Bayesian estimates of the parameter variability in the $\mathrm{k}-\varepsilon$ turbulence model

W.N. Edeling,[1, a)] P. Cinnella,[1] R.P. Dwight,[2] and H. Bijl[2]

[1)] *ENSAM ParisTech, DynFluid laboratory, 151 Boulevard de l'Hospital, 75013 Paris, France*

[2)] *Delft University of Technology, Faculty of Aerospace Engineering, Kluyverweg 2, Delft, The Netherlands*

(Dated: 22 January 2013)

In this paper we are concerned with obtaining estimates for the error in Reynolds-Averaged Navier-Stokes (RANS) simulations based on the Launder-Sharma $k-\varepsilon$ turbulence closure model, for a limited class of flows. In particular we search for estimates grounded in uncertainties in the space of model closure coefficients, for wall-bounded flows at a variety of favourable and adverse pressure gradients. In order to estimate the spread of closure coefficients which reproduces these flows accurately, we perform 13 separate Bayesian calibrations – each at a different pressure gradient – using measured boundary-layer velocity profiles, and a statistical model containing a multiplicative model inadequacy term in the solution space. The results are 13 joint posterior distributions over coefficients and hyper-parameters. To summarize this information we compute Highest Posterior-Density (HPD) intervals, and subsequently represent the total solution uncertainty with a probability-box (p-box). This p-box represents both parameter variability across flows, and epistemic uncertainty within each calibration. A prediction of a new boundary-layer flow is made with uncertainty bars generated from this uncertainty information, and the resulting error estimate is shown to be consistent with measurement data.

Keywords: Bayesian calibration; parameter variability; model inadequacy; RANS turbulence model

[a)] Electronic mail: W.N.Edeling@tudelft.nl

I. INTRODUCTION

Computational Fluid Dynamics (CFD) and Reynolds-averaged Navier-Stokes (RANS) simulations in particular form an important part of the analysis and design methods used in industry. These simulations are typically based on a deterministic set of input variables and model coefficients. However real-world flow problems are subject to numerous uncertainties, e.g. imprecisely known parameters, initial- and boundary conditions. For input uncertainties described as probability density functions (pdfs), established methods exist for determining the corresponding output uncertainty[1–3]. Furthermore, numerical predictions are affected by numerical discretization errors and approximate physical models (turbulence models in RANS). The former may be estimated and controlled by means of mesh refinement (e.g. Ref. 4), but no analogue exists for the latter. This error, which we call *model inadequacy* in the following, is therefore the only major source of simulation error that remains difficult to estimate. It is therefore the bottleneck in the trustworthiness of RANS simulations. This work is an attempt to construct an estimate of model inadequacy in RANS for a limited set of flows, and for a single turbulence closure model.

Within the framework of RANS, many turbulence models are available, see e.g. Ref. 5 for a review. There is general agreement that no universally-"best" RANS turbulence closure model is currently known; the accuracy of models is problem-dependent[6]. Moreover, each turbulence model uses a number of closure coefficients which are classically determined by calibration against a database of fundamental flows[7]. Model performance may strongly depend on these values, which are often adjusted to improve model accuracy for a given set of problems, or for a specific flow code. They are almost always assumed to be constant in space and time. For a given model there is sometimes no consensus on the best values for these coefficients, and often intervals are proposed in the literature[8].

Our approach is to represent model inadequacy by uncertainty in these coefficients. Summarized we proceed as follows: (1) we define the class of flows for which we wish to estimate the error, in this article turbulent boundary-layers for a range of pressure gradients. (2) We collect experimental data for a number of flows of this class. (3) We use Bayesian model updating to calibrate the closure coefficients against each flow in this data-set, resulting in posterior distributions on the coefficients for each flow[9]. (4) We summarize this information using highest posterior-density (HPD) intervals and p-boxes. This summary gives intervals

on the coefficients which represent both the spread of coefficients within the flow-class, as well as the ability of the calibration to provide information about the values these coefficients should take in each flow. (5) For a new flow of the class, for which there might be no experimental data, we then perform a simulation using the model with the specified coefficient uncertainties. The resulting interval on the model output is our probabilistic estimate of the true flow.

Representing model inadequacy by uncertainty in closure coefficients is reasonable since the coefficients are empirical, and not perfectly flow-independent. Furthermore each coefficient is involved in an approximation of the underlying physics, and therefore is closely related to some component of the model inadequacy. Finally an error estimate based on coefficient uncertainty has the virtue of being geometry-independent – that is we do not need to assume a particular flow topology to apply the estimate. We do not claim that it is possible to approximate all turbulence model inadequacy in this way. The method does rely on being able to approximate most of it, and we demonstrate that this is possible for the limited class of flows we consider.

The key step in the method is the calibration of the coefficients. For the calibration phase we follow the work of Cheung et al.[10], in which a Bayesian approach was applied to the calibration of the Spalart-Allmaras[11] turbulence model, taking into account measurement error[12]. In that work, for a given statistical model, the coefficients were calibrated once on all the available measured velocity profiles and wall-shear stress components. Model inadequacy was treated with a multiplicative term parameterized in the wall-normal direction with a Gaussian process, following the framework of Kennedy and O'Hagan[13]. In the present work, we perform an analysis by performing separate calibrations on multiple flows in our class, using the $k-\varepsilon$ model, with Launder-Sharma damping functions[14]. Using uniform priors and calibrating against a large, accurate data-set containing boundary-layer profiles at different pressure gradients, results in informative coefficient posteriors for each flow. The multiplicative model inadequacy term is retained to capture the part of the error which cannot be captured by the closure coefficients alone.

We choose the pressure gradient as the independent variable in our flow class because it is known to have a large impact on the performance of $k-\varepsilon$ model[15–17]. Approaching this problem in a Bayesian context allows us to estimate how much this deficiency can be reduced by choice of closure coefficients alone, and how much the coefficients have to vary

to match measurements at all pressure-gradients. The spread of coefficients is an indication of flow-independence of the model, and we expect better models to have smaller spreads.

The paper is laid out as follows: we briefly outline the framework of Bayesian data analysis in Section II, and the $k-\varepsilon$ model in Section III. Section V describes our calibration framework, in particular the statistical model and priors. The results, including verification, HPD analysis of calibration posteriors, and prediction using the obtained coefficient uncertainties are described in Section VI. Finally the limitations of the $k-\varepsilon$ model for strong adverse pressure gradients under any set of closure coefficients is investigated in Section VII.

II. GENERAL BAYESIAN DATA ANALYSIS

By Bayesian data analysis, we mean practical methods for making inferences from data using joint probability models for quantities we observe and for quantities we wish to learn about[9]. Bayesian data analysis is often applied to problems in uncertainty quantification, see e.g. Refs. 18–20. One type of uncertainty, namely *aleatoric* uncertainty, arises through natural random variations of the process. This type of uncertainty is irreducible (intrinsic to the system), such that more data or better models will not reduce it. *Epistemic* uncertainty on the other hand, arises from a lack of knowledge about the model, e.g. unknown model parameters or mathematical form, and can in principle be reduced. The Bayesian framework represents epistemic uncertainty using probability, which is often used to represent only aleatory uncertainty. Hence all sources of uncertainty are probabilistic, leading to a unified treatment exploiting the tools of uncertainty quantification and Bayes' theorem. The general process for Bayesian data analysis consists of four steps: 1) define a joint-probability distribution for both the observed and unobserved quantities in the problem, 2) calibrate the model against observations, 3) validate the model, and finally 4) use the calibrated model to make predictions.

Inferences are made for two kinds of unobserved quantities, namely

1. *Parameters* that govern the model, which we denote by the column vector $\boldsymbol{\theta}$, and are treated as random variables.

2. *Future predictions* of the model. If we let $\mathbf{z} = [z_1, z_2, \cdots, z_n]$ denote the observed data, then the currently unknown (but possibly observable) future predictions are denoted

by $\tilde{\mathbf{z}}$.

In addition to parameters subject to calibration, we have a class of *explanatory* variables $\mathbf{t}$. These are the variables that we do not bother to model as random, e.g. the pressure gradients in our boundary layer calibrations, but do affect the model predictions.

In short, the aim of Bayesian data analysis is to draw conclusions about $\boldsymbol{\theta}$ (calibration) through the conditional posterior distribution $p(\boldsymbol{\theta} \mid \mathbf{z})$, or about $\tilde{\mathbf{z}}$ (prediction) through $p(\tilde{\mathbf{z}} \mid \mathbf{z})$. Here, we let $p(\cdot)$ denote a probability distribution. We can achieve this via the application of *Bayes' theorem*

$$p(\boldsymbol{\theta} \mid \mathbf{z}) = \frac{p(\mathbf{z} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\mathbf{z})}, \tag{1}$$

where the law of total probability states that $p(\mathbf{z}) = \int p(\mathbf{z} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$. Since this denominator does not depend upon $\boldsymbol{\theta}$, it is often omitted to yield the unnormalized version of (1),

$$p(\boldsymbol{\theta} \mid \mathbf{z}) \propto p(\mathbf{z} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta}). \tag{2}$$

The term $p(\mathbf{z} \mid \boldsymbol{\theta})$, i.e. the distribution of the data given the parameters is called the *likelihood function*, and it provides the means for updating the model once more data becomes available. The term $p(\boldsymbol{\theta})$ is the *prior distribution* of $\boldsymbol{\theta}$, i.e. it represents what we know about the parameters before the data became available.

The posterior predictive distribution conditional on the observed $\mathbf{z}$ can be written as

$$p(\tilde{\mathbf{z}} \mid \mathbf{z}) = \int p(\tilde{\mathbf{z}}, \boldsymbol{\theta} \mid \mathbf{z})\, d\boldsymbol{\theta} = \int p(\tilde{\mathbf{z}} \mid \boldsymbol{\theta}, \mathbf{z}) p(\boldsymbol{\theta} \mid \mathbf{z})\, d\boldsymbol{\theta} = \int p(\tilde{\mathbf{z}} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathbf{z})\, d\boldsymbol{\theta}. \tag{3}$$

The last step follows because $\tilde{\mathbf{z}}$ and $\mathbf{z}$ are assumed to be conditionally independent given $\boldsymbol{\theta}$, i.e. $p(\tilde{\mathbf{z}} \mid \mathbf{z}, \boldsymbol{\theta}) = p(\tilde{\mathbf{z}} | \boldsymbol{\theta})$.

In general the relationship between $\boldsymbol{\theta}$ and $\mathbf{z}$ involves evaluation of a numerical computer code. Therefore the posterior distribution of coefficients, $p(\boldsymbol{\theta} \mid \mathbf{z})$, and integrals such as (3) can not be evaluated analytically. Evaluation of $p(\boldsymbol{\theta} \mid \mathbf{z})$ at a single value of $\boldsymbol{\theta}$ requires at least one flow-solve, so for moderate-dimensional $\boldsymbol{\theta}$ brute-force sampling of the posterior is computationally intractable. Markov-chain Monte-Carlo (McMC) methods[21,22] are used to provide samples $\boldsymbol{\theta}^j, j = 1, 2, \cdots, J$ from $p(\boldsymbol{\theta} \mid \mathbf{z})$ at a more acceptable cost. The form of the

right-hand side of (3) then suggests using these samples to draw samples $\tilde{\mathbf{z}}^j$ from the conditional distribution $p\left(\tilde{\mathbf{z}} \mid \boldsymbol{\theta}\right)$, at which point the integral can be approximated numerically by standard Monte-Carlo.

In the case of CFD applications, sampling $\boldsymbol{\theta}$ could be much too expensive such that even McMC sampling requires an unacceptably high overall computational time. This problem may be alleviated by replacing the CFD model with an approximated unexpensive analytical model, a so-called surrogate model, such as polynomial interpolation or Kriging (see e.g. Ref. 3). This is not used in the present work since the boundary layer code used for computing the numerical solution is cheap enough to be directly coupled to McMC.

Note the result of the analysis depends on the chosen stochastic model (which defines the joint pdf $p(\mathbf{z}, \boldsymbol{\theta})$), hence the necessity for a post-calibration validation step. A way to reduce the bias introduced by the choice of a single stochastic model is to apply Bayesian model averaging, in which a Bayesian analysis is performed using multiple stochastic models M_i from a set of competing model classes $M = \{M_1, \cdots, M_m\}$. This avoids having to choose one model, by creating a weighted average model constructed from the models in the set M. The weights of the models are unknowns to be calibrated, and they can be interpreted as the level of evidence for a given model[23].

