$, (c) and $z/H_{ref} = 0.19$, (d), respectively.

relation (5.1) as valid out of the inertial layer (scenario **C**) may lead to lower agreement with measurements. Indeed, this assumption may hold reasonably in flows where turbulence behaviour is isotropic. However, in the presence of bluff body, e.g flows with separation and recirculation dynamics, this assumption is undoubtedly unrealistic. Scenario **B** might be considered as the best optimisation choice to get better wind load representation on high-rise building given the considered turbulence closure (*i.e.* realizable $k-\epsilon$). Following the best optimisation scenario, 15% gain on the overall predicted loads are obtained. Furthermore, a comparison of the predicted pressure coefficient at the building facades (see figure 12) this gain is associated to the slight improvement observed especially along the side facades.

Nearly no change at the front facade and along the upfront corners is observed. As a matter of fact, this observations confirms what was earlier mentioned in the sensitivity analysis

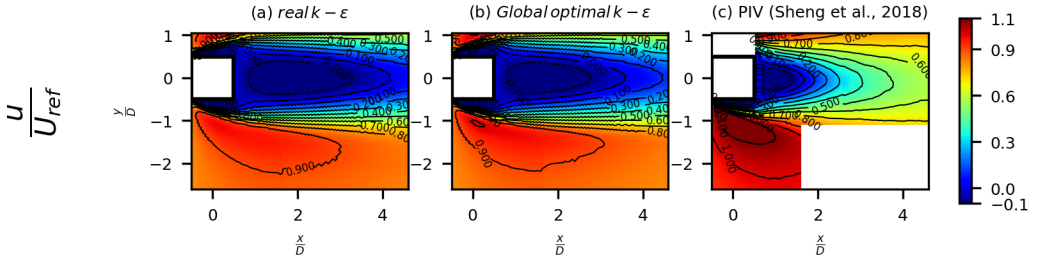


Figure 13: Comparison of mean stream-wise contour between CFD (scenario **B**) and Re. Sheng (PIV) experiments at horizontal plan with normalised height $\frac{z}{H_{ref}} = 0.19$.

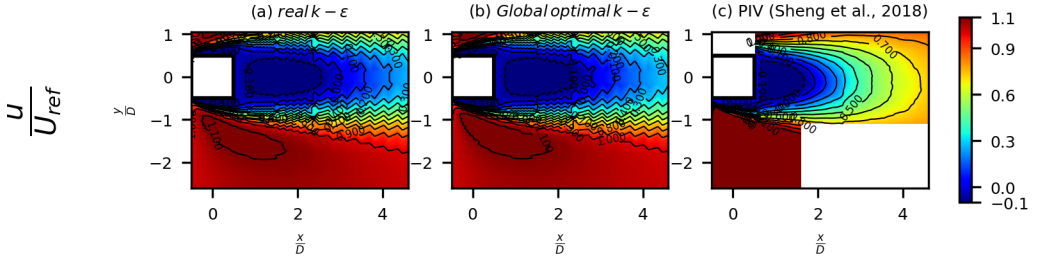


Figure 14: Comparison of mean stream-wise contour between CFD (scenario **B**) and Re. Sheng (PIV) experiments at horizontal plan with normalised height $\frac{z}{H_{ref}} = 1$.

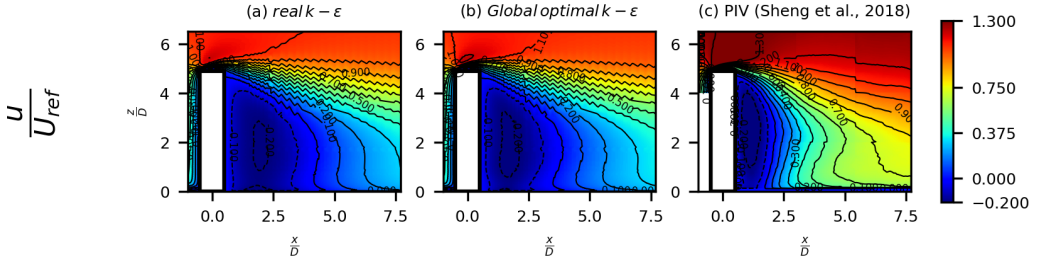


Figure 15: Comparison of mean stream-wise contour between CFD (scenario **B**) and Re. Sheng (PIV) experiments at symmetry plan with normalised height $\frac{z}{H_{ref}} = 1$.

where the rigidity of the considered turbulence model is shown to play a major role on the degree of improvement that can be achieved to fit with measurements.

With regard to the mean flow reconstruction, adopting the best optimisation scenario (scenario **B**), the contours of the mean velocity field are compared with the available PIV plans reported from the work of Sheng *et al.* (2018). It is a strong validation since these mesurements are not used in the data assimilation. Figures 13, 14 and 15 show the normalised streamwise velocity at the streamwise central plan (top) and at two horizontal plans, *i.e.* $z/H_{ref} = 0.19$ and $z/H_{\text{ref}} = 1$. The CFD with background values and optimised values following **B** are compared with the PIV measurements.

In order to show the effect of data assimilation, velocity contours are superimposed (right column) and thicker lines are plotted to track the size of the recirculation region. The reattachment length on the ground, is reported in table 3. After optimisation, velocity contours show a better estimation of the recirculation region length which is shorter compared to the non-optimised model. This improvement is more affirmed near the ground, where the relative error of reattachment length Y with respect to PIV, $\varepsilon_Y = \frac{\nabla Y_0}{\nabla Y_{\text{end}}}$, with the default

	Exp	*realizable* $k-\epsilon$	*optimised*
$Y^f(m)$	~ 50	139.8	117.1
$Y^r(m)$	-	13.6	11.7

Table 3: Comparison of the (dimensional) reattachment lengths on the roof and floor, CFD optimised with global constant calibration (scenario **B**).

	$\frac{\partial \mathcal{J}}{\partial f_\epsilon}(\times 10^8)$
AD	-3.55
FD ($\delta f_\epsilon = 10^{-1}$)	-14.03
FD ($\delta f_\epsilon = 10^{-2}$)	-4.03
FD ($\delta f_\epsilon = 10^{-3}$)	-4.45
FD ($\delta f_\epsilon = 10^{-4}$)	-4.43
FD ($\delta f_\epsilon = 10^{-5}$)	-4.19

Table 4: Sensitivity derivative w.r.t. f_ϵ computed from the proposed adjoint model (AD) and the finite difference (FD) with different perturbation sizes.

(Y_0) and the optimised (Y_{end}), is reduced by 25%. Despite this enhancement, it should be pointed out that CFD model still under-predicts the flow in the wake region. This is the best improvement of this specific turbulence model we obtained by assimilating the pressure measurements. The two limiting ingredients are the model rigidity and the partial sparse observations (pressure at the boundary) of the complex flow. In the next section, spatially distributed control parameters in the transport equation of dissipation are considered in order to relax this structural constraint.

