To cite this version :

Tom MOUSSIE, Paolo ERRANTE, Marcello MELDI - Statistical Inference of Upstream Turbulence Intensity for the Flow Around a Bluff Body with Massive Separation - Flow, Turbulence and Combustion - Vol. 113, p.853-889 - 2024

Flow, Turbulence and Combustion (2024) 113:853–889
https://doi.org/10.1007/s10494-024-00573-z

RESEARCH

Statistical Inference of Upstream Turbulence Intensity for the Flow Around a Bluff Body with Massive Separation

Tom Moussie[1] · **Paolo Errante**[1,2] · **Marcello Meldi**[1]

Received: 29 May 2024 / Accepted: 18 July 2024 / Published online: 10 August 2024

Abstract
The Benchmarck on the Aerodynamics of a Rectangular 5:1 Cylinder is studied using a data-driven technique which bridges numerical simulation and available experimental results. Because of intrinsic features of the tools used for investigation, in particular in terms of set-up and boundary conditions, significant discrepancies have been observed in the literature when comparing experimental and numerical results. An approach based on the Ensemble Kalman Filter is here used to optimize a synthetic turbulent inlet used as boundary condition in the numerical calculation, in order to reduce the discrepancy with the available experiments. The data-driven method successfully optimizes the boundary condition features, which produce a significant improvement of the accuracy in the prediction of the flow. These findings open perspectives of application towards the analysis of realistic cases, where boundary conditions are complex and usually unknown.

Keywords BARC · DDES · EnKF · Syntethic turbulence inlet

1 Introduction

Among the open challenges in wind engineering applications for urban settings, the accurate prediction of unsteady features and higher order statistical moments of the flow field represents a key element for technological advancement. Detailed information about the instantaneous organization of the flow can be essential to predict and control

---

Tom Moussie, Paolo Errante and Marcello Meldi have contributed equally to this work.

✉ Tom Moussie
tom.moussie@ensam.eu

Paolo Errante
paolo.errante@ensam.eu

Marcello Meldi
marcello.meldi@ensam.eu

[1] CNRS, ONERA, Arts et Métiers ParisTech, Centrale Lille, UMR 9014- LMFL- Laboratoire de Mécanique des fluides de Lille, Univ. Lille, Kampé de Feriet, 59000 Lille, France

[2] L2EP, Laboratoire d'Electrotechnique et d'Electronique de Puissance de Lille, Avenue Henri Poincaré, Bâtiment ESPRIT, 59655 Lille, France

the emergence of extreme events. In addition, the accurate representation of instantaneous features of the flow is a key element to obtain precise estimation of the higher-order statistical moments of the physical phenomena at play. Such statistical moments, which are not well described by stationary closures such as Reynolds-Averaged Navier–Stokes (RANS) models (Pope 2000; Wilcox 2006), play an important role in the mechanical stress of urban structures and have to be taken into account in the concept and design phases of production. For example, the second order moment (variance) of the pressure field is essential to measure the surface stress affecting buildings (Charisi et al. 2019) and is can also be tied to dangerous aeroelastic phenomena which can affect slender structures (Matsumoto and Tamura 1993). Experiments in wind tunnels can measure such instantaneous features of the flow. However, the positioning of velocity sensors, such as hot wires, and pressure taps can be precluded in sensitive physical region of the model and systematic usage in realistic applications is not realistically attainable. On the other hand, the development of new computational architectures and the availability of the resources of supercomputing centers provide reliable tools to investigate such problems using numerical strategies based on Computational Fluid Dynamics (CFD). In particular, scale-resolving simulations able to capture the large scales of the flow, such as wall-modeled Large Eddy Simulation (LES) (Pope 2000; Sagaut 2005) or hybrid RANS–LES (Pope 2000; Sagaut 2005; Shur et al. 2008), show promising features for the representation of unstationary flows. In fact, these techniques can naturally provide a complete volume description of the flow, unlike most experimental techniques. In addition, they capture the unstationary, three-dimensional features of High-Reynolds regimes while requiring reduced computational resources when compared with Direct Numerical Simulation (DNS). The main drawback of such reduced-order techniques is the need to introduce turbulence models/wall functions in the discretized numerical problem. These models are extremely sensitive to their parametric description and can introduce a bias in the numerical results.

While both experiments and numerical approaches show advantageous features and drawbacks for the analysis of unsteady flows, both families of tools have to face an additional challenge to provide an accurate prediction. This difficulty lies in the lack of knowledge about initial and boundary conditions. High-Reynolds regimes such as the ones observed in urban settings are extremely sensitive to minimal changes in upstream conditions, which can govern the instantaneous evolution and be responsible for the emergence of extreme events. Such realistic features are difficult to be replicated in an isolated environment such as a wind tunnel for experimental analysis. Numerical simulation has the potential to take into account these upstream flow details, even if non-linear interactions of error sources associated with discretization, modelling and boundary conditions must be well assessed to obtain an accurate prediction. This last difficulty is critical for scale-resolving methods and in particular for LES. Therefore, even in the case an exact match of inlet conditions can be established between twin experiments and numerical simulations, results in terms of bulk flow quantities and statistics of the flow field may exhibit large discrepancies due to the aforementioned problems.

In the present work, the capabilities of scale resolving numerical methods in terms of replicating experimental results available are investigated for the analysis of a well known flow configuration, the Benchmark on the Aerodynamics of a Rectangular 5:1 Cylinder (BARC) (Bruno et al. 2014) for $Re = 2 \times 10^5$. This test case is challenging to be analyzed via scale resolving CFD, because of the interaction of numerous physical aspects that must be captured and are which difficult to model. For this reason, a high sensitivity of such models to variations in the numerical set-up is observed (Mariotti et al. 2017; Rocchio et al. 2020). Within this framework, the hybrid method known as Delayed Detached Eddy

Simulation (DDES) (Gritskevich et al. 2012) will be improved to reproduce realistic upstream turbulent flow conditions (Mannini et al. 2017; Pasqualetto et al. 2021). To do so, a data-driven approach will be used to calibrate CFD inlet conditions to better match experimental data produced by CSTB in Nantes (Haffner et al. 2023). More generally, the objective of the present work is to investigate whether data-driven model can efficiently optimize the parametric description of CFD, creating a strong bridge between experimental and numerical approaches. This analysis aims to contribute to the identification of needed guidelines to move forward from simple comparison between experiments and numerics towards a better understanding of how these approaches can be synergistically integrated. While the present work focuses on inlet conditions, it must be considered that similar analyses can be replicated for different parametric descriptions of the CFD model. This analysis shows complementarity with the work of Lamberti et al. (2018), which used a gradient-based method to optimize the performance of a synthetic turbulent inlet for the analysis of an atmospheric boundary layer. The challenging aspect here is to obtain a suitable optimization when limited data is provided and when the flow configuration exhibit complexities such as flow separation, recirculation regions and turbulent wakes. To this purpose, the optimization problem will be performed using the Data Assimilation technique known as Ensemble Kalman Filter (EnKF) (Bocquet et al. 2016). This tool efficiently complements non-linear numerical models, such as CFD solvers, for the analysis of complex flows. In addition, it can handle the optimization of relatively large set of parameters, which can demand prohibitive computational requirements for techniques such as the generalized Polynomyal Chaos (gPC), which were previously used for stochastic analyses for the BARC (Mariotti et al. 2017).

The manuscript is organized as follows. In Sect. 2, an extensive discussion of the test case is performed and the experimental data available for the present analysis is described. In Sect. 3, the CFD solvers used in the present analysis are described and the numerical set-up of the test case is discussed. In Sect. 4, the data-driven technique chosen for this analysis, the EnKF, is detailed. In Sect. 5, the sensitivity of the numerical simulations to different parameters, including the behavior of the inlet condition, is investigated. In Sect. 6, the data-driven technique is used to optimize the inlet boundary conditions. The results are compared with preliminary CFD runs and the experiments available. In Sect. 7, conclusions are drawn and future perspectives are discussed.

2 Test Case of Investigation: The Benchmark on the Aerodynamics of a Rectangular Cylinder (BARC)

The Benchmark on the Aerodynamics of a Rectangular Cylinder with chord-to-depth ratio equal to 5 (BARC) (Bruno et al. 2014) which started officially in 2008, has the goal to analyze the aerodynamics of a rectangular bluff body. As mentioned by Bartoli et al. (2009), the benchmark scope is to contribute to the analysis of separated and turbulent flow around a fixed rectangular cylinder, to provide useful information for civil engineering applications, since this kind of geometry is often encountered in urban areas (e.g. long span bridges decks or high-rise buildings). Since then, this configuration has been extensively studied both experimentally and numerically (Rocchio et al. 2020; Le et al. 2009; Cimarelli et al. 2018). A qualitative flow representation and a sketch of the configuration are given in Fig. 1 with the computational domain used in this study.

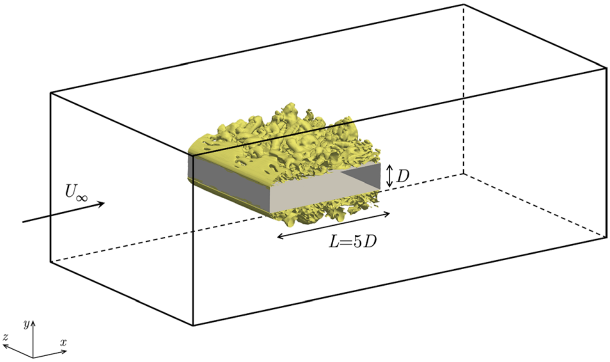


Fig. 1 Qualitative representation of the flow around the rectangular cylinder for the simulation DDES-I0-G2. The λ_2 criterion is shown to highlight the unstationary features of the flow around the immersed body

Despite its simple appearance, this configuration involves complex physical processes such as boundary layer separation, flow reattachment, recirculation zones, and von Kármán streets, which affect the propagation of acoustic waves and the flow's structural organization, as shown for instance in Le et al. (2009) and Ricciardelli and Marra (2008). Therefore, the complete representation of these mechanisms and their interactions is crucial to obtain an accurate prediction for similar cases of industrial and urban interests. One major challenge in studying this flow configuration via numerical methods is its sensitivity to the geometric characteristics and boundary conditions, particularly the velocity field imposed at the inlet. This is especially true for scale-resolving approaches like LES and hybrid RANS–LES, which are now widely used for analyzing bluff bodies (Minguez et al. 2008; Hesse and Morgans 2021; Haffner et al. 2022).

2.1 Review of the Analyses in the Literature

Numerical analyses of the BARC have mainly focused on the analysis of the sensitivity of this test case to variations in the set-up of the problem, in particular for geometric characteristics and features of turbulence modeling (Bruno et al. 2014). For the latter, unsteady RANS (URANS), LES and hybrid turbulence models have been tested by several authors. DNS for relatively low Reynolds numbers have also been proposed in the literature (Cimarelli et al. 2018; Chiarini and Quadrio 2021) which exhibit significant discrepancies in the statistical quantities investigated, confirming the difficulty for numerical approaches to capture the leading physical dynamics for this test case. URANS studies, such as the one by Mannini et al. (2010), indicate a strong sensitivity to the choice of the turbulence model. For LES, Timilsina (2015) showed that subgrid-scale (SGS) modeling can achieve satisfying predictions on drag coefficients and with a relatively small impact attributed to the closure used, while Rocchio et al. (2020) find that variations in the parametric choices for the SGS model can be dominant. Mariotti et al. (2017) performed a stochastic analysis of LES predictions of the flow around the BARC. The analysis is repeated for two different grid resolutions, which differ in terms of streamwise and spanwise resolution. The results indicate that for both numerical runs the most sensitive quantities are those which show the largest dispersion among the different BARC studies in the literature, such as the Strouhal number calculated with the lift coefficient and the pressure distribution over the lateral faces of the cylinder. Also, results obtained with the two grids exhibit remarkable differences, for the recirculation region for instance. Hybrid approaches have been also used by Mannini et al. (2011)

which highlight the importance of numerical dissipation introduced by the numerical schemes. It has been concluded that an excessive amount of dissipation damps out the turbulent structures directly resolved by the grid used, which are essential to capture the emergence and interactions of the numerous physical aspects which characterize this flow.

The analyses in the literature dealing with the sensitivity of the flow to geometric variations have mainly focused on two aspects. The first one deals with the alignment of the rectangular cylinder to the direction of the upstream flow. Patruno et al. (2016) investigated the statistical features of the flow with variations in the angle of attack. The analysis included large variations of the angle as well as minor variations which could be associated with uncertainty in the set-up of experiments. The second main topic of investigation deals with the shape of the edges of the BARC, which traditionally are considered to be sharp. The study by Rocchio et al. (2020), which consisted of a comparison of LES with and without edge rounding on the leading edge, indicated a significant discrepancy of the results. In particular, the length of the recirculation region is strictly connected to this geometric feature (Chiarini and Quadrio 2022; Mariotti et al. 2024).

The previous discussion highlights the very high sensitivity of the BARC test case to several parameters governing the set-up. This parameters can be optimized via data-driven strategies, which are now extensively used in fluid mechanics applications (Brunton et al. 2020). One recent example of application for the BARC is the usage of multigrid sequential data assimilation, which have been used to calibrate SGS models for LES by Moldovan et al. (2022). The results presented for the statistical moments of the velocity and pressure flow fields, which were obtained for the BARC with $Re = 4 \times 10^4$ show that data assimilation techniques based on the Ensemble Kalman Filter are able to improve the predictive capabilities of the CFD solver for reduced grid resolution. In addition, it has been observed that, despite the sparse and asymmetric distribution of observation adopted in the data-driven process, the data augmented results exhibit symmetric statistics and improved accuracy far from the sensor location.

Experimental studies in wind tunnels for the BARC (Bruno et al. 2014; Le et al. 2009; Ricciardelli and Marra 2008; Haan and Kareem 2009) also investigate the sensitivity of the flow to several parameters such as the Reynolds numbers, the shape/inclination of the rectangular cylinder and the turbulence intensity of the flow upstream. While important variations of the physical quantities are observed in experiments as well, the discrepancy with numerical simulation is important and significantly larger than the uncertainty in the set-up (Bruno et al. 2014). Therefore, comparison of results from different tools does not shed a light on the governing mechanisms driving interactions between the instantaneous/statistical features of the flow and the aerodynamic forces at play. A key question about the sensitivity of the BARC to variations in the set-up is associated with the upstream conditions of the flow. Experiments can naturally take into account this aspect, which is usually measured with the empty wind tunnel. However, once the rectangular cylinder is installed in the test tunnel, the characteristics of the flow are affected by its presence. For numerical simulations the turbulence intensity can be exactly imposed at the inlet, but it is usually difficult to be precisely estimated. Mariotti et al. (2016) investigated the sensitivity of the results to variations in inlet conditions for RANS models. While their results indicate that the sensitivity was weak, the scope of their analysis was limited to RANS modeling and the statistical physical quantities associated. It has been shown (Noda and Nakayama 2003) that inlet conditions for scale resolving simulations, which can account for instantaneous features of the flow, have the potential to govern the organization of the flow and to provide a more reliable picture of the sensitivity of the BARC to such aspect.

