



Data-driven Bayesian inference of turbulence model closure coefficients incorporating epistemic uncertainty

Daigo Maruyama¹ · Philipp Bekemeyer¹ · Stefan Görtz¹ · Simon Coggon² · Sanjiv Sharma²

Received: 20 September 2021 / Accepted: 1 November 2021
© The Author(s) 2022

Abstract

We introduce a framework for statistical inference of the closure coefficients using machine learning methods. The objective of this framework is to quantify the epistemic uncertainty associated with the closure model by using experimental data via Bayesian statistics. The framework is tailored towards cases for which a limited amount of experimental data is available. It consists of two components. First, by treating all latent variables (non-observed variables) in the model as stochastic variables, all sources of uncertainty of the probabilistic closure model are quantified by a fully Bayesian approach. The probabilistic model is defined to consist of the closure coefficients as parameters and other parameters incorporating noise. Then, the uncertainty associated with the closure coefficients is extracted from the overall uncertainty by considering the noise being zero. The overall uncertainty is rigorously evaluated by using Markov-Chain Monte Carlo sampling assisted by surrogate models. We apply the framework to the Spalart–Allmaras one-equation turbulence model. Two test cases are considered, including an industrially relevant full aircraft model at transonic flow conditions, the Airbus XRF1. Eventually, we demonstrate that epistemic uncertainties in the closure coefficients result into uncertainties in flow quantities of interest which are prominent around, and downstream, of the shock occurring over the XRF1 wing. This data-driven approach could help to enhance the predictive capabilities of computational fluid dynamics (CFD) in terms of reliable turbulence modeling at extremes of the flight envelope if measured data is available, which is important in the context of robust design and towards virtual aircraft certification. The plentiful amount of information about the uncertainties could also assist when it comes to estimating the influence of the measured data on the inferred model coefficients. Finally, the developed framework is flexible and can be applied to different test cases and to various turbulence models.

Keywords Turbulence modeling · Uncertainty quantification · Parameter calibration · Bayesian statistics · Surrogate-assisted methods · Spalart–Allmaras one-equation turbulence model · Large-scale industrial aircraft use-case

Abbreviations

x	A vector-valued input variable
y	A scalar-valued output variable
\mathbf{X}	A dataset of input variable x
\mathbf{Y}	A dataset of input variable y
\mathbf{a}	A parameter vector in general
σ	Standard deviation
θ	A hyperparameter vector in general
f	A function to produce the output value by CFD with a turbulence model
ε	A stochastic variable representing noise
p	A probability
N	Sample size of a dataset
\mathcal{N}	Normal distribution
\mathcal{D}	A dataset as $\mathcal{D} = (\mathbf{X}, \mathbf{Y})$

Executive Editor: Bernd Noack.

✉ Daigo Maruyama
daigo.maruyama@dlr.de; d.maruyama@tu-braunschweig.de

✉ Stefan Görtz
stefan.goertz@dlr.de

Philipp Bekemeyer
philipp.bekemeyer@dlr.de

¹ German Aerospace Center (DLR), Institute of Aerodynamics and Flow Technology, Lilienthalplatz 7, 38108 Braunschweig, Germany

² Airbus Operations Limited, Pegasus House, Aerospace Ave, Filton, Bristol BS34 7PA, UK

CFD	Computational fluid dynamics
RANS	Reynolds-Averaged Navier–Stokes
S-A	Spalart–Allmaras
UQ	Uncertainty quantification
QoI	Quantities of interest
MCMC	Markov-chain Monte Carlo
DoE	Design of experiment
POD	Proper orthogonal decomposition
CI	Credible interval
AoA	Angle of attack
2D	Two-dimensional
3D	Three-dimensional
C_d	Drag coefficient of a 2D profile
C_D	Drag coefficient of a 3D profile
C_p	Pressure coefficient

Subscripts

i	i th element of a vector
\mathcal{I}_a	Predictive distribution contributed only by the parameter \mathbf{a}
org	Original value
new	A new value to distinguish with the others
pri	A prior distribution

Superscripts

(i) i th sample point of a dataset

Oversets

- $\hat{\sim}$ Fixed by a point-estimate approach for a stochastic variable
- $\tilde{\sim}$ Approximated by a surrogate model for a function

1 Introduction

When using computational fluid dynamics (CFD) to investigate complex, industrial relevant aircraft configurations at high Reynolds numbers, the direct simulation of turbulent fluctuations is still computationally prohibitive. Thus, typically, the Reynolds averaged Navier–Stokes (RANS) equations are solved and turbulent behavior is modelled using so called turbulence models. Various different models are available. However, all of them are based on coefficients which need to be tuned theoretically and empirically based on experimental data in order to obtain accurate predictions. Traditionally, this was and is a rather manual process in which one tries to match numerical results with canonical experiments without accounting for uncertainties. Due to increasing computational resources and data being available,

data-driven approaches based on machine learning gained attention for such coefficient tuning tasks within the past few years [1]. The idea of these data-driven approaches is to re-infer and/or to adjust the established coefficients or terms in the turbulence model based on experimental data or numerical data computed by direct numerical simulations (DNS) or other higher-fidelity numerical techniques by using statistical methods.

In CFD computations, the issue of physical modeling inaccuracy is normally inevitable. Numerical errors due to spatial and temporal discretization of the governing equations and physical models are one of the most important issues in CFD computations [2]. These inaccuracy and errors are recognized as epistemic uncertainty, which can be reduced if more knowledge is gained. Turbulence models are one of the important issues here. In measurements, on the other hand, the observed data may contain measurement errors due to various unknown external factors. These unknown factors are classified as aleatory uncertainty, which is inherently regarded as irreducible.

Data-driven methods depend on our hypothesis in statistical inference. One example of a conceivable hypothesis is that the form of the turbulence closure model is accurate enough, but that the coefficients defined in empiricism that relate to specific flow properties are not always correct for a variety of flow cases due to insufficient knowledge. In this article these coefficients are treated as stochastic by Bayesian inference. Bayesian inference refers to statistical inference where uncertainty in inferences is quantified using probability. In general, statistical inference methods vary from problem to problem but can abstractly be grouped in two categories on how to treat the parameters (latent variables), which are the coefficients in closure models. On the one hand, if large data sets are available, so-called *point-estimate* approaches are commonly applied with deep learning being the arguably most promising and efficient methodology. On the other hand, the inference of the latent variables by regarding them as *stochastic* by enforcing a Bayesian perspective provides uncertainty quantification (UQ) capabilities associated with the latent variables and enables another perspective on the quantities of interest (QoIs) in terms of observable variables. These capabilities become especially remarkable when the amount of data is relatively small. However, effective numerical techniques for UQ are crucial especially in high dimensional cases. In this article, the closure coefficients are treated as stochastic latent variables. The expression latent variables refers to unobserved variables and is used as an antonym to observed variables such as input and output variables. The stochastic approach then enables to quantify uncertainties of the latent variables in the QoI in terms of credible intervals. The above-mentioned categories are rigorously classified in Sect. 2. In this article, to emphasize the differences between tuning parameters by a *point-estimate* and evaluating them

stochastically by an *interval-estimate*, we sometimes use the expressions *calibration* and *inference*, respectively. Here we start with reviewing some previous work on the point-estimate approach.

One of the examples of turbulence model calibration with the point-estimate approach is to minimize the error between experimental data and the output of RANS-based CFD computations for some QoIs. This is a direct approach. From a statistical inference perspective, this is known as a parametric regression process or parameter calibration: the turbulence model is considered as a *nonlinear regression model* whose inputs and outputs are parameters such as flow conditions, the numerical mesh, etc., and QoIs such as integral aerodynamic coefficients or pressure distributions, respectively. The closure coefficients are treated as the parameters (or globally speaking, latent variables) of the nonlinear regression model. Thus, least-squares optimization through minimizing the mean squared error (MSE) [3] provides turbulence closure coefficients calibration. In general, during the calibration process a larger number of RANS simulations are required causing a huge computational burden especially for industrial relevant cases. Hence, large computational costs during the optimization can be partly circumvented by relying on surrogate models such as Gaussian processes or neural networks [3, 4]. Another alternative is the use of adjoint flow solvers to efficiently compute the gradients of the QoIs with respect to the closure coefficients, enabling efficient gradient-based optimization for solving the least-square problems.

Another example of approaches using the point-estimate is calibration of the turbulence closure coefficients as the output of the statistical inference methods [5–7]. Since the coefficients to be inferred are now the output of regression models, not parameters in the models, any kind of supervised learning techniques are applicable. For example, the authors in Refs. [5–7] used neural networks to handle millions of data points by considering spatio-temporal parameters on flow fields.

These two approaches introduced above can theoretically be extended to an interval-estimate via the Bayesian perspective to evaluate the uncertainty associated with the closure coefficients. For example, in the direct approach using the least squares, the closure coefficients as the latent variables can be treated as stochastic to formulate a posterior distribution of the coefficients. As a result, the output QoIs also form a probability distribution, which is called predictive distribution. These probability distributions cannot be obtained analytically because the probabilistic models behind are inherently nonlinear regression models. The second approach mentioned in above may have further difficulties in extension to the interval-estimate. In case the parameters to be calibrated are high-dimensional, the use of Gaussian processes may become infeasible [7]. Bayesian neural networks would have to rely on approximation meth-

ods to compute the posterior distribution because of the high dimensionality of the parameters. In this article, our approach is based on the direct approach that the nonlinear regression model is a turbulence model equation with closure coefficients as parameters (the latent variables) to be inferred using the interval-estimate via the Bayesian perspective.

There are several methods to compute the posterior and the predictive distributions. The variational inference and the Laplace approximation [8] are versatile methods to approximate the posterior distribution by a parametric probability distribution. Since both of them require an iterative optimization process with many RANS computations, the direct simulation is still not as practical as the calibration process by the least squares. Markov-chain Monte Carlo (MCMC) methods [8, 9] are used to approximate the posterior distribution of a parameter of interest by random sampling in a parameter space. Unlike the above-mentioned parametric approximation methods, MCMC provides a theoretical guarantee that the histograms of the obtained sample points compose the posterior distribution when the sample size approaches infinity. Since each sample requires a RANS computation and the number of samples is usually greater than the number of iterations of the optimization, a direct implementation of MCMC is infeasible. It is essential for all MCMC methods to be supported by surrogate models in order to reduce the number of RANS computations. Once a surrogate model is constructed, the computational cost of MCMC can be negligible and MCMC can be regarded to have higher fidelity for computing the posterior than the other methods. However, it has to be noted that the application of MCMC to high-dimensional spaces remains a challenging topic. Also, from the point of view of surrogate model construction, it is desirable to keep the dimension of the parameters to be inferred as small as possible.