III. THE k $-$ ε TURBULENCE MODEL

The general simulation approach considered in this paper is the solution of the RANS equations for turbulent boundary layers, supplemented by a turbulence model. RANS equations remain up to now the most advanced and yet computationally acceptable simulation tool for engineering practice, since more advanced strategies, like Large Eddy Simulation (see e.g. Ref. 24) are yet too expensive for high-Reynolds flows typically encountered in practical applications. Under the assumption of incompressibility, the governing equations for a boundary-layer flow are given by

$$\frac{\partial \bar{u}_1}{\partial x_1} + \frac{\partial \bar{u}_2}{\partial x_2} = 0, \tag{4a}$$

$$\frac{\partial \bar{u}}{\partial t} + \bar{u}_1 \frac{\partial \bar{u}_1}{\partial x_1} + \bar{u}_2 \frac{\partial \bar{u}_1}{\partial x_2} = -\frac{1}{\rho}\frac{\partial \bar{p}}{\partial x_2} + \frac{\partial}{\partial x_2}\left[(\nu + \nu_T)\frac{\partial \bar{u}_1}{\partial x_2}\right], \tag{4b}$$

where ρ is the constant density, $\bar{u}_i$ is the mean velocity in x_i direction and ν is the kinematic viscosity. The eddy viscosity ν_T is meant to represent the effect of turbulent fluctuations on

the mean flow, and is calculated here through the $k-\varepsilon$ turbulence model:

$$\nu_T = C_\mu f_\mu \frac{k^2}{\tilde{\varepsilon}}, \tag{5a}$$

$$\frac{\partial k}{\partial t} + \bar{u}\frac{\partial k}{\partial x_1} + \bar{v}\frac{\partial k}{\partial x_2} = \nu_T \left(\frac{\partial \bar{u}}{\partial x_2}\right)^2 - \varepsilon + \frac{\partial}{\partial x_2}\left[\left(\nu + \frac{\nu_T}{\sigma_k}\right)\frac{\partial k}{\partial x_2}\right], \tag{5b}$$

$$\frac{\partial \tilde{\varepsilon}}{\partial t} + \bar{u}\frac{\partial \tilde{\varepsilon}}{\partial x_1} + \bar{v}\frac{\partial \tilde{\varepsilon}}{\partial x_2} = C_{\varepsilon 1} f_1 \frac{\tilde{\varepsilon}}{k}\nu_T \left(\frac{\partial \bar{u}}{\partial x_2}\right)^2 - C_{\varepsilon 2} f_2 \frac{\tilde{\varepsilon}^2}{k} + E + \frac{\partial}{\partial x_2}\left[\left(\nu + \frac{\nu_T}{\sigma_\varepsilon}\right)\frac{\partial \tilde{\varepsilon}}{\partial x_2}\right], \tag{5c}$$

see Ref. 5. Here, k is the turbulent kinetic energy and $\tilde{\varepsilon}$ is the isotropic turbulent dissipation, i.e. the term that controls the dissipation rate of k. The isotropic dissipation (which is zero at the wall) is related to the dissipation ε by $\varepsilon = \varepsilon_0 + \tilde{\varepsilon}$, where ε_0 is the value of the turbulent dissipation at $x_2 = 0$. The system (5a)-(5c) contains several closure coefficients and empirical damping functions, which act directly on these coefficients. Without the damping functions the $k-\varepsilon$ model would not be able to provide accurate predictions in the viscous near-wall region[5]. The Launder-Sharma $k-\varepsilon$ model[14] is obtained by specifying these damping functions as follows

$$f_\mu = \exp\left[-3.4/(1 + Re_T/50)\right],\ f_1 = 1,$$
$$f_2 = 1 - 0.3\exp\left[-Re_T^2\right],\ \varepsilon_0 = 2\nu\left(\frac{\partial\sqrt{k}}{\partial x_2}\right)^2,$$
$$E = 2\nu\nu_T\left(\frac{\partial^2 \bar{u}}{\partial x_2^2}\right)^2, \tag{6}$$

where $Re_T \equiv k^2/\tilde{\varepsilon}\nu$. In the case of the Launder-Sharma $k-\varepsilon$ model, the closure coefficients have the following values

$$C_\mu = 0.09, \quad C_{\varepsilon 1} = 1.44, \quad C_{\varepsilon 2} = 1.92,$$
$$\sigma_k = 1.0, \quad \sigma_\varepsilon = 1.3. \tag{7}$$

We do not expect these values to be generally applicable 'best' values, and other $k-\varepsilon$ models do use different values. For instance, the Jones-Launder model[25], which only differs from (6) by a slightly different f_μ, uses

$$C_\mu = 0.09, \quad C_{\varepsilon 1} = 1.55, \quad C_{\varepsilon 2} = 2.0,$$
$$\sigma_k = 1.0, \quad \sigma_\varepsilon = 1.3. \tag{8}$$

We refer to Wilcox[5] for further discussion on $k-\varepsilon$ type models and their limitations.

A. Classical identification of closure coefficients

The values of the closure coefficients in (7) are classically chosen by reference to fundamental flow problems. We illustrate how the nature of the coefficients leads to some ambiguity regarding their values, and how flow-independent single best values are unlikely to exist. One such a fundamental flow problem often considered is homogeneous, isotropic, decaying turbulence. In this case the k and ε equations (4a)-(5c) (without damping functions) simplify to

$$\frac{dk}{dt} = -\varepsilon, \tag{9}$$

$$\frac{d\varepsilon}{dt} = -C_{\varepsilon 2}\frac{\varepsilon^2}{k}. \tag{10}$$

These equations can be solved analytically to give

$$k(t) = k_0 \left(\frac{t}{t_0}\right)^{-n}, \tag{11}$$

with reference time $t_0 = nk_0/\varepsilon_0$ and $n = 1/(C_{\varepsilon 2} - 1)$. And thus,

$$C_{\varepsilon 2} = \frac{n+1}{n}. \tag{12}$$

The standard value for n is such that $C_{\varepsilon 2} = 1.92$. However, this is by no means a hard requirement and other models do use different values for $C_{\varepsilon 2}$. For instance, the RNG $k-\varepsilon$ model uses a modified $\tilde{C}_{\varepsilon 2} = 1.68$ and the $k-\tau$ model (essentially a $k-\varepsilon$ model rewritten in terms of $\tau = k/\epsilon$[26]) uses $C_{\varepsilon 2} = 1.83$[5]. Also, the experimental data from Ref. 27 suggests that most data agrees with $n = 1.3$, which corresponds to $C_{\varepsilon 2} = 1.77$.

The coefficient C_μ is calibrated by considering the approximate balance between production and dissipation which occurs in free shear flows, or in the inertial part of turbulent boundary layers. This balance can be expressed as

$$\mathcal{P} = \nu_t \left(\frac{\partial \bar{u}_1}{\partial x_2}\right)^2 = C_\mu \frac{k^2}{\varepsilon}\left(\frac{\partial \bar{u}_1}{\partial x_2}\right)^2 = \varepsilon. \tag{13}$$

Equation (13), can be manipulated together with the turbulent-viscosity hypothesis $-\overline{u_1' u_2'} = \nu_t \partial \bar{u}_1/\partial x_2$ to yield $-\overline{u_1' u_2'} = \varepsilon(\partial \bar{u}/\partial x_2)^{-1}$, which in turn yields

$$C_\mu = \left(\frac{\overline{u_1' u_2'}}{k}\right)^2. \tag{14}$$

The DNS data from Ref. 28 can be used to show that $\overline{u_1' u_2'} \approx -0.30k$ (except close to the wall), such that $C_\mu = 0.09$ is the recommended value. Again however, different models use different values for C_μ, such as $C_\mu \approx 0.085$ in the case of the RNG $k - \varepsilon$ model.

Another fundamental flow to be considered is fully developed (so $Dk/Dt = D\varepsilon/Dt = 0$) channel flow. The resulting simplified governing equations allows us to find the following constraint amongst several parameters[8]

$$\kappa^2 = \sigma_\varepsilon C_\mu^{1/2} \left(C_{\varepsilon 2} - C_{\varepsilon 1} \right), \tag{15}$$

where κ is the von-Karman constant. It should be noted that the suggested values (7) satisfy this constraint only approximately. Using (7) in (15) gives $\kappa \approx 0.43$, instead of the 'standard' value of 0.41.

The following constraint (between $C_{\varepsilon 1}$ and $C_{\varepsilon 2}$) can be found by manipulating the governing equations of uniform (i.e. $\partial \bar{u}_1 / \partial x_2 = \text{constant}$) shear flows[8]

$$\left(\frac{\mathcal{P}}{\varepsilon} \right) = \frac{C_{\varepsilon 2} - 1}{C_{\varepsilon 1} - 1}, \tag{16}$$

where the non-dimensional parameter $\mathcal{P}/\varepsilon$ is the ratio between the turbulent production $\mathcal{P}$ and dissipation ε. Tavoulakis et. al.[29] measured $\mathcal{P}/\varepsilon$ for several uniform shear flows. They reported values between 1.33 and 1.75, with a mean around 1.47. Note however, that (16) becomes 2.09 with the standard values for $C_{\varepsilon 1}$ and $C_{\varepsilon 2}$, which is significantly different from the mentioned experimental values.

The parameter σ_k can be considered as a turbulent Prandtl number, defined as the ratio of the momentum eddy diffusivity and the heat-transfer eddy diffusivity. These quantities are usually close to unity, which is why the standard value for σ_k is assumed to be 1.0. As noted in Ref. 30, no experimental data can be found to justify this assumption. And again, we see a range of recommended values amongst the different variations of the $k - \varepsilon$ model. For instance, the RNG $k - \varepsilon$ model uses $\sigma_k = 0.72$[5].

The parameter σ_ε controls the diffusion rate of ε, and its value can be determined by using the constraint (15), i.e.

$$\sigma_\varepsilon = \frac{\kappa^2}{C_\mu^{1/2} \left(C_{\varepsilon 2} - C_{\varepsilon 1} \right)}. \tag{17}$$

Finally, it should be noted that the 'constant' value of the von Karman constant (0.41) is being questioned. An overview of experimentally determined values for κ is given in Ref. 31, which reports values of κ in $[0.33, 0.45]$

B. Numerical solution of the k − ε model

To obtain efficient numerical solutions for the boundary-layer problem (4a)-(5c) we used the program EDDYBL of Ref. 32, which we modified slightly to make it more suitable for our purpose. EDDYBL is a two-dimensional (or axisymmetric), compressible (or incompressible) boundary-layer program for laminar, transitional and turbulent boundary layers. This program has evolved over three decades and is based on a code originally developed by Price and Harris in 1972[32]. The advantage of using a boundary-layer approximation rather than a full RANS code, is that a boundary-layer code allows for quicker numerical simulations, and thus avoids the need of a surrogate model.

Parabolic systems of equations such as the boundary-layer equations can, in general, be solved using unconditionally stable numerical methods. EDDYBL uses the variable-grid method of Blottner[33], which is a second-order accurate finite-difference scheme designed to solve the turbulent boundary-layer equations. This scheme uses a three-point forward-difference formula in the stream-wise direction, central differencing for the normal convection term and conservative differencing for the diffusion terms.

We verify that the discretization error is small enough such it does not dominate over the uncertainties we want to quantify. The rate at which the grid-point spacing increases in normal direction is set such that the first grid point satisfies $\Delta y^+ < 1$, which provides a good resolution in the viscous layer. Initially, the maximum number of points in the normal direction is set to 101, although EDDYBL is capable of adding more points if needed to account for boundary-layer growth. The maximum number of stream-wise steps is set high enough such that EDDYBL has no problems reaching the specified s_{stop}, i.e. the final arc length in stream-wise direction. Using this setup we verify that the discretization errors are substantially smaller than the uncertainties present in the model and data. To give an example of the magnitude of the discretization error, we computed the boundary layer over the curved airfoil-shaped surface of Ref. 34 with $s_{stop} = 20.0$ $[ft]$ for both our standard mesh with the first grid point below $y^+ = 1$, and on a finer mesh with the first 15 points below $y^+ = 1$. The maximum relative error between the two predicted velocity profiles was roughly 0.3%, which is well below the expected variance in the model output that we might see due to for instance the uncertainty in the closure coefficients. Discretization error is assumed to be negligible hereafter.

