Data-Free and Data-Driven RANS Predictions with Quantified Uncertainty

W. N. Edeling[1] (iD) · **G. Iaccarino**[1] · **P. Cinnella**[2]

Abstract For the purpose of estimating the epistemic model-form uncertainty in Reynolds-Averaged Navier-Stokes closures, we propose two transport equations to locally perturb the Reynolds stress tensor of a given baseline eddy-viscosity model. The spatial structure of the perturbations is determined by the proposed transport equations, and thus does not have to be inferred from full-field reference data. Depending on a small number of model parameters and the local flow conditions, a 'return to eddy viscosity' is described, and the underlying baseline state can be recovered. In order to make predictions with quantified uncertainty, we identify two separate methods, i.e. a data-free and data-driven approach. In the former no reference data is required and computationally inexpensive intervals are computed. When reference data is available, Bayesian inference can be applied to obtained informed distributions of the model parameters and simulation output.

Keywords RANS modeling · Uncertainty quantification · Bayesian inference · Return to eddy viscosity · Lag model

1 Introduction

Despite the increasing availability of high-performance computational resources, Reynolds-Averaged Navier-Stokes (RANS) closure models are projected to remain a widely used option for the prediction of turbulent flows [1]. However, it is well known that the resulting predictions are potentially sensitive to parametric and model-form uncertainty. The former

✉ W. N. Edeling
    wedeling@stanford.edu; wouteredeling@gmail.com

1   Center for Turbulence Research, Stanford University, 488 Escondido Mall, Stanford, CA 94305, USA

2   Laboratoire DynFluid, Arts et Métiers ParisTech, 151 Boulevard de l'Hopital, 75013 Paris, France

concerns imperfectly known closure coefficients, while the latter deals with the uncertainty due to assumptions made in the mathematical formulation of the model itself. A user of RANS models should therefore be aware of flow-dependent performance, and expect to obtain predictions which can be trustworthy for certain regions of a flow domain, yet erroneous elsewhere.

Over the years various attempts have been made to quantify the uncertainty in RANS closures. One might consider the closure coefficients as random variables, and perform a Bayesian inference procedure to learn their posterior distribution from reference data [2, 3]. To also obtain a measure for the model-form uncertainty, it is possible to combine multiple closure models in a single predictive framework using Bayesian model averaging [4]. Moreover, in traditional turbulence modeling practices, closure models are calibrated on a small set of canonical flow cases. The assumption is then made that these calibrated models can be applied to flows different from those in the training set. To relax this assumption, one can also average over different training scenarios as well in an attempt to broaden the predictive capability, see [5]. The downside of an ensemble approach is if all models in the chosen set are linear eddy-viscosity models, the underlying Boussinesq hypothesis is not challenged, thus resulting in a potential bias in the predictions.

We will build upon the work of [6, 7], in which perturbations in the eigenvalues of the Boussinesq anisotropy tensor are introduced. The perturbations were either homogeneous, or informed by ad-hoc marker functions. To make the perturbations representative of local flow features, we introduce two additional transport equations for linear combinations of these aforementioned eigenvalues. This limits the amount of information that must be extracted from reference data, since the spatial structure of the perturbations is now embedded in the mathematical form of the transport equations. The use of transport equations to prescribe the local perturbations also tends to impose a regularity in the solution, avoiding oscillatory Reynolds stress components. Our model form is similar to the lag model of [8], which can be derived from an argument involving a correction to the Boussinesq hypothesis. The location, magnitude and direction of the eigenvalue perturbations are now governed by the two model transport equations. The general behavior of our discrepancy model is determined by two coefficients, resulting in a low-dimensional Uncertainty Quantification (UQ) problem. Depending on the value of these coefficients and the local flow physics, our transport equations will revert to the underlying baseline Boussinesq state. In a manner similar to return to isotropy models, our transport equations therefore describe a 'return to eddy viscosity'.

The low-dimensional nature of our model allows for two distinct applications in a predictive setting, which we denote the data-free and data-driven predictions. If no reference data is available for the flow case at hand (or for a closely related flow), cheap uncertainty intervals can be computed. When we do have access to data, a Bayesian inference procedure can be applied in order to obtain informed probability distributions describing the uncertainty in the coefficients and Quantity-of-Interest (QoI). Although this data-driven approach is more expensive, accurate surrogate models can be readily constructed with $\mathcal{O}(100)$ RANS samples. Markov-chain Monte-Carlo (McMC) methods, or other common sampling techniques, are then available to obtain posterior distributions of the model coefficients. These distributions are easily interpreted from a physical point of view, as the values of the model coefficients can be directly related to a state of anisotropy. As validation cases we will consider a backward facing step and a subsonic jet, for which it is known that Boussinesq-type models struggle to accurately capture all relevant flow physics.

This paper is organized as follows. First, we will very briefly describe the RANS modeling framework. Sections 3 and 4 then deal with the eigenvalue perturbations, and describe

various aspects (e.g. realizability and effect of model coefficients) of the proposed model. Next, the data-free and data-driven methods are outlined, followed by the obtained results in Section 6. Finally, we give our conclusions in Section 7.

2 Reynolds-Averaged Navier-Stokes Equations

The well-known Reynolds-Averaged Navier-Stokes equations [9] are given by

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho \langle U_i \rangle}{\partial x_i} = 0,$$

$$\frac{\partial \rho \langle U_i \rangle}{\partial t} + \frac{\partial \rho \langle U_i \rangle \langle U_j \rangle}{\partial x_j} = \frac{\partial}{\partial x_j} \left[\sigma_{ij} - \langle p \rangle \delta_{ij} - \rho \langle u_i u_j \rangle \right]. \tag{1}$$

Here, ρ is the density, $\langle U_i \rangle$ is the mean velocity in x_i direction and μ is the viscosity. The three terms on the right-hand side represent the viscous stress tensor, the isotropic stress due to the average pressure field $\langle p \rangle$, and the apparent stress $\langle u_i u_j \rangle$ arising from the fluctuations in the velocity. To close (1), one must model the Reynolds stress tensor $\langle u_i u_j \rangle$ in terms of mean-field quantities. A common RANS closure is the Boussinesq hypothesis, i.e.

$$\langle u_i u_j \rangle \approx \langle u_i u_j \rangle^{(bl)} = \frac{2}{3} k \delta_{ij} - 2 \nu_T S_{ij}. \tag{2}$$

Here, k is the turbulent kinetic energy, δ_{ij} the Kronecker delta, and S_{ij} is the mean strain-rate tensor, defined as $S_{ij} := \frac{1}{2} \left(\partial \langle U_i \rangle / \partial x_j + \partial \langle U_j \rangle / \partial x_i \right) - \frac{1}{3} \partial \langle U_k \rangle / \partial x_k \delta_{ij}$ in compressible flows. The superscript (bl) stands for baseline, and denotes a quantity computed using the Boussinesq hypothesis.

The scalar quantity ν_T is the eddy-viscosity, which must be computed by means of some chosen turbulence model, e.g. the $k - \varepsilon$, $k - \omega$ or Spalart-Allmaras models, see [10] for an overview. Besides the approximate nature of these transport equations, several closure coefficients and corresponding uncertainties are also present, ultimately leading to an uncertain ν_T [3, 5]. Instead of focusing on one model-specific form of ν_T, we will inject uncertainty directly into $\langle u_i u_j \rangle$, by adding a tensorial discrepancy model D_{ij} to the right-hand side of Eq. 2, i.e. we set

$$\langle u_i u_j \rangle = \langle u_i u_j \rangle^{(bl)} + D_{ij}. \tag{3}$$

We therefore leave any chosen baseline turbulence model and its coefficients unperturbed, and instead focus on estimating D_{ij}. Specifically, D_{ij} will be specified (see Section 4) based on perturbations in the eigenvalues of the Reynolds-stress anisotropy tensor b_{ij}. It is important to note that we do not perturb the corresponding eigenvectors. As a result, we therefore enforce that $\langle u_i u_j \rangle$ is aligned with $\langle u_i u_j \rangle^{(bl)}$ throughout the flow domain.

To perform the simulations the commercial code ANSYS Fluent v16.1 was used. Our model was implemented into Fluent by means of a set of User-Defined Functions (UDFs). These UDFs are available from the authors upon request.

3 Reynolds Stress Anisotropy and the Barycentric Map

Let us define the normalized anisotropy tensor b_{ij} as

$$b_{ij} := \frac{\langle u_i u_j \rangle}{2k} - \frac{1}{3} \delta_{ij}; \quad -\frac{1}{3} \le b_{\alpha\alpha} \le \frac{2}{3}; \quad -\frac{1}{2} \le b_{\alpha\beta} \le \frac{1}{2}; \tag{4}$$

which is a symmetric, deviatoric (zero trace) tensor. Greek subscripts are excluded from the summation convention.

Banerjee et al. express the anisotropy tensor as a convex combination of three basis tensors $\hat{b}_{1c}$, $\hat{b}_{2c}$, $\hat{b}_{3c}$, i.e.,

$$\hat{b}_{ij} = C_{1c}\hat{b}_{1c} + C_{2c}\hat{b}_{2c} + C_{3c}\hat{b}_{3c}. \tag{5}$$

These basis tensors represent the three limiting states of componentality (relative strengths of components in $\widehat{\langle u_i u_j \rangle}$), i.e., they represent one-, two- and three-component turbulence. A circumflex $\widehat{(\cdot)}$ represents the tensor $(\cdot)$ expressed in principal axes, e.g.

$$\hat{b}_{ij} = \mathrm{diag}\,(\lambda_1, \lambda_2, \lambda_3), \tag{6}$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of b_{ij}.

The basis tensors of Eq. 5 are given by

–   One-component limiting state (1C): $\hat{b}_{1c} := \mathrm{diag}\left(\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}\right)$. One component (and one eigenvalue) of $\langle u_i u_j \rangle$ is nonzero. Note that a zero eigenvalue of $\widehat{\langle u_i u_j \rangle}$ leads to an eigenvalue of $-\frac{1}{3}$ for $\hat{b}_{ij}$.

–   Two-component axi-symmetric limiting state (2C): $\hat{b}_{2c} := \left(\frac{1}{6}, \frac{1}{6}, -\frac{1}{3}\right)$. Two components of $\widehat{\langle u_i u_j \rangle}$ are nonzero.

–   Three-component limiting state (3C): $\hat{b}_{3c} := (0, 0, 0)$. All three components of $\widehat{\langle u_i u_j \rangle}$ are equal, where the limiting basis tensor corresponds to isotropic turbulence.