There is a body of existing research on surrogate-assisted MCMC approaches [10–12]. In Ref. [12], output uncertainty is directly computed via the samples generated by MCMC using an efficient boundary layer code. In Ref. [11] MCMC is used to evaluate the probability in the Bayesian update process for the purpose of calibrating turbulence model parameters. Surrogate-assisted MCMC approaches are also used for sensitivity analysis in the context of UQ [3, 13]. A review of recent work related to turbulence models can be found in Ref. [14].

In this article, the uncertainty associated with the closure coefficients is rigorously evaluated with a help of MCMC. The quantified uncertainty in the QoIs can be identified intuitively as a lack of knowledge of the coefficients. On the other hand, there are other uncertainties lurking in flow physics as mentioned above. All the other uncertainties, no matter where they come from, are classified as noise in a Gaussian probabilistic model to focus on the uncertainty associated with the closure coefficients.

The probabilistic model is a parametric model governed by a model parameter that describes the set of closure coefficients and a parameter to control the noise. The main feature of the Bayesian perspective here is to regard the parameters (the latent variables) as stochastic variables. In order to fully execute the Bayesian approach to account for all the uncertainties that are present, all the parameters are treated as stochastic. This modeling approach can quantify the uncertainty in the QoIs that stems from the contribution of all the uncertainties, i.e., uncertainty due to the closure coefficients and due to noise. However, our interest is to quantify the uncertainties associated with the closure coefficients. From a physical perspective, this uncertainty needs to be clearly distinguished from the noise. However, the parameter to control the noise depends on the turbulence model parameters when data is observed (it forms a joint probability). A method for extracting the uncertainty associated with the closure coefficients is presented and demonstrated using various test cases. Note that modeling only the closure coefficients as stochastic variables yields different results which do not account for the effects of the uncertainty due to noise.

As mentioned earlier, the ability to quantify the epistemic uncertainty associated with closure coefficients has a number of potential applications, for example, feedback to experimental measurement techniques and propagation of the uncertainty to the output QoI for arbitrary flow types. It can be possible to identify at which domains in the input parameter space (such as angles of attack, Mach number, and the location on the wings) we need to measure the output QoI in experiments (sensor placement). Note that this process is basically the same concept as so-called Bayesian optimization and active learning, effectively employing the posterior uncertainty. There is, however, an important difference as the inferred parameters (the latent variables) are physically meaningful in our applications.

The target test case to demonstrate the data-driven Bayesian inference of the turbulence model closure coefficients is based on the Airbus XRF1 research aircraft configuration and corresponding wind tunnel data at flight Reynolds numbers. The RAE2822 airfoil is also used to verify the methodology, but with numerically generated measured data. The output QoIs are the drag and surface pressure coefficients, for which measured data is required to infer the model coefficients.

The flow solver and the turbulence model used in this article, the methodologies and the developed framework are introduced in Sect. 2. In Sect. 3, results are presented for two test cases and discussed. Section 4 provides a summary and conclusions.

2 Methodology

This section introduces a framework for data-driven turbulence modeling and provides a generalized perspective to showcase that the developed framework can be applied to other (turbulence) models as well as other test cases.

2.1 Flow solver and turbulence model

The flow solver employed is the DLR TAU-code. It is based on a finite volume scheme, where inviscid terms are computed via a first or second order scheme. Viscous terms are computed via a second order scheme. Details of the solver can be found in Ref. [15–17]. In this paper, matrix dissipation is used. For time integration a LUSGS implicit scheme is chosen. Convergence acceleration is achieved with a multi-grid algorithm based on agglomerated coarse grids. Various turbulence models are available and the negative Spalart–Allmaras turbulence model [18] (referred to as the negative S–A) is investigated and hence altered in this paper.

The equations of the negative S–A are briefly described next. The negative S–A is derived from the original Spalart–Allmaras turbulence model (referred to as the original S–A). Details of the derivations provided below are available in Ref. [18]. The original S–A is a linear eddy viscosity model based on Boussinesq approximation. The turbulent eddy viscosity ν_t is calculated by:

$$\nu_t = \tilde{\nu} f_{v1}. \quad (1)$$

Here, the transport equation to be solved to obtain $\tilde{\nu}$ is as:

$$\begin{aligned} \frac{D\tilde{\nu}}{Dt} = & \underbrace{c_{b1} (1 - f_{t2}) \tilde{S} \tilde{\nu}}_{\text{production}} - \underbrace{\left(c_{w1} f_w - \frac{c_{b1}}{\kappa^2} f_{t2} \right) \left(\frac{\tilde{\nu}}{d} \right)^2}_{\text{destruction}} \\ & + \dots \underbrace{\frac{1}{\sigma} \left[\nabla \cdot ((\nu + \tilde{\nu}) \nabla \tilde{\nu}) + c_{b2} (\nabla \tilde{\nu})^2 \right]}_{\text{diffusion}}, \end{aligned} \quad (2)$$

where \tilde{S} is the modified vorticity,

$$\tilde{S} = S + \bar{S}, \quad \bar{S} = \frac{\nu}{\kappa^2 d^2} \chi f_{v2}, \quad (3)$$

where S is the vorticity magnitude. The other coefficients in Eqs. (1–3) are defined as:

$$\left\{ \begin{aligned} c_{w1} &= \frac{c_{b1}}{\kappa^2} + \frac{1+c_{b2}}{\sigma}, \\ f_{v1} &= \frac{\chi^3}{\chi^3+c_{v1}^3}, \\ \chi &= \frac{\tilde{v}}{v}, \\ f_{v2} &= 1 - \frac{\chi}{1+\chi f_{v1}}, \\ f_w &= g \left(\frac{1+c_{w3}^6}{g^6+c_{w3}^6} \right)^{1/6}, \\ g &= r + c_{w2}(r^6 - r), \\ r &= \min\left(\frac{\tilde{v}}{\delta\kappa^2 d^2}, 10\right), \\ f_{v1} &= c_{t3} \exp(-c_{t4}\chi^2). \end{aligned} \right. \quad (4)$$

The original S–A contains two additional coefficients namely c_{t1} and c_{t2} . In the DLR TAU-code, the trip term in Eq. (2) is neglected. Therefore, 9 closure coefficients are accounted for.

In the negative S–A, there are two modifications to the original S–A. The first one is that the modified vorticity \tilde{S} in Eq. (3) is enforced to be positive. The other one is a modification of the diffusion coefficient $\nu + \tilde{\nu}$ in Eq. (2). Therefore \bar{S} in Eq. (2) is represented by two additional coefficients c_{v2} and c_{v3} when $\bar{S} \leftarrow c_{v2}S$ as:

$$\tilde{S} = S + \frac{S(c_{v2}^2 S + c_{v3}\bar{S})}{(c_{v3} - 2c_{v2})S - \bar{S}}. \quad (5)$$

Otherwise Eq. (3) is retained. The diffusion coefficient $\nu + \tilde{\nu}$ in Eq. (2) is modified to be $\nu + \tilde{\nu}f_n$, where f_n is expressed as follows by introducing one coefficient c_{n1} :

$$f_n = \frac{c_{n1} + \chi^3}{c_{n1} - \chi^3}. \quad (6)$$

Hence, the negative S–A includes 12 closure coefficients. These coefficients are regarded as a parameter vector which should be inferred. Moreover, constraints on the feasible intervals of the coefficients are introduced to $\pm 50\%$ of the default values. Table 1 shows the original coefficient values of the negative S–A and the valid intervals for the inference processes. Note that in Ref. [18], the modified vorticity \tilde{S} in the negative S–A should not below $0.3S$ posing an additional implicit constraint. Equations (3) and (5), might introduce further constraints on the intervals on c_{v2} and κ , but these restrictions are not considered.

Table 1 Default values and intervals of the negative Spalart–Allmaras turbulence model for the inference processes

	Min	Original	Max
σ_{SA}	0.333	0.667	1
κ	0.205	0.41	0.62
c_{b1}	0.06775	0.1355	0.20325
c_{b2}	0.311	0.622	0.933
c_{v1}	3.55	7.1	10.65
c_{v2}	0.35	0.7	10.5
c_{v3}	0.45	0.9	1.35
c_{w2}	0.15	0.3	0.45
c_{w3}	1	2	3
c_{ct3}	0.6	1.2	1.8
c_{ct4}	0.25	0.5	0.75
c_{n1}	8	16	24

2.2 Theory

This sub-section introduces a parametric statistical inference method as the cornerstone of the developed framework. The process of the statistical inference method is fundamentally composed of two steps, learning and prediction. These steps are conducted on a parametric probabilistic model that we define. The definition relies on two hypotheses as the first two steps. Eventually the process of the parametric statistical inference is in a broad way composed of the following four steps:

1. Define a regression model
2. Define a probabilistic model
3. Compute a posterior by given data (learning process)
4. Compute output distributions (prediction process)

The first two steps introduce the two hypotheses by each. The overall process here can be also regarded as extension of Bayesian linear regression to nonlinear regression models in general. A principal figure to illustrate the process is shown Fig. 1. The goal is to compute the posterior distribution $p(\mathbf{a}, \sigma | \mathbf{X}, \mathbf{Y})$ and the predictive distribution contributed only by the closure coefficients expressed as $p(y|x, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ in Fig. 1. Details of each step are described in the next sub-sections.