IV. EXPERIMENTAL DATA

EDDYBL comes with configuration files which mimic the experiments described in the 1968 AFOSR-IFP-Stanford conference proceedings[35]. From this data source, we selected one zero pressure-gradient flow, and 12 flows from other types of available pressure gradients, which range from favorable ($d\bar{p}/dx < 0$) to strongly adverse ($d\bar{p}/dx > 0$) gradients. These 13 flows are described in table I. The identification number of each flow is copied from Ref. 35. According to Ref. 17, the flows are identified as being 'mildly adverse', 'moderately adverse' etc, based upon qualitative observations of the velocity profile shape with respect to the zero-pressure gradient case. We plotted the experimentally determined, non-dimensional, streamwise velocity profiles in Figure 1. As usual, the normalized streamwise velocity is defined as $u^+ \equiv \bar{u}_1/\sqrt{\tau_w/\rho}$, where τ_w is the wall-shear stress. The normalized distance to the wall, displayed on the horizontal axis of Figure 1, is $y^+ \equiv x_2\sqrt{\tau_w/\rho}/\nu$. Too much weight should not be given to the classifications of the severity of the adverse gradients, since some flows (such as 2400) experience multiple gradient types along the spanwise direction. Also, when we try to justify the classification based upon the velocity profile shape we find some discrepancies. For instance, based upon the profile shape alone, we would not classify flow 1100 as mildly adverse, or 2400 as moderately adverse.

To obtain an estimate of the spread in closure coefficients, we calibrate the $k-\varepsilon$ model for each flow of table I separately, using one velocity profile as experimental data. We omit any experimental data in the viscous wall region. Since damping functions (6) dominate the closure coefficients in this region, little information is obtained from the measurements here.

V. CALIBRATION FRAMEWORK

Bayesian calibration requires selection of a prior and joint pdf (or statistical model). In our turbulence model calibration we have a large number of accurate observations, and a belief that model inadequacy will dominate the error between reality and prediction. In this situation we expect the prior on closure coefficients to be substantially less influential than the joint pdf. We therefore impose uniform priors on closure coefficients, on intervals chosen to: (i) respect mild physical constraints, and (ii) ensure the solver converges in most cases.

TABLE I. Flow descriptions, source[35].

Identification	Type	Description
1400	Zero	Equilibrium boundary layer at constant pressure
1300	Fav	Near-equilibrium boundary layer in moderate negative pressure gradient
2700	Fav	Equilibrium boundary layer in mild negative pressure gradient
6300	Fav	Near-equilibrium boundary layer growing beneath potential flow on model spillway
1100	Mild adv	Boundary layer in diverging channel
2100	Mild adv	Boundary layer on large airfoil-like body; pressure gradient first mildly negative, then strongly positive, with eventual separation
2500	Mild adv	Equilibrium boundary layer in mild positive pressure gradient
2400	Mod adv	Initial equilibrium boundary layer in moderate positive pressure gradient; pressure gradient abruptly decreases to zero, and flow relaxes to new equilibrium
2600	Mod adv	Equilibrium boundary layer in moderate positive pressure gradient
3300	Mod adv	Boundary layer, initially at constant pressure, developing into equilibrium flow in moderate positive pressure gradient
0141	Str adv	Boundary-layer with strong adverse pressure gradient, source[15]
1200	Str adv	Boundary layer in diverging channel with eventual separation
4400	Str adv	Boundary layer in strong positive pressure gradient

After the calibration we perform model checking, verifying that the posterior is not unduely constrained by the prior intervals.

To specify the joint pdf we start from the framework of Cheung et. al.[10], who use a multiplicative model inadequacy term, modeled as a Gaussian process in the wall-normal direction. By considering multiple different flows we have additional modeling choices. Unlike Cheung et. al., we choose to calibrate closure coefficients and model-inadequacy hyper-

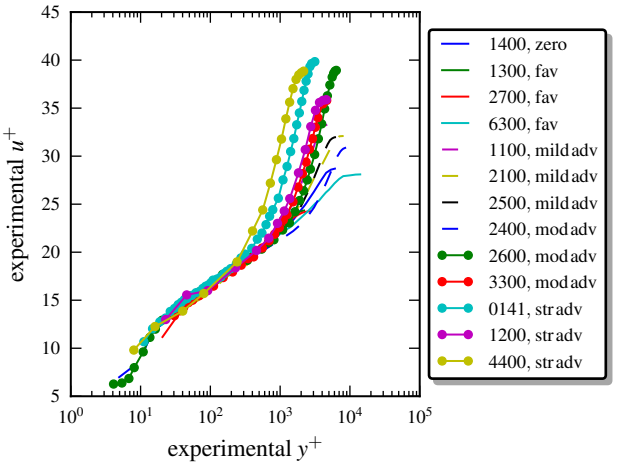


FIG. 1. The experimental data from Ref. 35.

parameters independently for each flow, and examine the variability between flows in a post-calibration step.

Let the experimental observations from flow-case $k \in \{1, \cdots, N_C\}$ be $\mathbf{z}_k = [z_k^1, \cdots, z_k^{N_k}]$. Here N_k is the number of scalar observations in flow-case k, and z_k^i is the scalar observation at location $y_k^{+,i} > 0$, where in the following we work in y^+-units. Following Ref. 10, we assume the observation noise $\boldsymbol{\lambda}_k = [\lambda_k^1, \cdots, \lambda_k^{N_k}]$ is known and uncorrelated at all measurement points. Furthermore, the closure coefficients and flow parameters for case k are denoted $\boldsymbol{\theta}_k$ and $\mathbf{t}_k$ respectively. The flow parameters include specification of the pressure-gradient as a function of the x-coordinate. The observation locations $\mathbf{y}_k^+$, noise $\boldsymbol{\lambda}_k$, and flow parameters $\mathbf{t}_k$ are modeled as precisely known explanatory variables. In the case that substantial uncertainties existed in the experiments these could be modeled stochastically as nuisance parameters.

A statistical model accounting for additive Gaussian noise in the observations and model inadequacy via a multiplicative term is: $\forall k \in \{1, \cdots, N_C\}$

$$\mathbf{z}_k = \zeta_k(\mathbf{y}_k^+) + \mathbf{e}_k, \tag{18a}$$

$$\zeta_k(\mathbf{y}_k^+) = \eta_k(\mathbf{y}_k^+) \cdot u^+(\mathbf{y}_k^+, \mathbf{t}_k; \boldsymbol{\theta}_k), \tag{18b}$$

where $u^+(\cdot, \cdot; \cdot)$ is the simulation code, and is applied element-wise to its arguments. Observational noise is modeled as

$$\mathbf{e}_k \sim \mathrm{N}\left(\mathbf{0}, \Lambda_k\right), \quad \Lambda_k := \mathrm{diag}(\boldsymbol{\lambda}_k),$$

and the model-inadequacy term $\eta_k(\cdot)$ is a stochastic process in the wall-distance y^+. Therefore (18a) represents the difference between the true process ζ_k and the measurement observations, and (18b) the difference between ζ_k and model predictions. Together they relate $\boldsymbol{\theta}_k$ to $\mathbf{z}_k$.

Cheung et. al. consider three models of this form, which differ only in the modeling of η. They compared the posterior evidence, and showed that modeling η as a correlated Gaussian process yielded by far highest evidence of the three models considered[10]. We therefore adopt the same strategy and model each η_k as a Gaussian process of unit mean (dropping the subscript k for convenience):

$$\eta \sim \mathrm{GP}(1, c_\eta), \tag{19}$$

and covariance function

$$c_\eta(y^+, y^{+\prime} \mid \boldsymbol{\gamma}) := \sigma^2 \exp\left[-\left(\frac{y^+ - y^{+\prime}}{10^\alpha l}\right)^2\right],$$

where $y+$ and $y^{+\prime}$ represent two different measurement points along the velocity profile, and l is a user-specified length scale. We fix this length scale to 5.0, which is the y^+ value that denotes the end of the viscous wall region. The smoothness of the model-inadequacy term is controlled by the correlation-length parameter α, and its magnitude by σ. This can be seen in Figure 2 which shows the mean and standard deviation of $\eta(\cdot)$, computed from 500 samples of (19). Both α and σ require to be calibrated from the data, and form a hyper-parameter vector $\boldsymbol{\gamma} := [\alpha, \sigma]$. Together (18b) and (19) imply a model inadequacy which increases with increasing velocity. A consequence is that the true process ζ is also modeled as a Gaussian process:

$$\begin{aligned} \zeta \mid \boldsymbol{\theta}, \boldsymbol{\gamma} &\sim \mathrm{GP}(\mu_\zeta, c_\zeta) \\ \mu_\zeta(y^+ \mid \boldsymbol{\theta}) &= u^+(y^+, \mathbf{t}; \boldsymbol{\theta}) \\ c_\zeta(y^+, y^{+\prime} \mid \boldsymbol{\theta}, \boldsymbol{\gamma}) &= u^+(y^+, \mathbf{t}; \boldsymbol{\theta}) \cdot c_\eta(y^+, y^{+\prime} \mid \boldsymbol{\gamma}) \cdot \\ &\qquad u^+(y^{+\prime}, \mathbf{t}; \boldsymbol{\theta}), \end{aligned} \tag{20}$$

which is still centered around the code output. The assumption of normality is made mainly for convenience, and more general forms are possible. A more boundary-layer specific model than (19), which takes the multi-scale structure of the boundary layer into account is de-

scribed in Ref. 36. It tries to achieve this goal by allowing the correlation length to vary in y^+ direction.

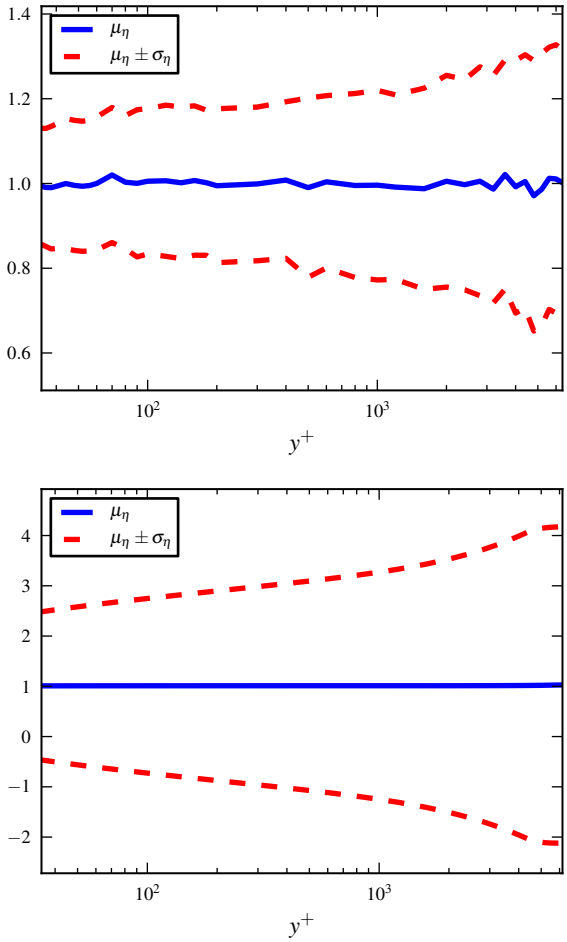


FIG. 2. The statistics of $\boldsymbol{\eta}$ for low correlation length and deviation (top, $\sigma = 0.01$, $\alpha = 0.1$), and for high correlation length and deviation (bottom, $\sigma = 0.1$, $\alpha = 3.5$).

The likelihood evaluated at the measurement locations $y^{+,i}$ can now be written for each flow case k independently as:

$$
\begin{aligned}
p(\mathbf{z} \mid \boldsymbol{\theta}, \boldsymbol{\gamma}) &= \frac{1}{\sqrt{(2\pi)^N |K|}} \exp\left[-\frac{1}{2}\mathbf{d}^T K^{-1} \mathbf{d}\right], \\
\mathbf{d} &:= \mathbf{z} - u^+(\mathbf{y}^+) \\
K &:= \Lambda + K_\zeta.
\end{aligned} \tag{21}
$$

where

$$[K_\zeta]_{ij} := c_\zeta(y^{+,i}, y^{+,j}|\boldsymbol{\theta}, \boldsymbol{\gamma}).$$

Since in general the computational grid does not coincide with measurement locations we linearly interpolate the code output at $y^{+,i}$ where needed.

Note that η includes no physics other than correlation length, whereas closure coefficients are related to specific physical approximations in the model derivation. The expectation is therefore that uncertainties on coefficients will represent model inadequacy more closely than η. Furthermore η is spatially dependent, and it is not obvious how to transplant it to make predictions of a flow with a different topology to the calibrated flow. Overall $\boldsymbol{\theta}$-like uncertainties are more general and useful than η-like uncertainties, though η is still necessary to capture remaining model inadequacy.