2.2 Reference Experiments for the Present Work

The *Centre Scientifique et Technique du Bâtiment* (CSTB) in Nantes, France, has recently open a new benchmarch for the BARC (Haffner et al. 2023). The main difference is the Reynolds number of investigation, which is higher than the previous studies summarized by Bruno et al. (2014). The researchers in CSTB have performed an experimental campaign in their *Jules Verne* wind tunnel, shown in Fig. 2, analyzing the flow features with the following variations of parameters:

- Reynolds number 2×10^5, 3.3×10^5
- Three configurations of the front edges: straight, chamfered and curved.
- Angle of attack of the rectangular cylinder: 0, ± 2, ± 5 degrees
- Turbulence intensity for the upstream flow $\approx 1\,\%$, $\approx 3.5\,\%$

The configuration considered for the present analysis is the one for $Re = 2 \times 10^5$, sharp front edges, 0° angle of attack and $I_T \approx 3.5\,\%$ inlet turbulence intensity. In this case, the rectangular cylinder with height $D = 0.2$ m and length $L = 5D$, is positioned in the middle of the wind tunnel whose cross section is $5\,\text{m} \times 6\,\text{m}$ and its total length is 12 m. The mean flow is aligned with the streamwise direction x and it is described by a bulk velocity of $U_\infty = 15\,\text{m} \cdot \text{s}^{-1}$. Measurements for the velocity and the pressure fields are performed.

For the latter, 150 sensors on the cylinder's surface are used to obtain time-resolved pressure samples at a frequency of 200 Hz. The velocity field is investigated using a large-scale Particle Tracking Velocimetry (PTV) on a plane which is normal to the spanwise direction and of size $\Delta_x \times \Delta_y = (2\,L \times L)$, in order to obtain time-resolved measurement of the velocity components u_x and u_y. The frequency of acquisition of the PTV is equal to 6.6 kHz. The size of the first plane presented is large enough to sample the flow under the rectangular cylinder as well as part of the wake. In addition to the measurements presented for the flow features, the time history of the lift and drag coefficients is also provided.

3 Numerical Tools

The numerical algorithms used in the present work are now introduced.

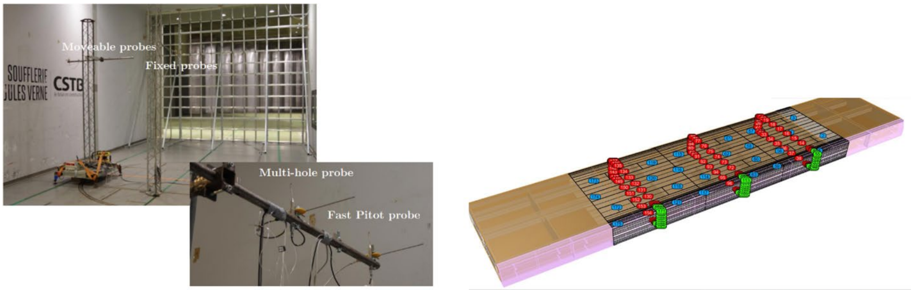


Fig. 2 Experimental mock up, wind tunnel and sensor positioning for the CSTB campaign (Haffner et al. 2023)

3.1 Dynamic Equations and Numerical Solvers

The numerical solvers used in this work rely on Finite-Volume discretization of the Navier–Stokes equations for incompressible flows and Newtonian fluids. As previously discussed in the Introduction, the analyses will be performed using coarse grained approaches based on turbulence/subgrid-scale modeling. Therefore, considering a reduced-order operator $\tilde{\cdot}$,which could represent a RANS average or an LES filtering, the dynamic equations can be written as:

$$\nabla \cdot \tilde{\mathbf{u}} = 0 \tag{1}$$

$$\frac{\partial \tilde{\mathbf{u}}}{\partial t} + (\tilde{\mathbf{u}} \cdot \nabla)\tilde{\mathbf{u}} = -\nabla \tilde{p} + \nu \nabla^2 \tilde{\mathbf{u}} - \nabla \cdot \boldsymbol{\tau}_t \tag{2}$$

where $\mathbf{u}$ is the velocity field, p is the pressure (normalized over the density ρ), ν is the kinematic viscosity and $\boldsymbol{\tau}_t$ is the tensor representing the effects of the turbulence closure. This term would be Reynolds stress tensor in the case of RANS approaches, while it would be the subgrid stress tensor for LES. For eddy-viscosity models (Pope 2000; Sagaut 2005), the components of this tensor are determined by the spatial gradients of the resolved velocity field and by the quantity ν_t, which is referred to as turbulent viscosity. Relying on the Boussinesq approximation, it is used to take into account the statistical moments of stresses due to turbulence as a diffusive effect. In this case, the components of the tensor $\boldsymbol{\tau}_t$ are:

$$\tau_{t,ij} = -2\nu_t \left(\tilde{S}_{ij} - \frac{1}{3}\mathcal{K}\delta_{ij} \right) = -\nu_t \left(\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i} \right) + \frac{2}{3}\mathcal{K}\delta_{ij} \tag{3}$$

where $\tilde{S}_{ij}$ are the components of the resolved rate of strain tensor, $\mathcal{K} = 0.5\left(\tau_{t,xx} + \tau_{t,yy} + \tau_{t,zz}\right)$ is the turbulent kinetic energy and δ_{ij} is the Kroneker symbol. Several proposals in the literature for eddy viscosity models provide different expressions for the quantity ν_t. In the context of this study, the hybrid RANS–LES method known as Delayed Detached Eddy Simulations (DDES) will be used (Spalart et al. 2006). This model, which behaves like a RANS closure in the proximity of the wall and transitions to an LES behavior moving in free stream turbulent regions, does not require the refinement of wall resolved LES and DNS at the body surface. Therefore, it has been selected because it can capture the three dimensional unsteady features of High-Reynolds flows with reasonable computational costs. A detailed description of the DDES model used in this work is reported in Appendix A.

Information about the CFD code is now provided. The simulations are performed using the C++ open-source framework OpenFOAM (Greenshields 2021; Tabor and Baba-Ahmadi 2010; Meldi et al. 2012; Selma et al. 2014; Constant et al. 2017). In this work, two flow solvers have been used, which are based on the Simple (Caretto et al. 1973) and Pimple (Caretto et al. 1973; Issa et al. 1986) algorithms. Both solvers rely on a recursive procedure, where the velocity field and the pressure gradient are iteratively updated until convergence. The solver SimpleFOAM is used in this analysis for stationary simulations, while the PimpleFOAM solver is used for unstationary calculations. For the unstationary simulations, the time discretization relies on a second-order backward scheme. The time step is adaptively tuned to obtain a maximum local Courant number equal to 1. The resulting value is $\Delta t \approx 5 \times 10^{-3}\, t_A$, where $t_A = D/U_\infty$ is the advection time scale. The time step exhibits very small variations in time during a full simulation. The space discretization is performed using second order schemes. In particular, the advective term is discretized using a native LUST scheme, which combines with a ratio 75–25% a second-order centered scheme and a second-order upwind scheme. For the stationary RANS simulation, a first-order upwind scheme is used for the advective term.

3.2 Numerical Set-up of the Test Case

The physical domain of investigation and the boundary conditions applied are shown in Fig. 3.

The axis are set so that x represents the streamwise direction, y is the normal direction and z is the spanwise direction. The resolution of the dynamic system is performed normalizing the physical quantities over the upstream mean velocity U_∞ and the height of the rectangular cylinder D. Therefore, the kinematic viscosity for the calculations is set to $\nu = Re^{-1} = \left(2 \times 10^5\right)^{-1}$. The size of the physical domain is now discussed. For the wind tunnel used at CSTB, the grid is located $27.5\,D$ upstream of the center of the rectangular cylinder. In the proximity of the grid the flow is still affected by lingering shear production mechanisms, which are not efficiently taken into account by the inlet boundary condition used in this analysis. This effect becomes negligible from a distance of $10\,D$ up to $20\,D$ downstream the grid, depending on its features (Valente and Vassilicos 2011). For all of these reasons, we decided to fix the upstream distance of the inlet to $7.5\,D$. Preliminary tests using RANS with a steady inlet have shown negligible differences for simulations using an upstream distance in the range $[7.5\,D,\ 17.5\,D]$, assessing the very small variability of the algorithm for this range. Periodic conditions are imposed on the lateral faces in the $x-y$ planes, while freestream boundary conditions are applied at the top and bottom faces on $x-z$ planes. A boundary condition conserving the mass flow rate is applied at the outlet. The parameters of this specific boundary condition, defined as *inletOutlet* in OpenFOAM, have been set so that reverse flow is excluded. For the inlet, two proposals have been considered. The first one is a classical set-up for which the velocity field is uniform and equal to U_∞ while a zero-gradient condition is used for the pressure. The second inlet condition used in the present analysis uses a synthetic turbulence generator based on the works by Poletto et al. (2013) and Shur et al. (2018). This condition is already implemented in OpenFOAM and determines the inlet velocity $\tilde{\mathbf{u}}$ as the sum of the bulk velocity $\mathbf{u}_b = [U_\infty, 0, 0]$ and a fluctuating velocity $\mathbf{u}'$. The methodology is based on the concept of synthetic eddies, which are tied with prescribed velocity fluctuations and are injected on the inlet plane. The eddies are defined by their center and a description of

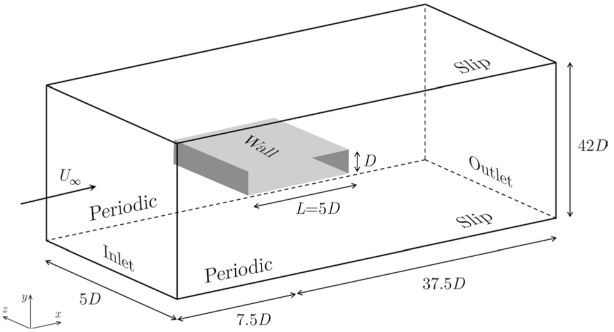


Fig. 3 Set-up and boundary conditions for the BARC test case

the velocity fluctuations distribution around them. These eddies are randomly generated at the inlet and advected in the physical domain. This process of random generation is constrained by physical features of the velocity field that can be imposed by the user, such as the components of the Reynolds stress tensor and the integral length scale ℓ. From Shur et al. (2018), the fluctuating velocity field is computed as follows:

$$\mathbf{u}'(\mathbf{x}) = C_1 \sum_{k=1}^{N} \mathbf{q} \circ \left(\mathbf{r}^k \times \boldsymbol{\alpha}^k \right) \tag{4}$$

where $\mathbf{r}^k = \frac{x - x^k}{\sigma^k}$ is the normalized distance from the eddy center. $\boldsymbol{x}$ is a position vector and $\boldsymbol{x}^k$ is the center of the k-th eddy. In addition, the vectors $\boldsymbol{\sigma}^k$ and $\boldsymbol{\alpha}^k$ represents respectively the length scales and the intensity of the k-th eddy in each direction. $\mathbf{q}$ is a shape function based on the distance from the eddy centre, whose formulation from Poletto et al. (2013) ensure divergence-free velocity field. The components of the eddy intensity $\boldsymbol{\alpha}$ and length scale $\boldsymbol{\sigma}$ are given as following:

$$\alpha_i^2 = \gamma \frac{\sum_j \left(\lambda_j / \sigma_j^2 \right) - 2\lambda_i / \sigma_i^2}{2C_2}; \qquad \sigma_i = \max\left(\delta, c\ell, \max\left(\Delta x, \Delta y, \Delta z\right)\right) \tag{5}$$

where γ is a random integer which can take the value of -1 or 1 with equal probabilities. The eddy intensity is based on λ, namely the eigenvalues of the user-provided Reynolds stress tensor. It is worth to notice that formulation of α_i is similar to an ellipsoid having the components of the length scale σ_i as semi-axis, and oriented along the local principal reference system to reproduce anisotropy. C_2 is a normalization constant that takes into account the magnitude of the shape function integral, whose values are suggested by Shur et al. (2018). Assessing the eddy length scale is not a trivial task. Synthetic Turbulence Generation (STG) involves a predefined model energy spectrum which comes from the definition of σ_i. In the present case, according to Eq. (5), one has to provide a reasonable estimation on the integral length scale ℓ and the domain characteristic length scale δ. ℓ is balanced by a constant value c, either the von Kármán or an user-defined constant.

In the context of hybrid RANS–LES modeling, an algebraic expression is often derived to approximate this length scale (Guo et al. 2023), although it may not adequately consider the impact from the flow history and boundary information without explicitly modeling the transport equation associated with this length scale. This deficiency can potentially influence the process of emulating synthetic turbulence to match realistic experimental conditions. In the next section, a Data-Assimilation approach is described with the objective to overcome this scarcity of details.

4 Data Assimilation - Ensemble Kalman Filter

The technique of Data Assimilation (DA) that will be used to infer the parametric behavior of the inlet is now introduced. The Kalman Filter (KF), proposed by Kalman (1960), is a sequential DA tool used to obtain an estimation of the physical variables and/or a set of parameters of a system at a given time. The estimation relies on multiple sources of information, which are characterized by a level of uncertainty supposedly known. The classical version of the KF relies on a set of observations $\boldsymbol{y}$ and a prior state $\boldsymbol{x}$ which is obtained via a linear model $\mathbf{M}$. The updated (*augmented*) state is obtained in an analysis step, where the discrepancy between model prediction and available observation is used to update the former:

$$\boldsymbol{x}^a = \boldsymbol{x}^f + \boldsymbol{K}\big(\boldsymbol{y} - \mathcal{H}(\boldsymbol{x}^f)\big) \tag{6}$$

Here the suffix f refers to the model forecast and the suffix a to the state obtained after the analysis phase. The projection operator $\mathcal{H}$ maps the prediction of the model over the space where the observation is sampled. Usually, $\mathcal{H}$ performs as an interpolator, providing the model solution in the location of the sensors. The update of the physical state is governed by the so called Kalman gain matrix $\boldsymbol{K}$, which is obtained via manipulation of the error covariance matrix $\boldsymbol{P} = \mathbb{E}\big((\boldsymbol{x} - \mathbb{E}(\boldsymbol{x}))(\boldsymbol{x} - \mathbb{E}(\boldsymbol{x}))^\top\big)$. The classical KF formulation is not designed to be used in CFD applications. Navier–Stokes equations include non-linear terms, therefore, the derivation of a linear model $\mathbf{M}$ may imply a significant loss of accuracy. Secondly, the size of the matrix $\boldsymbol{P}$ is proportional to the number of degrees of freedom of the model n i.e. number of mesh elements times number of physical variables considered for CFD. The extensive manipulation of $\boldsymbol{P}$ required to obtain $\boldsymbol{K}$, including a matrix inversion, would demand prohibitive computational resources for realistic CFD applications.

The two critical issues previously mentioned can be mitigated using the Ensemble Kalman Filter (EnKF) first proposed by Evensen (1994). Within this context, an ensemble of n_e prior states is advanced in time between consecutive analysis phases using a model $\mathcal{M}$ which can be non-linear. Let us assume to consider an analysis phase at the time step k.