2.2.1 The 1st Step: defining a regression model

The regression model here means a deterministic function that produces a scalar output y when an arbitrary input vector \mathbf{x} is given. The model characteristics are controlled by a parameter vector \mathbf{a} . These are deterministic variables currently. Normally the parameter \mathbf{a} is being determined by the

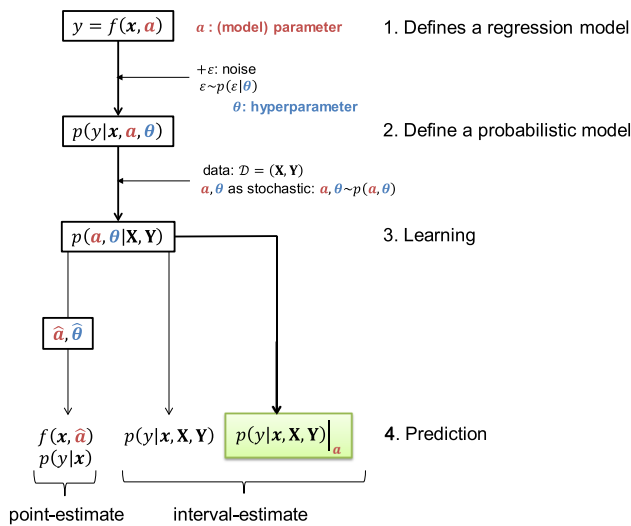


Fig. 1 Principal figure of the statistical inference process

learning process in the 3rd step with available data. Based on this concept, a CFD computation using a turbulence model can be represented as the deterministic function device:

$$y = f(\mathbf{x}, \mathbf{a}), \tag{7}$$

where \mathbf{a} represents a vector as a set of all the turbulence model coefficients. The input \mathbf{x} is a vector as a set of flow conditions and/or a spatial coordinate. The output y is a scalar value as an output aerodynamic value such as lift and drag or distributed quantities, e.g. pressure coefficient distributions. The function f represents then the CFD computation with a turbulence model as a mapping $f : (\mathbf{x}, \mathbf{a}) \mapsto y$. For example, in this paper a turbulence model here is the negative S–A. This is the first hypothesis in the statistical inference method. As this first hypothesis for example, other turbulence models could be set here by modeling it as $f(\mathbf{x}, \mathbf{a})$. The negative S–A has 12 independent coefficients, which are summarized as one parameter vector \mathbf{a} , whose dimensionality is 12, i.e. $\mathbf{a} = (\sigma_{SA}, \kappa, c_{b1}, c_{b2}, c_{v1}, c_{v2}, c_{v3}, c_{w2}, c_{w3}, c_{ct3}, c_{ct4}, c_{n1})$ by following Table 1. The parameter \mathbf{a} in Eq. (7) is called *model parameter* or simply *parameter* in this paper. Note that when our interest is only the input and the output as supervised learning, the function f is freely chosen (e.g. by model selection conducted in the third and fourth steps).

2.2.2 The 2nd Step: defining a probabilistic model

In the second step, now the output y is considered to be a stochastic variable. The process starts from assuming that the observed data y as the output has been generated by a probabilistic model as $y \sim p(y)$. Our interest is evaluating the epistemic uncertainty associated with the turbulence model $y \sim p(y)$ by diminishing the effects of all the other uncertainties, even the mesh and the scheme effects. Other

possible strategies for Bayesian inference can be found in Ref. [19]. First, the discrepancy between the output y and the model $f(\mathbf{x}, \mathbf{a})$ is represented by introducing a stochastic variable labelled ε . Thus, the output y can be defined as follows:

$$y := \underbrace{f(\mathbf{x}, \mathbf{a})}_{\substack{\text{the source term of} \\ \text{the uncertainty} \\ \text{associated with } \mathbf{a}}} + \underbrace{\varepsilon}_{\substack{\text{the source term of} \\ \text{all the other} \\ \text{uncertainties}}}. \tag{8}$$

The physical meaning of the stochastic variable ε is *noise* in the output value y . Since ε is stochastic, a probabilistic model is introduced into this variable. One example of the probabilistic model of ε is an univariate Gaussian distribution as $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, where \mathcal{N} and σ^2 represents the Gaussian distribution and its variance, respectively. This eventually brings the following conditional probability distribution of y :

$$p(y|\mathbf{x}, \mathbf{a}, \sigma) = \mathcal{N}(y|f(\mathbf{x}, \mathbf{a}), \sigma^2). \tag{9}$$

Now σ has been added as another parameter (called *hyperparameter* or *noise parameter* in this paper to be distinguished with the parameter \mathbf{a}). Equation (9) is the probabilistic model of the observed variable y as the objective here. Different from the observed variable y (and \mathbf{x}), the parameters \mathbf{a} and σ are not observed (the latent variables). In the fully Bayesian approach, the hyperparameter σ is also treated as a stochastic variable as $\sigma \sim p(\sigma)$. Thus, the defined probabilistic model can account for all the uncertainties that are present.

Following Eq. (8), two sorts of induced uncertainties can be easily separated by looking at each term individually. since the model parameter \mathbf{a} is the only parameter that can directly interact with the internal structures of the model f . All the other external uncertainties are dealt within the term ε governed by the noise parameter σ . No other functions (such as bias functions) are added in Eq. (8). The parameter \mathbf{a} is then treated as a stochastic variable in the Bayesian approaches as $\mathbf{a} \sim p(\mathbf{a})$. The function f as the first term also becomes stochastic. The stochastic variable $y \sim p(y)$ as the predictive distribution becomes a mixture of probabilities derived from the two stochastic terms after the learning process. After the learning process when data is given, all the parameters in Eq. (9) as the latent variables are not independent of each other and compose a joint probability distribution. In the 4th step, a rigorous form of the probability of y as $p(y)$ is presented. In that form, the contribution of the uncertainty associated with the parameter \mathbf{a} is not that obvious like Eq. (8) due to the above fact.

2.2.3 The 3rd Step: computing a posterior by given data (learning process)

The third step is the learning process of the parameters \mathbf{a} and σ using data. In this step, derivation to the Bayesian approach is employed. The process in this step is to evaluate the joint probability of all the parameters \mathbf{a} and σ as a posterior distribution. The posterior distribution can be evaluated using the Bayes' theorem by prior distributions of all the parameters \mathbf{a} and σ , and the likelihood function obtained by the probabilistic model and the data. How to compute the posterior function is described below. Some additional details about the definition of the prior distributions etc. are in Appendix 1.

The likelihood function can be defined by data and the probabilistic model defined in the second step as:

$$p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma) = \prod_{i=1}^N \mathcal{N}(y^{(i)}|f(\mathbf{x}^{(i)}, \mathbf{a}), \sigma^2), \quad (10)$$

where $\mathbf{X} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)})$ and $\mathbf{Y} = (y^{(1)}, y^{(2)}, \dots, y^{(N)})$ are observed data. N denotes the sample size of the data. The data is assumed to be independently drawn from the defined probabilistic model by Eq. (9) since the model is an univariate Gaussian distribution, hence no correlations exist. It has to be emphasized that direct evaluation of the likelihood function by Eq. (10) is practically available for the learning process instead of that of the posterior $p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y})$ (see Appendix 1). The computation of the marginal likelihood $p(\mathbf{Y}|\mathbf{X})$ used as the normalization constant is either not necessary in executing the MCMC sampling techniques.

There are mainly two perspectives in the learning process (also see Fig. 1). One is *point-estimate* while the other is *interval-estimate*. In practice, these names are driven from the properties of the parameters to be inferred using the posterior distribution. The commonly known least square (LSQ) approach can be categorized as a special case of the point-estimate. The interval-estimate simply corresponds to using all the information of the posterior itself. The point-estimate in this paper directly indicates maximum a posterior (MAP). The MAP can be achieved by literally maximizing a posterior function $p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y})$ with respect to the parameters \mathbf{a} and σ :

$$\hat{\mathbf{a}}, \hat{\sigma} = \operatorname{argmax}_{\mathbf{a}, \sigma} p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y}). \quad (11)$$

In practice, Eq. (11) is solved analytically when the regression model $f(\mathbf{x}, \mathbf{a})$ is a linear regression model. Since the regression model here is a nonlinear regression model, typically Eq. (11) is solved employing optimizers.

Under the above-mentioned condition that no specific prior distribution is set, the direct evaluation of the likelihood function by Eq. (10) is available in practice. When pursuing the point-estimate using an optimizer, error functions which are derived by taking the negative of the log-likelihood function (see Appendix 5 for further details) are practically used. When pursuing the interval-estimate, the sampling method using MCMC can be directly applied to the likelihood function in Eq. (10). This evaluation by MCMC provides an approximation of the posterior $p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y})$ by sample points. The posterior is composed by a histogram of the sample points in the parameters space \mathbf{a}, σ . This means that the frequency of the generated sample points directly forms a scaled function of the posterior. The mode of the posterior $p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y})$ represents a pair of the calibrated parameters $\hat{\mathbf{a}}, \hat{\sigma}$. By using the obtained sample points, the predictive distributions $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ shown in Fig. 1 as our goal are computed at the fourth step. Again note that same as requirement of optimizers in Eq. (11), the reason why numerical approaches are required is because the regression model $f(\mathbf{x}, \mathbf{a})$ is a nonlinear regression model.

Various established algorithms for MCMC are available and herein the random walk Metropolis-Hastings is applied. Details on MCMC algorithms can be found in many papers [8, 9]. As for optimizer for the point-estimate, we use differential evolution algorithm, which is a heuristic method. The reason why we do not use only the MCMC even for the point-estimate is for double-check of the accuracy of the tools. Therefore, the results by optimizers are not directly used in the prediction process.

2.2.4 The 4th Step: computing output distributions (prediction process)

In the last step the output y is computed by using the learned parameters: $\hat{\mathbf{a}}, \hat{\sigma}$ by the point-estimate, or $p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y})$ by the interval-estimate. First as for the point-estimate, the prediction of the output y for a given input \mathbf{x} is obtained by evaluating the regression model in Eq. (7):

$$y = f(\mathbf{x}, \hat{\mathbf{a}}). \quad (12)$$

Note that even for the point-estimate, the predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})$ is introduced following Eq. (9). However, this predictive distribution contains only the noise ε as identified by the point-estimate represented by the standard deviation $\hat{\sigma}$ as $\varepsilon \sim \mathcal{N}(0, \hat{\sigma}^2)$.