5.3. *Distributed closure parameter in the energy dissipation budget*

This section is dedicated to the results related to the investigation of the adjunction on $k-\epsilon$ (realizable) model of a distributed control parameter in the energy dissipation rate budget. First, we analyse the sensitivity fields to highlight the spatial locations where the closure form of ϵ budget appears inadequate to reproduce the measurements and would require a structural correction. Then, the data-assimilation results of this spatially corrected model are analysed.

5.3.1. *Sensitivity analysis*

For quantitative validation, adjoint sensitivities are compared with finite differences in table 4 for different perturbations, $\delta f_\epsilon = \{10^{-5}, \ldots, 10^{-1}\}$ inside Ω_{in} and zero elsewhere. Thus, the lower deviation obtained for a step size of 10^{-2}, for which rounding errors are less effective, confirms the validity of the implementation.

We analyse the parameter sensitivity fields given by the proposed closure model, which corresponds to the first iteration step of the data assimilation procedure. Indeed, we are interested in the gradient of the cost functional with respect to the distributed control parameters for $f_\epsilon = 0$. Figures 16, 17, and 18 compare the sensitivity maps for the added control parameter against a direct forcing (which corresponds to the adjoint variable on ϵ^*). Globally, sensitivity to the proposed parameter f_ϵ shows a strong response in a restricted flow area. In contrast with very diffused sensitivity maps for the direct forcing, the sensitivity maps of the additional forcing do highlight the regions of great relevance for the model

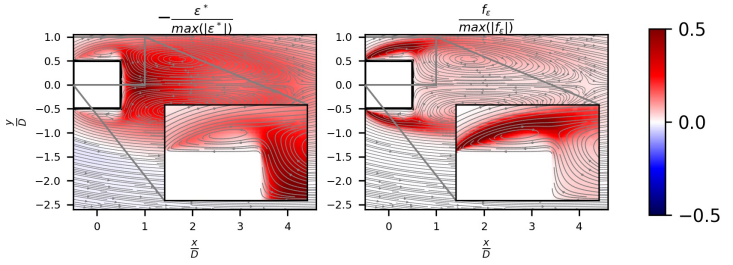


Figure 16: Adjoint turbulence dissipation and the constrained control at horizontal plan with normalised height $\frac{z}{H_{ref}} = 0.19$. Variables are normalised by their in-plan peak values.

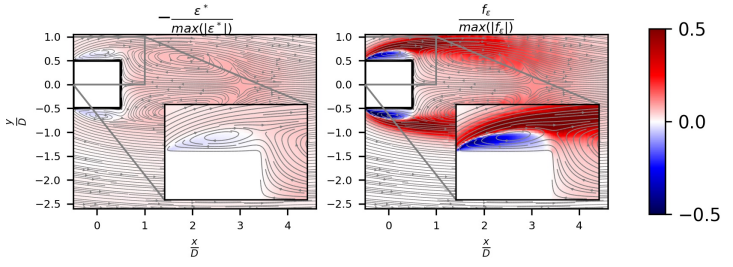


Figure 17: Adjoint turbulence dissipation and the constrained control at horizontal plan with normalised height $\frac{z}{H_{ref}} = 1$. Variables are normalised by their in-plan peak values.

improvement. They correspond to the same regions as those designated by the sensitivity maps of the global constants in the previous section. For instance, we note a tendency to bring significant dissipation rate adjustments starting from the leading edges and continuing into the lateral shear layers and more downstream at the wake region edges. Let us note that multiplication by the variable ϵ has damped sensitivties at regions nonrelevant for turbulence energy budget, such as the high peaks of sensitivity observed around the wake centerline for the direct forcing (see figure 16). As we span upward, as shown at height $z/H_{ref} = 1$ in figure 17, the maps actually reveal a step function tendency as we go from separated flow regions, *i.e.,* the lateral and top shear layers, toward the wake region. Furthermore, with regard to the sign, eventual contributions to the ϵ budget are interpreted as follows. A negative value of f_ϵ would tend to increase the dissipation rate, while a positive value would instead decrease it. Hence, on both lateral and top separated flows, the parameter suggests there an increase of the dissipation rate. The sensitivity analysis points here an over-production of turbulent kinetic energy, which is a known common default of the $k-\epsilon$ closure models in such flow configurations reported, for instance, in (Murakami 1990, 1997; Shirzadi *et al.* 2017). Moreover, along the outer edges of the lateral shears toward the wake edges, the sensitivity maps suggest reducing the dissipation rate. This tendency is consistent with a rather under-predicted turbulent mixing, resulting in the overly extended wake region behind the building (Shirzadi *et al.* 2017).

5.3.2. *Closure analysis through data-assimilation*

We consider now the solution of a data-model coupling using the modified closure equation (3.2). Regarding the assimilation procedure's setting, the steepest descent algorithm is used with the Sobolev gradient computed in (3.6) as a descent direction. Regarding the filtering choice, two values $l_{sob} = 0.1D$ and $0.2D$ were tested. Let us note that with $l_{sob} = 0$, *i.e* no smoothing of the L^2 gradient, the procedure was notably unstable, showing the need for

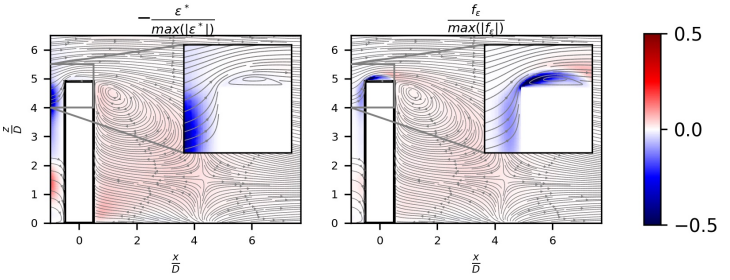


Figure 18: Adjoint turbulence dissipation and the constrained control at symmetry plan. Variables are normalised by their in-plan peak values.

regularisation. As for a higher value of $l_{sob} = 0.2D = 6\,\text{m}$, this choice yielded to an over smoothing. As the sensitivity varies by length scales that are quite small in comparison to this length, this advocates a value of $l_{sob} \sim$ lateral recirculation width. Therefore, a filtering length scale equivalent to 10% of the building's width seems to give a good compromise to filter the small-scales, as suggested also in Tissot *et al.* (2020).

