State Observer data assimilation for RANS with time-averaged 3D-PIV data

Saredi, E.; Tumuluru Ramesh, Nikhilesh; Sciacchitano, A.; Scarano, F.

DOI
10.1016/j.compfluid.2020.104827

Publication date
2021

Document Version
Final published version

Published in
Computers & Fluids

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.
State observer data assimilation for RANS with time-averaged 3D-PIV data

Edoardo Saredi\textsuperscript{a}\textsuperscript{*}, Nikhilesh Tumuluru Ramesh\textsuperscript{b}, Andrea Sciacchitano\textsuperscript{a}, Fulvio Scarano\textsuperscript{a}

\textsuperscript{a}TU Delft, Faculty of Aerospace Engineering, Delft, The Netherlands
\textsuperscript{b}University of Waterloo, Mechanical and Mechatronics Engineering, Waterloo, Canada

\textbf{ARTICLE INFO}

\textbf{Article history:}
Received 7 July 2020
Revised 4 December 2020
Accepted 22 December 2020
Available online 29 December 2020

\textbf{Keywords:}
Data assimilation
Robotic Volumetric PIV
RANS
state observer

\textbf{ABSTRACT}

State observer techniques are investigated for the assimilation of three-dimensional velocity measurements into computational fluid dynamics simulations based on Reynolds-averaged Navier–Stokes (RANS) equations. The method relies on a forcing term, or observer, in the momentum equation, stemming from the difference between the computed velocity and the reference value, obtained by measurements or high-fidelity simulations. Two different approaches for the forcing term are considered: proportional and integral-proportional. This technique is demonstrated considering an experimental database that describes the time-average three-dimensional flow behind a generic car-mirror model. The velocity field is obtained by means of Robotic Volumetric PIV measurements. The effects of the different forcing terms and the spatial density of the measurement input to the numerical simulation are studied. The state observer approach forces locally the solution to comply with the reference value and the extent of the region modified by the forcing input is discussed. The velocity distribution and flow topology obtained with data assimilation are compared with attention to the object wake and the reattachment point where the largest discrepancy is observed between the different approaches. The results show that the integral term is more effective than the proportional one in reducing the mismatch between simulation and the reference data, with increasing benefits when the density of forced points, or forcing density, is increased.

\textsuperscript{*}Corresponding author.
E-mail address: e.saredi@tudelft.nl (E. Saredi).

1. Introduction

The advances of computational fluid dynamics (CFD) coupled with the increasingly affordable computational power have made CFD one of the main tools for aerodynamic study and design optimization for industrial problems [40]. The computational cost of high-fidelity simulations (i.e. LES, DNS), instead, remains relatively high for their use in optimization studies, especially for high Reynolds number flows. As a result, CFD analysis makes most often use of the simplified approach offered by the Reynolds-averaged Navier-Stokes (RANS) formulation [1,8]. Considered to be the workhorse in aerodynamic engineering for many years to come [40], RANS solvers model the entire turbulent spectrum, achieving lower computational cost and higher robustness with respect to mentioned higher fidelity models. The cost of the amount of required modelling is the aforementioned lower accuracy [5]. Several sources of uncertainty are presented in literature; those stemming from the choice of the turbulence model and its parameter, nowadays based mainly on expert judgment, are reported to often dominate the overall uncertainty of the simulation [40].

Conversely, the experimental approach offers accurate measurements, directly from scaled models. When flow quantities are obtained such as velocity or pressure, the measurements usually suffer from limited spatial range and spatio-temporal resolution. For this reason, it should be retained in mind that the sole experiment is often insufficient for the analysis of the flow problem, which renders the simulations necessary for a complete description of the aerodynamic behaviour. Experimental data is often used “offline”, to verify the validity of a given simulation, or more specifically, the adequacy of turbulence models parameters (also referred to as “calibration” [6,27]). The latter approach is often iterative and several simulations are conducted until model calibration is achieved.

In order to overcome the above limitations, direct integration of experimental data within CFD simulations has been considered. The resulting process is called data assimilation (DA). The aim of this approach is to produce a more accurate flow simulation, which is compliant with the data gathered during an experiment.

Originally introduced in the fields of meteorology and oceanography, as reviewed by Navon [24], DA has seen a growing number of applications in fluid mechanics [11]. Among the DA method-
ologies presented in the literature, three main categories have emerged: variational methods ([9,35]; among others), Kalman filter (KF; [17,18]) and state observer methods (reviewed in [11]).

Variational methods are based on the application of optimal control theory to find the minimum of the error function between the simulation and a reference, herein represented by the experimental data. The calculation of the adjoint and the need to solve an optimization problem entails a relatively large computational cost. KF and state observer techniques are sequential algorithms, which means that the result of the previous iteration is used at each following iteration. The solution of the assimilated simulation is expected to asymptotically converge to the reference along the simulation [11]. A comparison between a variational method and a KF application can be found in Mons et al. [23]. If the equations of state observer and KF are compared, a certain similarity can be noted [36]. KF can be considered a type of state observer method too, but designed for stochastic systems, while state observer techniques are typically applied to deterministic systems. While in a deterministic system, the variables of interest are represented by a scalar or a vector at each point, in a stochastic system they are represented by probability density functions, with a correspondent mean and standard deviation. Between the three of them, state observer methods require both a computational cost and an implementation effort that are significantly lower compared to variational methods, making them attractive for aerodynamic design and optimization [11].