In the interval-estimate (with fully Bayesian perspective), the parameters \mathbf{a}, σ form a joint probability as $p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y})$ which is represented by the MCMC sample points. The predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})$ is obtained as follows by the sum and product rules of the probability theory:

$$p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}) = \int \int p(y|\mathbf{x}, \mathbf{a}, \sigma) p(\mathbf{a}, \sigma | \mathbf{X}, \mathbf{Y}) d\mathbf{a} d\sigma$$

$$\approx \frac{1}{N_{mcmc}} \sum_{i=1}^{N_{mcmc}} p(y|\mathbf{x}, (\mathbf{a}, \sigma)^{(i)}) \quad (13)$$

The parameters \mathbf{a}, σ are marginalized (integrated out) to be eliminated. This operation can be regarded as taking the expectation. The probability as the weight is already realized by the MCMC sampling result. Therefore, it can be computed by mixtures of the probabilistic model $p(y|\mathbf{x}, \mathbf{a}^{(i)}, \sigma^{(i)})$, where each sample point i is weighted with its respective probability $p(\mathbf{a}^{(i)}, \sigma^{(i)} | \mathbf{X}, \mathbf{Y})$. The constraint that the sum of all the weights has to be 1 to be a probability is accomplished by being divided with the sample size N_{mcmc} . Note that, the predictive distribution shown in Eq. (13) is a common way for a fully Bayesian approach [8]. No hyperparameters in priors of the parameters \mathbf{a} and σ are explicitly presented in Eq. (13) since an improper prior for both \mathbf{a} and σ is assumed. Note that investigation when using an empirical Bayes approach in contrast to the full Bayes approach is shown in Appendices 2 and 3. In Appendix 3, a comparison of results between $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})$ in Eq. (13) and $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ in Eq. (15a, 15b) using the empirical Bayes approach is also shown.

The predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})$ in Eq. (13) represents the contribution of all the sources of the uncertainties, one of which is associated with the model parameter \mathbf{a} , and all the others of which is associated with the noise parameter σ . These two sorts of uncertainties physically stem from different sources as shown in Eq. (8) and the predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})$ in Eq. (13) lost the information of their origins. Our goal is to quantify the epistemic uncertainty associated with the closure coefficients \mathbf{a} , without including the external factors represented by the noise $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. This is accomplished through the following two steps by using the sum and product rules of the probability theory. The first step is to obtain the marginal posterior $p(\mathbf{a} | \mathbf{X}, \mathbf{Y})$ by marginalizing the posterior $p(\mathbf{a}, \sigma | \mathbf{X}, \mathbf{Y})$ with respect to σ :

$$p(\mathbf{a} | \mathbf{X}, \mathbf{Y}) = \int p(\mathbf{a}, \sigma | \mathbf{X}, \mathbf{Y}) d\sigma. \quad (14)$$

Then, in order not to take the noise term in Eq. (8) into account, the probabilistic model of Eq. (9) where $\sigma \rightarrow 0$ is considered. Eventually the predictive distribution where only the contribution of the parameter \mathbf{a} is considered is represented as follows:

$$p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}} = \int p(y|\mathbf{x}, \mathbf{a}) p(\mathbf{a} | \mathbf{X}, \mathbf{Y}) d\mathbf{a}$$

$$\approx \frac{1}{N_{mcmc}} \sum_{i=1}^{N_{mcmc}} p(y|\mathbf{x}, \mathbf{a}^{(i)}) \quad (15a)$$

with $p(y|\mathbf{x}, \mathbf{a}) = \delta(y - f(\mathbf{x}, \mathbf{a}))$,

where $p(\mathbf{a} | \mathbf{X}, \mathbf{Y})$ indicates the marginal distribution obtained by Eq. (14), δ is the Dirac delta function as the limit of normal distributions as $\sigma \rightarrow 0$. In practice, the technical process itself becomes eventually very simple. The predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ in Eq. (15a) can be approximated by the output values of the MCMC sample points themselves in practice (by substituting the MCMC samples for the regression model $f(\mathbf{x}, \mathbf{a})$):

$$f : (\mathbf{x}, \mathbf{a}^{(i)}) \mapsto y^{(i)} (i = 1, 2, \dots, N_{mcmc}), \quad (15b)$$

where \mathbf{a}_i is the i th sample point $(\mathbf{a}, \sigma)_i$ obtained by the MCMC sampling with omitting the σ_i component in the sample point. The meaning of the predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ in Eq. (15a) is different from that of $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})$ in Eq. (13). Both of them are a probability of the output y after learning from the same data. However, the predictive distribution in Eq. (15a) represents the output uncertainty contributed only by the turbulence closure coefficients \mathbf{a} which have been learned by considering all the uncertainties that are present. The closure coefficients \mathbf{a} themselves were learned by taking the information of the uncertainty of the noise into account but the output y as QoI reflects only the epistemic uncertainty associated by \mathbf{a} eventually.

Overall for the entire steps, the process is common when a likelihood function is determined. The likelihood function in Eq. (10) is one example. Based on the concept of the four steps, the likelihood function to be evaluated (by either an optimizer or MCMC) can be naturally and flexibly determined for each problem setting. Note that further abstraction of the entire process of the four steps can be generalized by classifying all the stochastic variables into the latent variables and the observed variables (see Appendix 6).

2.3 Practical implementation details

This section provides insight into the practical implementation details with a special focus on computing the likelihood function. The likelihood function represented by Eq. (10) contains $f(\mathbf{x}^{(i)}, \mathbf{a})$, which is obtained through CFD computations. It is iteratively computed during the processes by the optimizer and MCMC, due to variations of the parameters \mathbf{a} and σ . The iteration number of this process could be the order of 10^3 to 10^6 . It means that this order of times of the CFD computations are needed. This is the most time-consuming part in the whole process. The iteration number is unable to be predicted and this process itself is infeasible in practice especially for 3D aircraft test cases. A surrogate-assisted strategy is used for efficiency. A series of the following surrogate-assisted strategy is conducted in a framework of the DLR's surrogate and reduced-order modeling toolbox called

SMARTy. Recent articles on research works using SMARTy are found in Refs. [20, 21].

Once a surrogate model is constructed, no additional CFD computations are needed. A surrogate model is constructed at each flow condition (and/or at each spatial coordinate) $\mathbf{x}^{(i)}$. The process of building a surrogate model is as follows. First, search domains of the turbulence model parameter \mathbf{a} are required to be decided. We set this as the ranges from 0.5 to 1.5 \mathbf{a}_{org} as already defined in Table 1, where \mathbf{a}_{org} is a vector of a set of the original turbulence model coefficients. Sampling locations where CFD computations are executed are determined by design of experiment (DoE). We used Sobol sequence, which is categorized as a kind of quasi Monte Carlo sampling methods [22, 23]. The Sobol sequence is able to keep the uniformity (the low-discrepancy) of the sampling filling the space even when the dimensionality of the input parameter is relatively high. It is also able to add sequentially low-discrepancy sampling points when the original sample size is not sufficient to achieve the accuracy of the constructed surrogate model. Based on the sample points generated by DoE, a surrogate model is constructed. Note that the process of constructing a surrogate model is theoretically the same as the statistical inference method in the previous subsection (a regression model by Eq. (12) or analytical solutions of the dual representation of Eq. (13) in e.g. Gaussian processes, where \mathbf{x} and \mathbf{a} in the previous subsection corresponds to \mathbf{a} and parameters to describe the surrogate model, respectively).

The constructed surrogate model eventually can be recognized as function approximation on the parameter space of \mathbf{a} at each flow condition (and/or at each spatial coordinate) as $\tilde{y}^{(i)} = \tilde{f}(\mathbf{a}; \mathbf{x}^{(i)})$, where noted that \mathbf{a} is now input but $\mathbf{x}^{(i)}$ as a set of parameters of the flow conditions is fixed. In practice, the new input of the surrogate model \tilde{f} corresponds to a new turbulence model parameter \mathbf{a}_{new} nominated during the process in the optimizer/MCMC. Thus, the output $\tilde{f}(\mathbf{a}_{new}; \mathbf{x}^{(i)})$, which is iteratively computed in the parameter space \mathbf{a} during the optimization or MCMC process, is efficiently obtained via the surrogate model at each fixed flow condition (and/or at each spatial coordinate) $\mathbf{x}^{(i)}$. Now the most time-consuming part has been replaced by the limited CFD computations on the DoE sample points, which are now predictive in terms of the cost. The required sample size to construct a surrogate model is normally dependent on the dimensionality of the input parameter \mathbf{a} . The sample size is shown later at each test case in Sect. 3. The surrogate models here used are Gaussian processes with hyperparameter tuning for prediction of drag coefficient as f , and additionally assisted by proper orthogonal decomposition (POD) for the cases of prediction of pressure coefficient. The kernel function commonly used in the Gaussian processes is a Gaussian

kernel. The hyperparameter tuning as point-estimate to determine the probability distribution is achieved by the MLE. In prediction of pressure coefficients, multiple output of $\mathbf{Y} = (y^{(1)}, y^{(2)}, \dots, y^{(N)})$ is assisted by the POD at once for multiple input $\mathbf{X} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)})$, where \mathbf{x} is a spatial coordinate in this case, N as the sample size can be the number of the surface output of a wing.

During the entire process, a variance-based sensitivity analysis can be additionally inserted for the purpose of selection of dominant model coefficients to QoI as the output. The results of the sensitivity analysis are expected to bring physical aspects of the turbulence model coefficients. It can even bring dimensionality reduction, which can benefit of the computational costs. The required sample size for the DoE is basically dependent on the dimensionality of the input parameter space \mathbf{a} . Therefore, it directly brings reduction of the required number of CFD computations. However, the sensitivity analysis itself requires the whole information using the original dimensionality of the parameter. Hence, it can be useful to use the results of a test case to other test cases to proceed the whole process with the selected coefficients. The global sensitivity of the input variables with respect to the output is computed by DoE samples using the Sobol sequence assisted by the constructed surrogate models. The execution of the Sobol indices itself was done by the Python library called SALib [24]. Details of the theory of the Sobol indices and sensitivity analysis methods in general can be found in Ref. [25].

Finally the flow chart of the data-driven method in this paper is shown in Fig. 2. This flow chart describes the direct process used in the application test cases in Sect. 3. The estimation of the modes by the optimizer can be a strategy for double check if the MCMC works properly.

3 Applications and results

Two case studies were performed to validate and demonstrate the methods presented above: (a) a two-dimensional RAE2822 airfoil and (b) a three-dimensional full aircraft model based on the Airbus XRF1. Analyses of both test cases targeted the transonic regime. For the 2D test case, artificial experimental data was used to validate the proposed framework. Then, for the 3D test case, several sets of proprietary wind tunnel data for XRF1 were used to compare the simulation results against the experimental results. In this article, actual values cannot be published for XRF1 due to reasons of confidentiality. Therefore, quantitative results and/or relative evaluations are shown in Sect. 3.2. The 3D case is, however, valuable to demonstrate that the framework can be applied to real, industrial-grade test cases.