In Section VI the calibrated values of $\boldsymbol{\theta}_k$ for each t_k, $k \in \{1, \ldots, N_C\}$, will be used to estimate the uncertainty in $\boldsymbol{\theta}$. An alternative to this independent calibration per case, is to model closure coefficients as random variables represented as polynomial chaos expansions with unknown coefficients, i.e. a hierarchical Bayesian model with a general form of probability density[37]. This holistic approach would make better use of the data, is closer to our goal of estimating uncertainties, and will be the subject of future work.

A. Priors for $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$

Unlike Cheung et. al., we do not treat all closure coefficients as independent random variables in the prior. Instead we use the physical relations described in Section III A to constrain the value of two closure coefficients. Specifically we fix $C_{\epsilon 1}$, by rewriting (16) as

$$C_{\epsilon 1} = \frac{C_{\epsilon 2}}{\mathcal{P}/\varepsilon} + \frac{\mathcal{P}/\varepsilon - 1}{\mathcal{P}/\varepsilon}, \tag{22}$$

where, similar to Ref. 30, we fix the ratio $\mathcal{P}/\varepsilon$ to 2.09. In our results, this choice locates the mode of the posterior for $C_{\varepsilon 2}$ relatively close to the standard value of 1.92. If we instead would have used a different (experimentally determined) value of $\mathcal{P}/\varepsilon$, the mode $C_{\varepsilon 2}$ would be located elsewhere. Whether or not our choice is reasonable has to be determined by the ability of the posterior distributions to capture the observed data, as outlined in Section VI D. Two other possibilities we do not employ are: (i) to move $\mathcal{P}/\varepsilon$ into $\boldsymbol{\theta}$ and calibrate it along with the other parameters with some suitable prior, or (ii) model $\mathcal{P}/\varepsilon$ as

an aleatory uncertainty, using the $\mathcal{P}/\varepsilon$ data from Ref. 29 to construct an approximate pdf $p(\mathcal{P}/\varepsilon)$. Also, we fix σ_ε using (15). Such a choice avoids running the boundary-layer code with non-physical parameter combinations.

All priors, for both the closure coefficients $\boldsymbol{\theta}$ and hyper-parameters $\boldsymbol{\gamma}$, are independent uniform distributions. The choice of interval end-points was made based on three factors: the spread of coefficients recommended in the literature, the range of coefficients for which the solver was stable, and avoidance of apparent truncation of the posterior at the edge of the prior domain. The range we used is specified in Table II. We chose uniform distributions

TABLE II. The empirically determined range (absolute and relative to nominal value) of the uniform prior distributions.

coefficient	**left boundary**	**right boundary**
$C_{\epsilon 2}$	1.15 (-40%)	2.88 (+50%)
C_μ	0.054 (- 40 %)	0.135 (+50 %)
σ_k	0.450 (-45 %)	1.15 (+50 %)
κ	0.287 (-30 %)	0.615 (+50 %)
σ	0.0	0.1
$\log \alpha$	0.0	4.0

because we lack confidence in more informative priors for these parameters. We note however that some reasonable, informative priors can be obtained using the classical framework for coefficient identification (c.f. Section III A) in combination with multiple experimental measurements from different sources[30].

To obtain samples from the posterior distributions $p(\boldsymbol{\theta} \mid \mathbf{z})$, we employed the Markov-chain Monte Carlo (McMC) method[22]. We subsequently approximated the marginal pdf of each closure coefficient using kernel-density estimation, using the last $5,000$ (out of a total of $40,000$) samples from the Markov chain. It was observed that at $35,000$ samples, the Markov chain was in a state of statistical convergence.

VI. RESULTS AND DISCUSSION

A. Marginal posterior pdf's

Calibration of the $k-\varepsilon$ model, using the experimental data described in Section IV, and the statistical model of Section V was performed. The marginal posterior pdfs of all four parameters in $\boldsymbol{\theta}$, for all of the 13 calibration cases are shown in Figures 3a-3b. There we see that the data has been informative for $C_{\varepsilon 2}$ in all cases, resulting in sharply peaked posteriors. This is in contrast to C_μ, which has been only weakly informed. Experience suggests that for these cases the level of informativeness is predicted by the sensitivity of the Quantity of Interest (QoI) $u^+(\boldsymbol{\theta})$ to the individual parameters. Parameters with the largest sensitivities are the best identified, see section VI B. The calibrations have provided us with a bit more information regarding σ_k, since more posterior distributions show clear modes compared to the C_μ results. The spread of coefficients between cases is also visible. For instance $C_{\varepsilon 2}$ values cluster around the center of the prior interval, while κ is sharply identified at distinctly different values for the different flow conditions.

To examine dependencies between variables in the posterior, we show a two-dimensional marginal plot of the posterior $\boldsymbol{\theta}$ McMC traces for flow 3300 in Figure 4. This is a typical of result of the considered flow cases. Any trending behavior between two coefficients indicates either a positive or negative correlation, depending on the sign of the trend. As can be seen, there is a weak negative correlation between $C_{\varepsilon 2}$ and C_μ, and $C_{\varepsilon 2}$ and σ_k show a weak positive correlation. However, overall the coefficients appear largely uncorrelated, as they were in the prior.

B. Sobol indices

Figures 3a-3b show that, for a given flow case k, there is significant variation in the amount of information contained in the posterior closure-coefficient distributions, even though they are calibrated on the same data. In an attempt to explain this behavior, we perform a global sensitivity analysis on $\boldsymbol{\theta}$. If the QoI is very sensitive to a particular parameter in $\boldsymbol{\theta}$, we expect the corresponding posterior distribution to be well informed and *vice versa*.

Therefore, we perform a variance-based, global-sensitivity analysis. In such an analysis, the total variance of the QoI is decomposed into partial variances of increasing dimension-

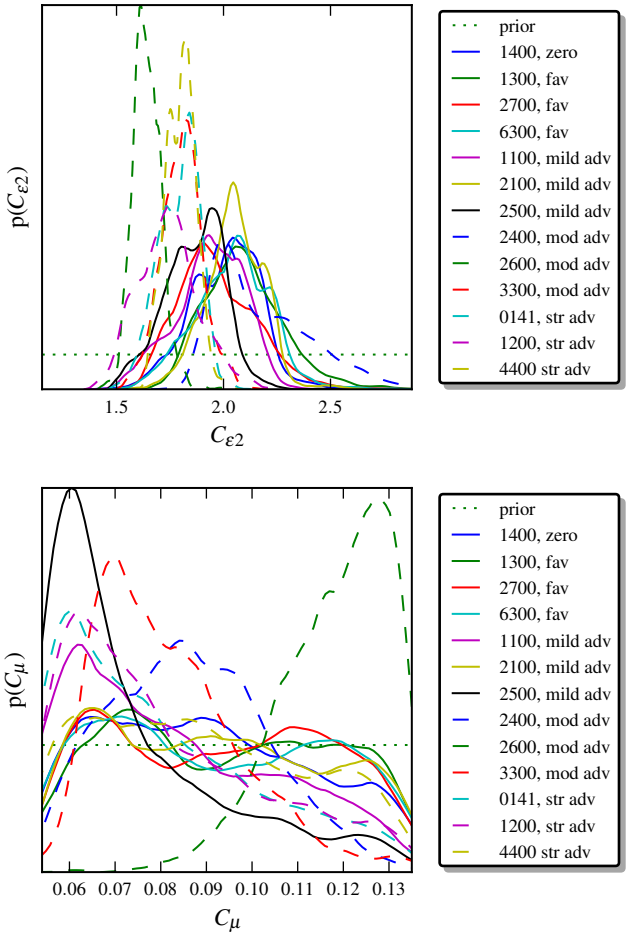


FIG. 3a. The marginal posterior distributions of the coefficients $C_{\varepsilon 2}$ and C_μ, for the 13 cases of Table I.

ality. This allows us to measure the contribution of a single parameter, or any combination of parameters, to the total QoI variance.

Let us define u as a multi-index $u \subset \mathcal{U}$ of the parameters, where $\mathcal{U} = \{1, 2, \cdots, Q\}$ is a set of integers and Q is the number of parameters in the model. Furthermore, the ratio of a partial variance over the total variance is defined as a Sobol sensitivity index. Thus, the Sobol index corresponding to the u parameters is given by

$$S_u = \frac{D_u}{D}. \tag{23}$$

Thus, D_u is the partial variance and D is the total variance. Our final note on Sobol indices is that they satisfy the property $\sum_{\mathcal{F}} S_u = 1$, where $\mathcal{F}$ is the collection of all u, i.e. $\mathcal{F} = \mathcal{P}(\mathcal{U})$, the power set of $\mathcal{U}$. For an overview of the theory behind the Sobol indices, we refer to

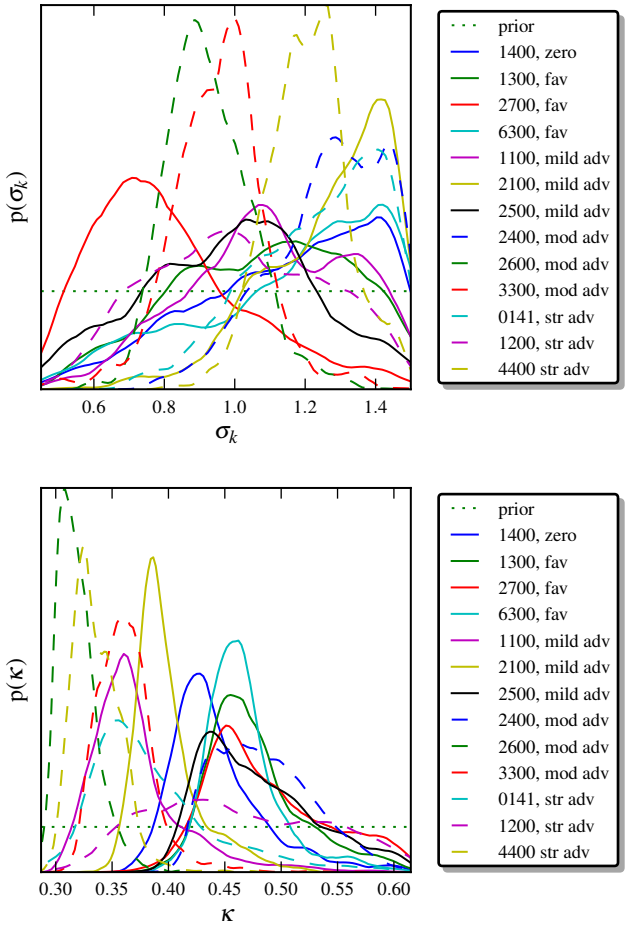


FIG. 3b. The marginal posterior distributions of the coefficients σ_k and κ, for the 13 cases of Table I.

Ref. 38.

In order to compute the partial variance D_u, several multi-dimensional integrals must be evaluated. These integrals can be approximated using MC techniques[38]. However, it is shown by Ref. 39 that the formulation of polynomial-chaos expansions is very amenable for the outlined sensitivity analysis. In similar work, Tang[40] shows that the stochastic-collocation expansion can be used for the same purpose, i.e. computing D_u at a fraction of the computational cost compared to a standard MC approach. We will apply the latter approach to the Probabilistic Collocation Method (PCM) of Ref. 41.

In the same way as the traditional MC method, the PCM is used for propagating input uncertainties through the model in order to obtain an uncertain response. The difference is that is does so at reduced computational cost, provided that the number of uncertain inputs

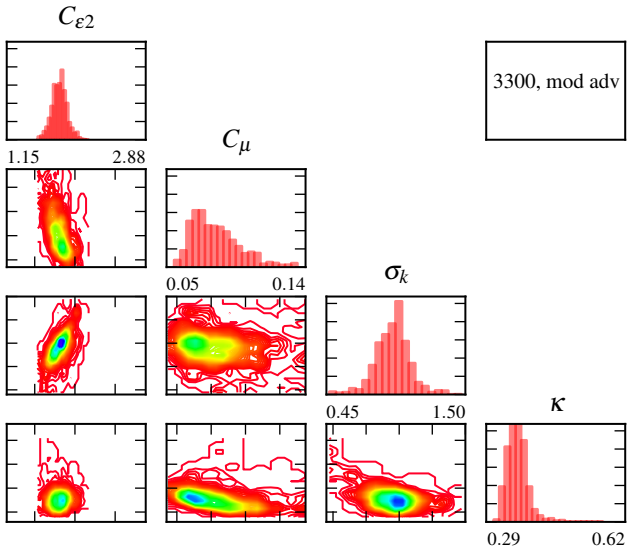


FIG. 4. A two-dimensional contour plot of the posterior $\boldsymbol{\theta}$ samples from flow 3300.