To assess the performance of the proposed regularisation technique, we compare the optimisation convergence using the Sobolev gradient with the cost penalisation approach. With an additional regularisation term given by the gradient of f_ϵ, the cost function would read

$$\tilde{\mathcal{J}}(P, f_\epsilon) = \mathcal{J}(P) + \gamma \langle \nabla f_\epsilon, \nabla f_\epsilon \rangle_\Omega$$

in which $\gamma = \zeta * \max(\epsilon\epsilon^*)$ is a free hyper-parameter controlling the amount of penalisation. Computing the Gateaux derivative w.r.t. f_ϵ brings an additional volume term, such that the L^2 optimality condition (3.4), which rewrites as

$$\frac{\partial \mathcal{L}}{\partial f_\epsilon} = \epsilon\epsilon^* - \gamma \nabla \cdot \nabla(f_\epsilon).$$

The comparison of the results obtained with both approaches are shown in figure (19) for $\zeta = 0.5, 1$, and 2.5. We can see that Sobolev gradient leads to the lowest discrepancy ($\frac{\mathcal{J}}{\mathcal{J}_0} \simeq 0.42$) compared to the best gradient penalisation ($\frac{\mathcal{J}}{\mathcal{J}_0} \simeq 0.56$ with $\zeta = 1$). Regarding the convergence speed, minimisation with the Sobolev gradient reaches the minimum in less than 200 iterations while convergence for the regularisation approach with $\zeta = 1$ requires more than 400 iterations. The convergence process through the projection in $H^1(\Omega)$ is therefore much faster. This might come because adding the gradient to the cost function changes the original descent direction prescribed by the discrepancy to data. Moreover, concerning the calibration, the cost functional reduction has doubled compared to the calibration results obtained using the best scenario **B** ($\frac{\mathcal{J}}{\mathcal{J}_0} \simeq 0.85$). Indeed, this difference is consistent because the additive parameter f_ϵ is less constrained by the parametric rigidity associated to the global coefficients. In what follows, we analyse the method's performance in terms of the flow reconstruction. This will further illustrate the adjoint approach's capability for data-coupling with local corrections of the turbulence closure.

Wind load profiles The reconstructed pressure loads are compared with the experimental data (Sheng *et al.* 2018) and the non-assimilated model in figure 20. We can see that, in comparison with the coefficient calibration, the modified closure model produces far better results in most of the building's wall regions. In terms of pressure discrepancy, the modified closure model manages to capture well suction at both top leading edge in the symmetry plan (point B in sub-figure(a)) and lateral leading edges (see figure 20(b), (c), and (d)). However, while a good agreement with the data is obtained along the lateral facades, minor

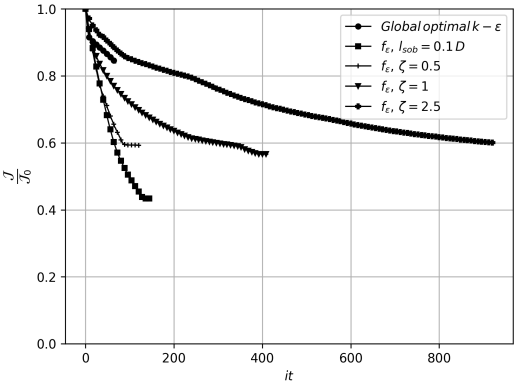


Figure 19: Cost function reduction. Comparison is made between an ϵ budget correction and global calibration using scenario **B**.

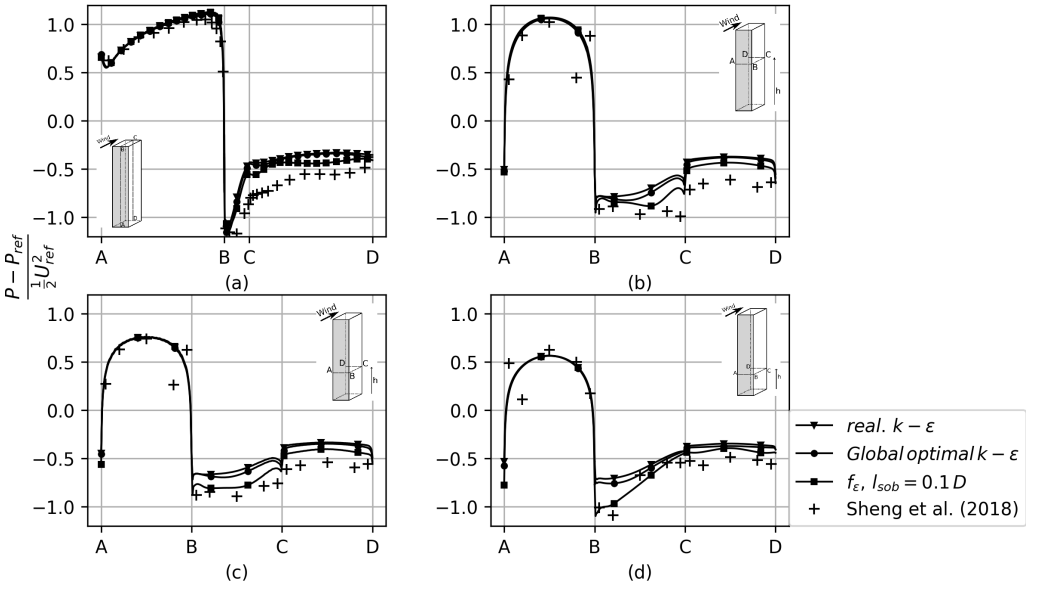


Figure 20: Pressure coefficient profiles along building facades. Comparison is made between an ϵ budget correction and global calibration using scenario **B**. Contours are token at building symmetry plan, (a) and three horizontal plans at $z/H_{ref} = 1$, (b), $z/H_{ref} = 0.27$, (c) and $z/H_{ref} = 0.19$, (d), respectively.

to important deviations are apparent as we get closer to the trailing edges and especially when we approach the high-end. This gradually leads to poorer pressure interpolation as the discrepancy reaches a maximum value at the upper back facade around H_{ref} (sub-figure(a) and (b)). Yet, the modified closure model shows a slightly better prediction near the top-trailing edge than the calibrated default model. Therefore, at this level of comparison, such closure model does improve the data-model capability. The remaining regions where no improvement is seen might reflect the limited controllability of such turbulence model with wall pressure measurement.

Flow topology Regarding the spanwise flow structure, figures 21 and 22 show sectional streamlines at both $z/H_{ref} = 0.19$ and $z/H_{ref} = 1$ height, respectively. Here it is noteworthy to mention that these sectional streamlines are computed for in-plane velocity components.

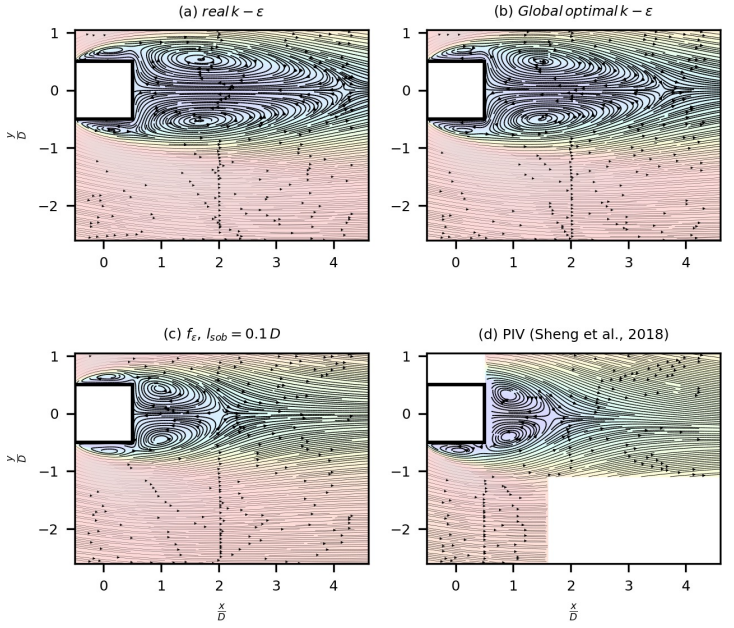


Figure 21: Flow topology (2D) on horizontal plan at normalised height $\frac{z}{H_{ref}} = 0.19$ with local constraint correction. Comparison is made between an ϵ budget correction and global calibration using scenario **B**.