The state observer algorithm, originally proposed by Luenberger [20] is nowadays a key concept of the control theory, based on a feedback term in the modelling equation. The intensity of the feedback is a function of the difference between the model results and the reference (viz. experimental) data multiplied by a gain factor. The latter is usually tuned to maximise convergence rate. The design of the feedback and its law of application is the key point of a state observer algorithm [11]. Hayase and Hayashi [12] applied a state observer algorithm to improve the simulation of a fully developed turbulent square duct. The authors modified the pressure boundary conditions at the inlet and the outlet with a proportional law in order to reduce the error between the simulation and the ground truth, which in that case was a pre-calculated numerical solution performed on the same mesh with a different initial condition. The application of the feedback gave an accelerated convergence of the simulation and a reduction of one order of magnitude in the final error across the whole domain. This approach alleviates the errors given by the erroneous boundary conditions but is not able to tackle the error given by the approximation performed in the turbulence modelling.

Imagawa and Hayase [13] performed a Measurement-Integrated (MI) simulation for the turbulent flow along a square duct and applied a feedback by means of a body force term in the momentum equation. The forcing term was linearly proportional (DASOP) to the difference between the velocity returned by the unsteady simulation and that from a reference simulation performed with the same numerical schemes on the same grid but with different initial conditions. The application of the feedback led to a reduction of the steady state error by four orders of magnitude when all the velocity components were forced at all the grid points. However, interrupting the assimilation during the simulation caused the system to convergence again towards the non-assimilated solution. The authors also considered the case of a limited density of forcing points, forcing from one plane every four down to only one plane. They concluded that, with their forcing term, tuning the gain factor, it was possible to obtain the same error reduction considering one plane every four. Furthermore, disabling the forcing term at a quarter of the total running time, the simulation was naturally converging towards the not assimilated simulation, with the effect of the forcing term effect that was vanishing. This shows a typical behaviour of a proportional controller, with the efficacy of the forcing term that reduces when the error reduces. Neeteson and Rival [25] applied the state observer algorithm to the Karman vortex shedding problem at Re = 10^2, using a computational reference. They introduced a Proportional-Integral-Derivative (PID) control in the pressure equation, which was formulated as feedback law. The latter was more elaborate than the sole proportional law used in most previous works and aimed at solving the reduction of effectiveness of the forcing term at the low error stages. The application of the PID control law (DASOPID) improved the results compared to those given by the sole proportional feedback law, returning a vortex shedding behaviour closer to that of the reference data in terms of shedding frequency. The state observer algorithm has been used also with experimental data as a reference. Yama- gata et al. [39] studied the unsteady behaviour of the Karman vortex street behind a truncated cylinder at Re = 1.2 × 10^3 performing a MI simulation with planar PIV data input. The assimilation allowed reproduction of the large-scale unsteadiness which are not obtained by ordinary simulations.

The discussion above shows that the state observer data assimilation has been considered to enhance the accuracy of RANS-based CFD simulations, mainly in the unsteady flow regime and at low Re number. However, many applications of industrial aerodynamics are primarily concerned with the accuracy of the steady-state flow solution. Moreover, many relevant problems involve high Reynolds number flows and a fully turbulent regime for which many questions about the applicability and the potential of DA remain unanswered.

This study considers steady-state simulations of the flow around a wall-mounted bluff obstacle, representing the geometry of a simplified car mirror presented by de Villiers [37]. The problem is studied at Re = 8 × 10^3. The reference time-averaged velocity field is obtained from experiments that make use of Robotic Volumetric PIV [15]. Two formulations of the State Observer forcing term are considered, namely Data-Assimilation-State-Observer-Proportional (DASOP) and Data-Assimilation-State-Observer-Proportional-Integral (DASOPI). Moreover, the work investigates the extent of the regions affected by the forcing point, with the aim of identifying a criterion for optimum data assimilation forced points density. Finally, a metric is introduced to quantify the effect of DA methods and parameters on the accuracy of the assimilated solutions with respect to the reference velocity.

2. State observer data assimilation for RANS

Reynolds-averaged Navier-Stokes (RANS) equations are considered here as the framework for the numerical simulation of incompressible turbulent fluid flows. The baseline simulation is intended as the process leading to a solution (baseline solution) with no use of a-priori information from the experiments. The working principle of data assimilation is to introduce a certain amount of reference data, indicated here as \( \bar{u}_{ref} \), and force the simulation to comply with it. In this work, the resulting process is called the assimilated simulation, which returns the assimilated solution. The objective is to drive the assimilated solution towards a more correct evaluation of the relevant features present in the flow field (e.g. flow topology, separation, reattachment, pressure gradient distribution, aerodynamic loads and forces), thus minimising simulation errors, such as the ones associated with the turbulence modelling.

The underlying concepts and equations of data assimilation through a state observer (DASO) algorithm are recalled here. Applying Reynolds decomposition and time averaging, the dynamic behaviour of a steady, incompressible, viscous and turbulent flow can be described by the RANS equations, formed by the combination of the conservation of mass and momentum, respectively:

\[
\nabla \cdot \bar{u} = 0
\]

(1)
imposing consistent initial conditions (IC) and boundary conditions (BC) and where $R_i = u_i^* u_j^*$ represents the Reynolds stress tensor. Due to the incompressible flow condition, $p$ represents the static pressure divided by the constant density. As shown by Imagawa and Hayase [13], the assimilation of the data through a state observer algorithm is achieved by introducing a body force term $f$ into the momentum equation:

\[
(\bar{u} \cdot \nabla) \bar{u} = -\nabla \tilde{p} + \nu \Delta \bar{u} - \nabla \cdot R + f
\]

Under the hypothesis of steady flow conditions, in the following discussion the overbar indicating the time-average has been omitted, although all the quantities are meant to be time-averaged. The relative strength of the body force term represents the feedback and it is proportional to the local difference $\epsilon$ between the simulation and the experimental time-averaged velocities:

\[
\epsilon = (u_{ref}^* - u)
\]

where $\epsilon_{ref} = \mathcal{P}(u_{ref})$, with $\mathcal{P}$ being the operator that projects velocity information from the experimental grid to the computational grid. Further information on $\mathcal{P}$ are given in Section 4.1.