A state matrix $\boldsymbol{X}_k \in \mathbb{R}^{n \times n_e}$ is constructed, where each column i corresponds to the physical state of an ensemble member $\boldsymbol{x}_i^f \in \mathbb{R}^n$. With the EnKF, the covariance matrix $\boldsymbol{P}$ is not advanced in time anymore and is obtained via a Monte-Carlo approximation, thus:

$$\boldsymbol{P}^f = \boldsymbol{\Gamma}_k^f \left(\boldsymbol{\Gamma}_k^f\right)^\top \tag{7}$$

where $\boldsymbol{\Gamma}_k^f \in \mathbb{R}^{n \times n_e}$ is the state anomaly matrix representing the normalized deviation of the state vectors from their ensemble means. The i^{th} column of the state anomaly matrix is obtained as:

$$\boldsymbol{\Gamma}_{i,k}^f = \frac{\boldsymbol{x}_{i,k}^f - \overline{\boldsymbol{x}}_k^f}{\sqrt{n_e - 1}}, \qquad \overline{\boldsymbol{x}}_k^f = \frac{\sum_{i=1}^{n_e} \boldsymbol{x}_{i,k}^f}{n_e} \tag{8}$$

To obtain a well-posed mathematical and numerical problem, an ensemble of n_o observations is obtained through the perturbation of the observation vector available at the time step k, $\boldsymbol{y}_k \in \mathbb{R}^{n_o}$. The result of this perturbation is an observation matrix $\boldsymbol{Y}_k \in \mathbb{R}^{n_o \times n_e}$. The n_e columns of the observation matrix are obtained with $\boldsymbol{y}_{i,k} = \boldsymbol{y}_k + \boldsymbol{\epsilon}_i$ for $i \in [1, n_e]$. The added random noise $\boldsymbol{\epsilon}_i$ is described as a Gaussian probability function $\boldsymbol{\epsilon}_i \sim \mathcal{N}\big(0, \varsigma_k\big)$, where $\varsigma_k \in \mathbb{R}^{n_o \times n_o}$ is the observation covariance matrix. It is worth to underline that ς_k should be constructed from the experimental uncertainties, hence, the use of Gaussian noise to reproduce experimental error is one of the assumption made due to scarcity of information. However, when this information is not available, the usage of a Gaussian perturbation is helpful to obtain a robust performance of the global algorithm (Bocquet et al. 2016). Analogously to what is done for $\boldsymbol{\Gamma}_k$, each column i of the anomaly matrix $\boldsymbol{\Lambda}_k \in \mathbb{R}^{n_o \times n_e}$ is computed. This matrix also takes into account the discrepancy between the model results and their ensemble average. However, it is defined on the solution space of the observation, and it relies on the projection operator $\mathcal{H}$:

$$\boldsymbol{\Lambda}^f_{i,k} = \frac{\mathcal{H}(\boldsymbol{x}^f_{i,k}) - \overline{\mathcal{H}(\boldsymbol{x}^f_k)}}{\sqrt{n_e - 1}}, \quad \overline{\mathcal{H}(\boldsymbol{x}^f_k)} = \frac{\sum_{i=1}^{n_e} \mathcal{H}(\boldsymbol{x}^f_{i,k})}{n_e} \tag{9}$$

The Kalman gain matrix $\boldsymbol{K}_k$, which measures the correlations between the observations and the state vector, is obtained as follows:

$$\boldsymbol{K}_k = \boldsymbol{\Gamma}^f_k \left(\boldsymbol{\Lambda}^f_k\right)^\top \left[\boldsymbol{\Lambda}^f_k (\boldsymbol{\Lambda}_k)^\top + \varsigma_k\right]^{-1} \tag{10}$$

Finally, the physical state for each ensemble member i is obtained updating the forecast results with a correction term controlled by the Kalman gain:

$$\boldsymbol{x}^a_{i,k} = \boldsymbol{x}^f_{i,k} + \boldsymbol{K}_k \left(\boldsymbol{y}_{i,k} - \mathcal{H}(\boldsymbol{x}^f_{i,k})\right) \tag{11}$$

The Kalman Filter can be used to augment the state prediction obtained via the model as well as to optimize its free coefficients. In this way, the discrepancy between the model prediction and the observations is naturally reduced. Several strategies are presented in the literature to this purpose, and the model chosen for the present analysis is the *extended state* (Bocquet et al. 2016). In this model, the free parameters of the model are organized in an array $\boldsymbol{\theta}$, which is combined with the state $\boldsymbol{x}$ to obtain an extended state $\boldsymbol{x}^\star$:

$$\boldsymbol{x}^\star = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\theta} \end{bmatrix} \tag{12}$$

The EnKF is then resolved for the extended state $\boldsymbol{x}^\star$, so that the parameters of the model are updated with the forecast of the solution. The steps of the stochastic EnKF used in the present analysis are reported in Algorithm 4.1. One of the critical points which determines the efficacy of the EnKF is the careful calibration of the uncertainty in the model and observation. The procedure targets a state that is admissible for both the model and the observation, therefore the final solution usually converges to a configuration in the intersection between all the potential solutions of the model and the range of variability of the observation. For the latter, a global low confidence (and therefore very high variability) may result in a final state very close to the model prediction, therefore with a limited increase in accuracy. Optimized values for these uncertainties has been extensively studied in the literature (see the review article by Tandeo et al. (2020) for a comprehensive discussion). The uncertainty in the observation is usually calibrated using information about the analysis techniques used and the acquisition system. For PTV, an estimate of 5 % error is reasonable, and therefore this quantity has been used to set the uncertainty for the observation. For deterministic models such as CFD, one way to augment the uncertainty is to play with the variability of the parameters $\boldsymbol{\theta}$, which in this case correspond to the free coefficients driving the synthetic turbulence inlet. The choices performed for this work are detailed in Sect. 6.

Despite its proven efficiency in various domains of application (Asch et al. 2016), the EnKF relies on a relatively large ensemble of realizations, which can be computationally expensive for very high-dimensional systems. Reducing the ensemble size can lead to sampling errors and inaccurate estimate of the state variance, potentially resulting in a lack of convergence of the algorithm. In addition, the performance of the EnKF is driven by many hyperparameters that must be carefully selected. Erroneous choices can lead the parametric inference procedure towards local optimized states, providing a sub-optimal result. Nonetheless, due to its robustness and capability to investigate non-linear physics, a strategy based on the EnKF is the best DA tool to perform the present analysis in the Authors' opinion.

Algorithm 1 Algorithm for the stochastic Ensemble Kalman Filter

```
Input: $\mathcal{M}$, $\mathcal{H}$, $\varsigma_{k+1}$, and prior states $\boldsymbol{x}^a_{i,0}$ for each ensemble member $i$, where
    $\boldsymbol{x}^a_{i,0} \sim \mathcal{N}\left(\boldsymbol{\mu}_x, \varsigma^2_x\right)$.
for $k = 0$ to $K-1$ do
    Advancement in time of the state vectors between two analysis phases for
      each ensemble member $i$:
      $\boldsymbol{x}^f_{i,k} = \mathcal{M}(\boldsymbol{x}_{i,k})$
    Creation of an observation matrix perturbing the observation with a
      Gaussian noise:
      $\boldsymbol{y}_{i,k} = \boldsymbol{y}_k + \boldsymbol{\varepsilon}_i$, with $\boldsymbol{\varepsilon}_i \sim \mathcal{N}(0, \varsigma_k)$
    Projection of the model solution in the observation space:
      $\mathcal{H}(\boldsymbol{x}^f_{i,k})$
    Calculation of the ensemble means:
      $\overline{\boldsymbol{x}}^f_k = \frac{1}{n_e}\sum_{i=1}^{n_e} \boldsymbol{x}^f_{i,k}$, $\overline{\mathcal{H}(\boldsymbol{x}_k)} = \frac{1}{n_e}\sum_{i=1}^{n_e} \mathcal{H}(\boldsymbol{x}_{i,k})$
    Calculation of the anomaly matrices:
      $\boldsymbol{\Gamma}_k = \frac{\boldsymbol{x}_{i,k} - \overline{\boldsymbol{x}}_k}{\sqrt{n_e - 1}}$, $\boldsymbol{\Lambda}_k = \frac{\mathcal{H}(\boldsymbol{x}^f_{i,k}) - \overline{\mathcal{H}(\boldsymbol{x}^f_k)}}{\sqrt{n_e - 1}}$
    Calculation of the Kalman gain:
      $\boldsymbol{K}_k = \boldsymbol{\Gamma}^f_k(\boldsymbol{\Lambda}_k)^\top \left[\boldsymbol{\Lambda}_k(\boldsymbol{\Lambda}_k)^\top + \varsigma_k\right]^{-1}$
    Update of the state matrix:
      $\boldsymbol{x}^a_{i,k} = \boldsymbol{x}^f_{i,k} + \boldsymbol{K}_k(\boldsymbol{y}_{i,k} - \mathcal{H}(\boldsymbol{x}_{i,k}))$
```

5 Assessment of the DDES Model and Sensitivity Analysis

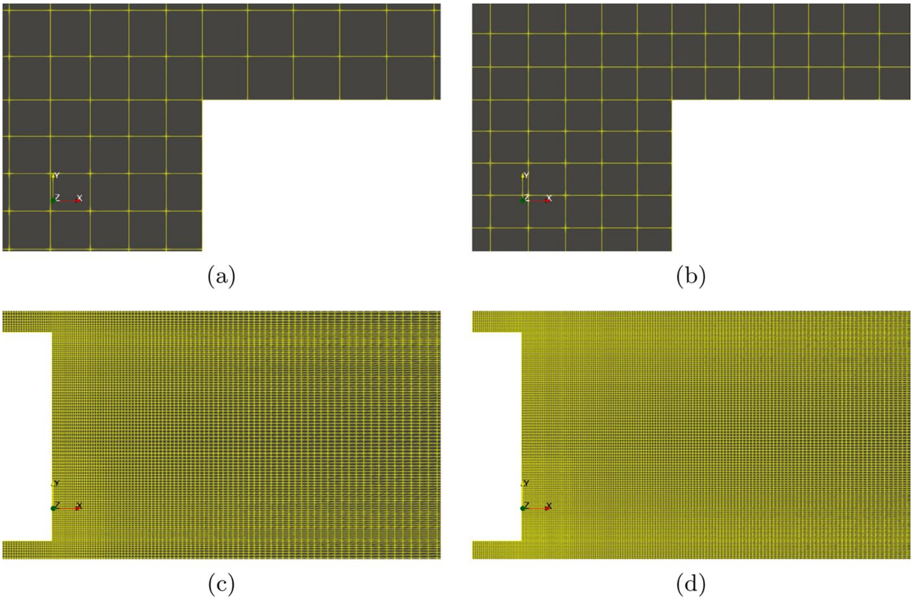
In this chapter, the DDES model used for numerical simulation is validated. In particular, the accuracy of the results will be tested assessing the sensitivity of the model to variations in the computational grid as well as to inlet conditions.

5.1 Sensitivity to Grid Refinement

Two DDES runs, referred to as DDES-I0-G1 and DDES-I0-G2, are performed using a constant velocity inlet and grids of different resolution. The grids, whose details are reported in Table 1, are referred to as G1 and G2, respectively, and their resolution in the proximity of the leading edge is shown in Fig. 4a, b. The main difference between the two grids is their global resolution, which sums up to 4 million elements and 30 million elements. On the other hand, the strategy for the distribution of the mesh elements is very similar for G1 and G2, and it follows indications of previous CFD works in the literature (Mariotti et al. 2017; Cimarelli et al. 2018; Rocchio et al. 2020). Owing to

Table 1 Characteristics of the grids used for the numerical simulations. The size of the elements is provided in D units

CFD run	$\Delta_x^\star$	$\Delta_y^\star$	$\Delta_z^\star$	r_x	r_y	Δ_x^d	Δ_y^d	N
G1	0.01	0.012	0.1	1.012	1.059	0.063	0.016	4×10^6
G2	0.0074	0.009	0.05	1.012	1.015	0.031	0.014	30×10^6



(a) (b)

(c) (d)

Fig. 4 Zoom of the grid resolution close to the leading edge (first row) and in the wake region (second row). Visualizations are proposed for for grids (left column) G1 and (right column) G2

the statistical homogeneity of the flow in the spanwise direction z, the size of the mesh elements is constant and equal to $\Delta_z^\star = \Delta_z/D$. In the streamwise direction x and the normal direction y, the smallest grid elements are in correspondence of the two leading edges and the two trailing edges. In these locations, the size of the mesh elements is $\Delta_x^\star$ and $\Delta_y^\star$, respectively. Moving away from the edges, the size of the mesh elements increases following a geometric progression of ratio r_x and r_y. In the wake region, for $x = 5.5$, $y = 0$ the resolution of the mesh is equal to Δ_x^d and Δ_y^d. A qualitative representation of the resolution in the wake region is shown in figure Fig. 4c, d. In this region, the ratio $\nu_t/\nu \approx 30$–40, which is in line with LES applications for industrial flows (Sagaut 2005) and is significantly smaller than what is classically observed for RANS, for which $\nu_t/\nu \in [100, 1000]$. A similar conclusion is drawn by the observation of the blending parameter f_d of the DDES model, which is equal to 1 in the first grid layer at the wall (full RANS mode) and it decays to zero (full LES mode) within the next two grid layers. In terms of resolution with respect to the Kolmogorov scale η, dimensional arguments manipulating the resolved flow field indicate that scales of size $\approx 70\eta$ are resolved in the wake region by the grid G1. These smallest resolved scales are within the inertial range for the *Re* investigated, confirming that LES is performed in the wake region.

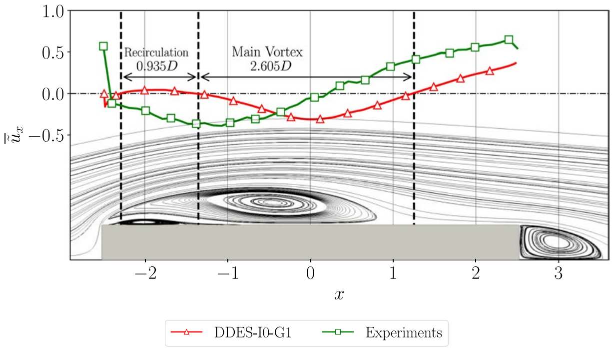
The results from the simulations DDES-I0-G1 and DDES-I0-G2 are compared with those obtained by a URANS using the $\mathcal{K} - \omega$ SST model and a RANS run using the Spalart–Allmaras model. The grid used for the last two calculations is identical to G1 in the x and y directions but it is two-dimensional. Isocontours of the time-averaged velocity magnitude are shown in Fig. 6. The four simulations are capturing the main physical aspects of the flow which include separation at the leading edge, the formation of a recirculation bubble, flow reattachment and evolution of a wake region downstream. However, results from the RANS calculation indicate a shorter recirculation bubble. The reattachment of the flow happens here at 65% of the total length L, against 75% for DDES-I0-G1, 76% for DDES-I0-G2 and 83% for the URANS $\mathcal{K} - \omega$ SST model. The results for the two DDES runs are qualitatively in line with experimental observations from CSTB for the same geometry, Reynolds numbers and low turbulence-intensity upstream conditions. Depending on the value of the turbulence intensity upstream, experiments show a flow reattachment in the range 50–80% approximately, with lower turbulence intensity providing larger recirculation bubbles (Bruno et al. 2014). The results obtained with the DDES here performed are towards the larger values of the experimental range. Considering that the inlet condition here used does not include any synthetic turbulence model, results obtained for this case, as it will shown further, are encouraging and constitute a good starting field to initialize the solution during the sequential Data Assimilation approach. A comprehensive review of the features of the recirculation bubble is reported in Table 2. One can see that the present results, such as the coordinate of the average reattachment point of the main structure x_{mr}, the coordinates of the center of the main structure x_{mv}, y_{mv} are in agreement with results reported in the literature. The results also include the center of the secondary structure at the leading edge, x_{sv}, which can be observed on Figs. 5 and 6).