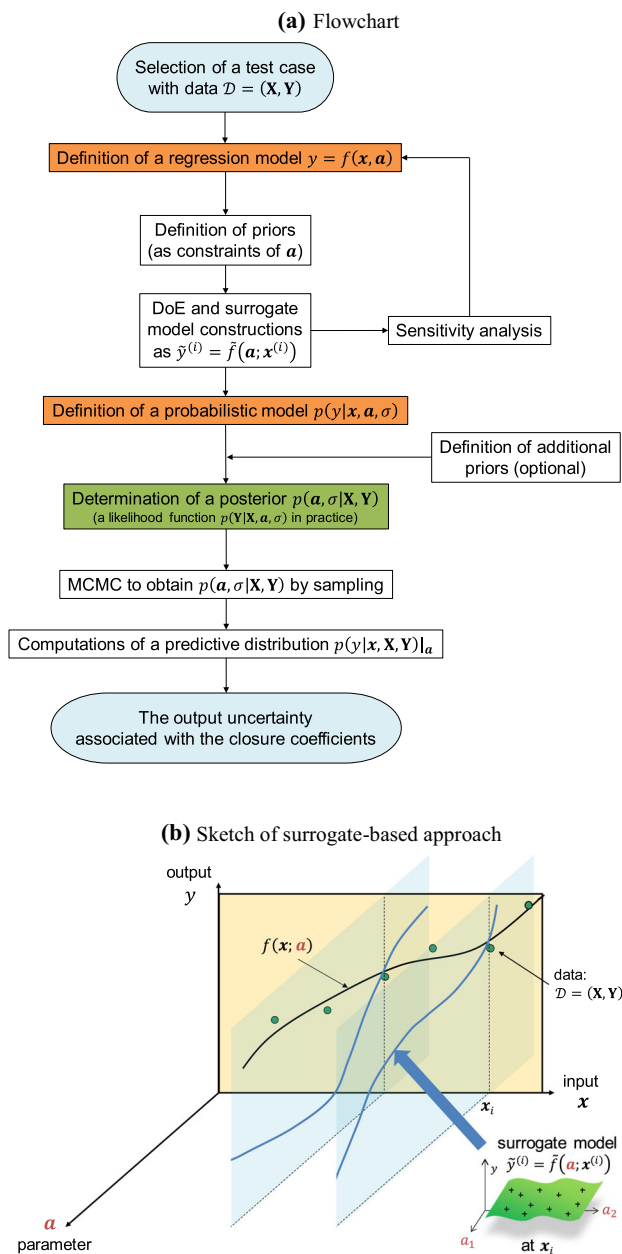


Fig. 2 Flow chart and simple sketch of surrogate-based approach

The parameter settings of the MCMC are common for all the test cases in this article. The sample size is 10^5 and the burn-in ratio is 0.8. The step length for \mathbf{a} normalized by \mathbf{a}_{org} and σ normalized by $\hat{\sigma}$ are 0.025 and 0.05, respectively. Here $\hat{\sigma}$ is the noise parameter obtained by the deterministic approach obtained by an optimizer (e.g. by Eq. (11)). Note just for visualization issues of the following figures that in the graphs of marginal posterior distributions of \mathbf{a} and σ obtained by the MCMC, the graphical ranges of \mathbf{a} and σ are common

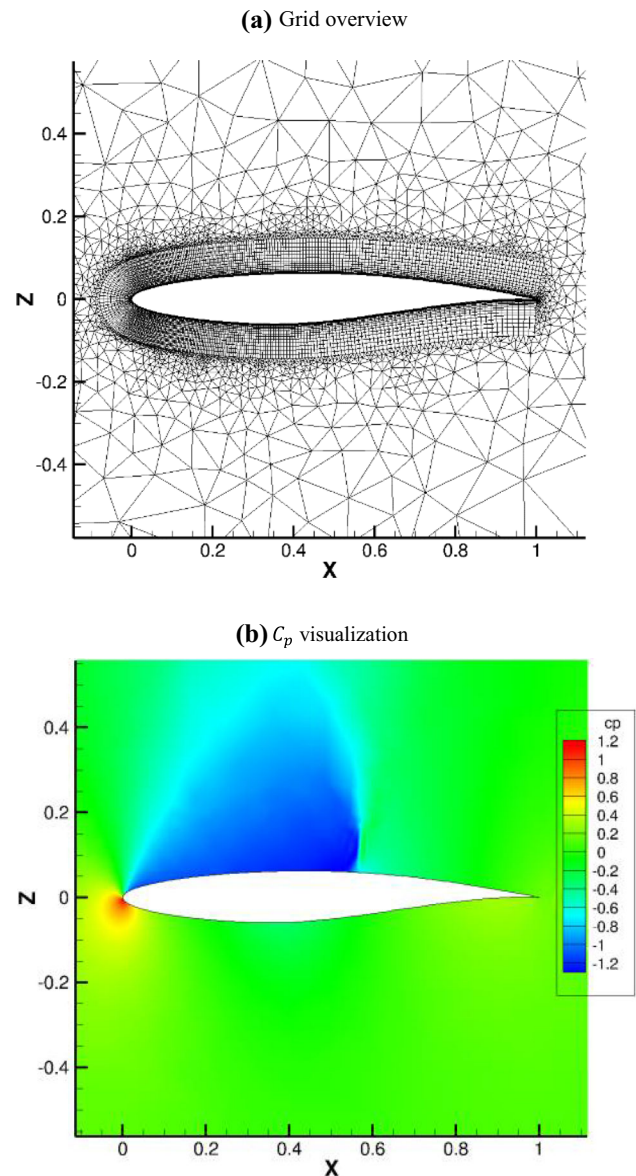


Fig. 3 RAE2822 airfoil configuration with grid overview and C_p visualization at Mach number of 0.734

for all the figures in this article. The range of \mathbf{a} is from 0.5 to $1.5 \mathbf{a}_{org}$. The center is \mathbf{a}_{org} . That of σ is from 0 to $5 \hat{\sigma}$ (only Fig. 7 is 0 to $2 \hat{\sigma}$). This indicates that the mode of the posterior of σ should be located one fifth from the left (or the middle in Fig. 7) when the MCMC works well by producing the same results as the optimizer. Note also that all the predictions depicted by lines, and predictive distributions depicted by filled colored areas, have only their valid points/ranges at the discrete input flow conditions since the surrogate models are available for those discrete input flow conditions.

3.1 2D RAE2822 Airfoil

The RAE2822 airfoil is introduced here to validate the methodology and to investigate the characteristics of the turbulence model closure coefficients. The baseline Mach number is 0.734 and it varies in the transonic regime. The angle of attack and the Reynolds number are fixed at 2.79 degree and 6.5 million, respectively. The mesh used for CFD computations and a C_p visualization at this flow condition are shown in Fig. 3. The number of nodes is around 29 thousand.

For the purpose of dimensionality reduction, dominant closure coefficients, with respect to aerodynamic coefficients, are selected based on their Sobol indices. These selected closure coefficients are subsequently used as parameter \mathbf{a} for the XRF1 test case to reduce the computational costs; the assumption being that the flow characteristics are similar to each other in the transonic regime. Note that the results of the Sobol indices depends on the QoI as the output.

3.1.1 Sobol indices as preprocessing

The required information to generate the Sobol indices is obtained through surrogate models, which contain all of the closure coefficients as input parameters and the QoI as the output value. In this case the drag coefficient (C_d) was selected as the QoI to generate the Sobol indices. The surrogate models of C_d were constructed at each flow condition where the Mach number ranged from 0.704 to 0.784 in 0.01 increment. Each surrogate model was built using 100 DoE sample points generated between 0.5 and 1.5 \mathbf{a}_{org} .

Table 2 shows results of the Sobol indices over the Mach number range. The percentages indicate the relative amount of contribution of the turbulence model coefficients to C_d at each flow condition. The mean values of the percentages at each coefficient are also shown in Table 2. It can be observed from the mean results that κ and c_{b1} are the most dominant followed by σ_{SA} and c_{v1} . This ranking of the result has similar characteristics to the works by Papadimitriou et al. [26] and Da Ronch et al. [2] when they investigated the original S–A. They listed κ , c_{b1} and c_{v1} as the most influential coefficients to the output uncertainty of their quantities of interest. From the result in Table 2, the coefficients resulting from the negative S–A model which are not present in the original S–A model are not influential.

The selection process to determine which coefficients are dominant is similar to the mindset enforced when POD is employed and relies on using eigenvalues. Here we decided to truncate the list of dominant coefficients at 90%. Eventually the coefficients, κ , c_{b1} and the σ_{SA} , were selected as the dominant coefficients. These selected coefficients are set to the components of the turbulence model parameter \mathbf{a} as $\mathbf{a} = (\sigma_{SA}, \kappa, c_{b1})$ to be inferred in the following test cases. The process described so far is a preprocessing process and leads

into the learning process using observed data. From now as a validation test of the framework, data-driven turbulence modeling results are shown by using artificial data.

3.1.2 Inference using drag coefficient

First the drag coefficient (C_d) is used as the QoI. Figure 4 shows marginal posterior distributions of each component of \mathbf{a} and σ after the learning process, and the obtained predictive distribution. As described in Sect. 2, the posterior distributions $p(\mathbf{a}, \sigma | \mathbf{X}, \mathbf{Y})$ and the predictive distributions $p(y | \mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ by Eq. (15) (contributed only by the closure model coefficient \mathbf{a}) were obtained directly by the MCMC sampling. Figure 4a also includes the result of $\hat{\mathbf{a}}$ obtained through the optimization for comparison.