In this work, two different forms of the feedback term have been considered: data assimilation based on state observer with proportional feedback (DASOP) and on proportional-integral feedback (DASOPI). The feedback control law used in the DASOP method reads as:

\[
f_{\text{DASOP}} = \frac{K_p}{D} \mathcal{P}(\epsilon) = \left[\frac{K_p}{D} \mathcal{P}(\epsilon)\right]^T
\]

where $K_p$ is the proportional feedback gain, the symbol $\mathcal{P}$ represents the Hadamard product [34] and $|\epsilon|$ is the component-wise absolute value of the error $|\epsilon| = |\epsilon_x|, |\epsilon_y|, |\epsilon_z|$. The forcing term for the DASOPI method reads as:

\[
f_{\text{DASOPI}} = \frac{K_p}{D} \mathcal{P}(\epsilon) + \frac{K_i}{N} \sum_{n=1}^{N-1} \mathcal{P}(\epsilon_n) \mathcal{P}(\epsilon_n)
\]

where $N$ is the current iteration and $K_i$ is the integral feedback gain and $\mathcal{P}(\epsilon)$ is the local error vector at the iteration $n$. While state observers are typically applied to unsteady systems and act in the time domain, in the methodology here proposed the state observer acts in the iteration domain, for the determination of a steady-state solution. The gains $K_p$ and $K_i$ are scalars as they are assumed equal for each component of the error and constant along the simulation and spatially in the entire simulation domain. The values corresponding to cells where $u_{ref}^*$ is not available are set to 0, thus disabling the forcing term. The quadratic term implemented in both the proposed feedback terms resembles an error-squared controller, as proposed in Shinskey [32]. Its advantage is that the forcing term is strengthened with respect to a linear controller when the error is large, penalizing the regions where the difference between the reference and the simulated velocity are lower.

The results obtained at the end of the simulation represents the assimilated solution. The presented methodology implies the usage of a single simulation to reach the final assimilated solution, contrary to ensemble Kalman filters and variational methods, which require multiple simulations to reach the final assimilated solution.

3. Setup of the reference dataset experiment

3.1. Wind tunnel and model

Experiments are performed at the TU Delft Aerodynamics Laboratories in an open-jet open-circuit low-speed wind tunnel (WTunnel). The tunnel features a 4:1 area contraction, after which the air flow reaches the free-stream velocity $U_{\infty} = 12 \text{ m/s}$ in a cross-section of $60 \times 60 \text{ cm}^2$. A flat plate 1.5 m long is installed at 10 cm height above the bottom edge of the exit (Fig. 1). The plate has a sharp leading edge and is equipped with a zig-zag tripping device at 5 cm past the leading edge that forces the boundary layer to the turbulent regime. The model is a half cylinder of diameter $D = 10 \text{ cm}$ topped by a quarter sphere (Fig. 1) with a total height of 15 cm, as that used in the investigations of de Villiers [37]. The Reynolds number based on the model diameter is $Re_D = 8 \times 10^4$. This geometry represents a benchmark for automotive aerodynamics for the study of the flow around appendices and in particular the side mirror. It reproduces the essential features of junction flow and bluff body aerodynamics, including horseshoe vortex, sharp separation at the back, and large-scale fluctuations of the wake and of the reattachment location [37].

3.2. Robotic Volumetric PIV

The three-dimensional velocity field is measured by Robotic Volumetric PIV [15]. An illustration of the system configuration during the experiments is shown in Fig. 1. The system is composed of a coaxial volumetric velocimeter (CVV), containing 4 high-speed cameras, and laser illumination through optic fiber. The CVV is installed on a robotic arm, UR5 from Universal Robots. The robot motion sequence is programmed through the proprietary software RoboDK and operated through the LaVision software DaVis 10. Synchronization of illumination and image acquisition is made through a programmable timing unit (LaVision PTU 9). Helium-Filled-Soap-Bubbles (HFSB) are used as flow tracers ([32], among others). The bubbles are nearly neutrally buoyant with a median diameter of approximately 0.3 mm [7]. A rake composed of 10 vertical elements hosting 200 bubble generators [4] is installed in the settling chamber of the wind tunnel. Soap, air and helium supply is controlled through a LaVision fluid supply unit (FSU). The flow is seeded at a concentration of approximately 0.3 bubbles/cm$^3$. Table 1 summarises the experimental parameters. The reader is referred to the works of Jux et al. [15,16] and Schneiders et al. [30] for a deeper discussion of the working principles of this.
measurement technique. Here the experimental procedure is discussed. Prior to the measurements, a calibration of the robot position w.r.t. the measurement domain (wind tunnel, plate and object) is performed. Each measurement volume spans approximately $30 \times 20 \times 40$ cm$^3$. At each robot position, 20,000 recordings are acquired in time-resolved mode at a frequency of 857 Hz ($T_{rec} \sim 23$ s). The time elapsed from one position to the subsequent is approximately 270 s. The overall measurement encompasses 15 views, where the measurement system is directed by robotic arm manipulation. At the end of the measurements, the raw data features a coverage of the measurement domain, with the datasets obtained from each viewing position. The procedure for data processing and reduction is discussed in the next section.