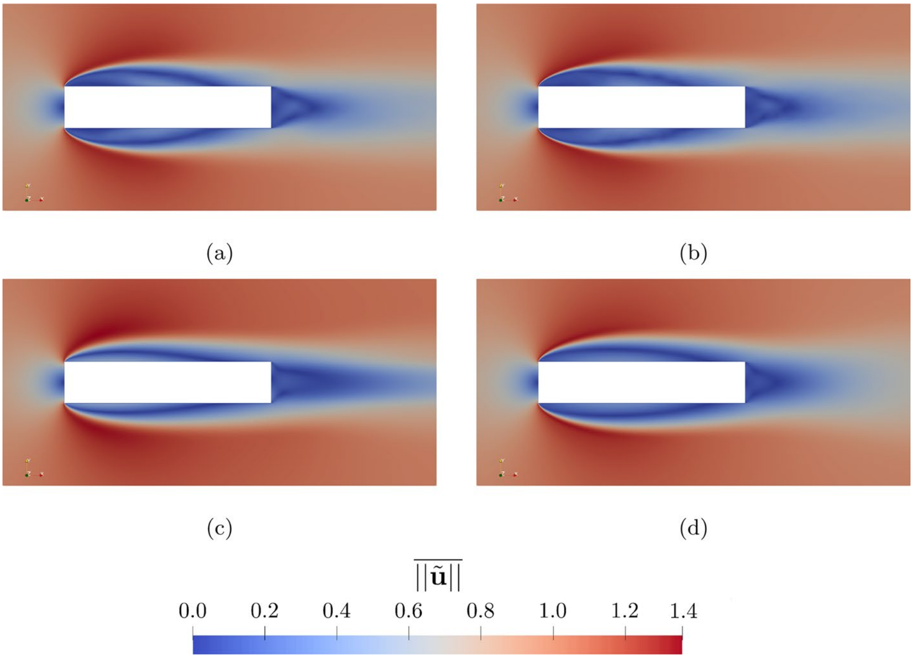
The pressure coefficient $C_p = 2(p - p_\infty)/\rho U_\infty^2$ is now analyzed. The time average and the variance of C_p at the body surface is shown for the four simulations in Fig. 7, along with experimental data reported in the literature (Bruno et al. 2014). The variance for the

Table 2 Features of the recirculation bubble observed on the side of the rectangular cylinder. With reference to Fig. 5, x_{sv} is the center of the secondary structure at the leading edge, $\left(x_{\mathrm{mv}}, y_{\mathrm{mv}}\right)$ are the coordinates of the center of the main structure and x_{mr} is the coordinate of the average reattachment point of the main structure

Simulation	*Re*	x_{sv}	$(x_{\mathrm{mv}}, y_{\mathrm{mv}})$	x_{mr}
Low turbulent intensity				
DDES-I0-G1	2×10^5	−2.29	(−0.45, 0.72)	1.23
Mannini et al. (2011)	2.64×10^4	–	([−1.44/ − 0.05], [0.77/0.88])	1.72–2.06
Grozescu et al. (2011)	2×10^4	–	(−0.17, [0.35/0.82])	1.64
Bruno et al. (2010)	4×10^4	–	(0.04, 0.8)	2.18
High turbulent intensity				
DDES-I1-G1, $I_T = 3.53\,\%$	2×10^5	−2.3	(−1.47, 0.75)	0.52
DDES-DA-G1, $I_T = 0.75\,\%$	2×10^5	−2.19	(−1.15, 0.75)	0.52
CSTB Experiments, $I_T = 3.53\,\%$	2×10^5	–	(−1.2, 0.76)	0.18
Ricci et al. (2017), $I_T = 2.9\,\%$	5×10^4	–	(−0.63, 0.81)	0.98
Ricci et al. (2017), $I_T = 13.6\,\%$	5×10^4	–	(−1.35, 0.72)	−0.2

Fig. 5 Time averaged velocity profile and streamlines in the proximity of the immersed cylinder for DDES-I0-G1 case




(a)

(b)

(c)

(d)

Fig. 6 Isocontours of the time-averaged velocity magnitude $\overline{||\tilde{\mathbf{u}}||}$. Results are shown for the simulations **a** DDES-I0-G1, **b** DDES-I0-G2, **c** a RANS calculation using the Spalart-Allmaras model and **d** an URANS calculation using the $\mathcal{K}-\omega$ SST model

RANS calculation is not available, as a steady-state simulation was performed. One can see again that a significant discrepancy is observed between the RANS and the DDES simulations for the time-averaged C_p. Differences are observed for the URANS simulation as well, even if they fit within the region of experiments reported in the literature (orange area). On the other hand, runs DDES-I0-G1 and DDES-I0-G2 exhibit very similar results and they also compare well with experimental results in the literature. The peak of the variance of C_p is observed for a streamwise position of $\approx$ 65–75% L i.e. around 10% before than the flow re-attachment. This result is consistent with experiments and numerical simulations in the literature (Bruno et al. 2014; Rocchio et al. 2020). However, it appears that the

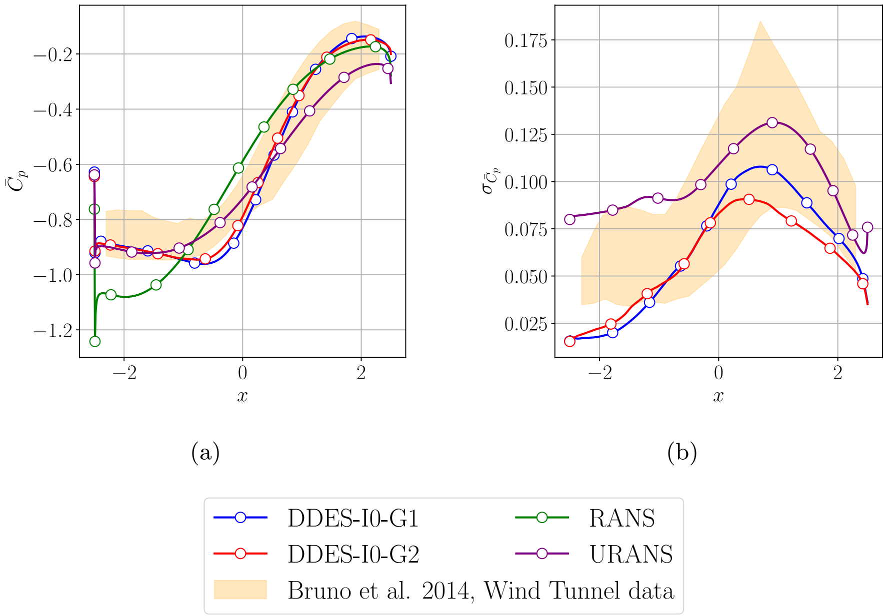


Fig. 7 Pressure coefficient C_p at the body surface. **a** Time-averaged and spanwise-averaged distributions for **a** $\bar{C}_p$ and **b** the variance $\sigma_{\bar{C}_p}$ are shown

URANS simulation exhibit a significantly larger variance for C_p in correspondence of the recirculation bubble, when compared with the DDES runs.

The bulk flow quantities (drag coefficient C_D, lift coefficient C_L, Strouhal number St) obtained using the DDES simulations are now compared with results reported in the literature. Data is provided in Table 3. One can see that the results here obtained with the runs DDES-I0-G1 and DDES-I0-G2 are extremely similar to the data reported in the literature

Table 3 Comparison of the main bulk flow quantities for low turbulent intensity and high turbulent intensity

	$\overline{C_D}$	C'_D	$\overline{C_L}$	C'_L	St
Low turbulent intensity					
DDES-I0-G1	1.02	0.024	0.018	0.22	~ 0.12
DDES-I0-G2	1.02	0.026	0.040	0.21	~ 0.12
Bruno et al. (2010)	~ 1	–	–	~ 0.7	~ 0.11
Mannini et al. (2011)	$[0.968, 1.04]$	–	$[0.0032, 0.047]$	$[0.42, 1.075]$	$[0.094, 0.102]$
Arslan et al. (2011)	$[0.984, 1.39]$	–	–	$[0.59, 0.84]$	$[0.107, 0.16]$
Wei et al. (2011)	$[1.165, 1.305]$	–	$[-0.33, 0.42]$	$[0.495, 1.465]$	–
High turbulent intensity					
DDES-I1-G1, $I_T = 3.53\,\%$	1.25	1.71	0.016	16.73	–
DDES-DA-G1, $I_T = 0.75\,\%$	1.21	0.36	−0.044	3.39	–
Ricci et al. (2017), $I_T = 2.9\,\%$	0.99	0.042	−0.07	0.55	–
Ricci et al. (2017), $I_T = 13.6\,\%$	1.105	0.145	−0.09	1.42	–

Table 4 Position of the sensors used to perform the Morlet transform of the velocity field obtained via DDES

DDES-I0-G1		DDES-I0-G2	
Probe (1)	(−2.374, 0.628)	Probe (1)	(−2.374, 0.616)
Probe (2)	(−1.874, 0.808)	Probe (2)	(−1.874, 0.787)
Probe (3)	(−1.374, 0.915)	Probe (3)	(−1.374, 0.903)
Probe (4)	(−0.873, 1.011)	Probe (4)	(−0.873, 0.999)
Probe (5)	(−0.373, 1.05)	Probe (5)	(−0.373, 1.053)
Probe (6)	(0.128, 1.129)	Probe (6)	(0.128, 1.085)
Probe (7)	(0.628, 1.172)	Probe (7)	(0.628, 1.120)

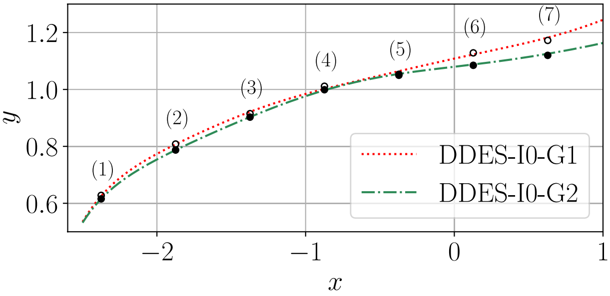


Fig. 8 Position of the sensors used to perform the spectral analysis of the velocity field obtained via DDES. Empty circles represent DDES-I0-G1 probes positioning, full circles represent DDES-I0-G2 probes positioning

for low intensity of the upstream turbulence, despite a significant difference in the Reynolds number.

At last, features of the instantaneous flow are analyzed. Time resolved velocity for the runs DDES-I0-G1 and DDES-I0-G2 is sampled in correspondence of 7 sensors. These sensors are located on the outer limit of the detaching shear layer i.e. where the mean velocity exhibits its maximum (see the procedure used in Rocchio et al. (2020); Moldovan et al. (2022)). The sensors locations are listed in Table 4. Figure 8 shows the detaching layers for runs DDES-I0-G1 and DDES-I0-G2. This sampled field is used to obtain spectra via a Morlet transform. The spectra, which are shown in Fig. 9, again exhibit minimal differences. One can see that a peak for the spectrum obtained for the first probe with $x < -2$ is clearly observed for $St \approx 1$, where St is the Strouhal number. This peak, which is associated with the frequency of Kelvin–Helmholtz instabilities, has also been observed in LES runs (Rocchio et al. 2020).

In summary, the DDES simulations provide a very similar prediction, which appears to be significantly improved when compared with the RANS calculation. For all of these reasons, the grid G1 is chosen to perform the simulations for the present work.

5.2 Sensitivity to the Inlet Conditions

A third numerical simulation, referred to as DDES-I1-G1, is run using the synthetic turbulence inlet presented in Sect. 3.2. The values of the eight coefficients are selected according to the experiments performed at CSTB. More precisely, the components of the Reynolds stress tensor $\tau_{t,ij}$ and the integral length scale ℓ are set according to measurements performed in the free wind tunnel i.e. before the rectangular cylinder was installed. The parameter δ, which is related to the mesh size in the proximity of the

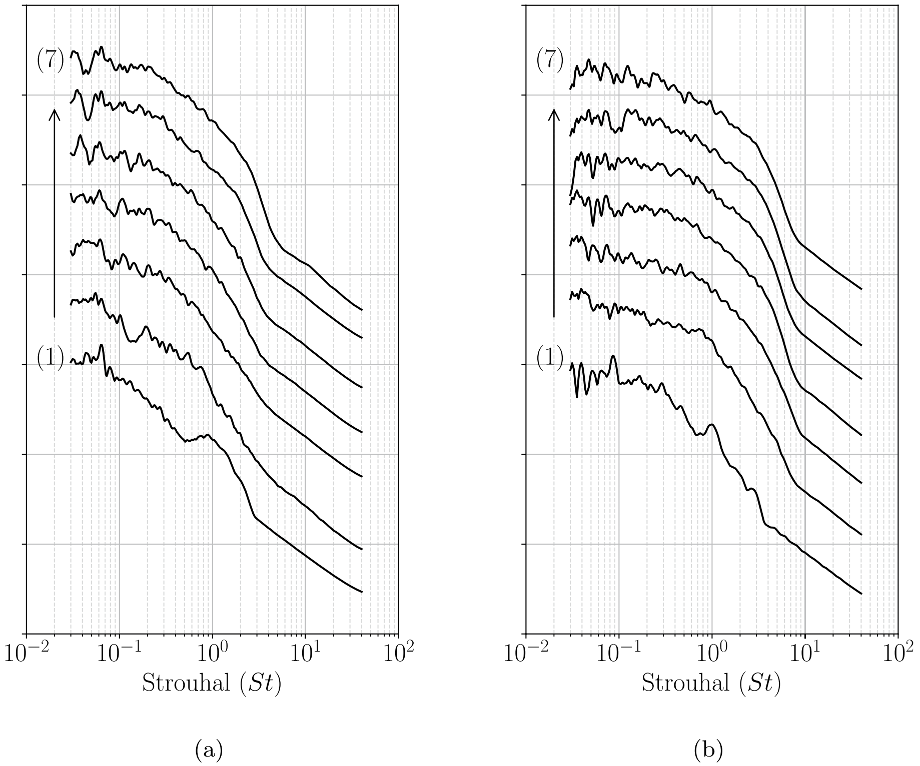


Fig. 9 Spectra obtained via Morlet transform for simulations **a** DDES-I0-G1 and **b** DDES-I0-G2. Results from different sensors are shown using an offset to improve the clarity of the representation

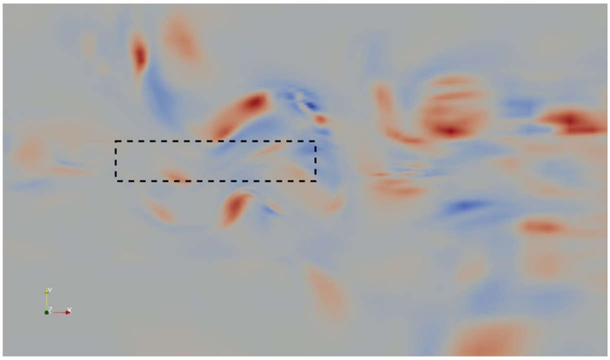
boundary condition, is set to an average value of 0.0922. The values for the eight parameters are summarized in line two of table Table 5. A preliminary simulation has been performed to check the flow obtained using the synthetic turbulence inlet condition. Similarly to what done at CSTB, this run uses the grid G1 but the rectangular cylinder is not immersed i.e. additional mesh elements are included in its place. The results shown in Fig. 10, where the isocontours of the λ_2 criterion are shown, clearly indicate that turbulent structures are advected from the inlet in the physical domain. The rectangular cylinder in the figure is included for reference to establish comparisons with the size of the turbulent structures. However, as previously stated, it is not taken into account in the simulation.