From Eq. 5 it is clear that the coefficients C_{1c}, C_{2c} and C_{3c} measure how close $\hat{b}_{ij}$ is to any of the three limiting states. Since Eq. 5 is required to be a convex combination, we have by definition:

$$C_{1c} + C_{2c} + C_{3c} = 1 \quad \text{and} \quad C_{1c} \geq 0 \quad C_{2c} \geq 0 \quad C_{3c} \geq 0. \tag{7}$$

$$
\begin{aligned}
C_{1c} &= \lambda_1 - \lambda_2 & \lambda_1 &= C_{1c} + \frac{C_{2c}}{2} + \frac{C_{3c}}{3} - \frac{1}{3} \\
C_{2c} &= 2\,(\lambda_2 - \lambda_3) \quad \Leftrightarrow & \lambda_2 &= \frac{C_{2c}}{2} + \frac{C_{3c}}{3} - \frac{1}{3} \\
C_{3c} &= 3\lambda_3 + 1 & \lambda_3 &= \frac{C_{3c}}{3} - \frac{1}{3}
\end{aligned}
\tag{8}
$$

Due to Eq. 7, again only two coefficients are needed to quantify the state of anisotropy.

To visualize the nature of the anisotropy implied by these coefficients, a barycentric map can be defined as ([11])

$$
\begin{aligned}
x_b &= C_{1c}x_{1c} + C_{2c}x_{2c} + C_{3c}x_{3c}, \\
y_b &= C_{1c}y_{1c} + C_{2c}y_{2c} + C_{3c}y_{3c}.
\end{aligned}
\tag{9}
$$

Here, (x_{1c}, y_{1c}), (x_{2c}, y_{2c}) and (x_{3c}, y_{3c}) are the three corner points corresponding to the limiting states of componentality. They can be chosen arbitrarily, but are commonly set to the corner points of an equilateral triangle. In Fig. 1 the barycentric map is depicted along with the variation of the coefficients along its edges. Each point $\mathbf{x} \in \mathbb{R}^N$ in the spatial flow domain has its own coefficients and thus can be mapped to a location (x_b, y_b) in the barycentric map using Eq. 9. Since Eq. 5 is a convex combination of limiting states, all possible realizable states of $\langle u_i u_j \rangle$ are therefore contained within the borders of the barycentric map.
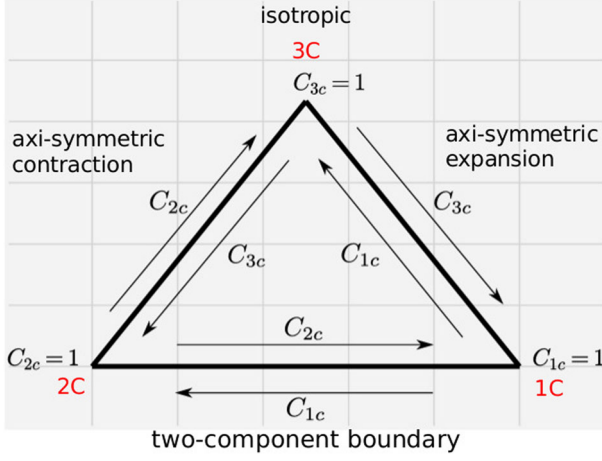
Fig. 1 The barycentric map with the one-, two- and three-component corners, and the variation of its coefficients. Arrows indicates directions along which the corresponding coefficient linearly decreases from one toward zero

4 Reynolds Stress Perturbation

4.1 Anisotropy tensor decomposition

The eigen decomposition of the anisotropy tensor reads

$$b_{ij} = v_{ik}\Lambda_{kl}v_{jl}, \tag{10}$$

where v_{ij} are the eigenvectors and Λ_{ij} the diagonal matrix of eigenvalues, ordered as $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Hence, from Eq. 4 we see that the Reynolds stress tensor can be written as

$$\langle u_i u_j \rangle = 2k \left(v_{ik}\Lambda_{kl}v_{jl} + \frac{1}{3}\delta_{ij} \right) \tag{11}$$

Here, the turbulent kinetic energy, eigenvectors and eigenvalues represent the magnitude, orientation and shape of the Reynolds stress respectively. As noted, we keep the alignment ($v_{ij} = v_{ij}^{(bl)}$), and perturb the eigenvalues Λ_{ij}, such that

$$\langle u_i u_j \rangle = \langle u_i u_j \rangle^{(bl)} + D_{ij} = 2k \left(v_{ik}^{(bl)}\Lambda_{kl}^{(bl)}v_{jl}^{(bl)} + \frac{1}{3}\delta_{ij} \right) + 2k \left(v_{ik}^{(bl)}\Delta\Lambda_{kl}v_{jl}^{(bl)} \right), \quad (12)$$

where $\Delta\Lambda_{ij}$ is the diagonal tensor containing the perturbations in the eigenvalues. To be clear, what we denote as uncertainty here is thus related to the uncertain shape of $\langle u_i u_j \rangle$. The Boussinesq hypothesis restricts this shape to a relative small subset of all possible tensor shapes, see Fig. 2. By including $\Delta\Lambda_{ij}$, we aim to break away from this particular restriction of the Boussinesq model. It should also be noted that what matters to the QoIs might not be reflected in the Reynolds stresses. Uncertainties in boundary- or inflow conditions can also significantly impact the solution, and may act in locations where the epistemic model form error is relatively small.

Emory and Iaccarino [6, 14] also used Eq. 11 to perturb the baseline eigenvalues. The perturbation direction and magnitude was either specified homogeneously or allowed to vary spatially through ad hoc sensors based on wall distance [14] or streamline curvature
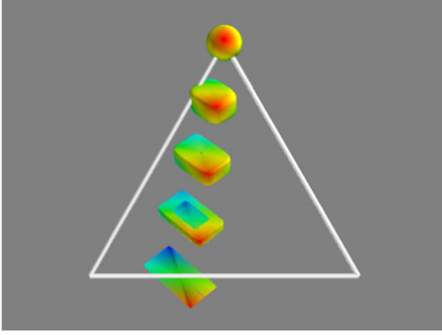
a Typical range of Boussinesq tensor shapes.
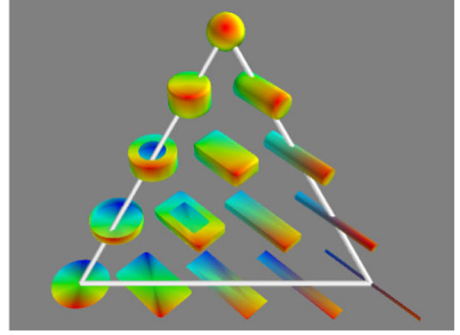
b All possible tensor shapes.

Fig. 2 Tensor shapes of $\langle u_i u_j \rangle / 2k$ visualized using super-quadratic tensor glyphs [12], figure generated with Mayavi [13]. The 1C, 2C and 3C limiting shapes are clearly visible on the right

[15]. Instead, we will propose transport equations to specify the magnitude, direction and spatial distribution of the eigenvalue perturbations.

4.2 Return to eddy viscosity model

Since the Boussinesq hypothesis incorporates a direct proportionality between $\langle u_i u_j \rangle$ and the mean strain-rate tensor S_{ij} (see Eq. 2), the turbulence reacts directly to changes in the mean flow. Strictly speaking, this is only a valid approximation under equilibrium conditions, for example when the time scale of the turbulence is much smaller than that of the mean flow. Our perturbation framework is based upon the lag-type models, which modulate the response through simple linear PDE's on ν_t [8] or $\langle u_i u_j \rangle$ [16], effectively slowing down the response when appropriate.

We prescribe the following lag equations for the coefficients of the barycentric map

$$\frac{\mathrm{D}C_{1c}}{\mathrm{D}t} = a_{1c}\frac{1}{\tau}\left(C_{1c}^{(bl)} - C_{1c}\right), \quad \frac{\mathrm{D}C_{2c}}{\mathrm{D}t} = a_{2c}\frac{1}{\tau}\left(C_{2c}^{(bl)} - C_{2c}\right). \tag{13}$$

It is important to note that this model is first and foremost an inspiration to construct spatially dependent eigenvalue perturbations suitable for the purpose of uncertainty quantification, rather than a formal framework to build a new turbulence closure. The quantity τ^{-1} is an inverse turbulent time scale (modeled as ε/k or ω) and $C_{1c}^{(bl)}$, $C_{2c}^{(bl)}$ are the coefficients computed using the eigenvalues of the Boussinesq anisotropy tensor, i.e. from

$$b_{ij}^{(bl)} = -\frac{\nu_T}{k}S_{ij}. \tag{14}$$

The right-hand side of Eq. 13 is similar to that of a linear return to isotropy model, see e.g. Rotta's model [17]. However, we consider a baseline eddy-viscosity model which can be erroneous at certain locations, but which is able to yield accurate predictions outside these regions. The transport equations for C_{1c} and C_{2c} are therefore meant to describe a 'return to eddy viscosity' when the Boussinesq hypothesis (2) is expected to be valid. Along a streamline, a perturbed barycentric coefficient will go to its baseline value with a time scale τ. The rate of return for each $C_{\alpha c}$ is independently controlled by two coefficients $a_{\alpha c}$. This enables (13) to explore a larger area of the barycentric map, making the lag concept amenable for uncertainty quantification, see the discussion in Section 4.3.

We define the boundary conditions following an approach similar to the original lag model [8]. At a solid smooth wall, we set the values of C_{1c} and C_{2c} equal to the baseline values $C_{1c}^{(bl)}$ and $C_{2c}^{(bl)}$. If the turbulence model equations are integrated down to the wall using damping functions, this essentially sets the values equal to zero. However, if wall functions are used, non zero values of C_{1c} and C_{2c} are obtained. At an inflow boundary, it is assumed that the freestream turbulence is isotropic, i.e. $\langle u_i u_j \rangle = \frac{2}{3} k_\infty \delta_{ij}$, yielding again $C_{1c} = C_{2c} = 0$. An isotropic inflow condition is equivalent to the option of specifying the turbulence intensity at an inlet, which can be done in many solvers. This is not expected to be universally valid, but this is also not the only option. If one has an informed guess, from higher-fidelity simulations for instance, specifying a profile is also possible. Outflow boundary values are extrapolated from internal cells.

Since the $C_{\alpha c}$ are such simple linear combinations of the eigenvalues (see Eq. 8), an equivalent approach would be to write lag equations for λ_α, $\alpha = 1, 2$. We prefer the use of the barycentric coefficients due to their intuitive properties. All 3 have realizable values in the same range ($[0, 1]$), and each one measures the distance from the corresponding limiting state of anisotropy, being 1 if at the limit and 0 if completely removed from it.

Finally, it should be noted that lag models have been further developed since the introduction of the lag concept, see for instance the stress-strain lag model of [18]. We use the original form proposed by [8] as a proof of concept for the use of the lag concept as a tool for uncertainty quantification.