In each sub-figure, streamlines predicted with the non-assimilated model (sub-figure (a)) is compared with the calibrated model (scenario **B**) (b), with f_ϵ closure correction (c) and PIV experiments from Sheng *et al.* (2018) (d). At both levels, the model's reconstructed flow (f_ϵ) still preserves the symmetry of the two distinct pairs of averaged vortex structure. Moreover, an excellent agreement with experiments is obtained in the wake transverse extension and the vortices focal point positions in comparison with the calibrated model.

Figure 23 shows time-averaged sectional streamlines at the $y/D = 0$ symmetry plane in the transverse-wise structure. As can be seen from figure 23, the two distinct types of average streamlines are also observed on both reconstructed flows ((b) and (c)). It is constituted by an upper recirculation starting at the roof-top and a lower recirculation region raised from the ground wall, separated by a saddle point. Thus, regarding the wake's extension, the modified closure model leads to a drastic reduction of the recirculating flow compared to the calibrated mode, thus reaching a realistic size. This can be quantified by the position of the saddle point ($x/D = 4$,$z/D = 2$) in the RANS model which has been moved around ($x/D = 2$,$z/D = 2$). This striking result is mitigated by the fact that this saddle point has been pulled slightly too far upstream. It should be recalled that only pressure measurements in the facade have been available and that PIV measurements are used here only for validation. This good agreement with external data proves that we are neither in overfitting nor in an over-constrained situation. Indeed, the two-dimensional vortices at both elevations ($\frac{z}{H_{ref}} = 0.19,\ 1$) are a transverse projections of the three-dimensional rolls, one on each side of the wake symmetry plan, which connects near the free end. Such structure is consistent with some model descriptions brought on wakes of finite length square cylinders, with similar height/width ratio, that are subject to boundary layer flows of various thickness (Kawamura *et al.* 1984; Wang & Zhou 2009). A global three-dimensional picture gathering the two-dimensional previous plots is shown in figure 24). Examining the optimal forcing fields in figure 25, we retrieve the same tendencies

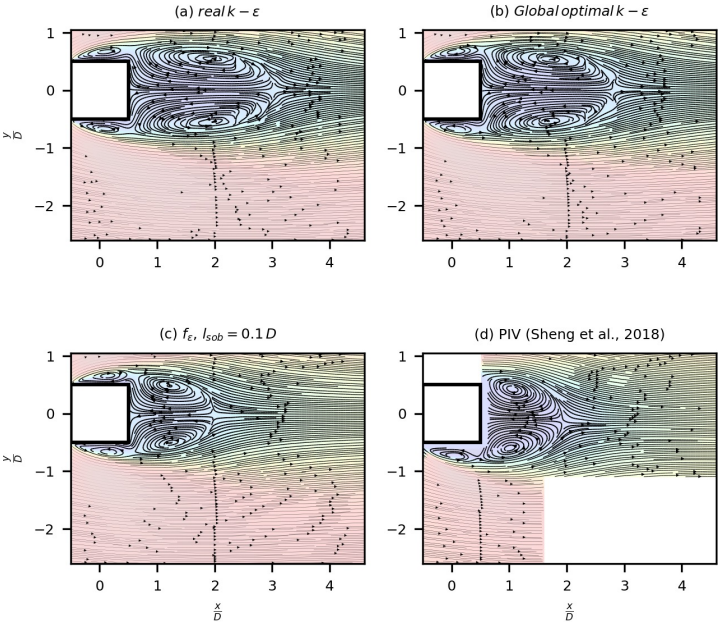


Figure 22: Flow topology (2D) on horizontal plan at normalised height $\frac{z}{H_{ref}} = 1$ with local constraint correction. Comparison is made between an ϵ budget correction and global calibration using scenario **B**.

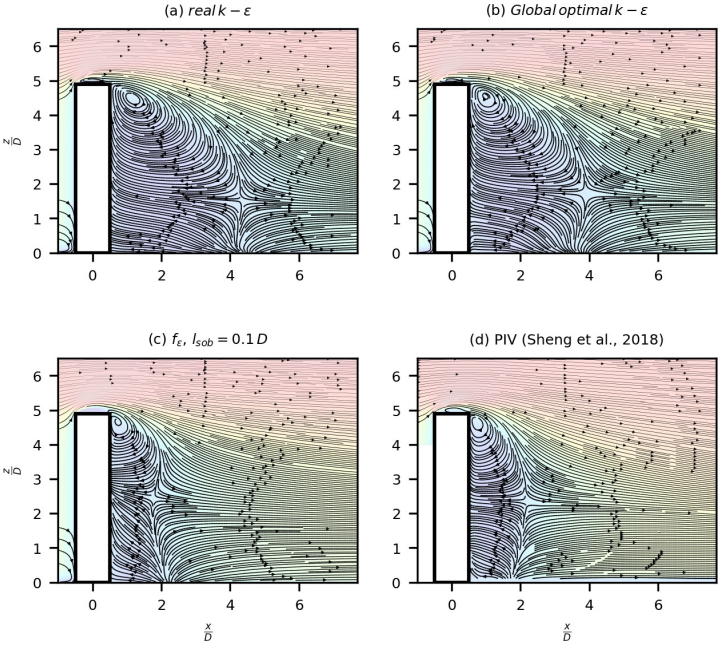


Figure 23: Flow topology (2D) on symmetry plan with local constraint correction. Comparison is made between an ϵ budget correction and global calibration using scenario **B**.