Q is not too high. In the PCM, our QoI u^+ is decomposed in a deterministic and stochastic part. More specifically, the deterministic part consists of so-called PCM coefficients, which in our case are u^+ profiles computed with the RANS solver. The stochastic dimension is spanned by Lagrange chaoses. The abscissas of the Lagrange chaoses are chosen such that they coincide with the Gaussian Quadrature (GQ) points that are used to approximate the (multi-dimensional) integrals arising in a Galerkin projection of the PCM expansion onto the Lagrange chaoses. For more information on the PCM, we refer to Ref. 41. Tang[40] shows that once the GQ weights and the PCM coefficients are calculated, the Sobol indices may be computed without any additional function evaluations.

Using the described setup, we calculate the main indices S_i with $i = 1, 2, 3, 4$ for our QoI. Here, $\{1, 2, 3, 4\}$ corresponds to $\{C_{\varepsilon 2}, C_\mu, \sigma_k, \kappa\}$. To create the required input uncertainties we use uniform distributions where we perturb each coefficient by $\pm 10\%$ from its standard value. The results are shown in Figure 5. The ranking from most sensitive parameter to least sensitive one for the velocity profile is $C_{\varepsilon 2}, \kappa, \sigma_k, C_\mu$. Notice that this is the same ranking that we get when we sort the coefficients from most informed posterior distribution to least informed one, see Figures 3a-3b. Thus, the very low sensitivity of our QoI to the value of C_μ is an explanation for the lack of information in the posterior C_μ distributions.

From a Bayesian viewpoint, the posterior distributions will get more informative if we increase the size of the experimental data set $\mathbf{z}_k$. This could be done by adding more velocity

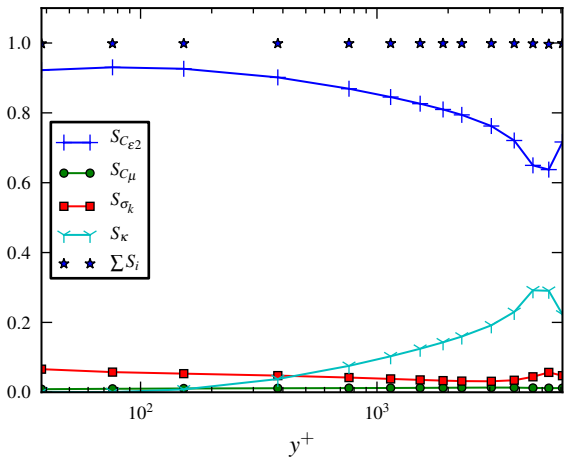


FIG. 5. The Sobol indices S_i for flow 1400, with the velocity profile as QoI. The horizontal axis represents the direction normal to the wall for the streamwise location $s = 16.3[ft]$.

profiles or by including other types of data, such as measured friction coefficients C_f. Using the Sobol indices for C_f, we might now attempt to answer the question if this will lead to better informed distributions, before we actually perform another calibration run. The S_i corresponding to the friction coefficient can be found in Figure 6.

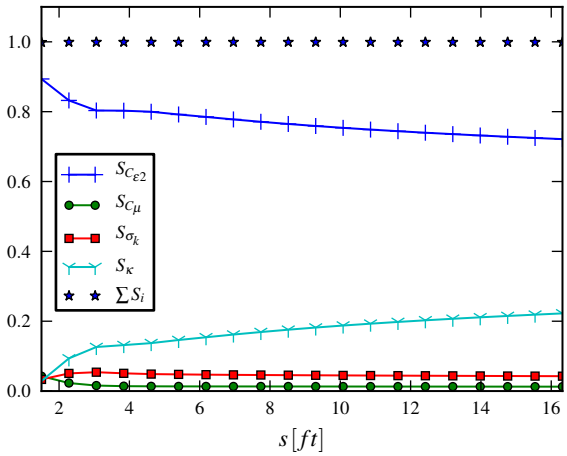


FIG. 6. The Sobol indices S_i for flow 1400, with C_f as QoI. The horizontal axis represents the streamwise direction.

Notice that the influence of C_μ is still very low. Therefore, if more informative posteriors were required, we would not choose to do an additional calibration run with added C_f data. This choice is consistent with the results from Cheung et. al.[10]. They actually did calibrate

the SA turbulence model using both velocity profiles and measured friction coefficients, and still some of their posterior distributions are quite uninformative as well.

In Figures (5)-(6) we also show $\sum S_i$, i.e. the sum of all displayed Sobol indices. This sum is very close, but not equal to 1. This indicates that for the considered range in the closure coefficients, the interaction effects are low, i.e. the Sobol indices corresponding a combination of closure coefficients are small.

C. Summary statistics

Our goal is to examine the spread of most-likely closure coefficient values. However, due to the large number of pdfs, it can be hard to obtain a clear picture of this spread in parameter space from Figures 3a-3b. We therefore use the Highest Posterior Density (HPD) statistic to summarize the results. An HPD interval is a Bayesian credible interval which satisfies two main properties, i.e.

1. The density for every point inside the interval is greater than that for every point outside the interval.

2. For a given probability content $1-\beta$, $\beta \in (0,1)$, the interval is of the shortest length.

We use the algorithm of Chen et. al.[42] to approximate the HPD intervals using the obtained McMC samples. To do so, we first sort the samples of the Q closure coefficients θ^q, $q = 1, 2, \cdots, Q$ in ascending order. Then, if we let $\{\theta^q_j,\, j = 1, 2, \cdots, J\}$ be the McMC samples from $p\left(\theta^q \mid \mathbf{z}\right)$, the algorithm basically consists of computing all the $1-\beta$ credible intervals and selecting the one with the smallest width. For a given j, we can use the empirical cumulative-distribution function to approximate the $1-\beta$ interval by computing the first θ^q_s which satisfies the inequality

$$\sum_{i=j}^{J} \mathbb{1}_{\theta^q_i \leq \theta^q_s} \geq \left[J\left(1-\beta\right)\right], \tag{24}$$

where $\mathbb{1}_{\theta^q_i \leq \theta^q_s}$ is the indicator function for $\theta^q_i \leq \theta^q_s$ and $\left[J\left(1-\beta\right)\right]$ is the integer part of $J\left(1-\beta\right)$. Secondly, if we let $\theta^q_{(i)}$ be the smallest of a set $\{\theta^q_i\}$, then the first θ^q_s for which (24) is satisfied simply is $\theta^q_{([J(1-\beta)])}$. Thus, the j^{th} credible interval is given by $\theta^q_{(j+[J(1-\beta)])} - \theta^q_{(j)}$ and the HPD interval for θ^q is found by solving

$$\min_j \ \theta^q_{(j+[J(1-\beta)])} - \theta^q_{(j)}, \quad 1 \leq j \leq J - [J\,(1-\beta)]\,. \tag{25}$$

The algorithm of Chen assumes a uni-modal posterior pdf, although it could possibly be extended to deal with multi-modal pdf's[42]. By examining Figures 3a-3b, it can be seen that our posterior distributions do not show strongly multi-modal behavior, so for sufficiently large β the above algorithm is sufficient.

The HPD results for $\boldsymbol{\theta}$ are shown in Figures 7a-7b. The spread of the posterior modes of $C_{\varepsilon 2}$ is quite concentrated, they all lie relatively close to the standard value of 1.92. The small width of the HPD intervals (compared to the prior range), indicates that the posterior distributions are informative. Also notice that a slight downward trend of the HPD intervals can be observed with increasing pressure gradient.

The spread of C_μ is relatively small, with most distributions centered close to 0.06, consistently to the left of the standard value (0.09). The only clear exception is flow 2600, which also showed some deviating behavior for $C_{\varepsilon 2}$. The parameter σ_k shows a larger spread, although in general values above the standard value of 1.0 are preferred.

Most individual pdf's of κ are quite well informed, but the modes are spread roughly between 0.31 and 0.46. Previous studies have looked at the spread of the von Karman constant. An overview is given in Ref. 31, which reports values of κ between $[0.33, 0.45]$, roughly similar to the spread that we have observed.

The spread of the κ HPD intervals in Figure 7b can be qualitatively explained by considering the deviation of the experimental velocity profiles of Figure 1 from the standard log law $1/\kappa \ln(y^+) + C$. As can be seen from Figure 1, from roughly $y^+ = 30$ the velocity profiles overlap onto the standard log law. However, around $y^+ = 200$ the first profiles start to deviate from this law. Qualitatively, the profiles which show a larger deviation from the log law, are also the ones which show a lower κ HPD interval compared to the rest.

The HPD intervals of the hyper-parameters σ and $\log \alpha$ can be found in Figure 8. Most posterior modes of σ are located near the bottom edge of the domain, indicating that for the cases we considered so far, most predictions do not deviate much from the true process. Thus, for the cases where most of the probability mass of σ is close to zero, the model inadequacy is not significantly high. The posterior modes for $\log(\alpha)$ all lie between 2.5 and 3.5, indicating that the model inadequacy is correlated over a large fraction of the boundary

layer[10]. In other words, a smooth model inadequacy term such as depicted on the right-hand side of Figure 2 is preferred over the ones depicted on the left.

Figure 8 shows three clear deviating σ HPD intervals, i.e. the ones from flows 2400, 2500 and 1200. These high HPD intervals correspond to a larger deviation of the posterior u^+ mean from the experimental data. This is an indication that in these cases the model inadequacy is becoming more severe compared to the other 10 flows.

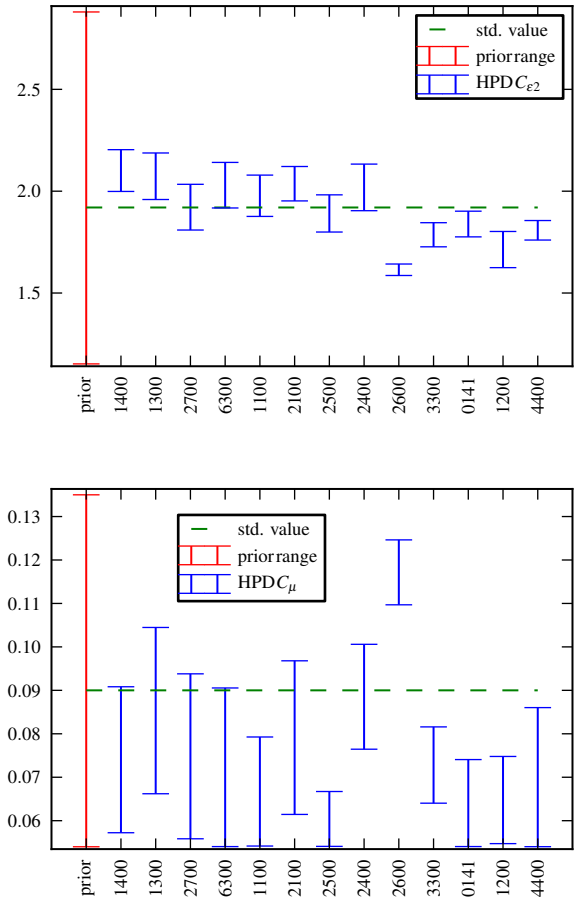


FIG. 7a. The 50 % HPD intervals of $C_{\varepsilon 2}$ and C_{μ}, for the 13 cases of Table I.

D. Posterior model check

In Bayesian analysis, it is good practice to assess the fit of the chosen model. We expect all observed data (used in the calibration) to lie within the range of the posterior predictive distribution of the true process ζ. It should be noted that this is not the same as validating

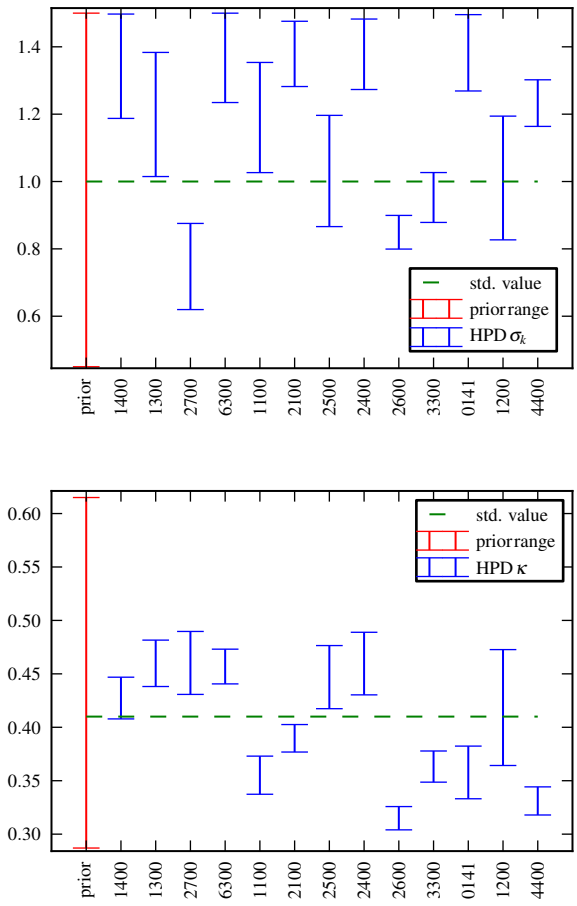


FIG. 7b. The 50 % HPD intervals of σ_k and κ, for the 13 cases of Table I.

the model, since it only ensures that the chosen model can predict the observed data. It does not ensure that it can also be used for extrapolative predictions. In our model the variability in ζ can be broken down into that due to the explicit model inadequacy term $\eta(\mathbf{y}^+)$, and that due to uncertainty in $\boldsymbol{\theta}$. The former can be obtained directly from (19) and the calibrated values of $\boldsymbol{\gamma} = [\sigma, \alpha]$. The latter is just the posterior of $u^+(\boldsymbol{\theta})$, and can be computed using the velocity traces stored during the McMC calibration run. This is equivalent to propagating posterior samples of $\boldsymbol{\theta}$ through the $k - \varepsilon$ model as in Monte-Carlo.