4.3 Effect of lag coefficients

The original default values of the lag coefficients, as defined by [8], are

$$a_{1c} = a_{2c} = 0.35. \tag{15}$$

Let us instead investigate the model behavior for three cases that represent limiting trajectories:

- Case 1 (plane-strain): $a_{1c} = a_{2c} = 0.35$,
- Case 2 (axi-symmetric contraction): $a_{1c} = 0$, $a_{2c} = 0.35$,
- Case 3 (axi-symmetric expansion): $a_{1c} = 0.35$, $a_{2c} = 0$.

The typical resulting trajectories in the barycentric map are shown in Fig. 3. A model will follow the plane strain turbulence line when at least one eigenvalue of b_{ij} is zero [11].
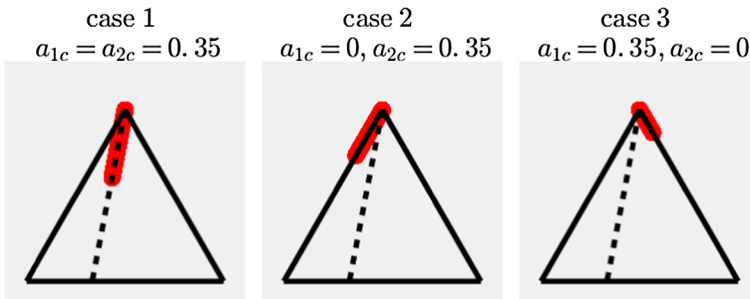
Fig. 3 The trajectories in the barycentric map of the three coefficient cases. The dashed line denotes the plane strain line where $\lambda_2 = 0$

Note that the Boussinesq anisotropy tensor (14) will yield a zero eigenvalue when one of the $S_{\mu\mu}$ equals zero. So all two-dimensional flow cases (or three-dimensional flows like the fully-developed square duct where $S_{11} = 0$) will yield trajectories along the plane stain line when a Boussinesq model is used, and $\hat{b}_{ij}^{(bl)} = \text{diag}(\lambda_1, 0, -\lambda_1)$. When $a_{1c} = a_{2c}$ (case 1), both transport equations in Eq. 13 return to the Boussinesq solution at the same rate. As a result, the trajectory in the barycentric map still lies along the plane strain line (see Fig. 3). Along this line the distribution is different than that of the unperturbed model, depending on the values of a_{1c} and a_{2c}.

To observe larger deviations from the results of the baseline model, unequal lag coefficients are required. Remember from Fig. 1 that $C_{1c} = 0$ along the axi-symmetric contraction boundary. In the second limiting case we enforce a homogeneous solution $C_{1c} = 0$ by setting $a_{1c} = 0$. This therefore forces the trajectory to follow the axi-symmetric contraction border (see Fig. 3). Similarly, setting $a_{2c} = 0$ results in a trajectory along the axi-symmetric expansion border, where $C_{2c} = 0$. Setting two nonzero, unequal, values for a_{1c} and a_{2c} will result in curved trajectories.

4.4 Spatial distribution of the perturbations

The error in any given baseline model will display a complex spatial distribution, which is dependent on the local flow physics [19]. The perturbations in the Boussinesq eigenvalues should reflect this localized behavior, and ideally be concentrated in those areas of the flow domain where a failure of the Boussinesq hypothesis might reasonably be expected. Besides determining the direction and the magnitude of the perturbations in $\lambda_\alpha^{(bl)}$, the lag terms also dictate the spatial distribution. The exact distribution will be dependent upon the values of a_{1c} and a_{2c}. As an example, consider the contour plots of $|\Delta\lambda_1| := |\lambda_1 - \lambda_1^{(bl)}|$ and $|\Delta\lambda_2| := |\lambda_2 - \lambda_2^{(bl)}|$, for an axisymmetric jet with an exit Mach number of approximately 0.51 [20], and a backward facing step with a Reynolds number based on the step height h of $Re_h = 5100$ in Fig. 4. Note that for the jet flow only at the nozzle exit and in the mixing layer do we find significant perturbations. Where the flow is essentially uniform (and in the potential core), no perturbations are made in λ_1. In the backstep case, the local perturbations are made near and after the separation point. Still, when moving away from the wall, the perturbations in λ_2 approach zero.

a $|\Delta\lambda_1|$ for $a_{1c} = a_{2c} = 0.011$ **b** $|\Delta\lambda_2|$ for $a_{1c} = 0.35, a_{2c} = 0.0$
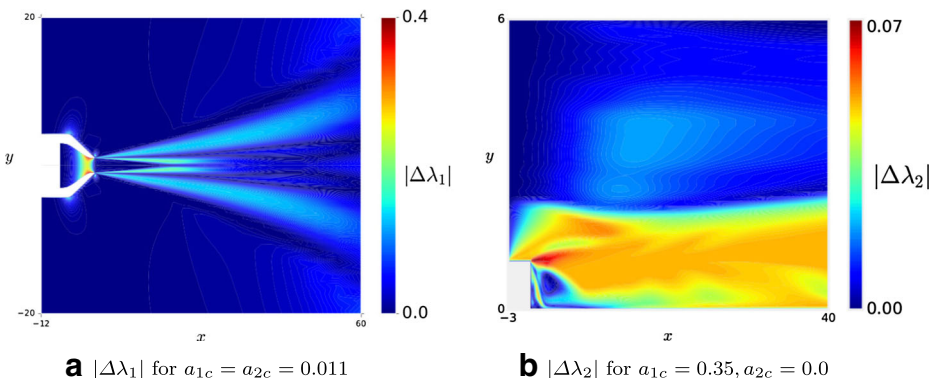
Fig. 4 Examples of $|\Delta\lambda_\alpha| := |\lambda_\alpha - \lambda_\alpha^{(eq)}|$ for a subsonic jet (left), and a backward facing step (right)

4.5 Symmetry around $a_{1c} = a_{2c}$

In two-dimensional flows, when the baseline state lies on the plane-strain line, certain QoIs will display symmetry in the stochastic domain around $a_{1c} = a_{2c}$. Specifically,

–   case (1): $a_{1c} = \alpha$, $a_{2c} = \beta$,
–   case (2): $a_{1c} = \beta$, $a_{2c} = \alpha$,

for any chosen values α, β will cause the so-called eigengaps $\delta_1 := \lambda_1 - \lambda_2$ and $\delta_2 := \lambda_2 - \lambda_3$ to switch between case (1) and (2). The following three statements are then all equivalent

$$\delta_1^{(1)} = \delta_2^{(2)} \quad \text{and} \quad \delta_2^{(1)} = \delta_1^{(2)},$$

$$C_{1c}^{(1)} = \frac{1}{2}C_{2c}^{(2)} \quad \text{and} \quad C_{2c}^{(1)} = 2C_{1c}^{(2)},$$

$$\lambda_1^{(1)} = -\lambda_3^{(2)}, \quad \lambda_2^{(1)} = -\lambda_2^{(2)} \quad \text{and} \quad \lambda_3^{(1)} = -\lambda_1^{(2)}. \tag{16}$$

Here, $\delta_1^{(2)}$ denotes the first eigengap of case (2), and likewise for the other quantities. To prove the above, use the relations $\delta_1 = C_{1c}$ and $\delta_2 = C_{2c}/2$ together with (13) to obtain the following transport equations

$$\frac{\mathrm{D}\delta_1^{(1)}}{\mathrm{D}t} = \alpha\omega\left(\delta_1^{(bl)} - \delta_1^{(1)}\right) \quad \text{and} \quad \frac{\mathrm{D}\delta_2^{(2)}}{\mathrm{D}t} = \alpha\omega\left(\delta_2^{(bl)} - \delta_2^{(2)}\right). \tag{17}$$

Clearly, both transport equations have the same form. In addition, the boundary conditions are equal and along the plane-strain line (where $\lambda_2^{(bl)} = 0$ by definition), both baseline eigengaps are always the same due to $\hat{b}_{ij}^{(bl)} = \mathrm{diag}(\lambda_1^{(bl)}, 0, -\lambda_1^{(bl)})$. Thus, Eq. 17 essentially displays the same right-hand side twice, implying $\delta_1^{(1)} = \delta_2^{(2)}$. A similar argument can be formed for the second equality $\delta_2^{(1)} = \delta_1^{(2)}$ of Eq. 16.

If Eq. 16 is satisfied, $\langle u_1 u_2 \rangle$ is the same for case (1) and (2), which in turn can cause very similar solutions for other QoI. To see this, denote $\alpha_1 \in [-\pi, \pi]$ be the minimum angle from the horizontal axis to the eigenvector $\mathbf{v}_1$, which corresponds to the largest eigenvalue λ_1. The angle α_3 is likewise defined in relation to $\mathbf{v}_3$. Since b_{ij} is real and symmetric, any two eigenvectors from different eigenspaces are orthogonal to each other [21]. Moreover, any two orthogonal eigenvectors in the x-y plane can be described by $\mathbf{v}_1 = (\cos\alpha_1, \sin\alpha_1, 0)^T$ and $\mathbf{v}_3 = (\cos\alpha_3, \sin\alpha_3, 0)^T$, where $\alpha_1 = \alpha_3 \pm \pi/2$. The remaining eigenvector must be orthogonal to the x-y plane, and thus $\mathbf{v}_2 = (0, 0, 1)^T$ is also an eigenvector. Now we can write b_{ij} as

$$b_{ij} = \begin{pmatrix} \cos\alpha_1 & 0 & \cos\alpha_3 \\ \sin\alpha_1 & 0 & \sin\alpha_3 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} \cos\alpha_1 & \sin\alpha_1 & 0 \\ 0 & 0 & 1 \\ \cos\alpha_3 & \sin\alpha_3 & 0 \end{pmatrix}. \tag{18}$$

From Eq. 18, we can extract the following expression for the Reynolds shear-stress component after some manipulation

$$\frac{\langle u_1 u_2 \rangle}{k} = 2\left(\lambda_3 - \lambda_1\right)\sin\left(2\alpha_3\right) = -\left(C_{1c} + \frac{1}{2}C_{2c}\right)\sin\left(2\alpha_3\right). \tag{19}$$

Provided that we do not perturb the eigenvectors, $\langle u_1 u_2 \rangle / k$ is determined by the coefficient $C_{1c} + C_{2c}/2$ only. From the relations in Eq. 16) it is readily determined that this coefficient is the same for both case (1) and (2). As a verification exercise we randomly picked values for α and β, and compared the $C_{1c} + C_{2c}/2$ and $\langle u_1 u_2 \rangle$ profiles from the

backward facing step flow of Fig. 4. We found that the results for case (1) and (2) indeed overlap.