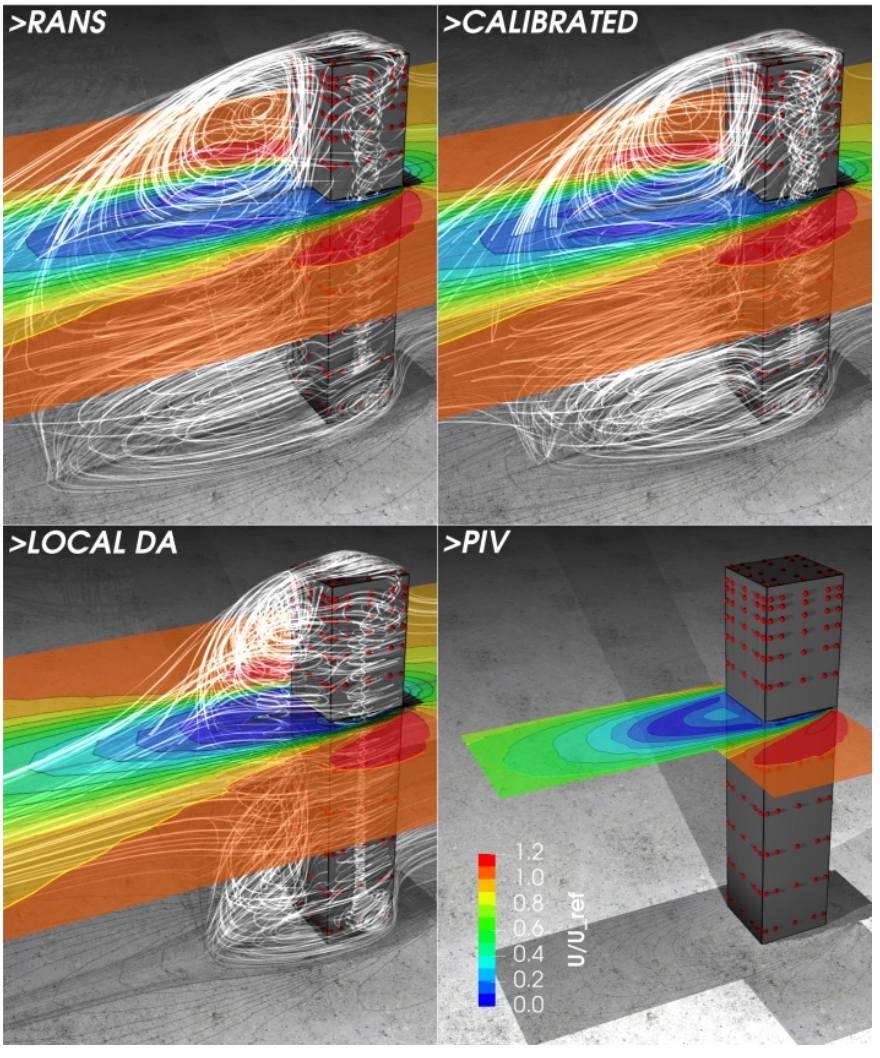


Figure 24: Flow topology (3D) of the realizable $k-\epsilon$ model, global coefficient calibration and ϵ budget correction against PIV plan at $z/H_{ref} = 1$.

observed in the sensitivity analysis before the reconstruction. After the optimisation, the parameter f_ϵ still keeps advocating less turbulence production into in the shear layers of the lateral separated flow (see 25(a) and (b)), while, conversely, more mixing at the edges of the wake downstream. Corrections are performed where strong turbulence inhomogeneity and anisotropy occur. It acts in a way to redistribute dissipation rate, by the means of sources and sinks, from the upstream region toward the downstream region. When considering only calibration of the global coefficients, the model structure prevents this redistribution. This suggests that some turbulence mechanisms related to anisotropy and inhomogeneity effects are not properly taken into account in the model closure and need to be included to represent accurately some key regions of the flow.

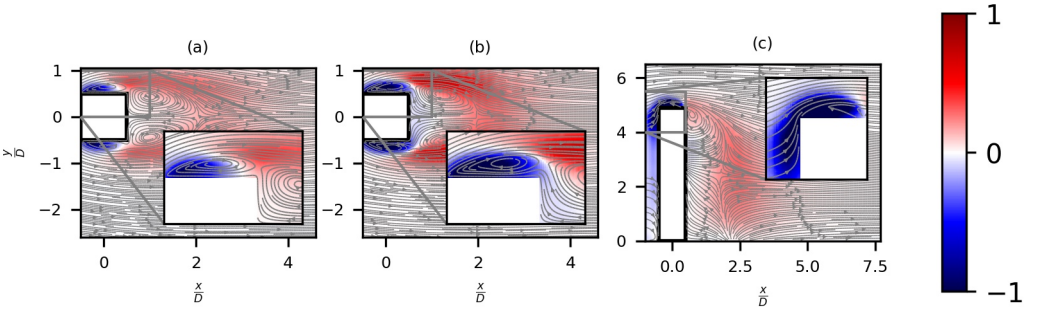


Figure 25: Converged parameter f_ϵ, with projection in H_1 ($l_{sob} = 0.1D$); (a), at horizontal plan at $z/H_{ref} = 0.19$, (b) $z/H_{ref} = 1$ and, (c) at symmetry plan $y/D = 0$.

6. Conclusions

The use of steady RANS models under the eddy viscosity hypothesis is known to be inaccurate for practical applications such as micro-climate studies (at urban scale). For instance, in the prediction of wind-loads on a high-rise building, most of state-of-the-art $k-\epsilon$ turbulence models (including the realizable revision studied here) tend to give poor wake flow accuracy estimations, as well as an inaccurate wall-pressure value, when compared to wind tunnel experiments. One way to tackle this deficiency consists in adopting data-model coupling techniques such as the variational DA approach based on optimal control. To set up such a framework in our context we devised a consistent analytical derivation of one of the most common turbulence models (*i.e.* realizable revision of $k-\epsilon$) coupled with near wall closure. This has resulted in the definition of a continuous adjoint model (together with its consistent boundary conditions) of the tangent linear operator of the RANS model. Given the dual description of the dynamics composed of the RANS direct model and the adjoint of its tangent linear representation, we have explored three methodological settings that provides an efficient sensitivity analysis and an in-depth diagnosis of the turbulence closure adopted on such flows.

The first tool consisted in the inspection of the adjoint state variables in relation with their physical meaning. The second one was dedicated to providing a better understanding of the model output's variabilities in terms of the model's closure global constants. With the last one, we went one step further. We considered the adjunction of a distributed parameter which enables the reanalysis of the closure at a structural level (such as the choice of the transport equation for the energy dissipation rate on the $k-\epsilon$ model as considered here). To conduct an efficient structural inspection, a distributed parameter is sought in a Sobolev space and further estimated through a data-assimilation procedure. As a sensitivity field is generally not very regular for distributed parameters, a projection onto the Sobolev space $H^1(\Omega)$ was proposed here for both a regularisation purpose and to define an improved descent direction for the minimising technique. These three settings have then been applied to a high-rise building case study.

Sensitivity maps of the $k-\epsilon$ global coefficients had revealed high interest in optimising them mainly at the shear layers resulting from flow separations at the leading lateral edges and on top of the building. Moreover, little sensitivity in the bulk of the recirculation wake region was observed. Despite all the spatial variability of the sensitivity fields, it was shown that the optimality condition drastically reduces the high dimensional dependency of the model to each coefficient. Regarding the model hypotheses which guided the choice for closed default values, a better data coupling is obtained by enforcing the constraint that equates turbulence energy mixing to momentum mixing, even if it offers fewer degrees of freedom in

the optimisation. This suggests that it is a physically valid hypothesis that structures the model and then helps for convergence. Conversely, by relaxing this constraint and establishing the relation (5.1), which dictates a strong bound limiting the production of energy dissipation rate to its redistribution (supposedly valid in the inertial layer near the wall), this leads to lower agreement with experiments. As both assumptions constitute a common practice for closure to most eddy viscosity models, it is expected that these results extend to several other models of similar forms. The limited performance of the DA procedure, achieved when controling global turbulence parameters, points out the rigidity of the considered turbulence model when used with realistic wall pressure measurements.