### 3.3. Data processing and reduction

The raw images are pre-processed to reduce background reflections using the Butterworth high-pass filter [31]. The tracer motion analysis is performed with the Lagrangian particle tracking algorithm Shake-The-Box [29]. Tracks containing more than six appearances of the particle tracer are accepted as valid. At each recording, a sparse measurement of the tracers’ velocity is obtained in the sub-volume.

The measurements from different robot positions are merged into a single dataset using the robot calibration data. The resulting domain is interrogated within cubic voxels (or bins) of 15 mm side length. Within each bin, the tracers’ velocity is ensemble-averaged, yielding the time averaged velocity vector distribution in a Cartesian grid. The averaging process inside the bin follows a spatially weighted algorithm with respect to the centroid of the bin. A Gaussian weighting function (with a width of half the bin size) is applied, following Agüera et al. [2]. Partial overlap by 3:4 of neighbouring voxels yields velocity vectors spaced by 3.75 mm. The result is rendered in a domain of $55 \times 30 \times 25$ cm$^3$ (Fig. 1) with a grid of $151 \times 84 \times 69$ data points describing the time-average velocity $\bar{u}_{\text{tot}}$ for use in the data assimilation procedure.

### 4. Numerical simulations

Both baseline and assimilated simulations are based on a RANS solver using the open-source C++ toolbox OpenFoam 1706 [14]. Eqs. (1) and (2) are discretised using the finite volume method on a collocated grid and solved using the SIMPLE algorithm [26]. For the baseline simulation, the steady solver simpleFoam implemented in OpenFoam is used. For the assimilated simulations, the feedback term is included within an in-house developed version of the same solver. Turbulence modelling is based on the $k$-$\omega$ SST model [22]. The computational domain is a square cuboid with dimensions $30 \times 20 \times 10$ diameters. The object is placed along the centreline of the domain, with the origin of the coordinate system posed at the intersection between the rear surface of the car mirror and the ground. The orientation of the coordinate system is shown in Fig. 1. An inlet condition is imposed at $X/D = -10$ and an outlet condition is imposed at $X/D = 20$. No-slip condition is imposed at the ground floor and the object surfaces, whereas slip conditions are applied to the top and side faces of the domain. A summary of the boundary and initial condition is reported in Table 2. The same hexahedral mesh is used for both the baseline and the assimilated simulations, formed by ~18.5 M cells and created with the commercial software CFMesh+. The region close to the surface of the object ($-7 < X/D < 15$ and $-7 < Y/D < 7$) is refined at several stages, as shown in Fig. 2, where the cell dimension reduces from 10 mm to 0.78 mm, corresponding to $y^+ < 5$. All the simulations advance for 5000 iterations and convergence is verified by reaching a relative variation of the averaged drag coefficient of the object $|C_{d,n} - C_{d,n-1}|$ below $10^{-3}$, where the average

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characteristics of the Robotic Volumetric system and the experimental parameters.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeding</td>
<td>Neutrally buoyant HSFL, $\sim 300$ µm diameter</td>
</tr>
<tr>
<td>Illumination</td>
<td>Quantronix Darwin-Duo Nd:YLF laser ($2 \times 25$ mJ @ 1 kHz)</td>
</tr>
<tr>
<td>Recording Device</td>
<td>LaVision MiniShaker Aero system: 4 x CCD cameras (540 × 452 @ 857 Hz)</td>
</tr>
<tr>
<td>Imaging</td>
<td>$f = 4$ mm, $f_r = 8$</td>
</tr>
<tr>
<td>Acquisition frequency</td>
<td>$f_r = 857$ Hz</td>
</tr>
<tr>
<td>Pulse separation time</td>
<td>Time-resolved: $\Delta t = 1/f_r = 1.17$ ms</td>
</tr>
<tr>
<td>Magnification factor</td>
<td>$M = 0.01$ at 30 cm distance</td>
</tr>
<tr>
<td>Number of recordings per region</td>
<td>$N = 20,000$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Boundary and initial conditions of the simulations for velocity and pressure.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inlet $X/D = [-10,20]; Y/D = [-10,10]; Z/D = [0,1]$</td>
<td>Dirichlet, $u = (12,0,0)$ m/s, $\bar{p}_w = 0$ Pa</td>
</tr>
<tr>
<td>Outlet $X/D = [-10,20]; Y/D = [-10,1]; Z/D = [0,1]$</td>
<td>Neumann, no backflow, $\frac{\partial p}{\partial n} = 0$</td>
</tr>
<tr>
<td>Object surface</td>
<td>no-slip condition</td>
</tr>
<tr>
<td>Lateral walls</td>
<td>slip condition</td>
</tr>
<tr>
<td>Top of the domain $Z/D = 10$</td>
<td>slip condition</td>
</tr>
<tr>
<td>Initial condition (entire domain)</td>
<td>$u = (12,0,0)$ m/s, $p = 0$ Pa</td>
</tr>
</tbody>
</table>

Fig. 2. Computational mesh around the object along the symmetry plane of the simulation domain.
is performed in the interval \([n=500...n]\). The same interval is used to average the velocity field to take out numerical oscillations.