4.6 Realizability with an unrealizable baseline model

To obtain physically feasible values for the components of the Reynolds stress tensor, the realizability constraints first defined by [22] must hold, i.e.

$$\langle u_\mu u_\mu \rangle \geq 0 \qquad \langle u_\mu u_\nu \rangle^2 \leq \langle u_\mu u_\mu \rangle \langle u_\nu u_\nu \rangle \qquad \det \langle u_i u_j \rangle \geq 0, \qquad \mu, \nu \in \{1, 2, 3\}. \tag{20}$$

Weak realizability conditions can be derived for second-order closures [23], and for Eq. 13 as well [24]. We now investigate weak realizability of our perturbation model on the backstep case, when the chosen baseline model is unrealizable. In this case $C_{1c}^{(bl)} > \frac{1}{3}$ along the plane-strain line for locations of unrealizable turbulence. This surely can occur, see for instance the trajectory of the baseline $k - \varepsilon$ model for the backward-facing step flow in Fig. 5a, and the corresponding spatial locations of unrealizable turbulence in Fig. 6. We selected the $k - \varepsilon$ model precisely because it shows such a pronounced violation of realizability for this particular flow case, making it a good model to showcase the ability of the lag equations to retrieve a realizable state. For the predictions of Section 5 we used the $k - \omega$ SST model.

Luckily, we can control the similarity to the Boussinesq solution through the lag coefficients, and limiting the values of a_{1c} and a_{2c} can prove sufficient to not follow an unrealizable baseline model across the two-component boundary. Hence, the approach does not only perturb the assumptions in the Boussinesq hypothesis, but also provides an easy means to correct violations of realizability. In Fig. 6 we show how much we have to increase the lag coefficients in order to obtain an unrealizable $\langle u_i u_j \rangle$ for the same backward-facing
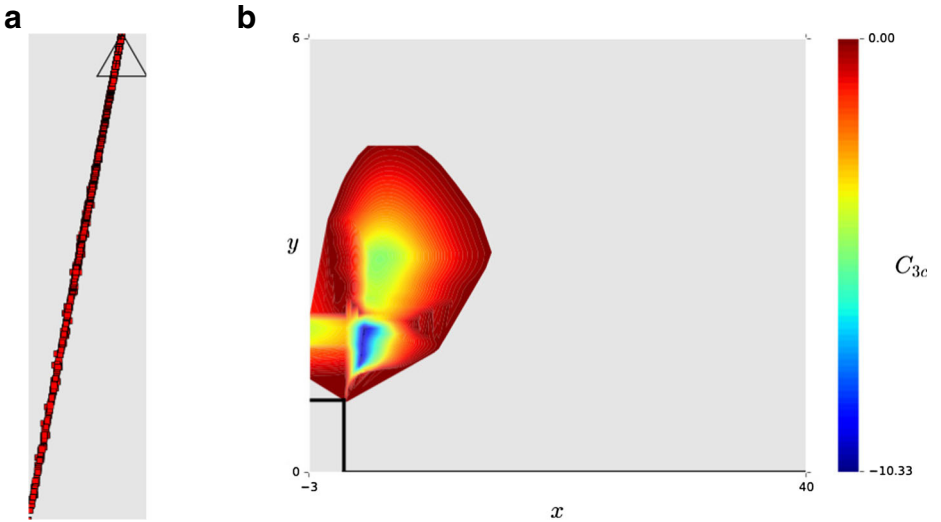
Fig. 5 Unrealizable turbulence. **a** The barycentric map for the baseline $k - \varepsilon$ model when simulating the backwards-facing step flow of [25] ($Re_h = 5100$, where h is the step height). **b** The spatial locations of unrealizable turbulence ($C_{3c} < 0$) for the $k - \varepsilon$ model
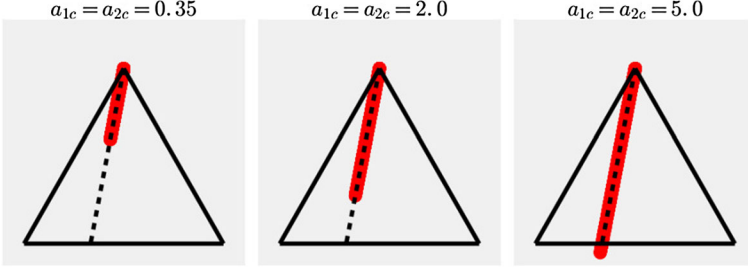
Fig. 6 Three simulations where $a_{1c} = a_{2c} \in [0.35, 2.0, 5.0]$ with an unrealizable baseline $k - \varepsilon$ model

step flow of Fig. 5a. Note that simply turning on model (13) with the default lag coefficient values yields a realizable result, even with the highly unrealizable baseline model. We have to increase a_{1c}, a_{2c} by an order of magnitude in order to force the result across the two-component boundary.

5 Predictions

The final goal of all preceding analysis, is to use Eq. 13 to make predictions with quantified uncertainty. We choose our method of prediction based on the availability of reference data. If no data is available, we will not attempt to extrapolate from other calibrated cases to the prediction scenario at hand. Instead, we compute conservative intervals, based on the method described in Section 5.1. Otherwise, we use the Bayesian inference procedure outlined in Section 5.2.

5.1 Data-free prediction

In order to compute the data-free intervals, let us determine the possibility of *a priori* identifying the lag coefficients a_{1c} and a_{2c} that will yield the maximum and minimum bound on the perturbed eigenvalues that can be achieved with model (13). As noted in Eq. 12, we decompose the anisotropy tensor as

$$b_{ij} = b_{ij}^{(bl)} + v_{ik}^{(bl)} \Delta \Lambda_{kl} v_{jl}^{(bl)}. \tag{21}$$

Where $\Delta \Lambda_{ij} := \mathrm{diag}\,(\Delta \lambda_1, \Delta \lambda_2, \Delta \lambda_3)$, and $\Delta \lambda_\alpha$ is the perturbation such that $\lambda_\alpha = \lambda_\alpha^{(bl)} + \Delta \lambda_\alpha$. Assuming again that the baseline model will follow the plane strain line in the barycentric map, $\Delta \Lambda_{ij}$ can be written as

$$\Delta \Lambda_{ij} = \begin{pmatrix} \frac{2}{3}\left[C_{1c} - \frac{3}{2}C_{1c}^{(bl)}\right] + \frac{1}{6}C_{2c} & 0 & 0 \\ 0 & -\frac{1}{3}C_{1c} + \frac{1}{6}C_{2c} & 0 \\ 0 & 0 & -\frac{1}{3}\left[C_{1c} - 3C_{1c}^{(bl)}\right] - \frac{1}{3}C_{2c} \end{pmatrix}. \tag{22}$$

By fixing the baseline coefficient to an arbitrary value $C_{1c}^{(bl)} \in [0, 1/3]$, we can plot the isocontours of the $\Delta \lambda_\alpha$ in order to identify the directions of maximum and minimum perturbation. The chosen value of $C_{1c}^{(bl)}$ will affect the scale but not the shape of the isocontours. The results are shown in Fig. 7. Clearly, $\Delta \lambda_1$ is maximized in the direction of 1C, and minimized toward 3C. The perturbation $\Delta \lambda_2$ is maximized at 2C and minimized
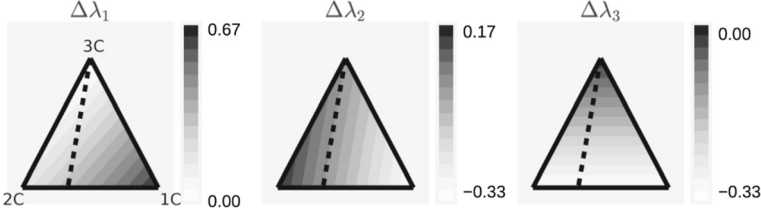
Fig. 7 Isocontours of eigenvalue perturbations for a given value of $C_{1c}^{(bl)}$, which this case is set to zero

at 1C. Finally, the maximum of $\Delta\lambda_3$ is located at 3C, and its minimum is found anywhere along the 2C boundary. Thus, straight lines in the direction of minimum/maximum $\Delta\lambda_\alpha$ are the axi-symmetric expansion, contraction and the plane-strain line. Trajectories along these lines can be obtained with *a priori* known values for a_{1c} and a_{2c} (see Fig. 3). The maximum values we use for $a_{\alpha c}$ are discussed in Section 6.1.2.

5.2 Data-driven prediction

5.2.1 Bayesian calibration

In a Bayesian approach all sources of uncertainty (parametric, model form and experimental) are condensed into a single expression, i.e Bayes' rule [26], given by

$$p\left(\boldsymbol{\theta} \,|\, \mathbf{z}\right) = \frac{p(\mathbf{z}|\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{p(\mathbf{z})}. \tag{23}$$

Here, $\boldsymbol{\theta} \in \mathbb{R}^C$ is a vector containing the model coefficients to be inferred, $\mathbf{z} \in \mathbb{R}^N$ is a vector of high-fidelity reference data, and $p\left(\boldsymbol{\theta}\right)$ is the prior distribution.

The likelihood function $p(\mathbf{z}|\boldsymbol{\theta})$ is found by specifying a relation between the true process ζ, the corresponding QoI q from the code, and the data $\mathbf{z}$ as [27]

$$\mathbf{z} = \zeta(\mathbf{x}) + \mathbf{e}. \tag{24}$$

Here, $\mathbf{x}$ are the spatial coordinates and $\mathbf{e}$ represents the uncertainty in $\mathbf{z}$ due to measurement noise, commonly modeled as a zero-mean Gaussian $\mathbf{e} \sim \mathcal{N}(0, \Lambda)$, $\Lambda := \mathrm{diag}(\lambda_i)$. In [28], the true process was modeled as

$$\zeta(\mathbf{x}) = (1 + \delta_m(\mathbf{x}; \ \boldsymbol{\gamma}))\,q(\mathbf{x}; \ \boldsymbol{\theta}), \tag{25}$$

but other forms are also possible. The quantity δ_m is a multiplicative model error term superposed on the code output, again distributed as a zero-mean Gaussian, and dependent upon a set of hyper parameters $\boldsymbol{\gamma}$. Note that without this term we would state that the true process equals our RANS code, which is unlikely to be an accurate assumption. In turn, this yields a Gaussian likelihood such that $\mathbf{z}|\boldsymbol{\theta} \sim \mathcal{N}(q, K)$. The covariance of the likelihood is $K = \Lambda + K_m$, where the entries of K_m are given by the function

$$k_m(\mathbf{x}', \mathbf{x}) = q(\mathbf{x}'; \boldsymbol{\theta})\left\langle\delta_m(\mathbf{x}'; \boldsymbol{\gamma})\delta_m(\mathbf{b}; \boldsymbol{\gamma})\right\rangle q(\mathbf{x}; \boldsymbol{\theta}). \tag{26}$$

Here, $\left\langle\delta_m(\mathbf{b}'; \boldsymbol{\gamma})\delta_m(\mathbf{b}; \boldsymbol{\gamma})\right\rangle$ is the covariance of the model error term, which remains to be specified.