Considering a distributed parameter to the ϵ budget, in order to complement the model in terms of local source/sink process, sensitivity maps highlight regions where global constants are not too sensitive (for instance, the wake region) and exhibit relatively less variability in term of sign changes. Maps had actually revealed binary tendencies, separating the lateral and top shear regions and the wake flow. Regarding the regularisation, a comparison of the cost reduction results with a conventional penalisation approach showed that projection in $H^1(\Omega)$ yields a much faster convergence and lower discrepancy levels. Let us note that, along with the $H^1(\Omega)$ projection, the robustness of the DA procedure was enabled as a first order numerical scheme to solve the dual dynamics. Regarding the reconstruction ability, compared with coefficient calibration, the modified closure model produced better results in most of the building's wall regions in terms of wind load profiles, yet, results suggest some remaining restrictions as reconstructed profiles tend toward the original model in some regions. An excellent agreement with PIV experiments was obtained in wake transverse extension. It should be recalled that only pressure measurements in the facade have been assimilated. This good agreement with measurements of different nature and that have not been used in the assimilation proved that we are neither overfitting the data nor in an over-constrained situation.

This work thus illustrates the capabilities of adjoint methods. Beyond providing a data-driven flow reconstruction, they enable an in-depth analysis of the turbulence closure. Indeed, by regarding adjoint fields as a physical forcing, rather than as a purely mathematical object, these data-driven reconstructed fields allows to highlight a misrepresentation of the turbulent flow by the RANS model, and hence, to address errors within a particular turbulence modelling form (*e.g.* energy production, backscatterring, redistribution or dissipation). Although the results presented were for a particular turbulence model and on a specific bluff-body-like case, over-estimation of the recirculation length in bluff bodies is a common features in RANS models, and the proposed methodology could be employed without loss of generality. Provided sparse wall pressure measurements, the technique can be directly applied to any complex wake flow embedded in the atmospheric boundary layer.

Acknowledgements

This work was supported by the Centre Scientifique et Technique du Bâtiment (CSTB) (Nantes, France) in collaboration with mixed research team Fluminance (INRIA, INRAE in Rennes, France). The authors wish to deeply thank Fabrice De Oliveira, for his help in providing us the experimental data.

Declaration of Interests

The authors report no conflict of interest.

APPENDIX

Under differentiability condition, it can be shown (Le Dimet & Talagrand 1986; Gunzburger 2003) that the problem of determining the optimal set of flow state variables,

$$\mathbf{X} = (U_x, U_y, U_z, P, k, \epsilon, \mu_t)$$

and the set of parameters α, of the cost function $\mathcal{J}(\mathbf{X}, \alpha)$ under the constraint

$$\mathbf{M}(\mathbf{X}, \alpha) = 0$$

is equivalent to the problem of determining the optimal set of theses variables in addition to an adjoint state

$$\mathbf{X}^* = (U_x^*, U_y^*, U_z^*, P^*, k^*, \epsilon^*, \mu_t^*)$$

of the Lagrangian functional $\mathcal{L}(\mathbf{X}, \mathbf{X}^*, \alpha)$. With the inner product defined as $\langle \psi, \phi \rangle_\Omega = \int_\Omega \psi^T \phi \, \mathrm{d}\Omega$ where ψ and ϕ are any two regular vectorial functions defined on the domain Ω, the Lagrangian, is

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

The first order variation $\delta\mathcal{L}$ resulting from perturbation $(\delta\mathbf{X}, \delta\mathbf{X}^*, \delta\alpha)$ of $(\mathbf{X}, \mathbf{X}^*, \alpha)$, in compact form, is equal to

$$\begin{aligned}
\delta\mathcal{L} = {} & \frac{\partial \mathcal{J}}{\partial P}\delta P + \frac{\partial \mathcal{J}}{\partial \alpha}\delta\alpha + \int_\Omega (\mathbf{X}^*)^T \left(\frac{\partial \mathbf{M}}{\partial U_x}\delta U_x\right) \mathrm{d}\Omega + \int_\Omega (\mathbf{X}^*)^T \left(\frac{\partial \mathbf{M}}{\partial U_y}\delta U_y\right) \mathrm{d}\Omega \\
& + \int_\Omega (\mathbf{X}^*)^T \left(\frac{\partial \mathbf{M}}{\partial U_z}\delta U_z\right) \mathrm{d}\Omega + \int_\Omega (\mathbf{X}^*)^T \left(\frac{\partial \mathbf{M}}{\partial P}\delta P\right) \mathrm{d}\Omega + \int_\Omega (\mathbf{X}^*)^T \left(\frac{\partial \mathbf{M}}{\partial k}\delta k\right) \mathrm{d}\Omega \\
& + \int_\Omega (\mathbf{X}^*)^T \left(\frac{\partial \mathbf{M}}{\partial \epsilon}\delta\epsilon\right) \mathrm{d}\Omega + \int_\Omega (\mathbf{X}^*)^T \left(\frac{\partial \mathbf{M}}{\partial \mu_t}\delta\mu_t\right) \mathrm{d}\Omega + \int_\Omega (\delta\mathbf{X}^*)^T \mathbf{M}(\mathbf{X}, \alpha) \, \mathrm{d}\Omega. \qquad (6.1)
\end{aligned}$$

Using the duality identity defined as

$$\int_\Omega (\mathbf{L}\phi)\, \psi \mathrm{d}\Omega = \int_{\partial\Omega} (\mathbf{B}\phi)\ (\mathbf{C}\psi) \, \mathrm{d}\partial\Omega - \int_\Omega \phi \, (\mathbf{L}^*\psi) \, \mathrm{d}\Omega$$

where $\mathbf{L}$ is a linear differential operator and ($\mathbf{B}$,$\mathbf{C}$) are lower order differential operators, resulting from the integration by part, that embed the natural boundary condition, $\delta\mathcal{L}$ becomes