The artificial data is also depicted in Fig. 4b. It was created based on CFD results obtained by using \mathbf{a}_{org} with independently and randomly generated noise from the Gaussian distribution $\mathcal{N}(0, \sigma_{org}^2)$, where σ_{org} is a constant value set to $\sigma_{org} = 0.001$ (10 drag counts), which was manually set by the authors. Since we know that the generated noise follows an independently identically distributed (i.i.d) Gaussian distribution, in this case, it is explicitly justified to use Eq. (10) as the likelihood function. It can be confirmed from Fig. 4a that σ_{SA} as a_1 (the first element of \mathbf{a}) has greater uncertainty than the other coefficients in the relative ranges defined by \mathbf{a}_{org} . In Fig. 4b, a reference result of the output obtained by \mathbf{a}_{org} and a result obtained by $\hat{\mathbf{a}}$ (by the point-estimate) are depicted by the dotted orange lines and the green lines, respectively. The values of $\hat{\mathbf{a}}$ and $\hat{\sigma}$ obtained by the optimizer is $\hat{\mathbf{a}} = [0.801, 0.417, 0.129]$ and $\hat{\sigma} = 0.000689$, respectively. Hence $\hat{\mathbf{a}}$ obtained by the optimizer agrees well with the modes of the marginal posterior distributions obtained by the MCMC in Fig. 4a. However, the obtained $\hat{\mathbf{a}}$ and $\hat{\sigma}$ could not detect \mathbf{a}_{org} and σ_{org} as the reference (each component of \mathbf{a}_{org} is at the center of the horizontal axis). Nevertheless, it can be confirmed from Fig. 4a that the marginal posterior distributions are wide and covers \mathbf{a}_{org} , σ_{org} with high probability density except for the first component of \mathbf{a}_{org} . Therefore, a lack of the data for this amount of noise has been revealed by the results of the posterior obtained by the MCMC. It is emphasized that this information is never detected by the point-estimate.

Thus, this test case using the artificial data is not only validating the developed method but moreover offers further insight into underlying flow physics. For example, it can be observed that the influence of the turbulence closure coefficients is less when predicting the drag coefficient for a Mach number around 0.74 in contrast to higher and lower Mach numbers.

Next, the same test case but using prior information is also investigated. Prior can be set to any arbitrary probabil-

Table 2 Contribution percentage of the turbulence model coefficients to drag coefficient (C_d) at each flow condition

Mach number	σ_{SA}	κ	c_{b1}	c_{b2}	c_{v1}	c_{v2}	c_{v3}	c_{cw2}	c_{w3}	c_{ct3}	c_{ct4}	c_{n1}
0.704	1.83	83.48	5.08	0.00	8.25	0.00	0.00	0.01	0.20	0.96	0.19	0.00
0.714	3.28	77.84	9.40	0.51	8.44	0.00	0.00	0.00	0.00	0.21	0.32	0.00
0.724	2.86	66.00	17.34	1.20	10.07	0.00	1.16	0.00	0.00	0.00	0.00	1.36
0.734	7.26	51.38	33.35	0.00	0.10	0.00	7.89	0.00	0.00	0.00	0.00	0.01
0.744	3.63	44.92	50.39	0.00	0.25	0.72	0.01	0.00	0.00	0.00	0.08	0.00
0.754	2.70	51.23	41.66	0.00	1.75	0.00	1.88	0.02	0.04	0.69	0.00	0.03
0.764	0.10	39.10	60.26	0.00	0.00	0.01	0.50	0.01	0.00	0.00	0.00	0.01
0.774	8.56	16.82	66.86	0.00	0.47	6.01	0.75	0.16	0.00	0.00	0.29	0.07
0.784	8.70	45.78	39.20	0.00	0.00	2.10	1.15	0.06	0.00	1.84	0.00	1.17
Mean	4.32	52.95	35.95	0.19	3.26	0.98	1.48	0.03	0.03	0.41	0.10	0.29

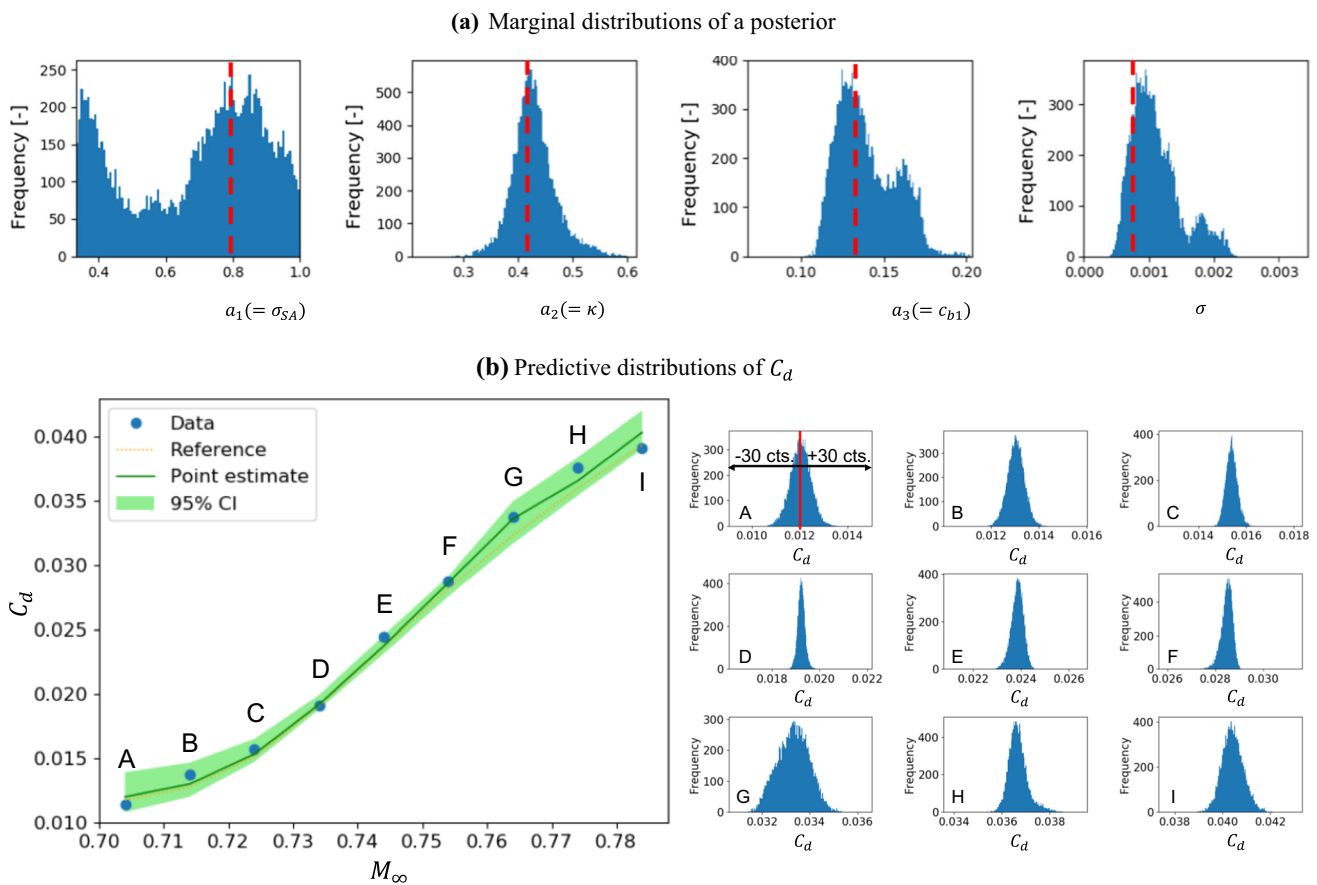


Fig. 4 Results of predictions using C_d . **a** Marginal distributions of a posterior of the turbulence model parameter (red dotted lines indicate \hat{a}), **b** predicted C_D distributions by the point-estimate (green line), and the predictive distributions of C_D by the interval-estimate (light green fill as 95% CI), which are also shown in the right side of the figure as distributions. The dotted orange line is C_D distributions by the original model

ity of the parameters \mathbf{a} and σ independently, e.g. as $p(\mathbf{a})$. So far, the constraints of the parameter as $\mathbf{a} \in [0.5 \mathbf{a}_{org}, 1.5 \mathbf{a}_{org}]$ have been used in the process of the optimization and MCMC. Hence, these constraints are implicitly employed as prior information in the form of uniform distributions for all the test cases (see Appendix 1). In contrast, we enforce a prior distribution for \mathbf{a} which is Gaussian and can be represented as:

$$p(\mathbf{a}) = \mathcal{N}(\mathbf{a} | \mathbf{a}_{org}, \text{diag}(\sigma_{pri}^2)), \tag{16}$$

where $\sigma_{pri} = (\sigma_{pri,1}, \sigma_{pri,2}, \dots, \sigma_{pri,M})$ and M is the dimensionality of the parameter \mathbf{a} . Therefore it is the number of the turbulence model coefficients to be inferred. $\sigma_{pri,j}$ indicates the variance of the j th element of the parameter \mathbf{a} .

Figures 5 and 6 show marginal posterior distributions and predictive distributions, respectively, when the specific pri-