Since Eq. 13 already injects perturbations in the Boussinesq assumption, it can be considered as an internal model error term. However, it is still dependent upon a_{1c} and a_{2c}. To prevent the posterior distribution of these coefficients from becoming overly confident, we

use a simple spatially uncorrelated model $\langle \delta_m(\mathbf{b}'; \boldsymbol{\gamma}) \delta_m(\mathbf{b}; \boldsymbol{\gamma}) \rangle := \sigma^2 \delta(\mathbf{x}' - \mathbf{x})$ when performing the Bayesian inference. Here, σ is a hyper parameter that is calibrated alongside a_{1c} and a_{2c}, and $\delta(\mathbf{x}' - \mathbf{x})$ is the Dirac delta function. The parameter σ has the effect of broadening the posterior $p(\boldsymbol{\theta}|\mathbf{z})$.

The Markov-chain Monte Carlo (McMC) method [29] is often used to draw samples $\boldsymbol{\theta}_i; \ i = 1, \cdots, N$, from the posterior density (23). The posterior density itself is therefore not inferred directly, and can only be estimated using a kernel density estimation after sufficient samples have been drawn, making McMC an expensive method to use in a RANS context.

A common practice is therefore to construct a polynomial approximation of the forward problem over the support of the prior distribution [30]. This approximation then replaces the full RANS code in the likelihood function, yielding a surrogate posterior distribution which can be sampled at minimal cost. We use the Stochastic Collocation (SC) expansion, formed by approximating the forward problem as the sum of N_p (multidimensional) Lagrange interpolation polynomials [31]. This technique is well known and not new, and many alternatives to the SC expansion exist, see e.g. [32–35]. We therefore refer to [31] for details.

Constructing a SC surrogate requires $N_p := m_1 \times m_2$ evaluations of the RANS code, which in our case corresponds to running the simulation on the collocation points of the tensor grid $a_{1c}^{i_1} \otimes a_{2c}^{i_2}$, $i_1 = 1, \cdots, m_1, i_2 = 1, \cdots, m_2$. Let $\tilde{q}_{N_p}$ denote the corresponding surrogate for some RANS output q. The hierarchical surplus error for the backstep case

$$E_{surplus} = \|\tilde{q}_{N_p}\left(a_{1c}^{i_1}, a_{2c}^{i_2}\right) - \tilde{q}_{N_{p-1}}\left(a_{1c}^{i_1}, a_{2c}^{i_2}\right)\|_2, \tag{27}$$

evaluated at the collocation points of $\tilde{q}_{N_p}$, was $\mathcal{O}(10^{-3})$ for mean velocity profiles and $\mathcal{O}(10^{-5})$ for $\langle u_i u_j \rangle$ components, when using $N_p = 144$ and $N_{p-1} = 121$. Note that a SC surrogate is exact ($\tilde{q}_{N_p} = q$), when evaluated at its collocation points. We selected the surrogate with $N_p = 144$ for use in the likelihood function.

6 Results

6.1 Application to a subsonic jet

6.1.1 Flow case description

This flow a well-known NASA validation case for which the settings and computational grids can be found in [20]. The jet exit Mach number is approximately 0.51. For computational stability, the jet does not flow into quiescent air, and a small reference Mach number $M_{ref} = 0.01$ is applied to the farfield. No-slip boundary conditions are applied to the walls of the nozzle, and at the nozzle inlet the total- to reference pressure ratio is $p_t/p_{ref} = 1.19671$, which yields the correct jet exit Mach number. Experimental Particle Image Velocimetry (PIV) data for several quantities, including uncertainty estimates on the data, are available for this flow case from [36]. A selection of experimental results, normalized by the centerline jet-exit velocity U_{jet} and nozzle diameter D_{jet}, are depicted in Fig. 8.
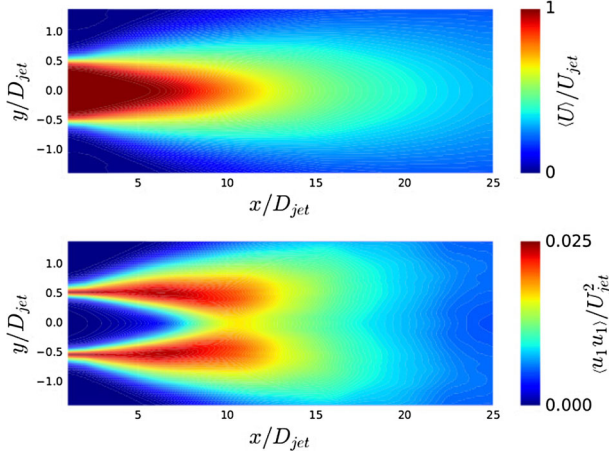
Fig. 8 The experimentally-determined normalized axial velocity (top) and Reynolds stress component (bottom). Source: [36]

6.1.2 Data-free prediction

Let us start by displaying the baseline results as computed with the $k-\omega$ SST closure model. Figure 9 shows the normalized axial velocity $\langle U \rangle / U_{jet}$ and Reynolds stress component $\langle u_1 u_1 \rangle / U_{jet}^2$. Note that for the axial velocity, the results agree fairly well with the reference data at $x/D_{jet} = 2$. However, the accuracy of the baseline prediction deteriorates when we move away from the nozzle exit. In the case of the Reynolds stress component, clear discrepancies can be observed along all profiles.

We now compute the intervals due to the 4 code solves that are predicted to yield the maximum and minimum eigenvalues. i.e. the solves corresponding to the tensor product $[a_{1c,min}, a_{1c,max}] \otimes [a_{2c,min}, a_{2c,max}]$. We select the support of the lag coefficients based on two limiting cases. The first case is the straightforward choice of isotropic turbulence, and thus $a_{1c,min} = a_{2c,min} = 0$. Second, for $a_{1c} = a_{2c} \gg 0$ the baseline state is approached. It makes sense that the baseline state should also be included in the interval. We therefore

a Axial velocity results. **b** Reynolds stress component results.
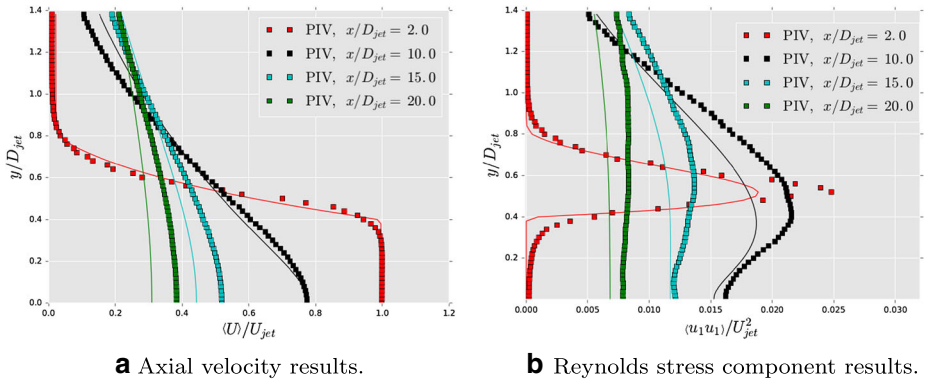
Fig. 9 Several SST baseline profiles (solid lines) along directions normal to the centerline, compared to the corresponding PIV data

a Radial velocity results.

b Reynolds stress component results.
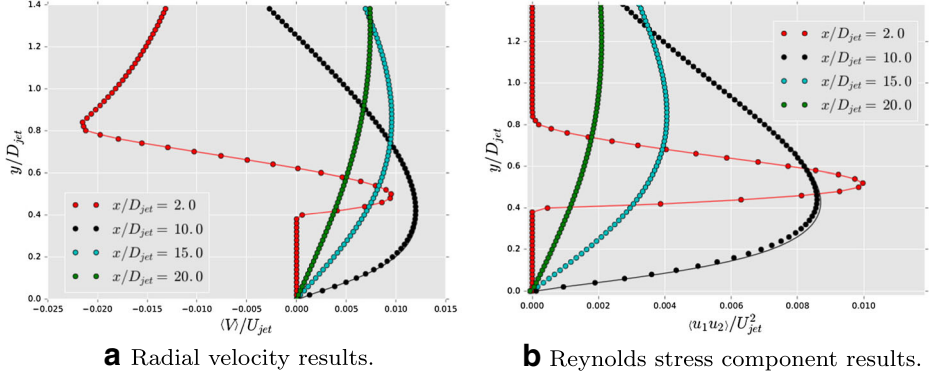
Fig. 10 A comparison between the baseline model (dots), and the results obtained when setting $a_{1c} = a_{2c} = 1.0$ (solid lines)

select a maximum cutoff value for $a_{\alpha c}$ for which a quantity of interest is sufficiently close to the Boussinesq result. Already for $a_{1c,max} = a_{2c,max} = 1.0$, the velocity components are almost indistinguishable from those obtained with the unperturbed SST model, and only minor differences are observed in the Reynolds stresses, see Fig. 10. Hence we compute the results for the combinations shown in Table 1, and form an envelope from the corresponding outputs. The results for the normalized axial velocity are shown in Fig. 11. Almost everywhere the envelope intersects the region of experimental uncertainty, and where there is no overlap the discrepancy is small. Moreover, as we move away from the nozzle exit (where the baseline model becomes less accurate, see Fig. 9a), the uncertainty bounds increase. We also show the results for the normal Reynolds stress component $\langle u_1 u_1 \rangle / U_{jet}^2$ in Fig. 12. The width of the intervals is more pronounced than in the case of the velocity profile, and the experimental data are completely enveloped. While it is certainly true that the Reynolds stress components have a very large interval, the width of velocity intervals is less pronounced. Thus, one way to interpret the Reynolds stress intervals is that, in the absence of reference data, large intervals can be advantageous in order to capture the model inadequacy in other quantities which are less sensitive to $\langle u_i u_j \rangle$, at least for this particular flow case.