In Figure 9 we show only the uncertainty due to $\boldsymbol{\theta}$ for two flows. The posterior prediction of u^+ encompasses all the experimental data, and this is true for all the flows described in Table I, even for the three flows with large σ HPD values. They are therefore all consistent (in the sense of Ref. 43, i.e. existence of an overlap between the predictions and the region of experimental uncertainty). In addition, the calibrated models approximate the data better

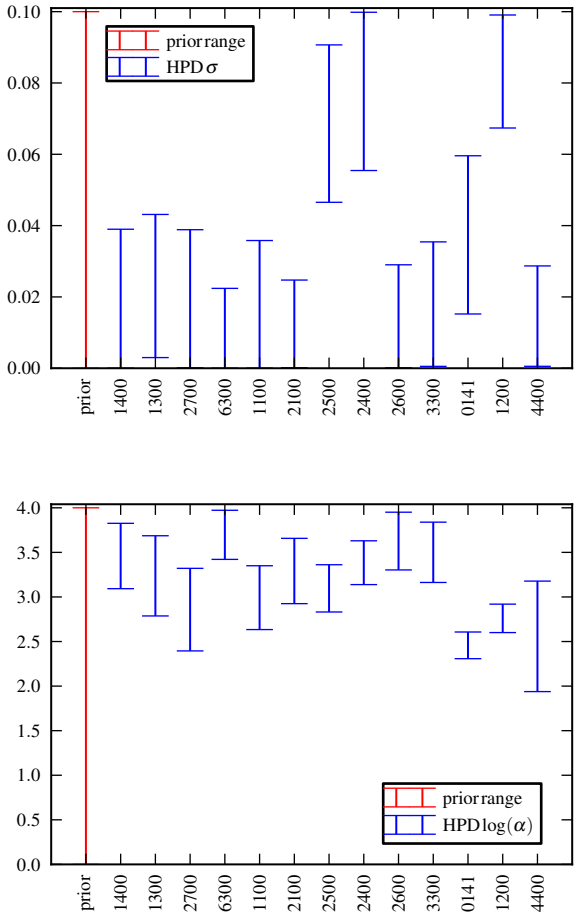


FIG. 8. The 50 % HPD intervals of $\boldsymbol{\gamma}$, for the 13 cases of Table I.

than the uncalibrated models in all cases. Based on this we judge the calibrations successful.

To illustrate the effect of using η, we compare posterior distributions of u^+ and ζ in Figure 10. The mean of both distributions is the same, which could be inferred from (20). Thus, including a model inadequacy term of the form (19) results in a posterior distribution of the true process with the same mean as the posterior u^+ distribution, but a larger variance. When making predictions with the model this contribution to the variance must be included, see Section VII.

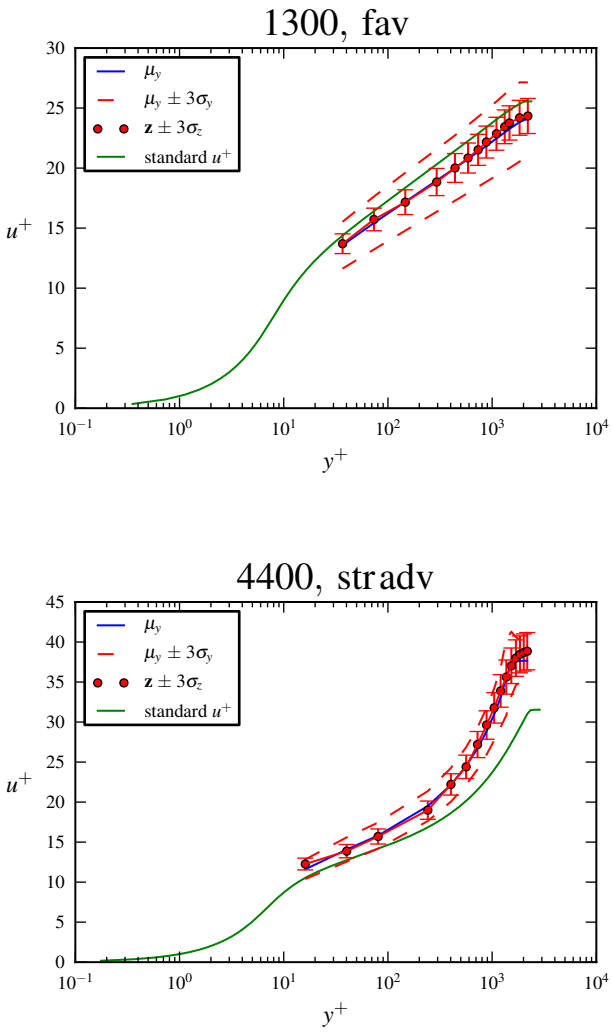


FIG. 9. The mean and 3 standard deviations of posterior $u^+(\mathbf{y}^+, \boldsymbol{\theta})$ samples of a favorable (top), and a strongly adverse flow (bottom). The green line indicates the solution of the $k-\varepsilon$ model using the standard values (7), and the red dots represent the experimental data with error bars.

E. Building a more general model for the uncertainty in the k − ε model

Performing a Bayesian calibration can tell us something about the posterior uncertainty present in the $k-\varepsilon$ model, but it does so only for the case considered in the calibration. And as can be seen from Figures 7a-8, there is significant case-to-case variability in the posterior uncertainty of the $k-\varepsilon$ model. This is true for both the closure coefficients and the hyper-parameters of the model inadequacy term.

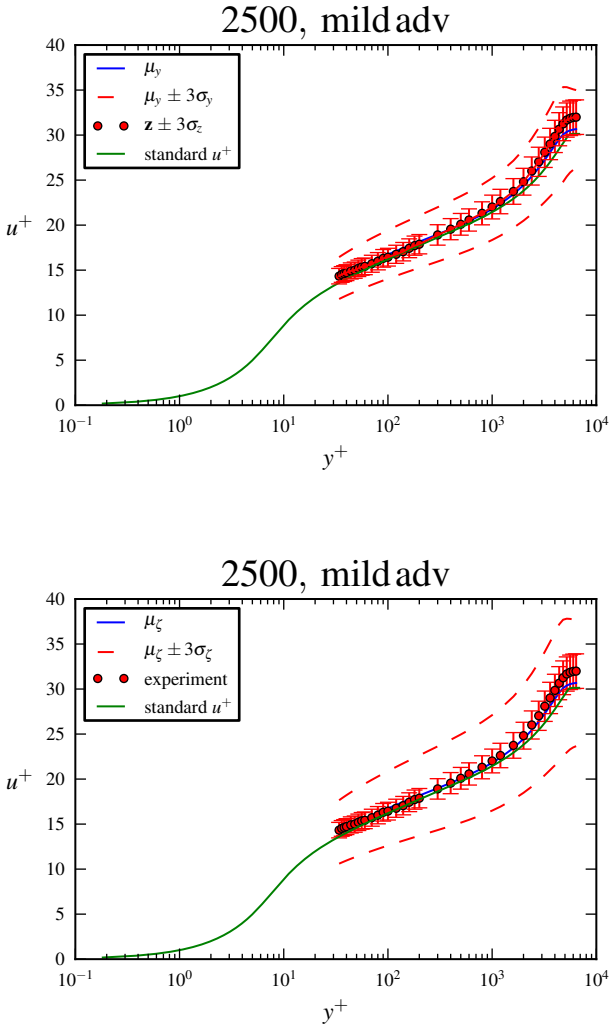


FIG. 10. The posterior distribution of $\mathbf{y}$ (top), and the posterior distribution of $\boldsymbol{\zeta}$ (bottom).

For this reason we propose to estimate the uncertainty in the QoI by combining the posterior uncertainty and the case-to-case variability of the coefficients, using all obtained posterior distributions $p_i(\boldsymbol{\theta} \mid \mathbf{z})$, $i = 1, 2, \cdots, N_C$. As a first step, we use the obtained HPD intervals to construct a probability box (p-box). A p-box is often used to visualize the range of possible outcomes in the QoI's due to epistemic uncertainty in the input parameters[44]. For a given new and uncalibrated flow case, we construct the p-box by plotting multiple distinct posterior cdfs of the QoI. In our case the distinction is due to the variability in closure coefficients over the N_C calibrated flow cases.

We approximate each of the N_C distinct posterior cdfs (all belonging to the new flow

case) by

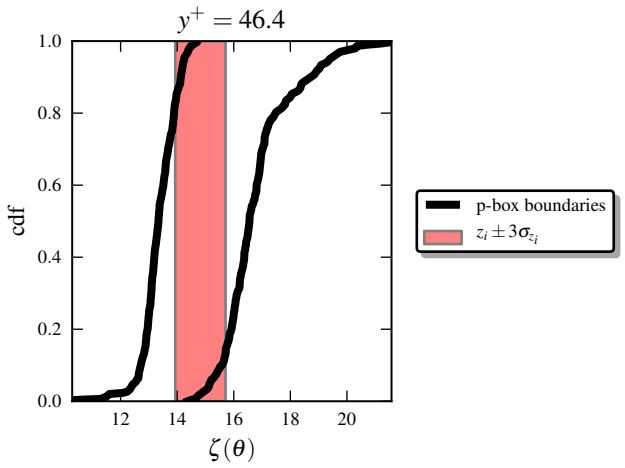
$$\mathrm{pr}\left(\zeta \leq \zeta^i\right) \approx \frac{1}{S}\sum_{j=1}^{S} \mathbb{1}_{\zeta^j < \zeta^i}, \tag{26}$$

i.e. we use empirical cumulative-distribution functions (ecdf's). Here, we use $\mathrm{pr}(\cdot)$ to denote a probability rather than a pdf, and ζ^i is a given value of ζ. Notice that the ecdf's in (26) are constructed based upon S samples from the true process ζ (20), rather than just the $k-\varepsilon$ output u^+. This way, the effect of the model inadequacy term is included in our final estimate of the uncertainty in the QoI. To obtain the S samples from ζ, we applied a straight-forward MC approach, using the HPD intervals corresponding to a certain flow case k as a uniform input distribution. We repeat this process over all N_C calibrated flow cases. Then, for a given y^+ station, we can use the S samples to construct N_C different ecdf's. The envelope formed by this collection of ecdf's is a p-box. For a given cumulative probability $\mathrm{pr}\left(\zeta \leq \zeta^i\right)$, the width of the p-box is an interval-valued QoI response.

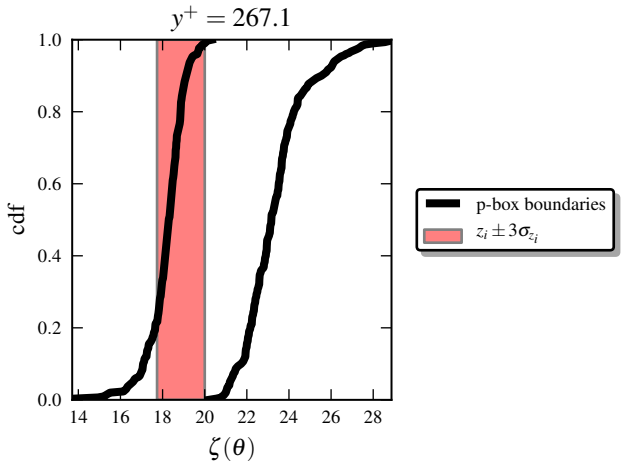
To illustrate this methodology, we applied it to a validation flow case not considered during the calibration. We used the data from Ref. 45, which is boundary-layer data on a cylinder in axially symmetric flow. This is flow 3600 from the 1968 AFOSR-IFP-Stanford conference[35]. The results for three y^+ stations ($46.2, 267.0$ and 1039.7) are given in Figure 11.