4.1. Assimilation procedure

After conducting the baseline simulation, several assimilated simulations are run to investigate the effects of assimilation. The analysis presented here aims at determining the effect of sparse forcing varying the concentration of forced cells. The latter is represented in terms of the normalised mean distance between neighbouring forcing cells \(γ = λ/D\), where \(λ\) is the mean distance between neighbouring forcing points. In order to evaluate \(λ\), the following equation is used:

\[
λ = \frac{2}{3πC}
\]

where \(C = N_f/V\), with \(N_f\) equal to the number of forced cells, selected randomly, and \(V\) the total forced volume.

The small positional mismatch between cells where forcing is applied and the locations where experimental data is available is accommodated by linear interpolation to the closest experimental data points, obtaining the velocity field \(\mathbf{u}_f\). This operation is then fully justified in cases in which the expected interpolation error are negligible, as in the case considered in this work. Furthermore, in order to damage the assimilated solution, outliers have to be removed from \(\mathbf{u}_f\) before the interpolation step. In the other cells, both the values of \(\mathbf{u}_f\) and \(E\) are set to zero. At each iteration of the assimilated simulation, the velocity field \(\mathbf{u}_f\) is compared with the velocity field \(\mathbf{u}\) of the previous iteration and the error field \(E\) is explicitly constructed according to Eqs. (5) and (6), respectively for the DASOP and DASOPI feedback law. The limited spatial resolution of the ensemble-averaged PIV velocity data close to the wall leads to prescribed reference velocity \(\mathbf{u}_r\), too high and not compatible with the no-slip condition assigned to the walls of the simulation. For this reason, PIV data closer than 3 cm (two averaging sub-volumes in the PIV analysis) have not been considered to build the forcing term \(f\). In relation to this, Fig. 3 shows a section along the symmetry plane of the regions where the forcing term was activated, with different forcing cell density, depending on the choice of \(γ\).

5. Results and discussion

5.1. Reference flow field

The measured velocity field \(\mathbf{u}_r\) is visualized in Fig. 4, where the streamwise velocity component \(\mathbf{u}_r\) is shown at the symmetry plane and in a wall-parallel plane at a height \(Z/D = 0.5\). The potential flow region in front of the object is visible, where the flow decelerates and eventually stagnates at the obstacle due to the adverse pressure gradient (Fig. 4 - bottom). The boundary layer that has developed along the plate is also visible in section \(Y = 0\) (Fig. 4 - top). Acceleration atop the obstacle is consistent with the streamlines curvature that follows the object head (Fig. 4 - top). The local velocity in this region exceeds the free-stream value by approximately 20%. The boundary layer developing along the object surface separates abruptly due to the sharp truncation of the object at the trailing edge. Within the separated zone, a reverse flow region is formed. This feature has been observed past a number of wall-mounted obstacles and reported frequently in the literature ([21,37,38]; among others). The separated region terminates downstream with a ground-reattachment line with its most downstream position at \(X_S\) in the symmetry plane. In the present experiment, the length of separation is observed to be \(X_S/D = 2.4\).

At the edge of the shear layer and the backflow region, from Fig. 4 (bottom) is also possible to visualize two counter-rotating vortices from the streamlines pattern. The correspondent foci are labelled \(F_2\) and \(F_3\) in Fig. 4 and in the rest of this work. As reported by de Villiers (2006) [37], these are part of a \(U\)-shaped structure that appears in the time-averaged velocity field, positioned upside-down w.r.t. the object and tilted towards the back surface of the mirror. The point in which this structure crosses the symmetry plane is detectable in Fig. 4 (top) by the foci named \(F_i\).

The analysis of the vorticity field, illustrated in Fig. 5, provides further insights on the reference flow topology. The adverse pressure gradient at the front of the object causes the boundary layer to detach from the ground and form the horseshoe vortex [33], which develops around the object, eventually aligning streamwise and further develop downstream of the object wake. The development from the front of the object of the horseshoe vortex is shown in Fig. 5 by the two portions of the iso-surface where the magnitude of the streamwise vorticity \(|ω_z| > 200\) Hz. As described by de Villiers [37], the time-averaged separation region is bounded by an arc-like free shear layer, which develops downstream and bounds a region of flow recirculation that also features arc-like shape.
5.2. Baseline solution

The flow field topology returned by the baseline simulation, as shown in Fig. 6, largely follows the flow pattern observed in the experiments. The boundary layer development under the adverse pressure gradient caused by the obstacle is visible, with the formation of a small recirculation ahead of the object (head of the horseshoe vortex) at approximately X/D = -0.8. The overall flow deceleration ahead of the object and subsequent acceleration towards the trailing edge are also visible by regions at a velocity exceeding the free-stream value. In the region upstream of the object, the result obtained by the baseline simulation complies rather well with the reference, as visible in Fig. 7. This is expected being RANS able to predict accurately flow feature in potential flow regions. A notable exception is instead the region close to the ground, where the simulation is not forced. The topological differences are discussed in the remainder. The flow separation at the sharp trailing
edge is also well reproduced. Whereas the most significant discrepancy with respect to the reference velocity field is produced in the separated flow region and the reattachment region in particular. The RANS data predict a longer wake, as clearly visible in Fig. 7, with a reattachment occurring at \( X_\ell/D = 3.1 \) along the symmetry plane. This behaviour has already been reported in previous works (199, among others), where RANS simulations tend to overestimate the length of the wake of a bluff body. Together with the elongation of the wake behind the object, other topological features in the wake are distorted. As visible in Fig. 6, combining the information obtained by the two sections, it is possible to infer the presence of the arc-shaped vortex. The latter has, however, a pronounced offset (judged from the location the two foci at \( Z/D = 0.5 \)) from \( X/D = 1.25 \) of the reference data, to \( X/D = 1.85 \). In the symmetry plane instead, the focus \( F_1 \) is located closer to the ground \( Z/D = 0.8 \) in the simulation, compared to \( Z/D = 1.05 \) of the reference data). The resulting arc-vortex in the baseline simulation is more elongated in the streamwise direction and with a lower position of its top portion. The latter is also visible considering the region of reverse flow (in Fig. 7 by the dashed line of the reference data is superimposed to the baseline simulation) appearing flattened towards the ground.