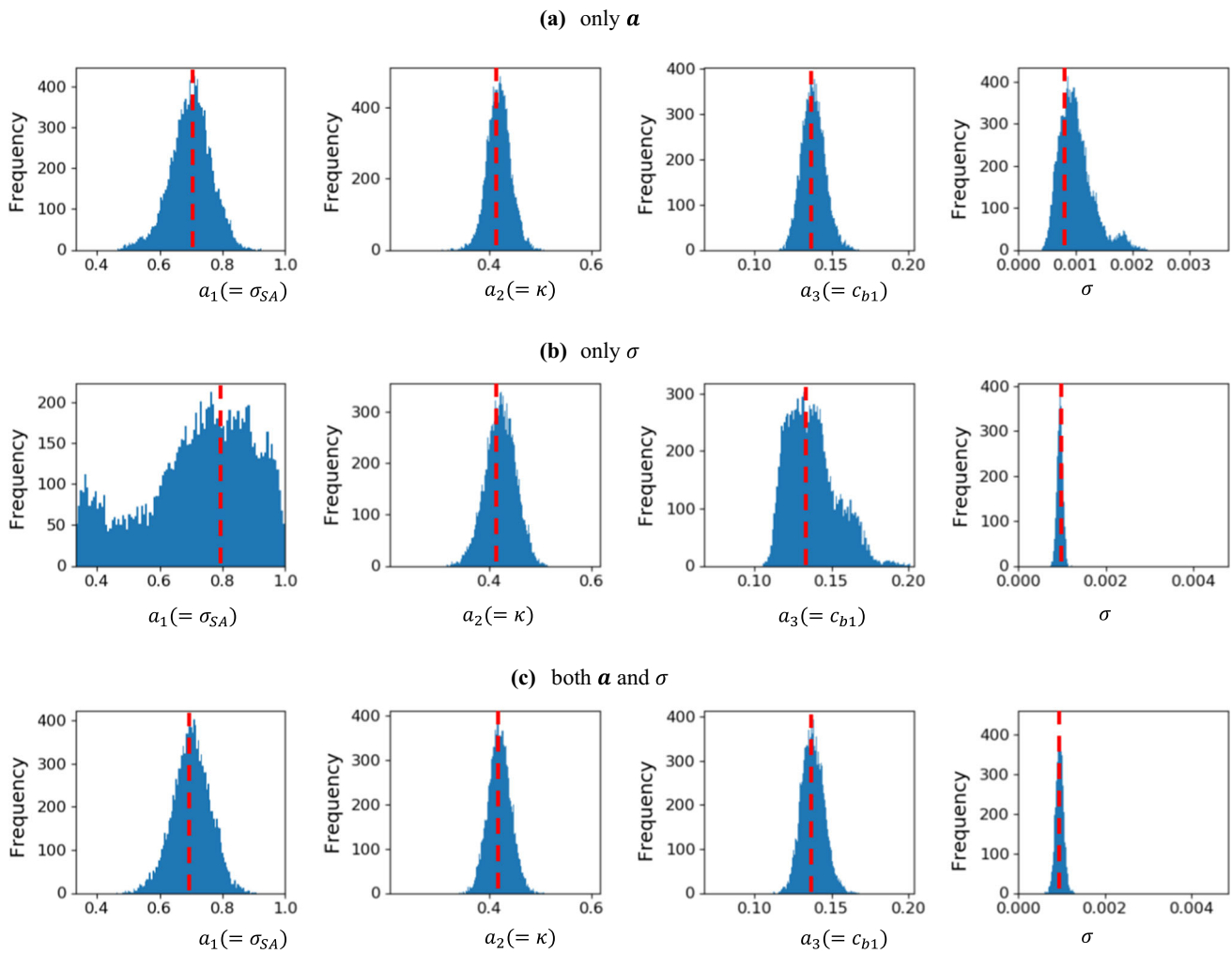


Fig. 5 Marginal distributions of a posterior of the turbulence model parameter \mathbf{a} when a prior is used (red dotted lines indicate $\hat{\mathbf{a}}$)

ors expressed by Eq. (16) are respected during the inference process. Priors for each element of the parameters can be independently set. Therefore, prior information is introduced at three different locations: (A) only for \mathbf{a} , (B) only for σ and (C) for both \mathbf{a} and σ . The function to be evaluated by the optimizer and the MCMC in case (A) is Eq. (16). Those for cases (B) and (C) are $p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)p(\mathbf{a})$ and $p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)p(\mathbf{a})p(\sigma)$, respectively, where $p(\sigma)$ can be obtained by as again following Eq. (16). The mode as $(\hat{\mathbf{a}}, \hat{\sigma})$ obtained by direct optimization for each case is also described in Fig. 5.

Due to the prior functions, the posteriors shown in Fig. 5 can be regarded as having updated from those in Fig. 4a. For example, in case (B), the uncertainty of σ as noise is reduced. In Fig. 5a, the uncertainty of \mathbf{a} is also reduced and the mode moved closer to \mathbf{a}_{org} . Moreover, the resulting predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ in Fig. 6a is smaller compared to Fig. 4b and is distributed around the reference result.

3.1.3 Inference using pressure coefficient

After using the global quantity C_d as the QoI, next the pressure coefficient (C_p) distributions is investigated. As for the drag coefficient analysis, artificial noise data is generated. Similar to before, the data was created based on CFD results using \mathbf{a}_{org} with independently and randomly generated noise from a Gaussian distribution $\mathcal{N}(0, \sigma_{org}^2)$, where $\sigma_{org} = 0.1$. The likelihood function is expressed as follows since the probabilistic model of the noise ε is assumed to be a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ so that the probabilistic model of the output y that we defined is $\mathcal{N}(f(\mathbf{x}^{(i)}, \mathbf{a}), \sigma^2)$. Therefore the likelihood function when the dataset $\mathcal{D} = (\mathbf{X}, \mathbf{Y})$ is:

$$p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma) = \prod_{j=1}^{N_{data}} \prod_{i=1}^{N_j} \mathcal{N}(y^{(i,j)} | f(\mathbf{x}^{(i,j)}, \mathbf{a}), \sigma^2), \quad (17)$$

where N_{data} and N_j indicate the number of data (equivalent to the number of the flow conditions) and the number of the

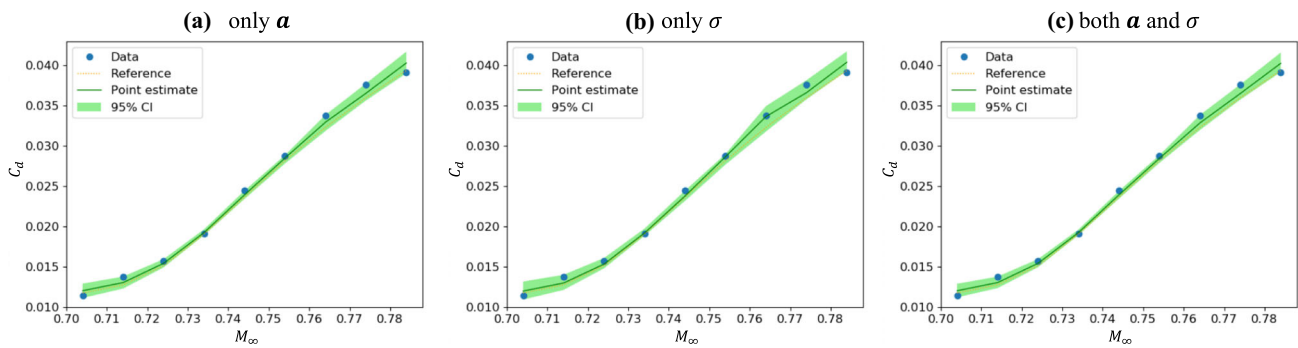


Fig. 6 Predicted C_d distributions when a prior is used, computed by a calibrated turbulence model parameter by the point-estimate, and the predictive distributions of C_d inferred by the interval-estimate

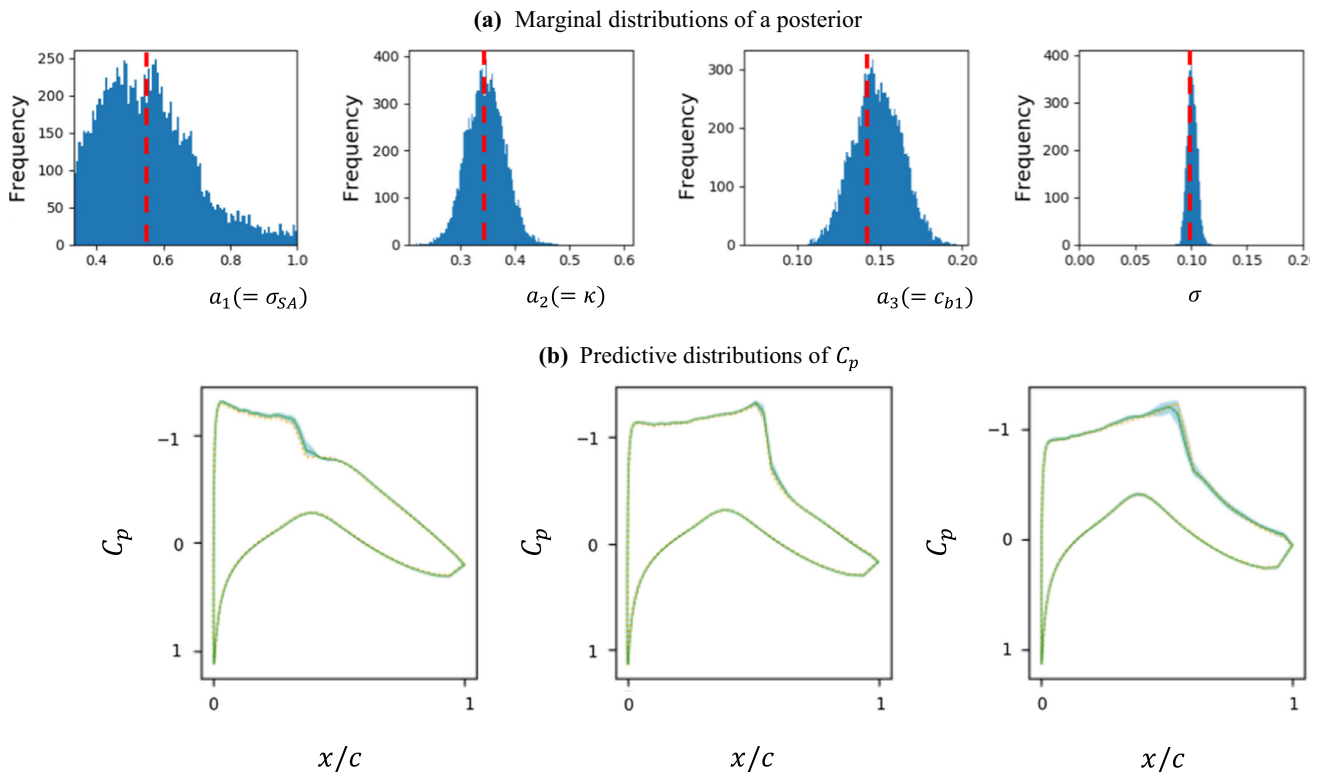


Fig. 7 Results of predictions using C_p distributions of 3 flow conditions (Mach numbers of 0.704, 0.734 and 0.764) when no prior is used. **a** Marginal distributions of a posterior of the turbulence model parameter \mathbf{a} (red dotted lines indicate $\hat{\mathbf{a}}$), **b** predicted C_p distributions by the point-estimate (green line), and the predictive distributions of C_p by the interval-estimate (light blue fill as 95% CI). The dotted orange is the reference

available pressure coefficient at each flow condition, respectively. Note that this equation means that one unique variance σ^2 to represent the output noise ε is assumed to be common at all the flow conditions, i.e. assuming that the noise level is the same at any flow conditions.