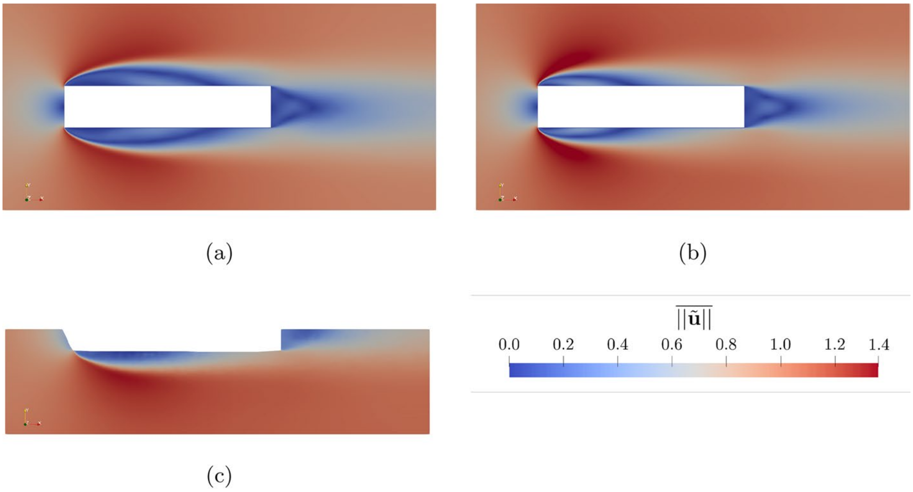
Results from the numerical simulations DDES-I0-G1 and DDES-I1-G1 are compared with the experimental results obtained at the wind tunnel of CSTB. Isocontours of the velocity magnitude are shown in Fig. 11 and the size of the recirculation region is shown in Fig. 12. One can see significant discrepancies between numerical and experimental results and in particular both numerical simulations do not provide an accurate representation of the recirculation bubble. This result was expected for the simulation with the classical, zero turbulence intensity inlet. However, one can see that the usage of the numerical values in Table 5 with the synthetic turbulent inlet implemented in OpenFOAM produces a higher level of turbulence intensity than expected. This result likely comes from a complex

Table 5 Parametric description of the synthetic turbulence inlet condition used in this work

	$\tau_{t,xx}$	$\tau_{t,yy}$	$\tau_{t,zz}$	$\tau_{t,xy}$	$\tau_{t,xz}$	$\tau_{t,yz}$	δ	ℓ
Exp. reference	1.25×10^{-3}	0.99×10^{-3}	0.93×10^{-3}	-2.00×10^{-5}	5.00×10^{-5}	-3.00×10^{-5}	–	1.38
DDES-I1-G1	1.25×10^{-3}	0.99×10^{-3}	0.93×10^{-3}	-2.00×10^{-5}	5.00×10^{-5}	-3.00×10^{-5}	0.0922	1.38
DA - prior	5.652×10^{-5}	5.105×10^{-5}	4.046×10^{-5}	-9.553×10^{-7}	2.446×10^{-6}	-1.411×10^{-7}	0.0922	1
DDES-DA-G1	5.34×10^{-5}	4.18×10^{-5}	7.13×10^{-5}	-7.04×10^{-7}	2.99×10^{-6}	-1.04×10^{-6}	0.0606	1.199

Fig. 10 Instantaneous λ_2 criterion isocontours using the turbulent synthetic inlet over an empty domain. The rectangular cylinder is indicated for reference





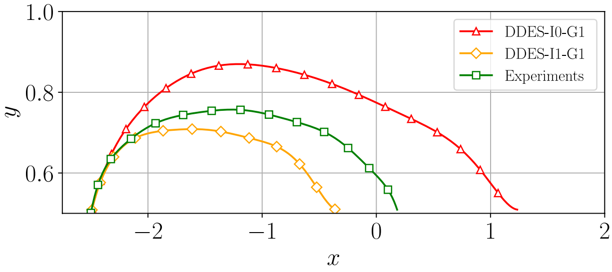
(a)

(b)

(c)

Fig. 11 Isocontours of the time-averaged velocity magnitude $\overline{||\tilde{\mathbf{u}}||}$. Results are shown for the simulations **a** DDES-I0-G1, **b** DDES-I1-G1 and **c** CSTB experiments

Fig. 12 Recirculation region at the top of the rectangular cylinder

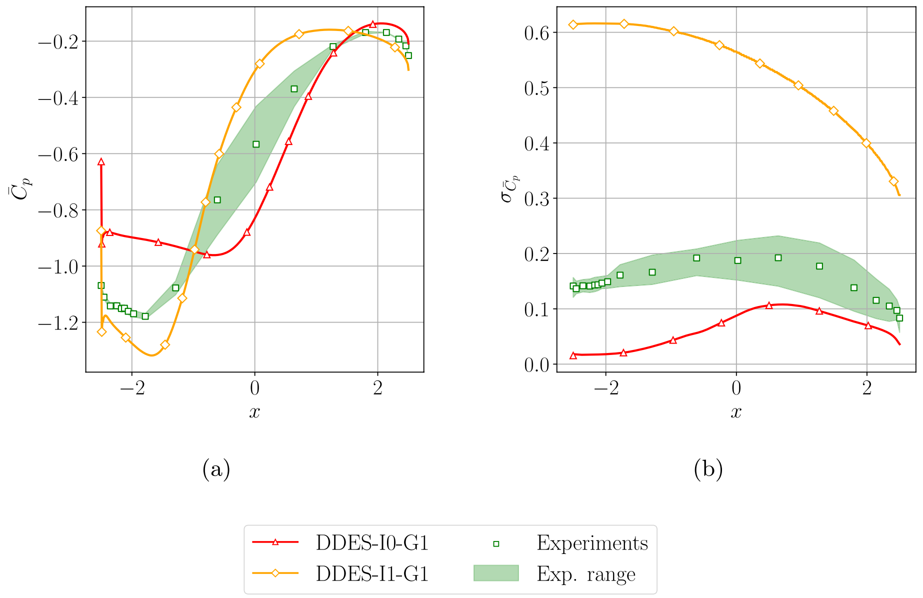


interplay between the modeling / test case choices performed for this analysis. The exact nature of this interaction remains unclear and requires further investigation. A minor effect could also be associated with uncertainties in the experimental measurements. One can also see that both numerical simulations develop a secondary bubble at the leading edge.

This result, which is not observed in the experiments, is actually pretty common in scale resolved numerical simulation for this test case (Bruno et al. 2014).

Similar conclusions can be drawn via the analysis of the time-averaged C_p and its variance, which are shown in Fig. 13. Significant discrepancies are observed for the distribution of the mean pressure coefficient. In addition, the magnitude of the variance of the pressure for the simulation DDES-I1-G1 is approximately four times larger than the experimental results. This observation is partially due to extreme peaks of the pressure that are locally observed. These peaks are most likely induced by the interaction of the numerical solver, the zero-gradient boundary condition at the wall and the unsteady behavior produced by the inlet. Some corrections have been proposed in the literature to mitigate this effect (Patruno and de Miranda 2020) which, however, are not implemented in the available version of OpenFOAM. However, results in Fig. 13b are filtered to exclude the larger peaks numerically observed. Therefore, one could argue that the distribution shown is mainly associated with the inlet boundary condition used for simulation DDES-I1-G1. The red line indicating the variance of the simulation DDES-I0-G1 is significantly lower than the experimental reference, in particular close to the leading edge. This result indicates that the classical inlet used for this simulation fails to capture phenomena at play close to the separation due to the sharp edges.

The analysis of the physical features of the recirculation bubble, which are reported in Table 2, confirms previous findings. The conclusions here drawn are consistent with findings in the literature (Bruno et al. 2010, 2012) which indicate that a shorter recirculation bubble is observed for higher levels of upstream turbulence intensity. When comparing present results with the ones by Ricci et al. (2017), one may think that a significant



(a)

(b)

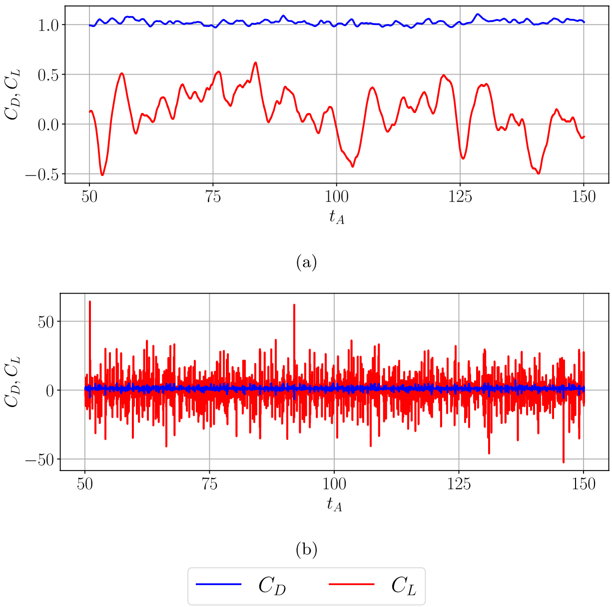
Fig. 13 Features of the pressure field at the top wall. Distributions of **a** the time-averaged pressure coefficient $\bar{C}_p$ and **b** its variance are shown. Numerical results obtained from the DDES simulations are compared with experimental data discrepancy is visible considering that the latter are obtained at a significantly higher inlet turbulent intensity. This discrepancy is actually a result of the position of the inlet, which is significantly more upstream ($80D$) in the work by Ricci et al. (2017). When transported downstream from the inlet, the turbulent kinetic energy $\mathcal{K} = 0.5\left(\tau_{t,xx} + \tau_{t,yy} + \tau_{t,zz}\right)$ and the integral length scale l evolve following a power law (Valente and Vassilicos 2011):

$$\mathcal{K}(x) \approx \mathcal{K}(x_0)\left(1 + \frac{x - x_0}{l(x_0)}\right)^{n_{\mathcal{K}}} \qquad l(x) \approx l(x_0)\left(1 + \frac{x - x_0}{l(x_0)}\right)^{n_l} \tag{13}$$

where x_0 is a virtual origin, $n_{\mathcal{K}} \in [-1.44, -1]$ and $n_l \in [0.28, 0.5]$ (Meldi et al. 2011). Therefore, a direct comparison using the turbulent intensity at the inlet would bring wrong conclusions in this case.

The analysis of the bulkflow quantities in Table 3 indicates that for higher upstream turbulence intensity the mean value of C_D increases with the standard deviation for both C_D and C_L. This result is consistent with the other works reported in the literature. In addition, present simulations seem to indicate that a characteristic Strouhal is not unabiguously identified anymore with the analysis of the time evolution of the lift coefficient C_L. For the present analysis this is probably due to the characteristic turnover time of the upstream turbulence, which is significantly faster than the dynamic wake instability observed for low upstream turbulent intensity. This feature can be qualitatively observed in Fig. 14, where the time history for C_D and C_L are reported for the runs DDES-I0-G1 and DDES-I1-G1. For the former, the governing dynamic instability can be clearly observed in the evolution of C_L, while for the latter the signal is significantly more noisy due to the complex multi-scale interaction at play in this case.

In summary, present results indicate that the usage of a synthetic turbulence inlet condition significantly affect the flow organization, and it can have the potential to produce accurate results in comparison with the experiments. However, the parametric set-up of such inlet can



(a)

(b)

Fig. 14 Drag and lift coefficient C_D, C_L for **a** DDES-I0-G1 and **b** DDES-I1-G1 not be directly extrapolated from the data sampled in the wind tunnels, because of strong non-linear interactions between several sources of error present in the CFD solver. In the following, a data-driven procedure will be used to calibrate the value of the model constants, in order to minimize the discrepancy observed with experiments.

6 Data-driven Augmentation of Inlet Boundary Conditions

Discussion in Sect. 5.2 highlighted the numerical difficulties in obtaining an accurate inlet condition and how this affects the comparison with experimental data. In this section, an EnKF approach is used to blindly optimize the parametric description of the synthetic turbulence inlet used in OpenFOAM in order to reduce the discrepancy between numerical and experimental results.

6.1 Selection of the EnKF Hyperparameters and DA Experiment

The EnKF relies on two main sources of information:

- A *model*, which provides a quasi-continuous description of the flow configuration investigated. The model chosen for the present investigation is the DDES CFD solver, which calculates the instantaneous evolution of the velocity and pressure fields.
- Some *observation*, which is usually high-fidelity data sparse in space and time. In this case, experimental results from CSTB are used. These results are obtained from 451 sensors which are positioned on the PTV plane normal to the spanwise direction z. The positioning of the sensors is shown in Fig. 15. On this location, the time-averaged values of the streamwise velocity $\tilde{u}_x$ and the normal velocity $\tilde{u}_y$ are sampled.

One can see that the CFD model and the experimental observation are different in nature i.e. the model prediction provides instantaneous information while experimental results are time-averaged. Therefore, the DA strategy must be adapted to take into account this difference. The algorithm is organized in five different steps:

(o) The $n_e = 30$ DDES calculations which constitute the numerical ensemble used in the EnKF are initialized with random values for the parametric description of the inlet. These parameters include the six component $\tau_{t,ij}$ of the Reynolds stress tensor at the inlet, as well as the integral length scale ℓ and the mesh characteristic length δ. These last two parameters are used to determine the length scale σ_i in Eq. (5). Using the formalism of Data Assimilation, the free coefficients are arranged in an array $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}_t = \begin{bmatrix} \tau_{t,xx} \\ \tau_{t,yy} \\ \tau_{t,zz} \\ \tau_{t,xy} \\ \tau_{t,xz} \\ \tau_{t,yz} \end{bmatrix} \qquad \boldsymbol{\theta}_l = \begin{bmatrix} \ell \\ \delta \end{bmatrix} \qquad \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_t \\ \boldsymbol{\theta}_l \end{bmatrix} \tag{14}$$

The random values for each of the coefficients of $\boldsymbol{\theta}$ and for each of the simulations is determined via sampling of a truncated Gaussian probability density function

$\mathcal{N}(\mu_{\theta_i}, \varsigma^2_{\theta_i})$. μ and ς are the average and the standard deviation provided for each parameter θ_i. These parameters, which are reported in Table 6, have been chosen taking into account the experimental results provided by CSTB and after a careful investigation of the preliminary CFD results. The choice of the prior values for the parametric description is important to obtain a fast and robust convergence of the EnKF optimization (Villanueva et al. 2024). Values for the parameters are accepted within the range $[\mu - 2\varsigma,\ \mu + 2\varsigma]$ and are re-sampled if they lay outside the prescribed interval. The instantaneous physical state for $t = 0$ is the same for all the ensemble members and it is obtained by a preliminary DDES run with the mean parameters μ in Table 6.

(i) The n_e DDES simulations are run for a total of $t = 25\, t_A$ times, where the advection time scale is $t_A = D/U_\infty$. This time window is divided into two phases. For $t < 10\, t_A$, no average is performed as the effects of the inlet parametric description are advected in the domain. The threshold $10\, t_A$ has been chosen observing the results of prior runs, for which aerodynamic coefficients would stabilize after $5\, t_A$ to $7\, t_A$. The evolution during this transient of the drag coefficient C_D is shown in Fig. 16a for one ensemble member. The size of this initial time window is associated with the distance of the inlet to the rectangular cylinder. In fact the turbulent structures at the inlet, whose size is $l \approx 1$ in D units, reach the immersed body for $t \approx 5\, t_A$. Increasing the upstream distance of the inlet would have resulted in an increase of the computational resources demanded in this case for this transient phase. In the second phase for $t \in [10, 25]\, t_A$, averages are performed in time and in the spanwise direction. This interval, which corresponds to the advection of approximately 15 turbulent structures through the body, has been chosen observing the behavior of a preliminary run. The rate of convergence of the averaged velocity field is measured via the quantity

$$\epsilon_U(t = i\, t_A) = \iint_{S_o} \sqrt{\frac{\left(\overline{\tilde{u}_i}^2 - \overline{\tilde{u}_{400}}^2\right)}{\overline{\tilde{u}_{400}}^2}}\, dS'$$

whose evolution in time is shown in Fig. 16b. This indicator measures the discrepancy between the averaged flow field at the time $t = i\, t_A$ and the converged solution for $t = 400\, t_A$ over the observation region S_o. One can see that for $t = 25\, t_A$ the value of $\epsilon_{25} \approx 0.02$ i.e. the average lack of convergence for the mean velocity field over this time window is roughly of the same order of magnitude of the uncertainty in experimental results. One can also see that for $t = 10\, t_A$ the spanwise average over 50 grid elements of an instantaneous snapshot produces a mean field with an average 0.07 discrepancy over the observation window. This phase corresponds to the *forecast* of the EnKF.