Further small differences are the upstream separation forming the head of the horseshoe vortex, a small recirculation region at the free end of the object and a small region of accelerated flow close to the object trailing edge (Fig. 6). These features are, however, not captured within the experiments due to the limited spatial resolution of the measurements.

An L-2 metrics is introduced to quantify the discrepancy between the baseline simulation and the reference data. The following relative error normalised with the free stream velocity \( U_\infty \) has been used:

\[
\varepsilon = \sqrt{\left( \frac{u - u_{ref}}{U_\infty} \right)^2 + \left( \frac{v - v_{ref}}{U_\infty} \right)^2 + \left( \frac{w - w_{ref}}{U_\infty} \right)^2}
\]

(7)

The reference data is available on a grid that does not coincide with that of the numerical simulations. The value of \( \varepsilon \) is therefore obtained by linear interpolation of the numerical simulations onto the grid of the reference data. The spatial distribution of \( \varepsilon \) at the symmetry plane is shown in Fig. 8. The relative error distribution can be clustered in four different regions. The widest region of error can be found for \( X/D > 2 \) around the centerline, where \( \varepsilon > 0.3 \) is reached. The source of this error lies on the length of the wake, having the reference flow already recovered momentum at that stage compared to the simulation. Another region of high relative error is centred at \( X/D = 1.25 \) and \( Z/D = 0.55 \). This is related to the height of the backflow region. Since the simulation predicts a backflow region shorter in height with respect to the reference, this causes a deficit of momentum and an increase of \( \varepsilon \) in that flow region. Furthermore, \( \varepsilon \) peaks in the close proximity of the object. This is due to the lack of resolution for the experimental data previously mentioned. Because the discrepancies in the latter region cannot be ascribed to the numerical simulations, an integral evaluation of the relative error will exclude the regions close to the object surface and the ground wall.

5.3. Assimilated simulations

Before presenting the results of the assimilated simulations, the local response to the forcing function at individual points is discussed. The effect of the forcing is expressed through the response function \( H = |(u_{bas} - u)/(u_{bas} - u_{ref})| \), where \( u_{bas} \) represents the baseline solution and \( u \) the assimilated solution. The analysis is performed on cells widely separated spatially (\( \gamma = \frac{5}{2} = 0.5 \) or 5 cm distance between forced cells), as illustrated in Fig. 3. This is done to consider the result only dependent on the local forcing, with no interference among neighbouring forced cells. For each forced cell with center \( X_c = (X_c, Y_c, Z_c) \), a \( 5 \times 5 \times 5 \) cm³ cube is considered, with its center on \( X_c \), and discretized with \( 101 \times 101 \times 101 \) elements. For each cube, the effect of the forcing is averaged over a local coordinate system \( X^* = (X - X_c, Y - Y_c, Z - Z_c) \) permitting the alignment of results among different points. The parameter \( H \) is evaluated around \( X_c \).

Because the forcing term is introduced in the momentum equation, the extent and shape of the region affected by the forcing
is expected to follow local diffusion and convection. To assess the region of influence of the forcing, three flow regimes are considered, namely: a) the outer (potential) flow at approximately free-stream velocity; b) the shear layer emanating from the object trailing edge; c) the recirculating flow inside the separated region.

Fig. 9 shows the spatial distribution of $H$ for the three considered regions obtained applying DASOP. The local effect of the forcing algorithm extends to a region with length not exceeding 2 cm (0.2 D). Such localised effect of the forcing has been also observed in previous works, namely, in the study by Imagawa and Hayase [13]. This result suggests in advance that data assimilation at values of $\gamma > 0.2$ may result in localised rather than global modifications of the simulation result.

The extent and shape of the region neighbouring the forced location appear to be affected by the local flow properties. In the outer flow region, the effect of the forcing is comparatively weak and highly elongated along the convection direction. In the regions of lower local velocity and high turbulent diffusion, such as the shear and the recirculation regions, the spatial response function is close to isotropic. In Fig. 9, also the local value of the turbulent kinetic energy $k$ normalised by the local kinetic energy is presented. The extent and shape of $H$ follow that of $k$. A similar analysis performed with the DASOP method yields no visible change in velocity, with values of $H$ typically below 0.1%.

In the remainder of this work, the location of the forcing points is chosen randomly with a distribution controlled through the density parameter $\gamma = [0.01...0.45]$; the corresponding mean distance $\lambda$ between forcing points varies between 4.5 cm (0.45 D) and 1 mm (0.01 D).