6.1.3 Data-driven predictions

We will compute the posterior distribution of $\boldsymbol{\theta} = [a_{1c}, a_{2c}]^T$, using Eq. 23. As mentioned, the surrogate $\tilde{q}$ will replace the full RANS code in the likelihood function such that we have: $\mathbf{z}|\theta \sim \mathcal{N}(\tilde{q}, K)$. Here, $K = \Lambda + K_m$, and $\mathbf{z}$ will be experimental $\langle U_1 \rangle$ data from [36]. The reported experimental accuracy is used to specify the diagonal Λ. K_m is given by Eq. 26, where $\langle \delta_m(\mathbf{b}'; \boldsymbol{\gamma}) \delta_m(\mathbf{b}; \boldsymbol{\gamma}) \rangle = \sigma^2 \delta(\mathbf{x}' - \mathbf{x})$. The hyper parameter σ (with prior

Table 1 Combinations of parameters used for computing the intervals in the jet case

| $[a_{1c}, a_{2c}]$	state
$[0, 0]$	isotropic
$[1, 1]$	baseline
$[0, 1]$	max axisymmetric contraction
$[1, 0]$	max axisymmetric expansion,

**Fig. 11** Experimental axial PIV velocity (squares), and the uncertainty intervals (shaded)
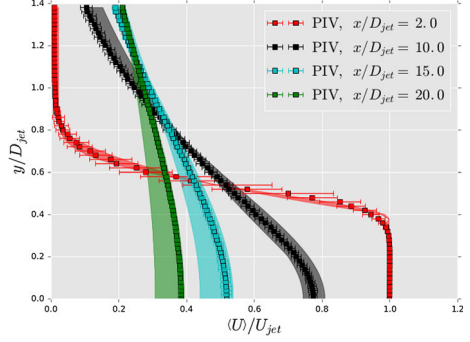
$\sigma \sim \mathcal{U}[0, 3])$ is calibrated alongside $\boldsymbol{\theta}$. Since in general it will be difficult to *a-priori* say which coefficients will be most likely, we use a uniform prior $p(\boldsymbol{\theta})$ which is defined on the same range as displayed in Table 1.

We performed a single calibration using the $\langle U_1 \rangle$ PIV data at $x/D_{jet} = \{2, 12, 22, 32, 42\}$. The resulting posteriors are depicted in Fig. 13. The axisymmetric state is preferred by the data. As a result, $p(\alpha_{1c}|\mathbf{z})$ resembles a delta distribution centered at 0. The posterior $\langle U_1 \rangle$ results are shown in Fig. 14, the variance of which is greatly reduced compared to the interval results of Fig. 11.
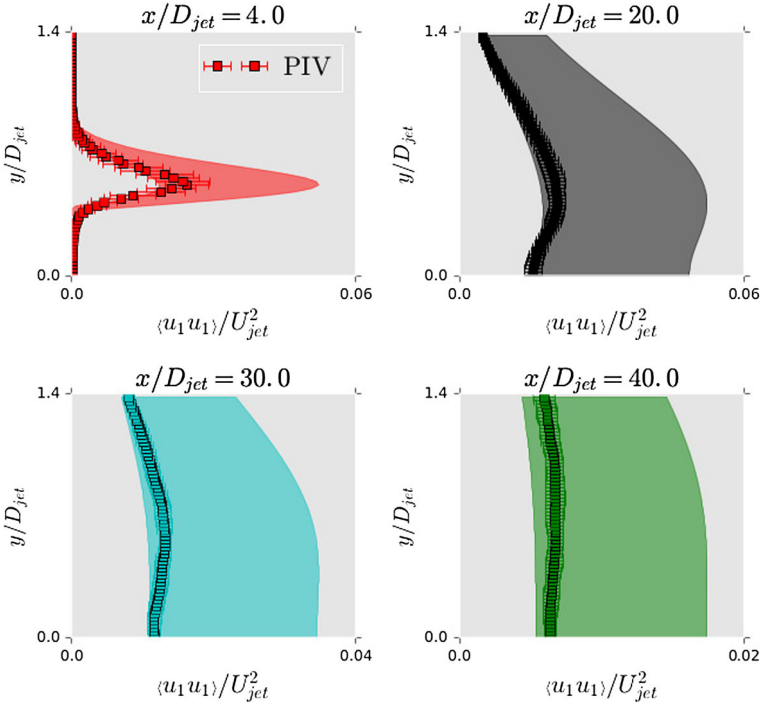
Fig. 12 The uncertainty intervals and PIV data of $\langle u_1 u_1 \rangle / U_{jet}^2$ along 4 centerline-normal directions. The squares represent the PIV data plus the experimental error, and the shaded area is the RANS uncertainty interval
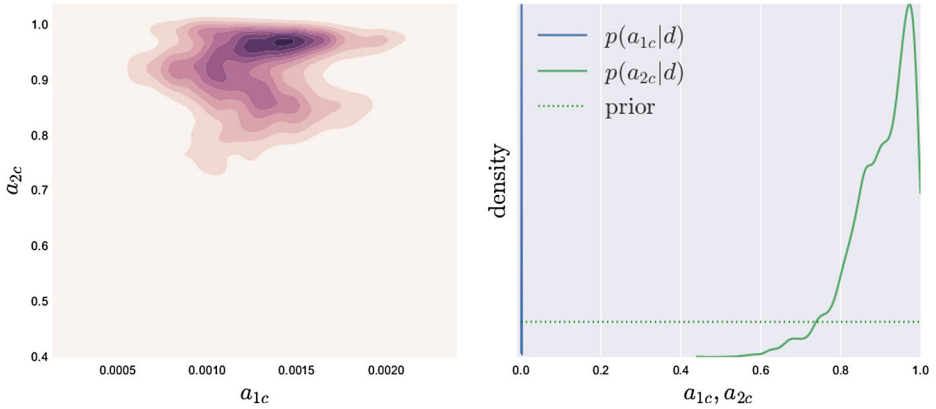
Fig. 13 The posterior distributions of the jet case

6.1.4 Cost overhead

Finally, we use the jet case to briefly give an indication of the cost overhead of our approach. Table 2 displays the relative cost overhead of the four cases of Table 1, compared to the runtime of the jet case using just the baseline SST model. All cases were run for 6000 iterations, and roughly add a 30 % increase to the baseline runtime.

Generally speaking, no convergence issues were encountered for all considered flows and a_{1c}, a_{2c} combinations, with the exception of $a_{1c} = a_{2c} = 0$. Depending on the flow case, perturbing all points to the 3C corner can require an increase in the under-relaxation factor of Eq. 13 in order to reach a converged solution. If reducing the under-relaxation is not enough to stabilize the solution, one might employ $\langle u_i u_j \rangle^{(bl)} + \alpha(\langle u_i u_j \rangle - \langle u_i u_j \rangle^{(bl)})$ with $\alpha < 1$ as a momentum source term, after applying the divergence theorem. Since this is not true under relaxation (due to the $\langle u_i u_j \rangle^{(bl)}$ terms), it is important to keep γ as close to 1 as possible.

Fig. 14 The posterior $\langle U_1 \rangle / U_{jet}$ trace of the jet case

[a_{1c}, a_{2c}]	Δt
[0, 0]	+32.3%
[0, 1]	+30.8%
[1, 0]	+28.1%
[1, 1]	+32.6%
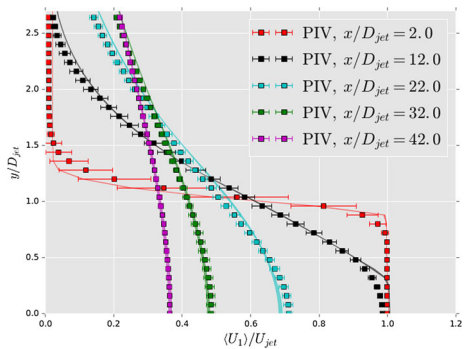
Table 2 The relative computational overhead Δt compared to the runtime of the baseline SST model

6.2 Application to a backward facing step

6.2.1 Flow case description

This is a well-known backward-facing step case with $Re_h = 5100$, where h represents the step height. DNS data from [37] is available for this flow, as well as experiments performed by [38]. At the bottom wall no-slip boundary conditions are applied, and the top boundary a parallel slip condition is prescribed in order to match the DNS setup. The computational domain extends to 40 step heights after the backstep, which is located at $x/h = 0$.
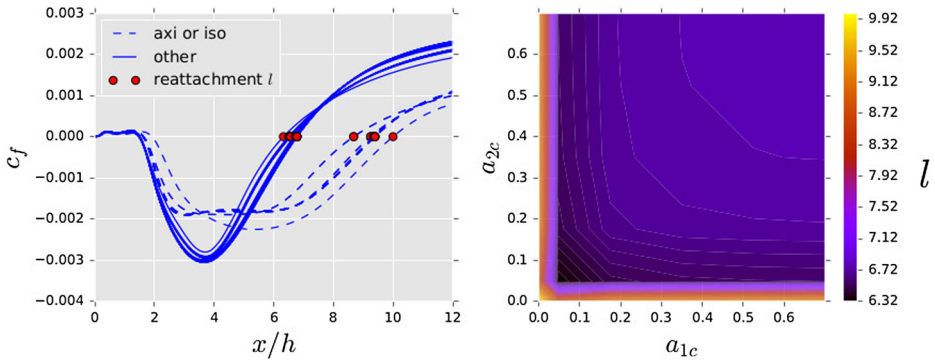
6.2.2 Reattachment length

Eddy-viscosity models are known to provide inaccurate predictions of the reattachment location, see e.g. [39]. The mean calculated reattachment length from the DNS data is $l = 6.28h$ [37]. We now examine the range in l when using our return to eddy viscosity framework. The baseline model is the SST model, and we estimate the reattachment length by the location for which we have a zero skin-friction coefficient, i.e. $c_f = 0$. Figure 15 shows a contour plot of the reattachment length l versus a_{1c} and a_{2c}, as well as various skin-friction profiles. From this figure it becomes clear that two different states exist. Fairly similar results are obtained throughout most of the [a_{1c}, a_{2c}] domain. However, close to the axi-symmetric borders, distinctly different results are obtained. Here, the reattachment length is highly over predicted, leading to significantly larger recirculation bubbles. The maximum occurs at the 3C corner, with a reattachment length close to $x/h = 10$. For comparison, we show the streamlines in Fig. 16 of the velocity field for representative samples of the aforementioned two states.

The transition between the two states is very sharp. Let us examine this transition by visualizing the difference between trajectories for $a_{2c} = \epsilon$ and $a_{2c} = 0$, where ϵ is a small parameter and $a_{1c} > 0$ is kept constant. Figure 17 shows the trajectories in the barycentric map for $\epsilon = 10^{-3}$ and $\epsilon = 10^{-4}$, computed from the flow in wall-normal direction at $x/h = 4$. For $a_{2c} = \epsilon = 10^{-3}$, the source term for DC_{2c}/Dt is still active, and the corresponding trajectory is different from that of the full axi-symmetric state, which is also depicted in Fig. 17. This effect is still present for $\epsilon = 10^{-4}$, although much less pronounced. Hence, a sharp gradient exists in the space of the $a_{\alpha c}$ coefficients as an axi-symmetric boundary is approached.