The width of the p-box denotes the range of interval valued response in the QoI, for a given cumulative probability $\mathrm{pr}\left(\zeta \leq \zeta^i\right)$. In our case, the width is determined by an inextricable mix between the case-to-case variability and the posterior uncertainty in most-likely closure coefficient values corresponding to different $d\bar{p}/dx$, and the structural uncertainty in the $k-\varepsilon$ model. If we would not have included the strong-adverse pressure gradients ecdf's, the width of the p-box would be significantly smaller.

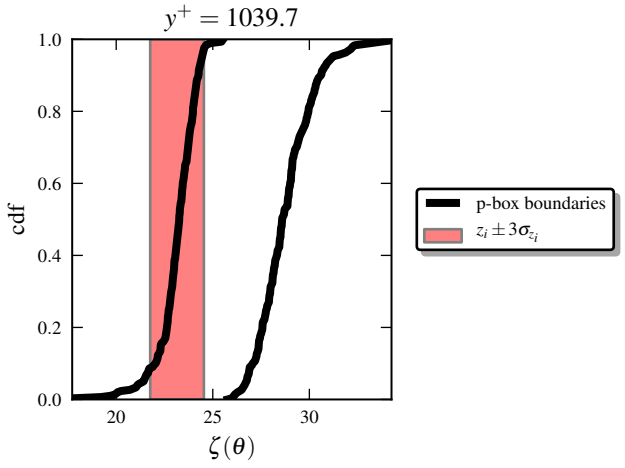
We can of course extract confidence intervals from the p-boxes of Figure 11. For instance in Figure 12 we show the 90 % confidence intervals obtained by selecting the ζ_i values corresponding to $\mathrm{pr}\left(\zeta \leq 0.05\right)$ from the left, and $\mathrm{pr}\left(\zeta \leq 0.95\right)$ from the right p-box boundaries at all experimental y^+ stations. Note that all error bars extracted from the p-boxes are consistent with the experimental data.



(a)The ζ p-box for flow 3600 at $y^+ = 46.2$.



(b)The ζ p-box for flow 3600 at $y^+ = 267.0$.



(c)The ζ p-box for flow 3600 at $y^+ = 1039.7$.

FIG. 11.

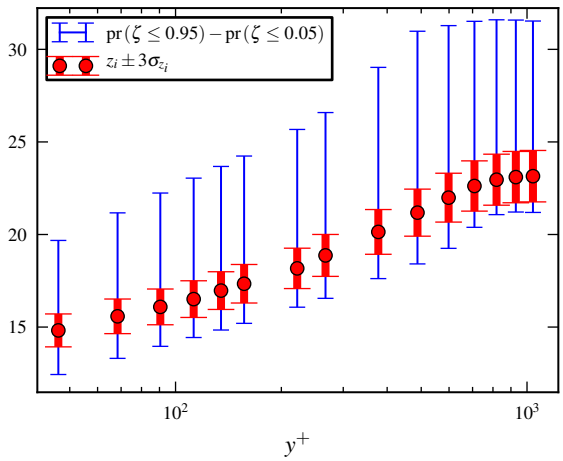


FIG. 12. The 90 % confidence intervals extracted from the N_C p-boxes.

VII. EXPLORING STRUCTURAL LIMITATIONS OF $\mathbf{k-\varepsilon}$ FOR ADVERSE PRESSURE GRADIENTS

In most cases considered so far the model inadequacy of the $k-\varepsilon$ model was relatively small, as can be seen by the HPD intervals of σ in Figure 8. In other words, we captured a significant fraction of the total uncertainty by means of the posterior closure-coefficient distributions. However, if we consider flow cases with more extreme adverse pressure gradients, the structural inadequacy of the $k-\varepsilon$ model becomes more dominant.

The $k-\varepsilon$ model is known to provide inaccurate predictions in the case of adverse pressure gradients, even for relatively simple flow problems[15–17]. To illustrate this we focus on flow 2100, which is the boundary-layer flow over a large airfoil-shaped body, investigated by Schubauer et. al.[34]. The pressure gradient is initially favorable, but increases progressively and becomes adverse in the aft part of the airfoil, with eventual separation near the trailing edge. The pressure coefficient C_p along with the standard u^+ solutions along three spanwise stations is depicted in Figure 13. Notice that the $k-\varepsilon$ model performs quite well at the favorable and zero $d\bar{p}/dx$ stations. For these cases we are confident that we would be able to obtain consistent posterior predictions by calibrating the model without a model inadequacy term, i.e. $\eta = \mathbf{1}$. For the strong adverse case however, there is a very large discrepancy between the computed solution and the experimental data. However, the profiles shown in Figure 13 are computed using the standard coefficients (7). Therefore, we do not know how much of this discrepancy is due to parametric uncertainty and how much due to model

inadequacy. Since we compute both types of uncertainty jointly using (18b), it is still hard to estimate from our Bayesian analysis how much of the observed discrepancy can be attributed to either one.

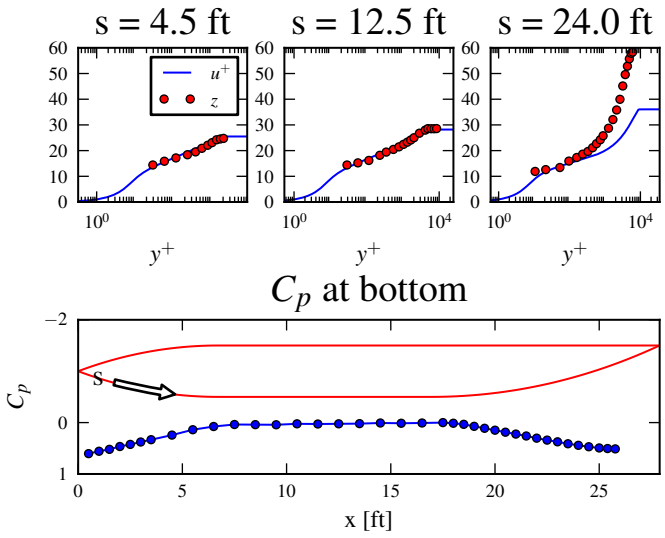


FIG. 13. The $k-\varepsilon$ solutions, using (7), on three spanwise locations of the Schubauer airfoil.

Therefore, to estimate whether the discrepancy is largely due to an inadequacy in the mathematical structure of the $k-\varepsilon$ model, we look for cases of severe model inadequacy. For this reason we calibrated flow 2100 at two stations with extremely-adverse pressure gradients. In anticipation that the model inadequacy would become more severe, we broadened the prior range of the σ hyper-parameter to $[0.0, 0.3]$. The posterior distributions of both u^+ and ζ for the first calibration (at arc length $s_{stop} = 23.5\,[ft]$), are shown in Figure 14. As can be seen from this figure, without a model inadequacy term the (calibrated) predictions of u^+ still fail to be completely consistent. The last part of the predicted defect layer has no real overlap with the region of experimental uncertainty.

The situation deteriorates when we move further downstream. We show the same results for the calibration at $s_{stop} = 24.0\,[ft]$ in Figure 15. Note that in both cases, the statistical model with a model-inadequacy term does result in a posterior distribution which captures the experimental data. This indicates the superiority of a statistical model of the form (20) in the case of extremely adverse pressure gradients.

The downside of the correlated-Gaussian model-inadequacy of Section V, is that it can only capture the extreme-adverse pressure-gradient data with the tail of the $p\,(\zeta \mid \boldsymbol{\theta})$ distri-

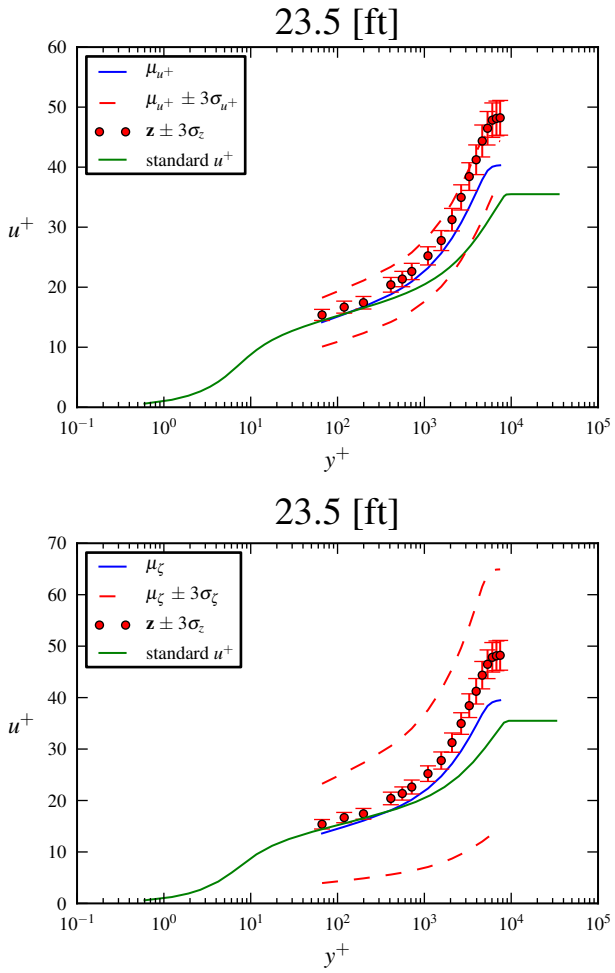


FIG. 14. The posterior distribution of u^+ (top), and the posterior distribution of ζ (bottom) at $s_{stop} = 23.5\,[ft]$.

bution. Since its mean equals u^+, this model is not capable of representing any bias in the $k-\varepsilon$ predictions. Or in other words, since the stochastic part (i.e. η) of ζ has a unit mean, any failure of the posterior ζ mean to capture the data can be attributed to a structural inadequacy of the $k-\varepsilon$ model. Since we assumed that ζ is normally distributed, its posterior mean is also the posterior mode, i.e. the 'best-fit' of $u^+(\boldsymbol{\theta})$ under the assumed form of the statistical model (20). The consequence is that the posterior uncertain range of ζ becomes very large. Especially notice the $\mu_\zeta - 3\sigma_\zeta$ boundary of figure 15, which is extremely close to zero.

To capture the data with the tail, the mode of the posterior σ distribution (and thus the corresponding HPD interval), must be shifted to higher values. This can be seen in Figure

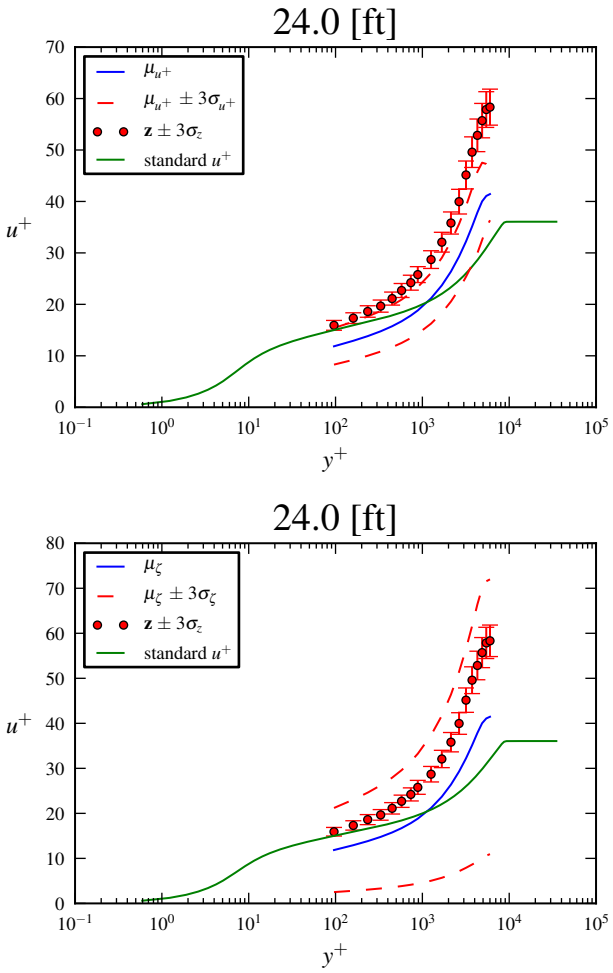


FIG. 15. The posterior distribution of u^+ (top), and the posterior distribution of ζ (bottom) at $s_{stop} = 24.0\,[ft]$.