Before comparing the results obtained by the two proposed algorithms, the effect of the two parameters $K_p$ and $K_t$ is investigated. The study has been performed using the DASOP algorithm for the case of $\gamma = 0.03$. In order to evaluate the effectiveness of the chosen values of $K_p$ and $K_t$, the error $\varepsilon$ has been averaged in the volume where experimental data were available; the symbol $\bar{\varepsilon}$ indicates the result of this operation. Since the results have been found to be mostly insensitive to the value of $K_p$, Fig. 10 shows the obtained $\bar{\varepsilon}$ varying $K_t$ and averaging the results obtained for $0.001 < K_p < 10$. The bars at each data point represents the variation obtained varying $K_p$ for a given $K_t$. The results show a decay of $\bar{\varepsilon}$ when $K_t$ is increased. This is expectable, since when $K_t$ is increased, the integral part of the forcing term becomes stronger, with a subsequent stronger correction towards the bias errors of the simulation. When $K_t = 1$, while the simulation results exhibit a lower error with respect to the reference, artificial edge effects start to appear at the boundary of the forced region. For values of $K_t > 1$, the simulations show numerical instability and diverge. For this reason, it has been decided to set the parameters as $K_p = 1$ and $K_t = 0.1$ for all the simulations presented in the remaining part of the manuscript.

Fig. 11 shows the results obtained by both the forcing methods DASOP and DASOPI when $\gamma = 0.03$. As for the baseline simulation, in the potential flow region, both DASOP and DASOPI assimilated results agree well with the reference measurement. Larger differences between the two forcing methods are found in the wake region. The application of the proportional forcing term (DASOP) does lead to a reduction in the length of the recirculation region, and furthest point of the reattachment line is moved to $X/D = 2.9$, closer to the reference than that of the baseline simulation. The height of the backflow region, however, remains lower compared to the reference as it can be seen in the section at $Z/D = 0.5$ in Fig. 11; furthermore, the backflow is underestimated by more than 50%.
The use of the DASOPI algorithm visibly improves the fidelity of the velocity field obtained by assimilation. The reattachment point upstream to X/D = 2.4, practically matching the reference data. The DASOPI result also recovers the topology of the recirculation region given in the reference data for what concerns the position of the foci (Fig. 12).

The effect of further refining the density of forcing points (the distance between forcing points is reduced to 0.01 D) is illustrated in Fig. 13. In this case, all grid cells in the selected volume are forced. The DASOPI algorithm with a greater number of forced points yields some improvement in terms of similarity to the reference data. While in the potential flow region no dramatic changes are visible, the length of the wake is shortened, with the reattachment point found at X/D = 2.28 at the symmetry plane. The shortening of the wake is also coupled with an increase of the backflow with respect to the simulation obtained by DASOP with γ = 0.03. However, at the plane Z/D = 0.5, the backflow magnitude peak is still underestimated with respect to the reference.

For what concerns the comparison between DASOP and DASOPI, Fig. 13 confirms the trend observed with γ = 0.03, with DASOPI overperforming. In the region where the forcing is applied, the contour lines of the assimilated solution with DASOPI overlap with the ones representing the reference. It is noticed, however, that the simulation attempts to converge everywhere to the reference data, even where the latter is affected by outliers. This behaviour can be noticed in the top-left corner of the symmetry plane of Fig. 13. Furthermore, when this density of forcing point is used, the extension of the volume where the forcing is activated becomes directly visible in the velocity field, with artefacts at its edges.

The global behaviour of the error relative to the reference data is quantified for both algorithms with its spatial average $\overline{\varepsilon}$ defined in Eq. 7. The variation of $\overline{\varepsilon}$ as a function of the forcing density γ is represented in Fig. 14. In order to capture the most relevant source of error shown in Fig. 8, the average has been obtained taking into consideration only the volume downstream of the body shown in Fig. 3. The result given by the baseline has been plotted with a dotted line as reference. The assimilated simulations appear to be too coarsely forced as long as γ ≥ 0.1. Under this condition, both DASOP and DASOPI methods produce an error comparable to the baseline solution. The distance between the forcing points is too high to have a global reduction of the error compared to the baseline and the effect of the forcing remains local. When γ < 0.1, the behaviour of the two methods differs. DASOP method is not able to reduce the error until γ < 0.05. If γ is further reduced, the error reduces, reaching $\overline{\varepsilon} = 0.06$ when γ = 0.01. Compared to the baseline error, an error reduction of 28% is obtained.

The DASOPI algorithm becomes effective for γ < 0.1. As also shown in Figs. 10 and 12, DASOPI is more effective than DASOP, yielding lower error values in a wider range of values for γ. When γ = 0.05, the relative mean error becomes $\overline{\varepsilon} = 0.055$. Further reducing γ to 0.01, yields $\overline{\varepsilon} = 0.023$, corresponding to a reduction of 72% compared to the baseline.

The reduction of $\overline{\varepsilon}$ at decreasing γ can be explained by the response function shown by Fig. 9. When γ decreases, the amount

![Fig. 13. Streamwise velocity contours of DASOP (left) and DASOPI (right) in the assimilated simulation. Contours of reference data are superimposed (dashed lines for negative values). Y/D = 0 in the top row; Z/D = 0.5 in the bottom row. Forcing density γ = 0.01.](image)

![Fig. 14. Spatial-averaged relative error $\overline{\varepsilon}$ as a function of the relative mean forced point distance γ. The horizontal dashed line represents the relative error of the baseline simulation. The vertical dashed-dotted line represents the vector spacing of the PIV measurements ($\gamma_{PIV} = 0.023$).](image)
of the flow field that is affected by the assimilation increases. The dimension of the volume affected by the forcing proportional-integral forcing term shown in Fig. 9 explains the behaviour of $\mathcal{F}$ for what concerns DASOPI. The majority of the source of $\mathcal{F}$ is concentrated in wake of the object. Since the average volume of effectiveness can be represented as a sphere with diameter $\approx 0.1 \text{ cm}$, the method starts being effective when such volumes start overlapping, corresponding to the condition $\gamma < 0.1$.