6.2.3 Data-free prediction

Let us briefly discuss the data-free results for the backward facing step, which were computed in the same way as discussed in Section 6.1.2. The results for several QoI are displayed in Fig. 18. For $\langle U_2 \rangle / U_{ref}$, there is a region where there is no overlap between our interval

Fig. 15 Left: Skin-friction profiles and the computed reattachment length l. Note the different shape of the axi-symmetric and isotropic profiles compared to those computed with $a_{1c} > 0$ and $a_{2c} > 0$. Right: a contour plot of l
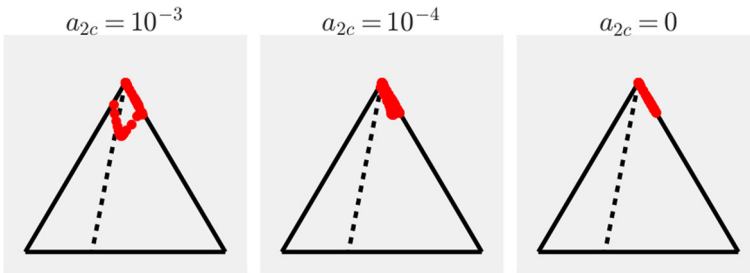
and the experimental error bars. This indicates that, given the current SST baseline model, other uncertainties besides the shape of the Reynolds stresses must be taken into account if $\langle U_2 \rangle / U_{ref}$ is considered important. For the other QoI the overlap is reasonable, although it occurs at the edges of the computed intervals. We can therefore already expect the posterior distributions to be skewed toward the edge of the prior domain.

6.2.4 Data-driven predictions

For this case, **z** will be experimental $\langle u_1 u_2 \rangle$ data from [38]. We performed 5 calibrations, at $x/h \in \{4, 6, 10, 15, 19\}$, and the posterior distributions are plotted in Fig. 19. Despite the fact that we calibrated using data from both in- and outside the recirculation bubble, the posterior distributions are very similar for all x/h. This is a desirable result, since we do not want the most-likely value of the $a_{\alpha c}$ to change from one station to the next. Figure 19 shows that regarding $\langle u_1 u_2 \rangle$, the baseline state is favored by the data since the joint posteriors are located at the top-right corner of the stochastic domain, see also Table 1. This is an indication that the use of a simple eddy-viscosity model is a reasonable choice for this particular flow, and that similar results might be obtained by only calibrating the closure coefficients of the

a The recirculation bubble at the 3C corner.

**b** A typical recirculation bubble away from the axi-symmetric borders.
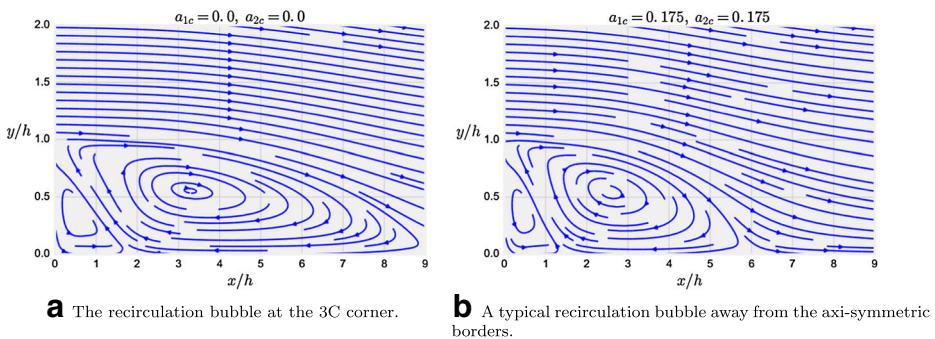
Fig. 16 Two distinctly different recirculation bubbles

Fig. 17 The trajectories for zero a_{2c} (left) and almost zero a_{2c} (right), along the wall-normal direction at $x/h = 4$. In both cases $a_{1c} = 0.3$

SST model. However, other flow scenarios can lead to different parameter estimates, see for instance the results of Section 6.1.3.

In Fig. 20, we computed the forward problem for a number of QoI, using the posterior distribution of $x = 4$ as input. As anticipated, since $\mathbf{z}$ consisted of $\langle u_1 u_2 \rangle$ data, the posterior prediction of $\langle u_1 u_2 \rangle$ is in good agreement with the data. The streamwise velocity $\langle U_1 \rangle / U_{ref}$ and the skin friction coefficient $c_f := \tau_w / (\rho U_{ref}^2 / 2)$ also display distributions which capture the validation data of [38] reasonably well. Note that the posterior mean for these three
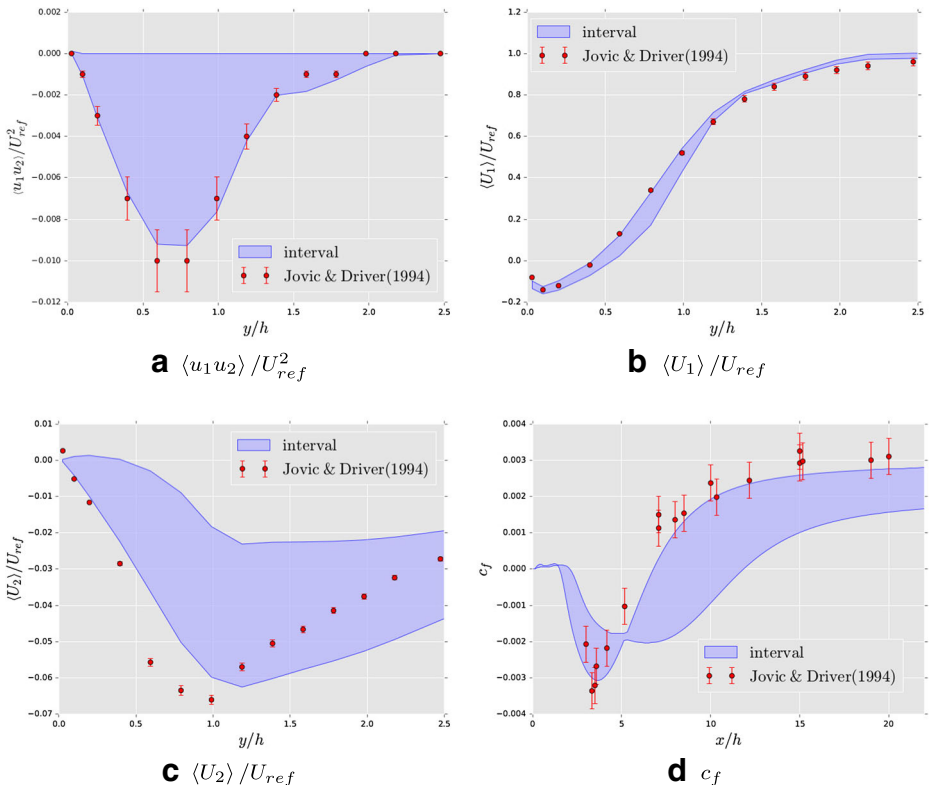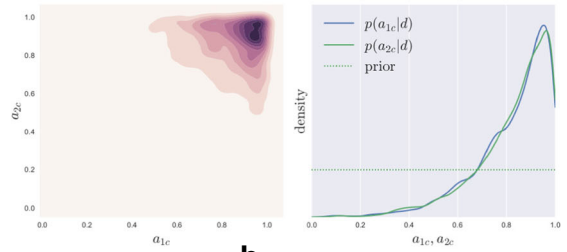

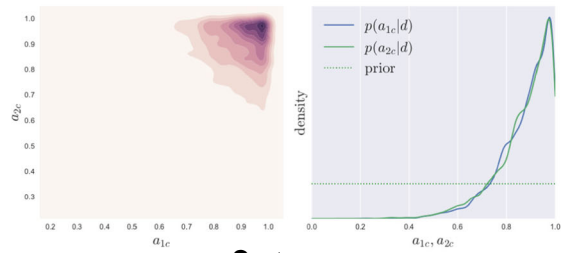**Fig. 18** Data-free intervals for several QoI, computed at $x/h = 4$

Fig. 19 The marginal and joint posterior distributions for the backward-facing step at various x/h stations, obtained by calibrating (13) on the experimental $\langle u_1 u_2 \rangle$ profiles from [38]
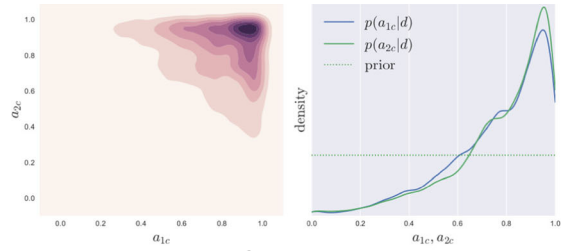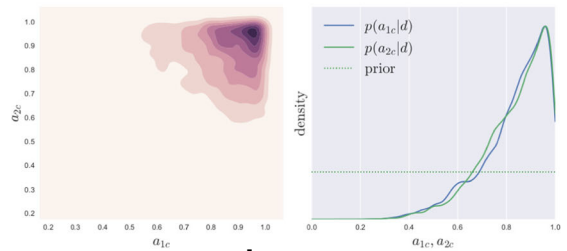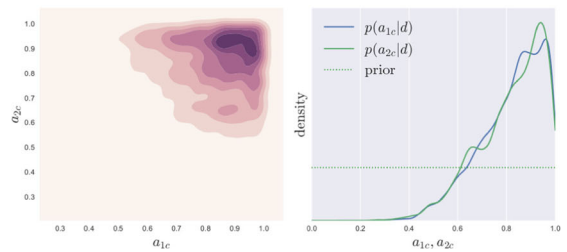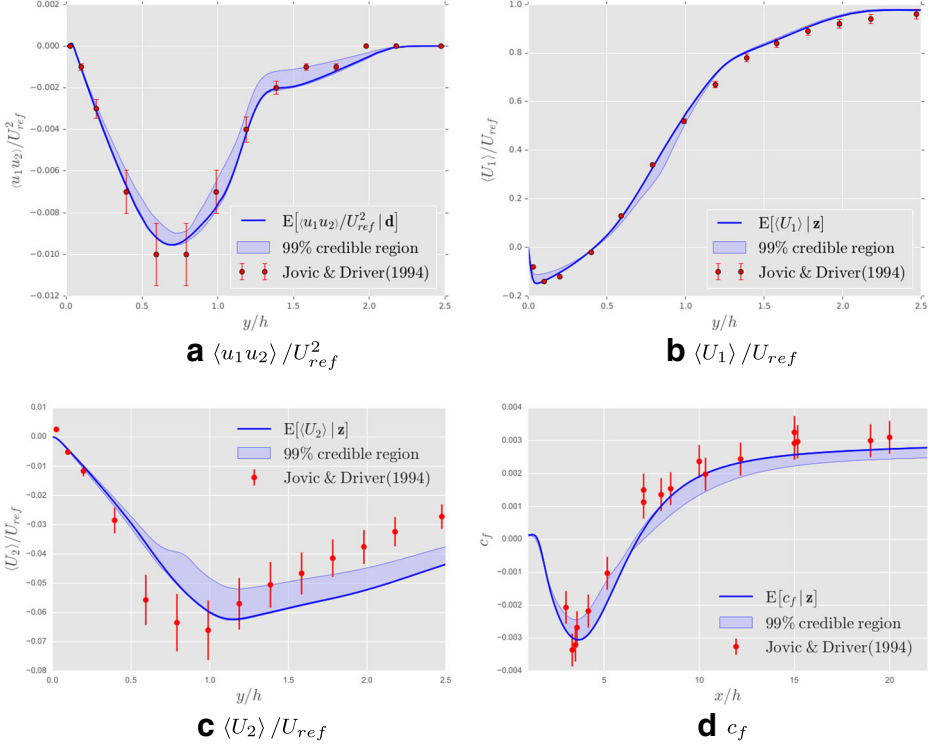