16. The modes in Figure 16 are almost one order of magnitude higher than the σ modes in Figure 8.

Thus, as expected, a high posterior σ mode indicates a region of significant model inadequacy, one where the posterior u^+ distribution might not be consistent with the data.

VIII. CONCLUSION

Using a Bayesian framework, we performed 13 separate calibrations of the closure coefficients in the standard $k-\varepsilon$ model. The experimental data on which we calibrated consisted of velocity profiles from 13 boundary-layer flows, each subject to a different pressure gradi-

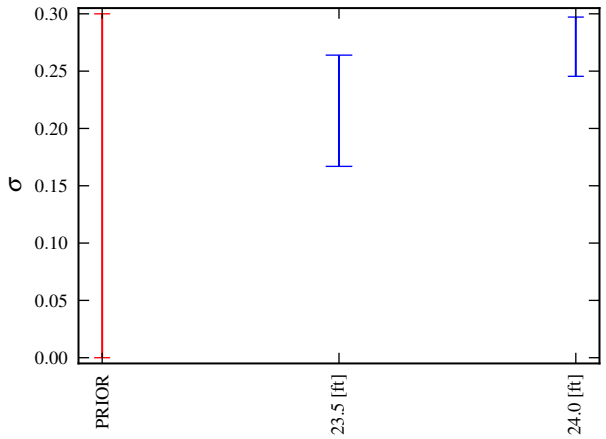


FIG. 16. The posterior 50 % σ HPD intervals for $s_{stop} = 23.5\,[ft]$ and $s_{stop} = 24.0\,[ft]$.

ent. The considered gradients ranged from favorable to ones labeled as 'strongly adverse'. This allowed us to investigate the resulting spread of the posterior parameter distributions, caused by the large range of considered pressure gradients. To summarize the spread we perform a Highest Posterior Density (HPD) analysis on all posterior distributions, which gives us 13 credible intervals of most-likely values for both the closure coefficients and the hyper-parameters used to parameterize the model-inadequacy term.

These results show a significant variation in the most-likely closure-coefficient values due to the different pressure gradients. Also, not all posterior distributions are equally informative. The closure coefficient which is the most informed by the calibrations is $C_{\varepsilon 2}$, which a sensitivity analysis based on Sobol indices shows is also the parameter with highest impact on the computed velocity. On the other hand, coefficients which have a small impact on our quantity of interest (e.g. σ_k), show a larger posterior variance. It should be noted that informative posteriors can also show a large spread. For instance κ is well informed by the data, but the individual posterior modes are not grouped closely together.

The case-to-case variability in the closure coefficients means that an individual calibration does not provide us with a general representation of the uncertainty present in the $k - \varepsilon$ model. To remedy this, we use all obtained HPD results to build a more general measure of error, which we represent through a probability-box (p-box) of the quantity of interest. The uncertainty displayed in the p-box results from the spread and uncertainty in the best-fit closure coefficients values, and the uncertainty introduced by the model-inadequacy term.

In the initial 13 calibrations, we tried to capture most of the output uncertainty with the posterior closure-coefficient distributions. However, in the final part of this article, we show that in operating conditions of extremely-adverse pressure gradients the uncertainty in the output cannot be captured by the closure coefficients alone. In such cases, the inclusion of a model-inadequacy term becomes crucial. Even in this case, if the model is applied to flow characterized by significant departures from the model range of applicability, the average calibrated solution can be far from the experimental data, which are only captured with the tail of the posterior solution distribution. This is true for the model-inadequacy term used in this paper, which is able to capture the data, even in the case of the extremely-adverse pressure gradients. However, it is always centered at the (possibly highly incorrect) calibrated $k-\varepsilon$ output, i.e. it is not able to correct for any bias in the RANS output with respect to the experimental data. The consequence is that the posterior uncertain range of the quantity of interest can become very large in cases of high model inadequacy.

Future work might include the investigation of the parameter spread for other turbulence models, as well as the dependence of their predictive quality upon the inclusion of a model inadequacy term. Alternatively, the already obtained parameter spread for the $k-\varepsilon$ model might be used to generate uncertainty bars on the predictions of different, (more complex) flow topologies, in order to investigate the generality of our error estimate.

Acknowledgements

The present work was supported by the French 'Agence Nationale de la Recherche' (ANR) under contract ANR-11-MONU-008-002.

REFERENCES

[1]B. J. Debusschere, H.N. Najm, P. P. Pébay, O. M. Knio, R. G. Ghanem, and O. P. Le Maître, "Numerical challenges in the use of polynomial chaos representations for stochastic processes," SIAM, J. Sci. Comput. **26**, 698–719 (2001).

[2]N. Zabaras and B. Ganapathysubramanian, "A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach," Journal of Computational Physics **227**, 4697–4735 (2008).

[3]Richard Dwight and Zhong-Hua Han, "Efficient uncertainty quantification using gradient-

enhanced Kriging," in *Proceedings of 11th AIAA Conference on Non-Deterministic Approaches, Palm Springs CA. AIAA-2009-2276* (AIAA, 2009) pp. 1–21.

[4]J.H. Ferziger and M. PERIĆ, "Further discussion of numerical errors in cfd," International Journal for Numerical Methods in Fluids **23**, 1263–1274 (1996).

[5]D.C. Wilcox, American Institute of Aeronautics, and Astronautics, *Turbulence modeling for CFD*, Vol. 3 (DCW industries La Canada, CA, 2006).

[6]Z. Zhang, W. Zhang, Z.J. Zhai, and Q.Y. Chen, "Evaluation of various turbulence models in predicting airflow and turbulence in enclosed environments by cfd: Part 2—comparison with experimental data from literature," Hvac&R Research **13**, 871–886 (2007).

[7]T. Sjögren and A.V. Johansson, "Development and calibration of algebraic nonlinear models for terms in the reynolds stress transport equations," Physics of Fluids **12**, 1554 (2000).

[8]S.B. Pope, *Turbulent flows* (Cambridge Univ Pr, 2000).

[9]A. Gelman, JB Carlin, HS Stern, and DB Rubin, "Bayesian data analysis. 2004," Chapman&Hall, London **2** (2004).

[10]S.H. Cheung, T.A. Oliver, E.E. Prudencio, S. Prudhomme, and R.D. Moser, "Bayesian uncertainty analysis with applications to turbulence modeling," Reliability engineering and systems safety **96**, 1137–1149 (2011).

[11]P.R. Spalart and S.R. Allmaras, "A one-equation turbulence model for aerodynamic flows," La recherche aérospatiale **1**, 5–21 (1994).

[12]L. Lourenco and A. Krothapalli, "On the accuracy of velocity and vorticity measurements with piv," Experiments in Fluids **18**, 421–428 (1995).

[13]M.C. Kennedy and A. O'Hagan, "Bayesian calibration of computer models," Journal of the Royal Statistical Society: Series B (Statistical Methodology) **63**, 425–464 (2001).

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

[15]SJ Kline, "The 1980-81 afosr-httm-stanford conference on complex turbulent flows: Comparison of computation and experiment," in *Objectives, evaluation of data, specifications of test cases, discussion and position papers*, Vol. 1 (1981).

[16]W. Rodi and G. Scheuerer, "Scrutinizing the k-ε turbulence model under adverse pressure gradient conditions," Journal of fluids engineering **108**, 174 (1986).

[17]D.C. Wilcox, "Comparison of two-equation turbulence models for boundary layers with

pressure gradient," AIAA Journal (American Institute of Aeronautics and Astronautics);(United States) **31** (1993).

[18]M. Christie, V. Demyanov, and D. Erbas, "Uncertainty quantification for porous media flows," Journal of Computational Physics **217**, 143–158 (2006).

[19]M. Oijen and A. Thomson, "Toward bayesian uncertainty quantification for forestry models used in the united kingdom greenhouse gas inventory for land use, land use change, and forestry," Greenhouse Gas Inventories, 55–67(2011).

[20]G. Freni and G. Mannina, "Bayesian approach for uncertainty quantification in water quality modelling: The influence of prior distribution," Journal of Hydrology **392**, 31–39 (2010).

[21]N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *et al.*, "Equation of state calculations by fast computing machines," Journal of Chemical Physics **21**, 1087 (1953).

[22]W.K. Hastings, "Monte carlo sampling methods using markov chains and their applications," Biometrika **57**, 97–109 (1970).

[23]J.A. Hoeting, D. Madigan, A.E. Raftery, and C.T. Volinsky, "Bayesian model averaging: a tutorial," Statistical science, 382–401(1999).

[24]P. Sagaut, *Large eddy simulation for incompressible flows: an introduction* (Springer, 2005).

[25]WP Jones and B.E. Launder, "The prediction of laminarization with a two-equation model of turbulence," International Journal of Heat and Mass Transfer **15**, 301–314 (1972).

[26]C.G. Speciale, R. Abid, and Anderson E.C., "A critical evaluation of two equation model for near-wall turbulence," AIAA Journal **90** (1990).

[27]M.S. Mohamed and J.C. Larue, "The decay power law in grid-generated turbulence," Journal of Fluid Mechanics **219**, 195–214 (1990).

[28]J. Kim, P. Moin, and R. Moser, "Turbulence statistics in fully developed channel flow at low reynolds number," Journal of Fluid Mechanics **177**, 133–166 (1987).

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

[30]PDA Platteeuw, GJA Loeven, and H. Bijl, "Uncertainty quantification applied to the k-e model of turbulence using the probabilistic collocation method," in *AHS Adaptive Struc-*

tures Conference; 10 th AIAA Non-Deterministic Approaches Conference; 9 th AIAA Gossamer Spacecraft Forum and the 4 th AIAA Multidisciplinary Design Optimization Specialist Conference (American Institute of Aeronautics and Astronautics, 1801 Alexander Bell Drive, Suite 500, Reston, VA, 20191-4344, USA,, 2008).

[31]E.S. Zanoun, F. Durst, and H. Nagib, "Evaluating the law of the wall in two-dimensional fully developed turbulent channel flows," Physics of Fluids **15**, 3079 (2003).

[32]D.C. Wilcox, *Companion Software:Turbulence Modeling for CFD Third Edition* (2006).

[33]FG Blottner, "Variable grid scheme applied to turbulent boundary layers," Computer Methods in Applied Mechanics and Engineering **4**, 179–194 (1974).

[34]G.B. Schubauer and PS Klebanoff, "Investigation of separation of the turbulent boundary layer," Tech. Rep. (DTIC Document, 1950).

[35]DE Coles and EA Hirst, "Computation of turbulent boundary layers," in *Proceedings of AFOSR-IFP Stanford Conference*, Vol. 2 (1968).

[36]T.A. Oliver and R.D. Moser, "Bayesian uncertainty quantification applied to rans turbulence models," in *Journal of Physics: Conference Series*, Vol. 318 (IOP Publishing, 2011) p. 042032.

[37]R.G. Ghanem and A. Doostan, "On the construction and analysis of stochastic models: Characterization and propagation of the errors associated with limited data," Journal of Computational Physics **217**, 63–81 (2006).

[38]GEB Archer, A. Saltelli, and IM Sobol, "Sensitivity measures, anova-like techniques and the use of bootstrap," Journal of Statistical Computation and Simulation **58**, 99–120 (1997); I.M. Sobol′, "Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates," Mathematics and computers in simulation **55**, 271–280 (2001).

[39]B. Sudret, "Global sensitivity analysis using polynomial chaos expansions," Reliability Engineering & System Safety **93**, 964–979 (2008).

[40]G. TANG, MS ELDRED, and L.P. SWILER, "Global sensitivity analysis for stochastic collocation expansion," CSRI SUMMER PROCEEDINGS 2009, 100(2010).

[41]GJA Loeven, JAS Witteveen, and H. Bijl, "Probabilistic collocation: an efficient non-intrusive approach for arbitrarily distributed parametric uncertainties," in *Proceedings of the 45th AIAA Aerospace Sciences Meeting*, Vol. 6 (2007) pp. 3845–3858.

[42]M.H. Chen and Q.M. Shao, "Monte carlo estimation of bayesian credible and hpd inter-

vals," Journal of Computational and Graphical Statistics **8**, 69–92 (1999).

[43]P.E. McSharry and L.A. Smith, "Consistent nonlinear dynamics: identifying model inadequacy," Physica D: nonlinear phenomena **192**, 1–22 (2004).

[44]W.L. Oberkampf and C.J. Roy, *Verification and validation in scientific computing* (Cambridge University Press, 2010).

[45]H.L. Moses, "The behavior of turbulent boundary layers in adverse pressure gradients," Tech. Rep. (DTIC Document, 1964).