$$\begin{aligned}
\delta\mathcal{L} = {} & \frac{\partial \mathcal{J}}{\partial P}\delta P + \frac{\partial \mathcal{J}}{\partial \alpha}\delta\alpha - \int_\Omega \left(\frac{\partial \mathbf{M}}{\partial U_x}^* \mathbf{X}^*\right)^T \delta U_x \, \mathrm{d}\Omega - \int_\Omega \left(\frac{\partial \mathbf{M}}{\partial U_y}^* \mathbf{X}^*\right)^T \delta U_y \, \mathrm{d}\Omega \\
& - \int_\Omega \left(\frac{\partial \mathbf{M}}{\partial U_z}^* \mathbf{X}^*\right)^T \delta U_z \, \mathrm{d}\Omega - \int_\Omega \left(\frac{\partial \mathbf{M}}{\partial P}^* \mathbf{X}^*\right)^T \delta P \, \mathrm{d}\Omega - \int_\Omega \left(\frac{\partial \mathbf{M}}{\partial k}^* \mathbf{X}^*\right)^T \delta k \, \mathrm{d}\Omega \\
& - \int_\Omega \left(\frac{\partial \mathbf{M}}{\partial \epsilon}^* \mathbf{X}^*\right)^T \delta\epsilon \, \mathrm{d}\Omega - \int_\Omega \left(\frac{\partial \mathbf{M}}{\partial \mu_t}^* \mathbf{X}^*\right)^T \delta\mu_t \, \mathrm{d}\Omega + \int_\Omega (\delta\mathbf{X}^*)^T \mathbf{M}(\mathbf{X}, \alpha) \, \mathrm{d}\Omega \\
& + \int_{\partial\Omega} (\delta\mathbf{X})^T (\mathbf{C}\mathbf{X}^*) \, \mathrm{d}\partial\Omega + \int_{\partial\Omega} (\mathbf{B}\delta\mathbf{X})^T \mathbf{X}^* \, \mathrm{d}\partial\Omega.
\end{aligned}$$

Since the perturbations are arbitrary, setting the first variation of $\mathcal{L}$ with respect to the Lagrangian arguments equal to zero leads to an optimality system. With respect to an arbitrary variation of the adjoint state, we recover the constraint equations; while for an arbitrary variation of the state $\mathbf{X}$ all the terms that include the product of adjoint state to the

tangent linear of the constraint has to vanish. Further, with respect to the set of parameters, vanishing the total variation leads to an optimality condition that enclose the optimality system. Collecting these results yields to

$$\begin{aligned} state\ equations \quad &\Rightarrow \quad \mathbf{M}(\mathbf{X}, \alpha) = 0 \\ adjoint\ equations \quad &\Rightarrow \quad \left(\frac{\partial \mathbf{M}}{\partial \mathbf{X}}\right)^* \mathbf{X}^* = 0 \\ optimality\ condition \quad &\Rightarrow \quad \frac{\partial \mathcal{J}}{\partial \alpha} + \left(\frac{\partial \mathbf{M}}{\partial \alpha}\right)^* \mathbf{X}^* = 0, \end{aligned}$$

where $\left(\frac{\partial \mathbf{M}}{\partial \alpha}\right)^*$ is the adjoint of the model derivative with respect to the parameters. If it is possible to solve this coupled optimality system through one-shot methods, then optimal states and parameters can be obtained without an optimisation iteration. However, due to non linearity and the very large size of this system ($\sim 3 \times size(\mathbf{X})$) one still have to iterate in order to solve the optionality system. Thus, having solved the state equations for $\mathbf{X}$ and then $\mathbf{X}^*$ solution of the adjoint system, model parameters can be iterated by a gradient based optimisation algorithm until optimality condition is satisfied. In a steepest descent algorithm, the parameter is updated at an iteration n according to:

$$\alpha^{n+1} = \alpha^n - \lambda_n \, d_n \tag{6.2}$$

where d_n is the descent direction which is defined recursively by:

$$d_n = \frac{\partial \mathcal{L}}{\partial \alpha} = \left(\frac{\partial \mathbf{M}}{\partial \alpha}\right)^* \mathbf{X}^* + \frac{\partial \mathcal{J}}{\partial \alpha}, \tag{6.3}$$

Concerning adjoint based optimisation methods, we refer the reader to (Gunzburger 2003; Gronskis *et al.* 2013).

REFERENCES

Bryson, Arthur Earl 1975 *Applied optimal control: optimization, estimation and control*. CRC Press.

Cermak, Jack E & Koloseus, Herman J 1954 Development of a miniature air velocity indicator: experimental study to determine sensitivity of new designs. *CER; 54-23* .

Chandramouli, Pranav, Memin, Etienne & Heitz, Dominique 2020 4D large scale variational data assimilation of a turbulent flow with a dynamics error model. *Journal of Computational Physics* p. 109446.

Cochran, Leighton & Derickson, Russ 2011 A physical modeler's view of computational wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics* **99** (4), 139–153.

Duynkerke, PG 1988 Application of the $E-\epsilon$ turbulence closure model to the neutral and stable atmospheric boundary layer. *Journal of the Atmospheric Sciences* **45** (5), 865–880.

Edeling, WN, Cinnella, Pasquale, Dwight, Richard P & Bijl, Hester 2014 Bayesian estimates of parameter variability in the $k-\epsilon$ turbulence model. *Journal of Computational Physics* **258**, 73–94.

EN, NF 2005 1-4: Eurocode 1: Actions sur les structures–partie 1-4: Actions générales–actions du vent.

Errico, Ronald M 1997 What is an adjoint model? *Bulletin of the American Meteorological Society* **78** (11), 2577–2592.

Etling, D, Detering, HW & Theunert, F 1985 On the simulation of wind-driven currents in shallow water. *Archives for meteorology, geophysics, and bioclimatology, Series A* **33** (4), 355–363.

Foures, Dimitry PG, Dovetta, Nicolas, Sipp, Denis & Schmid, Peter J 2014 A data-assimilation method for Reynolds-Averaged Navier-Stokes-driven mean flow reconstruction. *Journal of Fluid Mechanics* **759**, 404–431.

Franceschini, Lucas, Sipp, Denis & Marquet, Olivier 2020 Mean-flow data assimilation based on minimal correction of turbulence models: Application to turbulent high reynolds number backward-facing step. *Phys. Rev. Fluids* **5**, 094603.

Giles, Michael B & Pierce, Niles A 2000 An introduction to the adjoint approach to design. *Flow, Turbulence and Combustion* **65** (3-4), 393–415.

Gronskis, Alejandro, Heitz, Dominique & Mémin, Etienne 2013 Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. *Journal of Computational Physics* **242**, 480–497.

Gunzburger, Max D 2003 *Perspectives in flow control and optimization*, , vol. 5. SIAM.

Hall, Matthew CG & Cacuci, Dan G 1983 Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. *Journal of the Atmospheric Sciences* **40** (10), 2537–2546.

Irwin, Peter A & Kochanski, Will W 1995 Measurement of structural wind loads using the high frequency pressure integration method. In *Restructuring: America and Beyond*, pp. 1631–1634. ASCE.

Jasak, H 1996 Error analysis and estimation in the finite volume method with applications to fluid flows. PhD thesis, Imperial College, University of London. 1028.

Jensen, Martin 1958 The model-law for phenomena in natural wind. *Ingenioren* **2** (2), 121–128.

Kawamura, Takao, Hiwada, Munehiko, Hibino, Toshiharu, Mabuchi, Ikuo & Kumada, Masaya 1984 Flow around a finite circular cylinder on a flat plate: Cylinder height greater than turbulent boundary layer thickness. *Bulletin of JSME* **27** (232), 2142–2151.

Launder, Brian Edward & Sharma, BI 1974 Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc. *Letters in Heat and Mass Transfer* **1** (2), 131–137.