Fig. 15 compares the spatial distribution of the relative error $\varepsilon$ for the baseline simulation, DASOP and DASOPI, with $\gamma = [0.1, 0.03, 0.01]$. The spatial distribution of $\varepsilon$ illustrates that the error is confined in specific regions of the flow: in close proximity of the object; around the reattachment region; within the recirculating flow. For $\gamma = 0.1$, the spatial distribution of $\varepsilon$ presented by the assimilated simulations is close that given by the baseline, with a small reduction shown by DASOPI in the recirculation region and the far wake. When $\gamma$ is reduced to 0.03, both DASOP and DASOPI yield a reduction of $\varepsilon$ compared to the baseline. While DASOP still shows areas characterized by $\varepsilon > 0.2$ in the recirculation region and the far wake, DASOPI achieves $\varepsilon < 0.05$ everywhere except in proximity of the object and in the shear layer close to separation. Finally, for $\gamma = 0.01$, a further reduction of $\varepsilon$ is noticed in the entire field. It can be noted that, for what concerns DASOPI, some regions of high error outside the forcing area arise, as introduced in the previous paragraph. The optimisation of the location of the forcing points to minimise the error of the assimilated simulation will be investigated in future works.

Overall considerations on the effect of data assimilation to the flow topology can be done considering the horseshoe and the arc-vortex axis lines, alongside the specific velocity contour $u = 0$ at $Z/D = 0.5$ (Fig. 16).

Considering the horseshoe vortex, none of the simulations reproduces the path showed by the reference. With a larger separation distance [10], the horseshoe vortex showed by the simulations presents an offset toward negative $X$ values with respect to the reference. The assimilation, both through DASOP and DASOPI, does not influence the path of the horseshoe vortex before $X/D < 1$. It must be noted that, being the horseshoe vortex close to the ground, it lays outside the region where the forcing is active, as represented by Fig. 3. For $X/D > 1$, the assimilation through DASOPI of the flow above the vortex influences its path, reducing its distance with respect to the reference. The continuous line in Fig. 16 represents the contour line $u = 0$ at $Z/D = 0.5$. At this lo-
cation, inside the forced region, the assimilations performed both with DASOP and with DASOI reduces the distance between simulation and reference. As shown by Fig. 14, DASOI overcomes DASOP in reproducing the wake shape of the reference. The shortening of the wake brought by the assimilation causes also a reduction in the distance between the foci $F_2$ and $F_1$ between reference and assimilated simulations, foci that are represented by the crosses in Fig. 16.

The assimilation has also an effect on the convergence rate of the simulation. A simulation has been considered converged when the moving standard deviation (kernel equal to 500 iterations) of the model drag coefficient $C_D$ falls below a threshold value chosen equal to $10^{-3}$. Fig. 17 shows the number of iterations required to reach convergence with respect to the relative mean forced point distance $\gamma$. For both DASOP and DASOI, the amount of iterations for convergence is lower than in the baseline simulation (dashed line in Fig. 17). Also in this case, decreasing $\gamma$ has a positive effect, yielding a reduction of the number of iterations required. It must be noted however that for $\gamma < 0.03$, the number of iterations needed by DASOI does not decrease anymore, which is ascribed to the appearance of the aforementioned edge effects.

6. Conclusions

In this study, a data assimilation framework based on a state observer algorithm has been presented. Two forcing terms, proportional (DASOP) and integral-proportional (DASOI), have been considered. The performances of the proposed algorithms are based on an experimental dataset consisting of the time-average 3D velocity field around a simplified car mirror geometry. The measurements have been conducted with Robotic Volumetric PIV.

Both the forcing terms proposed are functions of the difference between the simulated velocity and the reference experimental data. The local response function of the forcing term has been evaluated, confirming the results of Imagawa and Hayase [13], where the effect is limited to the region close the forced location; the shape and extent of the affected region depending on the local convection and diffusion.

The data assimilation of the entire domain of interest is parametrised with respect to the spatial concentration of forced points of forcing density $\gamma$. Both DASOP and DASOI produce a reduction of the error compared to the baseline simulation. DASOP requires $\gamma < 0.05$ to produce significant effects, DASOI yields a comparable error reduction at already $\gamma = 0.1$, as a result of the higher strength, given the integral formulation of the forcing term. For $\gamma < 0.1$, DASOI reduces progressively the error of the assimilated simulation. The maximum error reduction is obtained at the maximum forcing density ($\gamma = 0.01$), where the error of the assimilated simulation is approximately 25% of the one of the baseline simulation.

The topological analysis based on the reattachment point, the arc-vortex and horseshoe vortex axes confirms that the data assimilation by DASOI produces a realignment of the simulation towards the experimental reference, also in those regions where the forcing is not applied locally. The latter suggests that the convective-diffusive nature of the forcing mechanism may extrapolate the effects of the assimilated region beyond the domain where experimental data is available.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Credit authorship contribution statement

Edoardo Saredi: Conceptualization, Methodology, Software, Writing - original draft. Nikhilesh Tumuluru Ramesh: Investigation, Writing - review & editing. Andrea Sciacchitano: Supervision, Writing - review & editing, Project administration. Fulvio Scaran: Supervision, Writing - review & editing, Project administration.

References