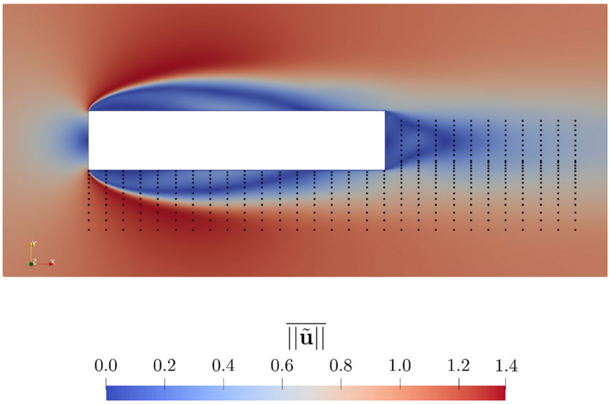
(ii) After the forecast is performed, the mean velocity field for each ensemble member i is post-processed to obtain $\mathcal{H}(\boldsymbol{x}_i^f)$ i.e. the velocity field is interpolated over the sensors where experimental observation is available. Because of the homogeneity in the spanwise direction, only the components in the streamwise direction $\overline{\tilde{u}_x}$ in the normal direction $\overline{\tilde{u}_y}$ are interpolated. Therefore, the vector including the model information used for DA for the ensemble member i is:

$$\mathcal{H}(\boldsymbol{x}_i) = \left(\overline{\tilde{u}_{x_1}} \ \dots \ \overline{\tilde{u}_{x_n}} \ \overline{\tilde{u}_{y_1}} \ \dots \ \overline{\tilde{u}_{y_n}} \right) \in \mathbb{R}^{2n_o} \tag{15}$$

Table 6 Features of the probabilistic distribution for the prior parameters of the CFD ensemble members

	$\tau_{t,xx}$	$\tau_{t,yy}$	$\tau_{t,zz}$	$\tau_{t,xy}$	$\tau_{t,xz}$	$\tau_{t,yz}$	δ	ℓ
$\boldsymbol{\mu}_x$	5.652×10^{-5}	5.105×10^{-5}	4.046×10^{-5}	-9.553×10^{-7}	2.446×10^{-6}	-1.411×10^{-7}	9.22×10^{-2}	1
ς_x^2	1.72×10^{-10}	9.02×10^{-11}	6.4×10^{-11}	4.13×10^{-14}	2.07×10^{-13}	9.34×10^{-14}	1.39×10^{-3}	3.98×10^{-2}

Fig. 15 Position of the sensors used to sample the experimental data. The sensors are shown with the velocity magnitude obtained by one of the DDES ensemble members



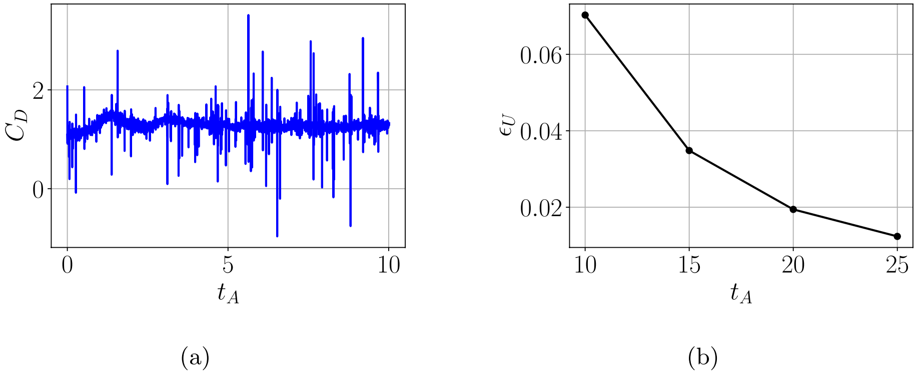


Fig. 16 **a** Evolution of C_D for one ensemble member of the EnKF in the initial time window $[0, 10]t_A$. **b** Convergence of the mean velocity profiles in the time window $[10, 25]t_A$

where the number of sensors is $n_o = 451$. The complete matrix used for ensemble approximation is $\mathcal{H}(\boldsymbol{x}) \in \mathbb{R}^{902\times30}$. The position of the sensors over the average velocity magnitude of one of the ensemble members is shown in Fig. 15.

(iii) The other numerical ingredients used in the EnKF are assembled. The observation vector, which is composed by the time-averaged velocity field sampled over the sensors from the experiments, is perturbed adding a Gaussian noise $\mathcal{N}(0, \varsigma_y)$, where ς_y accounts for the estimated 5% uncertainty over the experimental data. Using the classical hypothesis that observation from different sensors is not correlated (Bocquet et al. 2016), the additive noise is used to obtain n_e observation vectors which are combined in the observation matrix $\boldsymbol{Y}^{902\times30}$. The physical state obtained via model realizations is also used to calculate the anomaly matrices according to Eq. (8). All these numerical elements are combined to obtain the Kalman gain $\boldsymbol{K}$ using Eq. (10).

(iv) The *analysis* phase takes place, where the augmented state $\boldsymbol{x}_i^a$ for each ensemble member i is obtained with Eq. (11) from its forecast $\boldsymbol{x}_i^f$, the Kalman gain $\boldsymbol{K}$ and the discrepancy between experiments and model prediction. The latter is measured via the difference of the i^{th} column of $\boldsymbol{Y}$ and $\mathcal{H}(\boldsymbol{x})$. The physical state correction obtained via DA, which is derived for a steady configuration, is discarded. On the other hand,

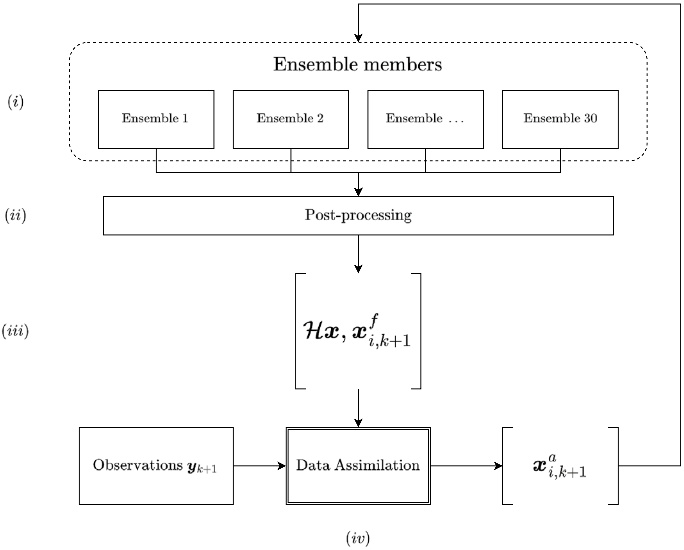


Fig. 17 Recursive procedure for the optimization of the coefficient of the inlet boundary condition based on the EnKF

> the EnKF update to the parameters is conserved. However, in order to avoid spurious oscillations of the coefficients which can be produced by the Kalman update for the analysis of strongly unsteady configurations (Villanueva et al. 2023), the maximum/minimum value accepted for the following forecast are truncated to the extremes of the previous parametric range. In order to obtain a sufficient variability of the ensemble, the parameters are then inflated using the same variance that was used for the prior state. Once the parameters are updated, the new forecast is ready to be submitted without any change of the initial conditions for bulk velocity and pressure.

This algorithm, which is exemplified in the scheme in Fig. 17, has been here repeated for three complete cycles i.e. three forecasts and three analyses. At the end of the last forecast, the member exhibiting the lowest discrepancy with the experiments, in terms of features of the recirculation bubble, has been selected. A full run has been performed for this member for a total of $300\,t_A$. The results obtained with this optimized DDES run, which will be referred to as DDES-DA-G1 in the following, are going to be compared with simulation DDES-I0-G1, DDES-I1-G1 and the experiments from CSTB. It is important to stress the general impact that DA strategies can have on numerical studies in fluid mechanics and therefore the importance of deriving clear guidelines of application. While DA is here used to infer the behavior of upstream inlet conditions, these findings are more generally important to identify efficient practices for these techniques for a broad range of investigations.

6.2 Results

The physical prediction of the flow obtained with the DA algorithm is now analysed. First, the isocontours of the velocity magnitude are presented in Fig. 18 along with the experimental results from CSTB. Here, the PTV results are not complete around

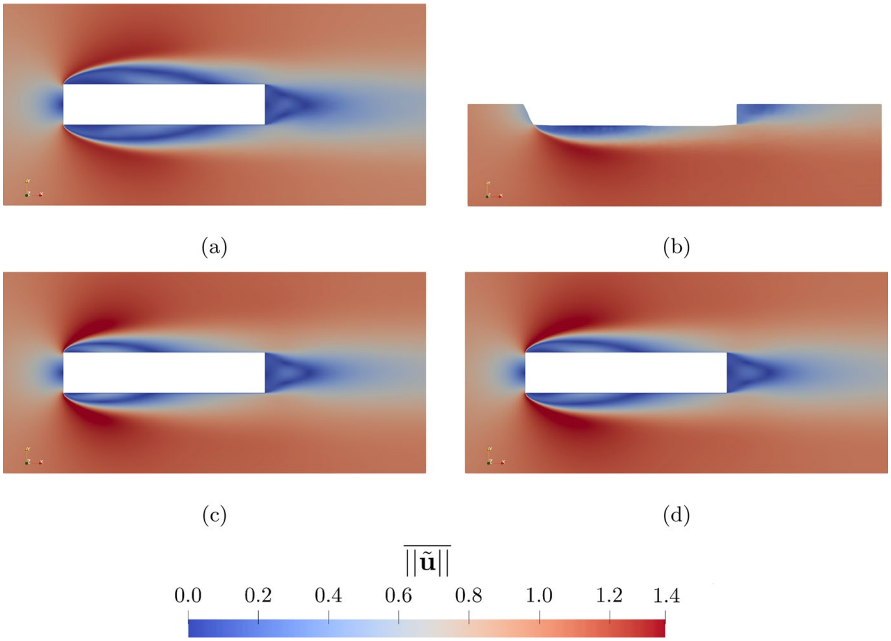


Fig. 18 Isocontours of the time-averaged velocity magnitude $\overline{||\tilde{\mathbf{u}}||}$. Results are shown for **a** the simulation DDES-I0-G1, **b** experimental data from CSTB's wind tunnel **c** the simulation DDES-I1-G1 and **d** the simulation DDES-DA-G1

the body surface, in particular close to the stagnation point at the leading edge, and it should not be interpreted as a different geometric shape. One can see that the data-driven run DDES-DA-G1 convincingly captures the flow topology like the prior simulations DDES-I0-G1 and DDES-I1-G1. However, the flow features in the proximity of the recirculation region appear to be in better agreement with the experimental data. This qualitative analysis is supported by the investigation of the velocity profiles over the surface of the cylinder which is reported in Fig. 19. One can see that the results obtained by the run DDES-DA-G1 exhibit an excellent agreement with experimental data, while large discrepancies are observed for simulations DDES-I0-G1 and DDES-I1-G1.

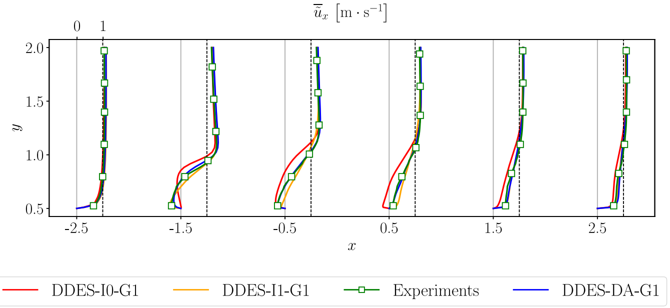


Fig. 19 Distribution of the time averaged streamwise velocity $\overline{\tilde{u}_x}$ over the surface of the cylinder. Numerical results for the runs DDES-I0-G1, DDES-I1-G1 and DDES-DA-G1 are compared with experiments from CSTB

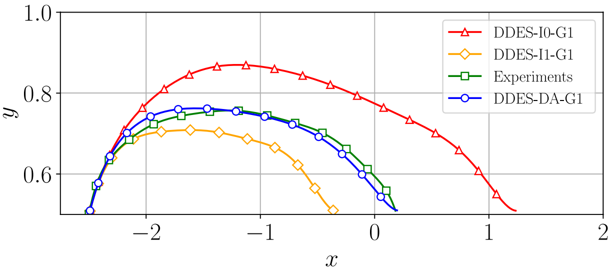


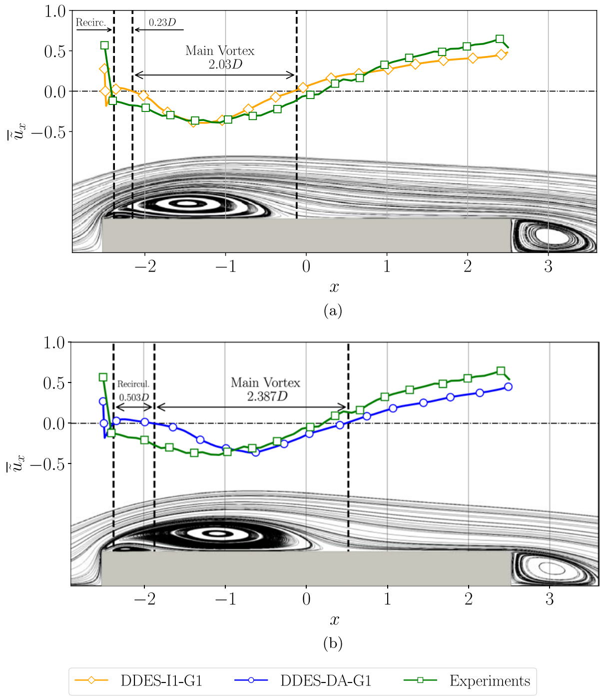
Fig. 20 Size of the recirculation bubble for the numerical simulations performed, which are compared with available experimental results

Further confirmation is obtained by the analysis of the size of the recirculation bubble, which is shown in Fig. 20. The curve obtained by the simulation DDES-DA-G1 almost exactly superposes with the experimental data with a significant improvement in the accuracy when compared with the two DDES preliminary simulations. Further analysis on this point is provided in Fig. 21, where the streamlines of the time-averaged velocity field are shown for the simulation DDES-I1-G1 and DDES-DA-G1. One can see that the increase of the main recirculation region obtained with the parametric optimization comes with the increase in size of a secondary recirculation region, which is observed at the leading edge. This result is responsible for a small degradation of the results in the near wall prediction for $x \in [-2.5, -1]$, even if a global improvement is observed.