Le Dimet, François-Xavier & Talagrand, Olivier 1986 Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A: Dynamic Meteorology and Oceanography* **38** (2), 97–110.

Li, Yi, Zhang, Jianlei, Dong, Gang & Abdullah, Naseer S 2020 Small-scale reconstruction in three-dimensional Kolmogorov flows using four-dimensional variational data assimilation. *Journal of Fluid Mechanics* **885**.

Li, Zhiyong, Zhang, Huaibao, Bailey, Sean CC, Hoagg, Jesse B & Martin, Alexandre 2017 A data-driven adaptive Reynolds-Averaged Navier–Stokes $k-\omega$ model for turbulent flow. *Journal of Computational Physics* **345**, 111–131.

Margheri, Luca, Meldi, Marcello, Salvetti, MARIA VITTORIA & Sagaut, Pierre 2014 Epistemic uncertainties in rans model free coefficients. *Computers & Fluids* **102**, 315–335.

Meldi, Marcello & Poux, Alexandre 2017 A reduced order model based on Kalman filtering for sequential data assimilation of turbulent flows. *Journal of Computational Physics* **347**, 207–234.

Meroney, Robert N & Derickson, Russ 2014 Virtual reality in wind engineering: the windy world within the computer. *Journal of Wind Engineering and Industrial Aerodynamics* **11** (2), 11–26.

Mons, Vincent, Chassaing, J-C, Gomez, Thomas & Sagaut, Pierre 2016 Reconstruction of unsteady viscous flows using data assimilation schemes. *Journal of Computational Physics* **316**, 255–280.

Mons, V, Margheri, Luca, Chassaing, J-C & Sagaut, Pierre 2017 Data assimilation-based reconstruction of urban pollutant release characteristics. *Journal of Wind Engineering and Industrial Aerodynamics* **169**, 232–250.

Moukalled, Fadl, Mangani, L & Darwish, Marwan 2016 *The finite volume method in computational fluid dynamics*, , vol. 113. Springer.

Murakami, Shuzo 1990 Computational wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics* **36**, 517–538.

Murakami, Shuzo 1997 Current status and future trends in computational wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics* **67**, 3–34.

Othmer, C 2008 A continuous adjoint formulation for the computation of topological and surface sensitivities of ducted flows. *International Journal for Numerical Methods in Fluids* **58** (8), 861–877.

Plessix, R-E 2006 A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International* **167** (2), 495–503.

Pope, Stephen B 2001 Turbulent flows.

Protas, Bartosz, Bewley, Thomas R & Hagen, Greg 2004 A computational framework for the regularization of adjoint analysis in multiscale pde systems. *Journal of Computational Physics* **195** (1), 49–89.

Rhie, CM & Chow, W Li 1983 Numerical study of the turbulent flow past an airfoil with trailing edge separation. *AIAA journal* **21** (11), 1525–1532.

Richards, PJ & Hoxey, RP 1993 Appropriate boundary conditions for computational wind engineering models using the $k-\epsilon$ turbulence model. In *Computational Wind Engineering*, pp. 145–153. Elsevier.

Sheng, R, Perret, L, Calmet, I, Demouge, F & Guilhot, J 2018 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**, 391–403.

Shih, Tsan-Hsing, Liou, William W, Shabbir, Aamir, Yang, Zhigang & Zhu, Jiang 1995 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* (95N11442).

Shirzadi, Mohammadreza, Mirzaei, Parham A & Naghashzadegan, Mohammad 2017 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**, 366–379.

Singh, Anand Pratap & Duraisamy, Karthik 2016 Using field inversion to quantify functional errors in turbulence closures. *Physics of Fluids* **28** (4), 045110.

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

Symon, Sean, Dovetta, Nicolas, McKeon, Beverley J, Sipp, Denis & Schmid, Peter J 2017 Data assimilation of mean velocity from 2D PIV measurements of flow over an idealized airfoil. *Experiments in Fluids* **58** (5), 61.

Tavoularis, S & Karnik, U 1989 Further experiments on the evolution of turbulent stresses and scales in uniformly sheared turbulence. *Journal of Fluid Mechanics* **204**, 457–478.

Tissot, Gilles, Billard, Robin & Gabard, Gwenael 2020 Optimal cavity shape design for acoustic liners using helmholtz equation with visco-thermal losses. *Journal of Computational Physics* **402**, 109048.

Tominaga, Yoshihide, Mochida, Akashi, Yoshie, Ryuichiro, Kataoka, Hiroto, Nozu, Tsuyoshi, Yoshikawa, Masaru & Shirasawa, Taichi 2008 AIJ guidelines for practical applications of cfd to pedestrian wind environment around buildings. *Journal of Wind Engineering and Industrial Aerodynamics* **96** (10-11), 1749–1761.

Tominaga, Yoshihide & Stathopoulos, Ted 2010 Numerical simulation of dispersion around an isolated cubic building: model evaluation of RANS and LES. *Building and Environment* **45** (10), 2231–2239.

Tominaga, Yoshihide & Stathopoulos, Ted 2017 Steady and unsteady RANS simulations of pollutant dispersion around isolated cubical buildings: Effect of large-scale fluctuations on the concentration field. *Journal of Wind Engineering and Industrial Aerodynamics* **165**, 23–33.

Tschanz, T & Davenport, Alan Garnett 1983 The base balance technique for the determination of dynamic wind loads. *Journal of Wind Engineering and Industrial Aerodynamics* **13** (1-3), 429–439.

Tsuchiya, M, Murakami, S, Mochida, A, Kondo, K & Ishida, Y 1997 Development of a new $k-\epsilon$ model for flow and pressure fields around bluff body. *Journal of Wind Engineering and Industrial Aerodynamics* **67**, 169–182.

Versteeg, Henk Kaarle & Malalasekera, Weeratunge 2007 *An introduction to computational fluid dynamics: the finite volume method*. Pearson education.

Wang, HF & Zhou, Yu 2009 The finite-length square cylinder near wake. *Journal of Fluid Mechanics* **638**, 453–490.

Yakhot, VSASTBCG, Orszag, SA, Thangam, Siva, Gatski, TB & Speziale, CG 1992 Development of turbulence models for shear flows by a double expansion technique. *Physics of Fluids A: Fluid Dynamics* **4** (7), 1510–1520.

Yang, Yin, Robinson, Cordelia, Heitz, Dominique & Mémin, Etienne 2015 Enhanced ensemble-based 4d-var scheme for data assimilation. *Computers & Fluids* **115**, 201–210.

Yoshie, R, Mochida, A, Tominaga, Y, Kataoka, H, Harimoto, K, Nozu, T & Shirasawa, T 2007 Cooperative project for CFD prediction of pedestrian wind environment in the architectural institute of Japan. *Journal of Wind Engineering and Industrial Aerodynamics* **95** (9-11), 1551–1578.

Zymaris, AS, Papadimitriou, DI, Giannakoglou, Kyriakos C & Othmer, C 2010 Adjoint wall functions: A new concept for use in aerodynamic shape optimization. *Journal of Computational Physics* **229** (13), 5228–5245.