a $x/h = 4$

b $x/h = 6$

c $x/h = 10$

d $x/h = 15$

e $x/h = 19$

Fig. 20 Posterior results for several QoI, computed at $x/h = 4$

QoIs is skewed toward the data. However, discrepancies between the validation data and the posterior distribution of $\langle U_2 \rangle / U_{ref}$ can be observed, which comes as no surprise in light of the interval results of Section 6.2.3.

Since we take care to include the baseline state within our prior, a better baseline model would be able to capture the data with less bias. Thus there can be an influence of the chosen baseline model, especially if the baseline state is preferred by the data. Given that the SST model is not perfect, other uncertainty measures might help to envelop the data better. We reiterate that we considered perturbations in the eigenvalues of b_{ij} only. A more general UQ framework will also perturb the orientation of the Reynolds stress tensor. A study which introduced uniform perturbations in the eigenvectors can be found in [40, 41]. Introducing non-uniformity through transport equations is the subject of future work.

7 Conclusion

We introduced two model transport equations for the coefficients of the barycentric map, for the purpose of epistemic model-form uncertainty quantification of Boussinesq type eddy-viscosity closures. The model effectively adds a tensorial discrepancy function to the Boussinesq hypothesis, based on perturbations made in the eigenvalues of the anisotropy tensor. The discrepancy tensor is assumed to be transported in a similar fashion as other turbulent transport equations such as k and ω, yielding spatially varying perturbations. Through

the inclusion of so-called lag terms, these eigenvalue perturbations are made with respect to a baseline eddy-viscosity model. Depending on the local flow physics, our model allows for the baseline state to be recovered.

The global behavior of the perturbation model is determined by two lag coefficients. These control the amount of perturbation from the baseline state, as well as the direction of perturbation in the barycentric map. This results in a low-dimensional inference problem, which is amenable to physical reasoning. This low-dimensional structure also provides us with two distinct UQ strategies, i.e. data-free and data-driven prediction. In the former we employ bounds on the eigenvalue perturbations to compute cheap intervals on the quantity of interest. In the data-driven approach, we use Bayesian inference to infer the posterior distribution of the two lag coefficients from any available data. Although more expensive, again due to the low dimensional nature of the UQ problem we can easily construct accurate surrogate models which can replace of the full RANS code in the likelihood function. With this in place, informed distributions of the lag coefficients are readily available.

References

1. Slotnick, J., Khodadoust, A., Alonso, J., Darmofal, D., Gropp, W., Lurie, E., Mavriplis, D.: Cfd vision 2030 study: a path to revolutionary computational aerosciences (2014)
2. Cheung, S.H., Oliver, T.A., Prudencio, E.E., Prudhomme, S., Moser, R.D.: Bayesian uncertainty analysis with applications to turbulence modeling. Reliab. Eng. Syst. Safe **96**(9), 1137–1149 (2011)
3. Edeling, W.N., Cinnella, P., Dwight, R.P., Bijl, H.: Bayesian estimates of parameter variability in the k–ε turbulence model. J. Comp. Phys. **258**, 73–94 (2014)
4. Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T.: Bayesian model averaging: a tutorial . Stat. Sci. **14**, 382–401 (1999)
5. Edeling, W.N., Cinnella, P., Dwight, R.P.: Predictive rans simulations via bayesian model-scenario averaging. J. Comp. Phys. **275**, 65–91 (2014)
6. Emory, M., Larsson, J., Iaccarino, G.: Modeling of structural uncertainties in reynolds-averaged navier-stokes closures. Phys. Fluids **25**(11), 110822 (2013)
7. Gorlé, C., Iaccarino, G.: A framework for epistemic uncertainty quantification of turbulent scalar flux models for reynolds-averaged navier-stokes simulations. Phys. Fluids **25**(5), 055105 (2013)
8. Olsen, M.E., Coakley, T.J.: The lag model, a turbulence model for non equilibrium flows. In: AIAA Computational Fluid Dynamics Conference, 15 th, Anaheim, CA (2001)
9. Pope, S.B.: Turbulent flows (2001)
10. Wilcox, D.C.: Turbulence modeling for cfd (1998)
11. Banerjee, S., Krahl, R., Durst, F., Zenger, C.: Presentation of anisotropy properties of turbulence, invariants versus eigenvalue approaches. J. Turbul. **8**, N32 (2007)
12. Kindlmann, G.: Superquadric tensor glyphs. In: Proceedings of the 6th Joint Eurographics-IEEE TCVG conference on Visualization, pages 147–154. Eurographics Association (2004)
13. Ramachandran, P., Varoquaux, G.: Mayavi: 3d visualization of scientific data. Comput. Sci. Eng. **13**(2), 40–51 (2011)
14. Emory, M., Pecnik, R., Iaccarino, G.: Modeling structural uncertainties in reynolds-averaged computations of shock/boundary layer interactions. AIAA Paper **479**, 2011 (2011)
15. Gorlé, C., Emory, M., Larsson, J., Iaccarino, G.: Epistemic uncertainty quantification for rans modeling of the flow over a wavy wall. Center for Turbulence Research, Annual Research Briefs (2012)
16. Lillard, R.P., Oliver, A.B., Olsen, M.E., Blaisdell, G.A., Lyrintzis, A.S.: The lagrst model: a turbulence model for non-equilibrium flows. AIAA Paper **444**, 2012 (2012)
17. Rotta, J.C.: Statistische theorie nichthomogener turbulenz. Zeitschrift für Physik **129**(6), 547–572 (1951)

18. Revell, A.J., Benhamadouche, S., Craft, T., Laurence, D.: A stress–strain lag eddy viscosity model for unsteady mean flow. Int. J. Heat Fluid Fl. **27**(5), 821–830 (2006)
19. Ling, J., Templeton, J.: Evaluation of machine learning algorithms for prediction of regions of high reynolds averaged navier stokes uncertainty. Phys. Fluids **27**(8), 085103 (2015)
20. Langley Research Center Turbulence Modeling Resource. Axisymmetric subsonic jet. https://turbmodels.larc.nasa.gov/jetsubsonic_val.html, 2017. [Retrieved February-2017]
21. Lay, D.C.: Linear algebra and its applications (2006)
22. Schumann, U.: Realizability of reynolds-stress turbulence models. Phys. Fluids **20**(5), 721–725 (1977)
23. Durbin, P.A., Speziale, C.G.: Realizability of second-moment closure via stochastic analysis. J. Fluid Mech. **280**, 395–407 (1994)
24. Edeling, W.N., Iaccarino, G., Cinnella, P.: A return to eddy viscosity model for epistemic uq in rans closures. Center for Turbulence Research, Annual Research Briefs (2016)
25. Le, H., Moin, P., Kim, J.: Direct numerical simulation of turbulent flow over a backward-facing step . J. fluid Mech. **330**(1), 349–374 (1997)
26. Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B.: Bayesian data analysis, vol. 2. Chapman & Hall/CRC, Boca Raton (2014)
27. Kennedy, M.C., O'Hagan, A.: Bayesian calibration of computer models. J. Roy. Stat. Soc. B **63**(3), 425–464 (2001)
28. Oliver, T.A., Moser, R.D.: Bayesian uncertainty quantification applied to rans turbulence models. In: J Phys Conf Ser, volume 318, page 042032. IOP Publishing (2011)
29. Chib, S., Greenberg, E.: Understanding the metropolis-hastings algorithm. Am. Stat. **49**(4), 327–335 (1995)
30. Marzouk, Y., Xiu, D.: A stochastic collocation approach to bayesian inference in inverse problems. Commun. Comput. Phys. **6**, 826–847 (2009)
31. Witteveen, J.A.S., Doostan, A., Pecnik, R., Iaccarino, G.: Uncertainty quantification of the transonic flow around the rae 2822 airfoil. Center for Turbulence Research, Annual Briefs, Stanford University (2009)
32. Edeling, W.N., Dwight, R.P., Cinnella, P.: Simplex-stochastic collocation method with improved scalability. J. Comp. Phys. **310**, 301–328 (2016)
33. Ghanem, R., Spanos, P.D.: Polynomial chaos in stochastic finite elements. J. Appl. Mech. **57**(1), 197–202 (1990)
34. Witteveen, J.A.S., Iaccarino, G.: Simplex stochastic collocation with random sampling and extrapolation for nonhypercube probability spaces. SIAM J. Sci. Comput. **34**(2), A814—A838 (2012)
35. Dongbin, X., Karniadakis, G.E.: The wiener–askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput. **24**(2), 619–644 (2002)
36. Bridges, J., Wernet, M.P.: The nasa subsonic jet particle image velocimetry (piv) dataset (2011)
37. Le, H., Moin, P.: Direct numerical simulation of turbulent flow over a backward-facing step, Center for Turbulence Research, Annual Briefs, Stanford University (1992)
38. Jovic, S., Driver, D.M.: Backward-facing step measurements at low reynolds number, $re_h = 5000$, NASA Technical Memorandum, (108807) (1994)
39. Thangam, S., Speziale, C.G.: Turbulent flow past a backward-facing step-a critical evaluation of two-equation models. AIAA J. **30**(5), 1314–1320 (1992)
40. Iaccarino, G., Mishra, A.A., Ghili, S.: Eigenspace perturbations for uncertainty estimation of single-point turbulence closures. Phys. Rev. Fluids **2**(2), 024605 (2017)
41. Mishra, A., Iaccarino, G.: Rans predictions fo high-speed flows using enveloping models, Center for Turbulence Research, Annual Research Briefs (2016)