The features of the pressure field on the top wall of the rectangular cylinder are now analysed. These quantities are of particular interest, because experimental results are available but not used as available information for the DA analysis step. Therefore, this comparison can highlight the global improvement in accuracy of the data-driven method. The time-averaged $\overline{C}_p$ and its variance $\sigma_{\overline{C}_p}$ are shown in Fig. 22. The blue shaded area corresponds to the variability of the solution of the DA final forecast, with the simulation DDES-DA-G1 in dark blue. On the other hand, the green shaded area corresponds to the confidence in the local experimental measurements. One can see that the prediction of $\overline{C}_p$ for the DA model, shown in Fig. 22a, exhibits a lower discrepancy with experimental data when compared with prior DDES runs. In addition, one can see a good superposition of the experimental and data-driven shaded areas, which indicates that the application of the EnKF was robust. Similarly to the time-averaged $\overline{C}_p$, the variance $\sigma_{\overline{C}_p}$ for the DDES-DA-G1 run shows closer values to the experiments. The DA run, again represented in dark blue, is in very good agreement with the experimental data from approximately the re-attachment point of the recirculation bubble to the trailing edge. In the recirculation bubble region, $\sigma_{\overline{C}_p}$ is significantly higher than the experiments, but still provides the lower normalized discrepancy for the three numerical simulations performed. This discrepancy can be again qualitatively associated with peaks of the pressure field which results from the interaction of boundary conditions and the numerical solvers. Contrarily to simulation DDES-I1-G1, the filtering of the results affects the DA profile for around 2–3% of the magnitude, indicating that this issue is not the main mechanism at play for the determination of the pressure variance for the run DDES-DA-G1.

At last, second-order statistics are analyzed. The shear component of the Reynolds stress tensor $\tau_{t,xy}$ is shown in Fig. 23. As for the pressure field, experimental results for this quantity are available but they were not used in the optimization process. One can see that none of the simulations is able to capture the high intensity turbulent shear close to the leading

Fig. 21 Comparison of time averaged velocity profiles and streamlines along the cylinder length. Simulations **a** DDES-I1-G1 and **b** DDES-DA-G1 are shown



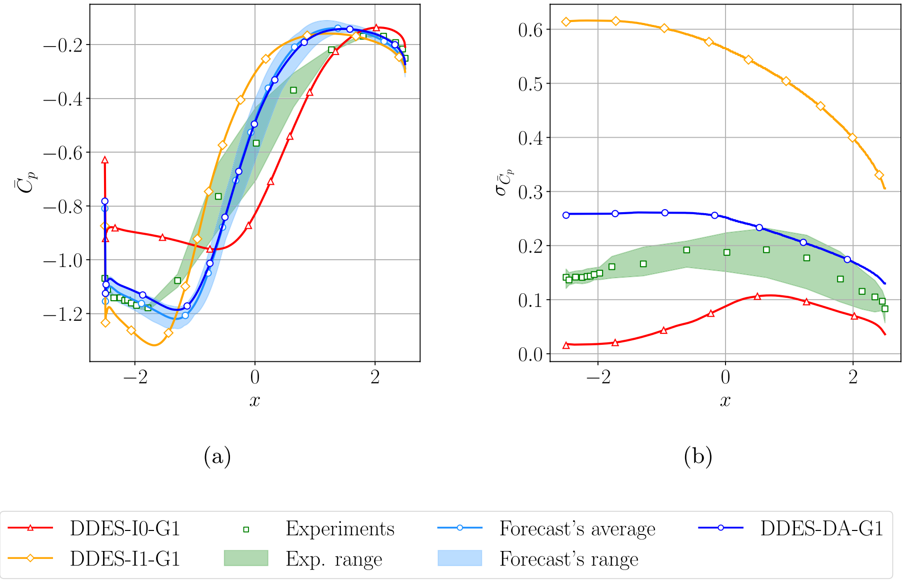


Fig. 22 Features of the pressure field at the top wall. Distributions of **a** the time-averaged pressure coefficient $\overline{C}_p$ and **b** its variance are shown. Numerical results obtained from the DDES simulations are compared with experimental data and the optimized DA run

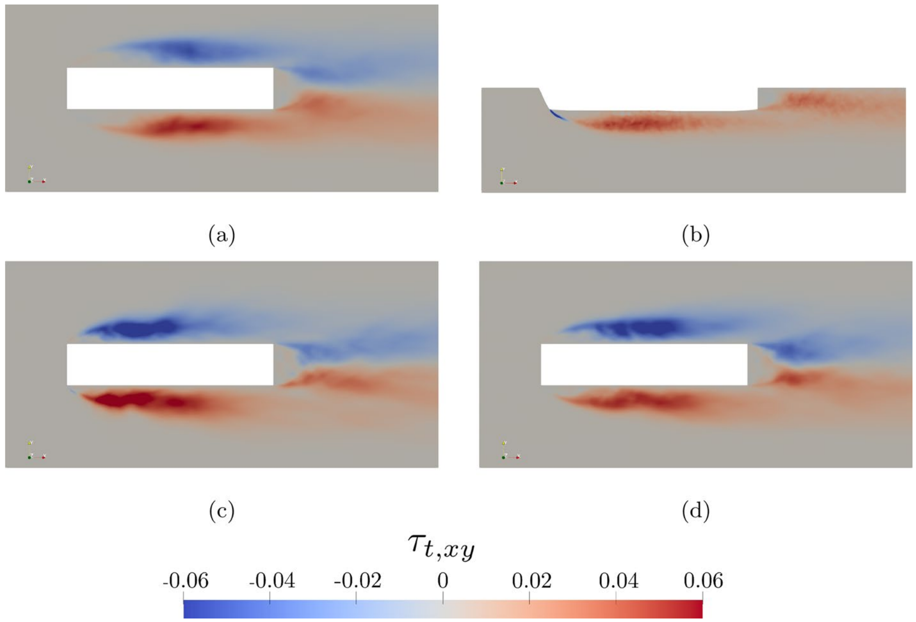


Fig. 23 Shear component $\tau_{t,xy}$ of the Reynolds stress tensor. Results are shown for **a** run DDES-I0-G1, **b** the experimental data, **c** run DDES-I1-G1 and **d** run DDES-DA-G1

edge which is observed in the experiments. This problematic aspect, which is tied to the mesh resolution and the interaction of the discretization error with the turbulence model, is probably among the governing elements in the difficulty observed to match experimental data. The worst numerical results for this quantity are obtained for the simulation DDES-I0-G1, which globally produces very little turbulent shear in the separation region of the leading edge. This is probably due to the features of the inlet used, to which no synthetic turbulence was added. A confirmation can be found observing the profiles for the simulation DDES-I1-G1 in Fig. 23c. In this case $\tau_{t,xy}$ is also underpredicted in the proximity of the leading edge, but it rapidly increases and largely exceeds the experimental prediction at around one-third of the length. Again, the DA results in Fig. 23d appear to be the closest to the experiments. Despite the underprediction at the leading edge, the distribution and intensity of $\tau_{t,xy}$ are very similar to the experimental reference. Overall, DA results exhibit as expected a significant improvement in accuracy for all the physical quantities investigated, when compared with the preliminary DDES runs. This result confirms that the usage of a model based on the Navier–Stokes equations within the DA formalism provides an optimization for all the physical variables even if the data sampled is restricted to some observable quantities.

Similar conclusions can be drawn by the analysis of the profiles of $\sqrt{\tau_{t,xx}}/U_\infty$ reported in Fig. 24. Again, results for the run DDES-DA-G1 exhibit the best agreement with experiments, even if the latter exhibit convergence issues in the proximity of the wall for $y < 0.7$.

In summary, the application of the EnKF to this case was able to obtain a suitable parametric description of the inlet, which results in a good agreement between numerical and experimental data. Looking at the optimized inlet coefficients reported in Table 5, one can draw a number of conclusions. The physical and numerical length scales l and δ were not dramatically modified by the optimization process, with variations of the order of 20–50%.

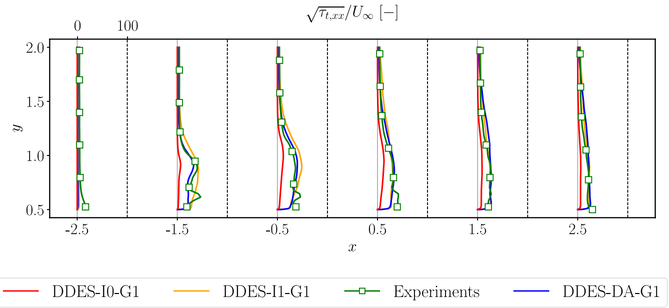


Fig. 24 Distribution of $\sqrt{\tau_{t,xx}}/U_\infty$ over the surface of the cylinder. Numerical results for the runs DDES-I0-G1, DDES-I1-G1 and DDES-DA-G1 are compared with experiments from CSTB

However, one can see that values for the diagonal components of the Reynolds stress tensor have been reduced of around 20 times when compared with the experiments. One possible reason for this result is connected with resolved turbulent dissipation. This quantity, which is proportional to the intensity of the resolved velocity gradients, is significantly decreased by the mesh coarsening used in hybrid CFD and it is probably not sufficiently compensated upstream by the SGS closure. This implies that, in order to obtain similar features of the flow, the energy of the structures injected at the inlet must be scaled to balance this issue. Also, a visible departure from isotropy is observed, considering that the component $\tau_{t,zz}$ is almost 50% larger than $\tau_{t,xx}$ and $\tau_{t,yy}$. One could argue that this observation may be connected with the size of the grid elements in the z direction, which are more coarse than the dimensions in the x and y directions in the proximity of the body and in the wake area.

7 Conclusion

The predictive capabilities of the hybrid CFD method Delayed Detached Eddy Simulation, here used for the simulation of the BARC test case, have been improved via a data-driven technique relying on Data Assimilation. More precisely, the parametric description of a synthetic turbulent inlet condition has been optimized to reduce the discrepancy with available experimental data. A total of eight parameters have been updated, six of them representing the components of the Reynolds stress tensor while the other two are related to physical / numerical length scales. The EnKF algorithm has been adapted to bridge the available experimental observation, which consists of time-averaged features of the velocity field, and the model prediction, whose fields are intrinsically unstationary. The results obtained show that the optimization of the inlet improves the accuracy of the quantities observed in the DA process, such as the velocity field, but also of the other variables, such as the pressure coefficient at the wall and the shear component of the Reynolds stress tensor. All of these results suggest that data-driven approaches can obtain a more realistic representation of the boundary conditions in CFD problems, as also shown by Lamberti et al. (2018) for the atmospheric boundary layer. In a more general context, DA techniques are showing potential to create solid bridges between experimental and numerical realizations and they open perspectives of application to several fields of investigation dealing with complex flows. The main constraints that affect the methodology used in the present work are associated with the limits in the optimization process. In fact, the number of parameters that can be manipulated is tied with the number of ensemble members available. Advancement on this topic is currently investigated by the team relaying on multilevel / multifidelity approaches (Moldovan et al. 2022, 2022), penalization of the parametric space (Hou et al. 2021) and joint application of EnKF with machine learning (Brajard et al. 2021).

Appendix A DDES model details

In DDES, the turbulent viscosity ν_t is obtained via an algebraic relation with a newly introduced quantity, the viscosity-like variable $\dot{\nu}$:

$$\nu_t = \dot{\nu} f_{v1} \tag{16}$$

with:

$$f_{v1} = \frac{\chi^3}{\chi^3 + C_{v1}^3}, \qquad \chi = \frac{\dot{\nu}}{\nu} \tag{17}$$

C_{v1} is a constant to be provided by the user. The variable $\dot{\nu}$ is obtained resolving a transport equation, like in classical one-equation eddy viscosity models:

$$\frac{\partial \dot{\nu}}{\partial t} + (\tilde{\mathbf{u}} \cdot \nabla)\dot{\nu} = \overbrace{\mathcal{P}_{\dot{\nu}}}^{\text{Production}} + \overbrace{\mathcal{T}_{\dot{\nu}}}^{\text{Diffusion}} - \overbrace{\mathcal{D}_{\dot{\nu}}}^{\text{Destruction}} \tag{18}$$

where:

$$\mathcal{P}_{\dot{\nu}} = C_{b1}[1 - f_{t2}]\dot{S}\dot{\nu} \tag{19}$$

$$\mathcal{T}_{\dot{\nu}} = \frac{1}{\sigma}\left(\nabla \cdot [(\nu + \dot{\nu})\nabla\dot{\nu}] + C_{b2}|\nabla\dot{\nu}|^2\right) \tag{20}$$

$$\mathcal{D}_{\dot{\nu}} = \left[C_{w1}f_w - \frac{C_{b1}}{\kappa^2}f_{t2}\right]\left(\frac{\dot{\nu}}{\dot{d}}\right)^2 \tag{21}$$

One can see that Eqs. (19)–(21) are governed by a number of free coefficients (C_{b1}, C_{b2}, C_{w1}, σ), physical parameters (von Kármán constant κ) and functions (f_w, f_{t2}) which can be chosen or modified by the user. In particular, the terms $\dot{S}$ and $\dot{d}$ determine the production and destruction terms of Eq. (18). $\dot{S}$ is obtained by a series of algebraic equations, which include new model coefficients to be selected:

Table 7 Values chosen for the model constants for the DDES closure used in the present analysis. These values are the ones provided as default in the implementation of the code

σ	C_{b1}	C_{b2}	κ	C_{w1}	C_{w2}	C_{w3}	C_{v1}	C_{t1}	C_{t2}	C_{t3}	C_{t4}
2/3	0.1355	0.622	0.41	3.2391	0.3	2	7.1	1	2	1.1	2

$$\begin{aligned} \dot{S} &\equiv S + \frac{\dot{\nu}}{\kappa^2 d^2} f_{v2}, & f_{v2} &= 1 - \frac{\chi}{1 + \chi f_{v1}} \\ S &= \sqrt{2 W_{ij} W_{ij}}, & W_{ij} &= \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right) \\ f_{t2} &= C_{t3} \exp\left(-C_{t4}\chi^2\right), & f_w &= g\left[\frac{1 + C_{w3}^6}{g^6 + C_{w3}^6}\right]^{1/6} \\ g &= r + C_{w2}(r^6 - r), & r &\equiv \frac{\dot{\nu}}{\dot{S}\kappa^2 d^2} \end{aligned} \tag{22}$$

The definition used for the parameter $\dot{d}$, which represents a characteristic length, is intrinsically associated with the hybrid RANS–LES chosen technique. In the context of the DDES model, $\dot{d}$ is defined as:

$$\dot{d} = d - f_{\mathrm{d}} \max\left(0, d - C_{\mathrm{DES}}\Delta\right) \tag{23}$$

where d is the distance from the closest wall, $C_{\mathrm{DES}} = 0.65$ and $\Delta \equiv \max(\Delta x, \Delta y, \Delta z)$. In addition:

$$f_{\mathrm{d}} \equiv 1 - \tanh\left(\left(8 \cdot r_{\mathrm{d}}\right)^3\right); \quad r_{\mathrm{d}} = \frac{\nu_t + \nu}{\kappa^2 d^2 \cdot \max\left(\sqrt{(\nabla\tilde{\mathbf{u}}) \cdot (\nabla\tilde{\mathbf{u}})}, 10^{-10}\right)} \tag{24}$$

The algebraic expression for $\dot{d}$ provided in Eq. (23) can blend the model behavior between full RANS and full LES modes. In particular, a dominant RANS behavior will be obtained in proximity of the wall. Increasing the wall distance d, $r_{\mathrm{d}} \to 0$ and therefore $f_{\mathrm{d}} \to 1$. In these conditions, a dominant LES behavior is observed. Moreover, one can see that $\dot{d}$ is not exclusively grid-dependent, as r_{d} exhibits an explicit dependence to $\tilde{\mathbf{u}}$, ν_t and ν. This choice alleviates some constraints about grid-dependency of the solution, even if a grid sensitivity analysis is always recommended. For sake of completeness, the value of the model constants are reported in Table 7.

Acknowledgements The team would like to acknowledge the help from Yann Haffner from CSTB who provided the experimental data used for Data Assimilation and validation of the results.

Author Contributions Tom Moussie: Methodology, Validation, Simulations, Analysis, Data Curation, Writing Original Draft & Revision, Visualization Paolo Errante: Methodology, Validation, Analysis, Data Curation, Writing Original Draft & Revision, Visualization Marcello Meldi: Conceptualization, Methodology, Analysis, Resources, Writing Original Draft & Revision, Supervision, Project Administration, Funding acquisition

Funding Open access funding provided by Arts et Metiers Institute of Technology. The present research work has been developed in the framework of the project ANR JCJC 2021 IWP-IBM-DA. Computational resources from the cluster CASSIOPEE have been used for the calculations.

Data Availability Numerical data from simulations will be shared on request.

Code Availability The EnKF algorithms here used are included in the platform CONES freely available on GitHub https://github.com/MiguelMValero/CONES

Declarations

Conflict of interest The authors declare no competing interests.

References

Arslan, T., Pettersen, B., Andersson, H.I.: Calculation of the flow around two interacting ships. In: International Conference on Computational Methods in Marine Engineering (2011)

Asch, M., Bocquet, M., Nodet, M.: Data assimilation: methods, algorithms, and applications. SIAM (2016)

Bartoli, G., Bruno, L., Buresti, G., Ricciardelli, F., Salvetti, M.V., Zasso, A.: BARC: a benchmark on the aerodynamics of a rectangular 5:1 cylinder. (2009). https://api.semanticscholar.org/CorpusID:108730058

Bocquet, M., Asch, M., Nodet, M.: Data Assimilation: Methods, Algorithms, and Applications. Fundamentals of algorithms, vol. 11. Society for Industrial and Applied Mathematics, Philadelphia (2016)

Brajard, J., Carrassi, A., Bocquet, M., Bertino, L.: Combining data assimilation and machine learning to infer unresolved scale parametrization. Philos. Transa. R. Soc. A Math. Phys. Eng. Sci. **379**(2194), 20200086 (2021). https://doi.org/10.1098/rsta.2020.0086

Bruno, L., Fransos, D., Coste, N., Bosco, A.: 3d flow around a rectangular cylinder: a computational study. J. Wind Eng. Ind. Aerodyn. **98**(6), 263–276 (2010). https://doi.org/10.1016/j.jweia.2009.10.005

Bruno, L., Coste, N., Fransos, D.: Simulated flow around a rectangular 5:1 cylinder: spanwise discretisation effects and emerging flow features. J. Wind Eng. Ind. Aerodyn. **104–106**, 203–215 (2012). https://doi.org/10.1016/j.jweia.2012.03.018

Bruno, L., Salvetti, M.V., Ricciardelli, F.: Benchmark on the aerodynamics of a rectangular 5:1 Cylinder: an overview after the first four years of activity. J Wind Eng. Ind. Aerodyn. **126**, 87–106 (2014). https://doi.org/10.1016/j.jweia.2014.01.005

Brunton, S.L., Noack, B.R., Koumoutsakos, P.: Machine learning for fluid mechanics. Ann. Rev. Fluid Mech. **52**(1), 477–508 (2020)

Caretto, L.S., Gosman, A.D., Patankar, S.V., Spalding, D.B.: Two calculation procedures for steady, three-dimensional flows with recirculation. In: Cabannes, H., Temam, R. (eds.) Proceedings of the Third International Conference on Numerical Methods in Fluid Mechanics, pp. 60–68. Springer, Berlin, Heidelberg (1973)

Charisi, S., Thiis, T.K., Aurlien, T.: Full-scale measurements of wind-pressure coefficients in twin medium-rise buildings. Buildings (2019). https://doi.org/10.3390/buildings9030063

Chiarini, A., Quadrio, M.: The turbulent flow over the barc rectangular cylinder: a dns study. Flow Turbul. Combust. **107**, 875–899 (2021)

Chiarini, A., Quadrio, M.: The importance of corner sharpness in the barc test case: a numerical study. Wind Struct. **34**, 43–58 (2022)

Cimarelli, A., Leonforte, A., Angeli, D.: Direct numerical simulation of the flow around a rectangular cylinder at a moderately high Reynolds number. J Wind Eng. Ind. Aerodyn. **174**, 39–49 (2018). https://doi.org/10.1016/j.jweia.2017.12.020

Constant, E., Javier, F., Meldi, M., Meliga, P., Serre, E.: An immersed boundary method in Tabor: verification and validation. Comput. Fluids **157**, 55–72 (2017)

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. **99**(C5), 10143 (1994). https://doi.org/10.1029/94JC00572

Greenshields, C.: OpenFOAM V9 User Guide. In: The OpenFOAM Foundation, London, UK (2021). https://doc.cfd.direct/openfoam/user-guide-v9

Gritskevich, M., Garbaruk, A., Schütze, J., Menter, F.: Development of ddes and iddes formulations for the k-omega shear stress transport model. Flow Turbul. Combust. **88**, 431 (2012)

Grozescu, A.N., Salvetti, M.V., Camarri, S., Buresti, G.: Variational multiscale large-eddy simulations of the barc flow configuration. In: Proceedings of the thirteenth International Conference on Wind Engineering (2011)

Guo, Q., Liu, P., Li, C., Sun, D., Yuan, X.: On the characteristic length scale for the synthetic turbulence based on the Spalart-Allmaras model (2023). Preprint at https://arxiv.org/abs/2304.02805

Haan, F.L., Kareem, A.: Anatomy of turbulence effects on the aerodynamics of an oscillating prism. J. Eng. Mech. **135**(9), 987–999 (2009). https://doi.org/10.1061/(ASCE)EM.1943-7889.0000012

Haffner, Y., Heitz, D., Perret, L., Braud, P.: Experimental Dataset of Turbulent Flow Around a 5:1 Rectangular Cylinder with Geometric Variations. (2023) https://doi.org/10.57745/WDIKG9

Haffner, Y., Li, R., Meldi, M., Borée, J.: Drag reduction of a square-back bluff body under constant crosswind conditions using asymmetric shear layer forcing. Int. J. Heat Fluid Flow **96**, 109003 (2022)

Hesse, F., Morgans, A.S.: Simulation of wake bimodality behind squareback bluff-bodies using les. Comput. Fluids **223**, 104901 (2021). https://doi.org/10.1016/j.compfluid.2021.104901

Hou, E., Lawrence, E., Hero, A.O.: Penalized ensemble Kalman filters for high dimensional non-linear systems. PLOS One **16**(3), 0248046 (2021). https://doi.org/10.1371/journal.pone.0248046

Issa, R.I., Gosman, A.D., Watkins, A.P.: The computation of compressible and incompressible recirculating flows by a non-iterative implicit scheme. J. Comput. Phys. **62**(1), 66–82 (1986). https://doi.org/10.1016/0021-9991(86)90100-2

Kalman, R.E.: A new approach to linear filtering and prediction problems. J. Basic Eng. **82**(1), 35–45 (1960). https://doi.org/10.1115/1.3662552

Lamberti, G., García-Sánchez, C., Sousa, J., Gorlé, C.: Optimizing turbulent inflow conditions for large-eddy simulations of the atmospheric boundary layer. J Wind Eng. Ind. Aerodyn. **177**, 32–44 (2018). https://doi.org/10.1016/j.jweia.2018.04.004

Le, T.-H., Matsumoto, M., Shirato, H.: Spanwise coherent structure of wind turbulence and induced pressure on rectangular cylinders. Wind Struct. **12**(5), 441–455 (2009)

Mannini, C., Schewe, G.: Numerical study on the three-dimensional unsteady flow past a 5:1 rectangular cylinder using the DES approach. In: ICWE13 - 13th International Conference on Wind Engineering (2011)

Mannini, C., Šoda, A., Schewe, G.: Unsteady rans modelling of flow past a rectangular cylinder: investigation of reynolds number effects. Comput. Fluids **39**(9), 1609–1624 (2010). https://doi.org/10.1016/j.compfluid.2010.05.014

Mannini, C., Marra, A.M., Pigolotti, L., Bartoli, G.: The effects of free-stream turbulence and angle of attack on the aerodynamics of a cylinder with rectangular 5:1 cross section. J. Wind Eng. Ind. Aerodyn. **161**, 42–58 (2017). https://doi.org/10.1016/j.jweia.2016.12.001

Mariotti, A., Salvetti, M.V., Shoebi-Omrani, P., Witteveen, J.A.S.: Stochastic analysis of the impact of freestream conditions on the aerodynamics of a rectangular 5:1 cylinder. Comput. Fluids **136**, 170–192 (2016)

Mariotti, A., Siconolfi, L., Salvetti, M.V.: Stochastic sensitivity analysis of large-eddy simulation predictions of the flow around a 5:1 rectangular cylinder. Eur. J. Mech. -B Fluids **62**, 149–165 (2017). https://doi.org/10.1016/j.euromechflu.2016.12.008

Mariotti, A., Lunghi, G., Salvetti, M.V.: Experimental investigations on the effect of upstream-edge rounding on the barc configuration. J Wind Eng. Ind. Aerodyn. **251**, 105787 (2024). https://doi.org/10.1016/j.jweia.2024.105787

Matsumoto, M., Tamura, Y.: Survey for the aeroelasticity of structures. In: Murakami, S. (ed.) Computational Wind Engineering 1, pp. 873–877. Elsevier, Oxford (1993). https://doi.org/10.1016/B978-0-444-81688-7.50101-7 . https://www.sciencedirect.com/science/article/pii/B9780444816887501017

Meldi, M., Sagaut, P., Lucor, D.: A stochastic view of isotropic turbulence decay. J. Fluid Mech. **668**, 351–362 (2011)

Meldi, M., Salvetti, M.V., Sagaut, P.: Quantification of errors in large-eddy simulations of a spatially evolving mixing layer using polynomial chaos. Phys. Fluids **24**, 035101 (2012)

Minguez, M., Pasquetti, R., Serre, E.: High-order large-eddy simulation of flow over the Ahmed body car model. Phys. Fluids **20**(9), 095101 (2008). https://doi.org/10.1063/1.2952595

Moldovan, G., Mariotti, A., Cordier, L., Lehnasch, G., Salvetti, M.-V., Meldi, M.: Multigrid sequential data assimilation for the large-eddy simulation of a massively separated bluff-body flow (2022)

Moldovan, G., Lehnasch, G., Cordier, L., Meldi, M.: Optimized parametric inference for the inner loop of the Multigrid Ensemble Kalman Filter. J. Comput. Phys. **471**, 111621 (2022). https://doi.org/10.1016/j.jcp.2022.111621

Noda, H., Nakayama, A.: Free-stream turbulence effects on instantaneous pressure and forces on cylinders of rectangular cress section. Exp. Fluids **34**, 332–344 (2003). https://doi.org/10.1007/s00348-002-0562-0

Pasqualetto, E., Lunghi, G., Rocchio, B., Mariotti, A., Salvetti, M.V.: Experimental characterization of the lateral and near-wake flow for the BARC configuration. Wind Struct. **11**, 1–22 (2021)

Patruno, L., de Miranda, S.: Unsteady inflow conditions: a variationally based solution to the insurgence of pressure fluctuations. Comput. Methods Appl. Mech. Eng. **363**, 112894 (2020). https://doi.org/10.1016/j.cma.2020.112894

Patruno, L., Ricci, M., Miranda, S., Ubertini, F.: Numerical simulation of a 5:1 rectangular cylinder at non-null angles of attack. J. Wind Eng. Ind. Aerodyn. **151**, 146–157 (2016)

Poletto, R., Craft, T., Revell, A.: A new divergence free synthetic eddy method for the reproduction of inlet flow conditions for LES. Flow Turbul. Combust. **91**(3), 519–539 (2013). https://doi.org/10.1007/s10494-013-9488-2

Pope, S.B.: Turbulent Flows. Cambridge University Press, Cornell University, New York (2000). https://doi.org/10.1017/CBO9781316179475 . https://www.cambridge.org/core/books/turbulent-flows/69322053C06F73F7EB7124915F9256BD

Ricci, M., Patruno, L., de Miranda, S., Ubertini, F.: Flow field around a 5:1 rectangular cylinder using les: influence of inflow turbulence conditions, spanwise domain size and their interaction. Comput. Fluids **149**, 181–193 (2017). https://doi.org/10.1016/j.compfluid.2017.03.010

Ricciardelli, F., Marra, A.: Sectional aerodynamic forces and their longitudinal correlation on a vibrating 5:1 rectangular cylinder, (2008)

Rocchio, B., Mariotti, A., Salvetti, M.V.: Flow around a 5:1 rectangular cylinder: effects of upstream-edge rounding. J Wind Eng. Ind. Aerodyn. **204**, 104237 (2020). https://doi.org/10.1016/j.jweia.2020.104237

Sagaut, P.: Large-eddy Simulation for Incompressible Flows. Introd. (2005). https://doi.org/10.1007/978-3-662-04416-2

Selma, B., Désilets, M., Proulx, P.: Optimization of an industrial heat exchanger using an open-source CFD code. Appl. Therm. Eng. **69**(1–2), 241–250 (2014)

Shur, M., Strelets, M., Travin, A., Probst, A., Probst, S., Schwamborn, D., Deck, S., Skillen, A., Holgate, J., Revell, A.: Improved Embedded Approaches. In: Mockett, C., Haase, W., Schwamborn, D. (eds.) Go4Hybrid: Grey Area Mitigation for Hybrid RANS-LES Methods vol. 134, pp. 51–87. Springer, Cham (2018). Series Title: Notes on Numerical Fluid Mechanics and Multidisciplinary Design

Shur, M.L., Spalart, P.R., Strelets, M.K., Travin, A.K.: A hybrid RANS-LES approach with delayed-DES and wall-modelled LES capabilities. Int. J. Heat Fluid Flow **29**(6), 1638–1649 (2008). https://doi.org/10.1016/j.ijheatfluidflow.2008.07.001

Spalart, P.R., Deck, S., Shur, M.L., Squires, K.D., Strelets, M.K., Travin, A.: A new version of detached-eddy simulation, resistant to ambiguous grid densities. Theor. Comput. Fluid Dyn. **20**(3), 181–195 (2006). https://doi.org/10.1007/s00162-006-0015-0

Tabor, G.R., Baba-Ahmadi, M.H.: Inlet conditions for large eddy simulation: a review. Comput. Fluids **39**(4), 553–567 (2010)

Tandeo, P., Ailliot, P., Bocquet, M., Carrassi, A., Miyoshi, T., Pulido, M., Zhen, Y.: Joint estimation of model and observation error covariance matrices in data assimilation: a Review. Mon. Weather Rev. **148**, 3973–3994 (2020). https://doi.org/10.1175/MWR-D-19-0240.1

Timilsina, P.: Benchmark on the aerodynamics of a 5:1 cylinder using nek5000. (2015). https://api.semanticscholar.org/CorpusID:116292564

Valente, P.C., Vassilicos, J.C.: The decay of turbulence generated by a class of multiscale grids. J. Fluid Mech. **687**, 300–340 (2011)

Villanueva, L., Truffin, K., Meldi, M.: Synchronization and optimization of Large Eddy Simulation using an online Ensemble Kalman Filter. Preprint at arxiv:2310.18016 (2023)

Villanueva, L., Valero, M.M., Glumac, A., Meldi, M.: Augmented state estimation of urban settings using on-the-fly sequential data assimilation. Comput. Fluids **269**, 106118 (2024). https://doi.org/10.1016/j.compfluid.2023.106118

Wei, Z., Kareem, A.: A benchmark study of flow around a rectangular cylinder with aspect ratio 1:5 at Reynolds number 1.E5. In: Proceedings of the thirteenth International Conference on Wind Engineering, Amsterdam (2011)

Wilcox, D.C.: Turbulence Modeling for CFD, 3rd edn. DCW Industries Inc., La Canada CA (2006)