Figure 7 shows marginal posterior distributions and their predictive distributions $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_a$ at Mach numbers of 0.704, 0.734 and 0.764. The reference results obtained by \mathbf{a}_{org} and the results by the point-estimate are described as the dotted orange lines and the green lines, respectively. It can be seen that the shock waves around the mid-chord when the

Mach number is at 0.704 and 0.734 can be accurately detected by the deterministic approach (the green line match the reference depicted by the orange dotted lines). The uncertainty shown as 95% credible interval (CI) is wider at the Mach number of 0.704. This uncertainty is the epistemic uncertainty associated with the turbulence closure coefficients. On the other hand, the point-estimate could not detect the reference at the Mach number of 0.764 but the resulting wide uncertainty (shown as 95% CI) by the interval-estimate covers the reference. These results mean that the point-estimate cannot detect the reference but the interval-estimate can as

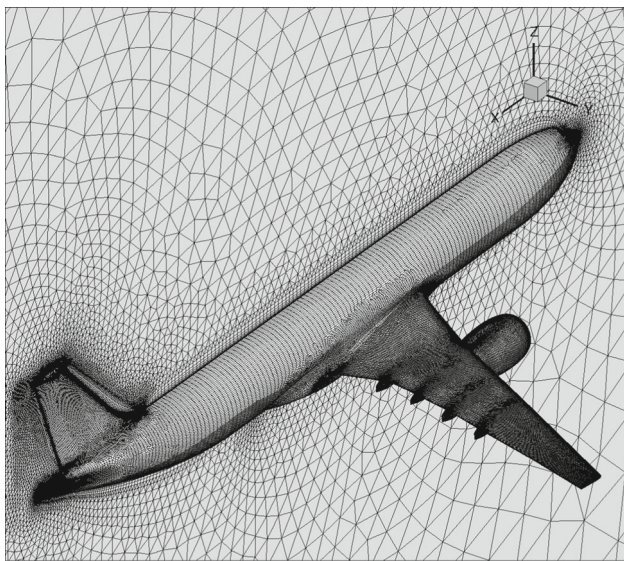


Fig. 8 XRF1 configuration at Case 4 and grid overview

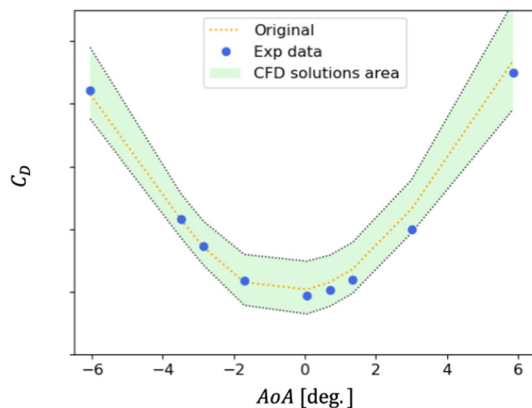


Fig. 9 Drag coefficient distribution of experimental data, CFD solutions using the original coefficients, and feasible CFD solutions range spanned by DoE samples

a probability and furthermore can indicate sensitive flow domains/conditions accordingly, which was not included in the given artificial data as information.

3.2 XRF1 generic long-range transport aircraft

In this section the XRF1 is investigated relying on high-quality but proprietary wind tunnel data at transonic flow conditions and flight Reynolds numbers. XRF1 is an Airbus provided industrial standard multi-disciplinary research testcase representing a typical configuration for a long-range wide body aircraft. The XRF1 research testcase is used by Airbus to engage with external partners on development and demonstration of relevant capabilities / technologies.

The XRF1 wind tunnel data set at hand consists of drag (C_D) and pressure coefficients (C_p) at nine flow conditions,

namely Case 1 to Case 9. The Mach number and the Reynolds number are fixed between the cases while the angle of attack (AoA) is varied. The Mach number and the Reynolds number are around 0.86 and 25 million, respectively. All the flight conditions are within the transonic regime and subsequently shock waves are present on the wing. In this test case, no specific prior is used (the improper prior is implicitly used) to learn the coefficients only from the wind tunnel data.

The inference is executed on two problems:

1. use C_D data from all the nine cases at once,
2. use C_p data available at three sections of the wing focusing on one Case only.

At each flow condition, the aircraft configuration is deformed due to its aeroelastic response. However, in this paper this influence is neglected and a fixed aircraft configuration is used for CFD computations for all nine cases. This fixed configuration has the surface deformation measured for Case 4. Hence for the second analysis focusing on C_p no deformation inconsistency is present. For the C_D analysis the inconsistency in surface deformations is expected to emerge to some extent as the variance σ^2 in Eq. (9). Figure 8 shows the corresponding grid used for RANS computations. The number of grid nodes is around 6.3 million. As one reference showing a validation result of the grid (which also causes an epistemic uncertainty but evaluated in the noise parameter when it exists), we take an example that the difference of C_D between the corresponding data and a RANS computation using the original turbulence model coefficient a_{org} is 3 drag counts (where one drag count is 10^{-4}).

The results of the Sobol indices for the RAE2822 airfoil are used to downselect the turbulence model parameters which are inferred for the XRF1 assuming that both cases share underlying physics due to the transonic flow regime. Hence, the turbulence model coefficients κ , c_{b1} and σ_{SA} are the inferred coefficients group in the parameter \mathbf{a} ($\mathbf{a} = (\sigma_{SA}, \kappa, c_{b1})$). All neglected coefficients are fixed at the original S-A model values given in Table 1.

Surrogate models for the drag coefficient are constructed at each flow condition while for the pressure coefficient distributions one model for each wing section is computed. Since the dimensionality of the turbulence model parameter \mathbf{a} is 3, the required number of CFD computations has been reduced compared to the 12-dimensional airfoil case. A minimum of 70 DoE sample points to construct a surrogate model at each flow condition is used. This varying number of DoE samples is caused from diverging CFD simulations which occur once a combination of closure coefficients has been selected which violate the underlying RANS modeling assumptions.

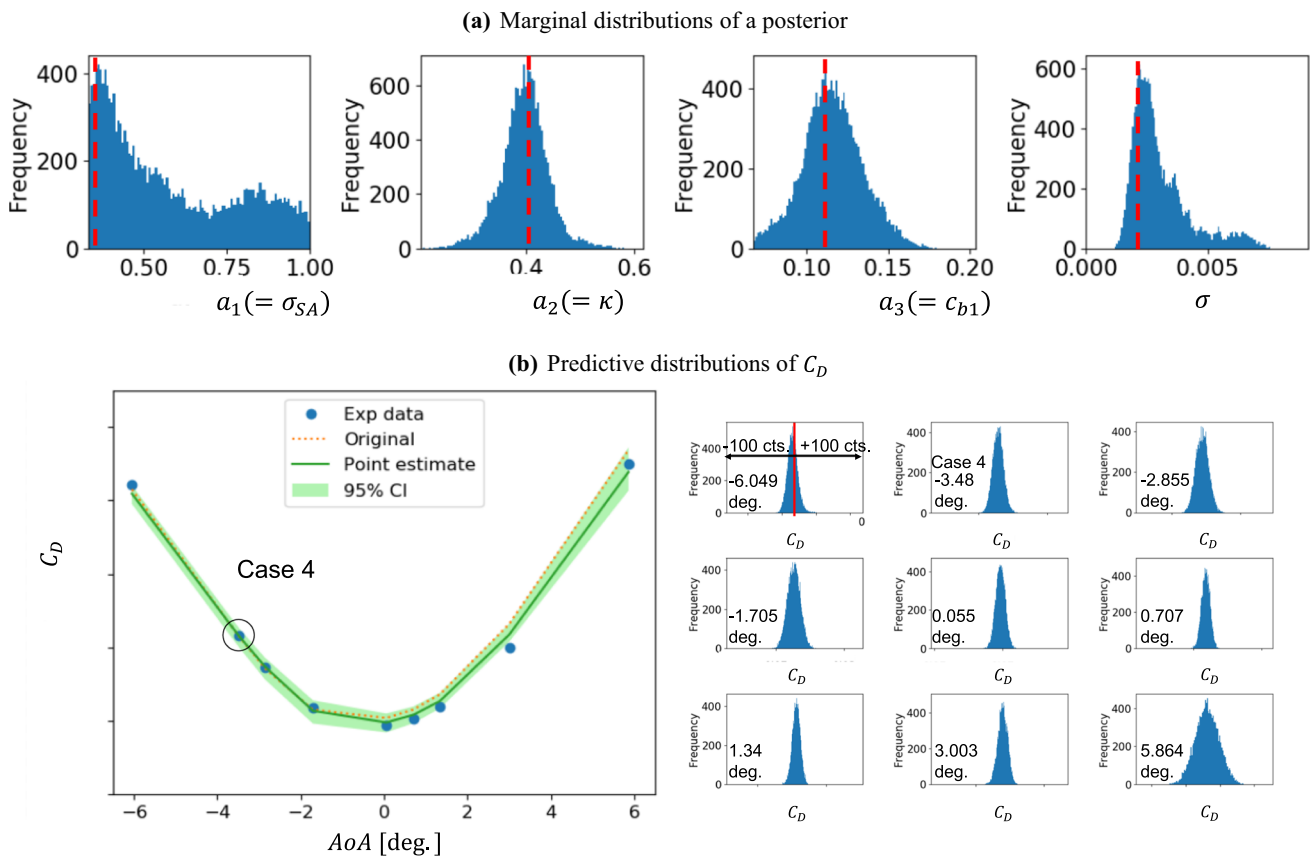


Fig. 10 Results of predictions using C_D by the experimental dataset at nine flow conditions. **a** Marginal distributions of a posterior of the turbulence model parameter (red dotted lines indicate \hat{a}), **b** predicted C_D distributions by the point-estimate (green line), and the predictive distributions of C_D by the interval-estimate (light green fill as 95% CI), which are also shown in the right side of the figure as distributions. The dotted orange line is C_D distributions by the original model

3.2.1 Inference using drag coefficient

First, turbulence model coefficients are inferred based on the drag coefficient C_D . The dataset used consisting of the input \mathbf{X} and the output \mathbf{Y} at all the nine flow conditions as $\mathbf{X} = (\mathbf{x}^{(case1)}, \mathbf{x}^{(case2)}, \dots, \mathbf{x}^{(case9)})$ and $\mathbf{Y} = (\mathbf{y}^{(case1)}, \mathbf{y}^{(case2)}, \dots, \mathbf{y}^{(case9)})$, where \mathbf{x} and \mathbf{y} are the AoA and C_D , respectively. The experimental data for C_D is shown in Fig. 9. As aforementioned, the deformation used for all computations is fixed and belongs to Case 4 which is the case of the second lowest angle of attack. Figure 9 also depicts the possible C_D region in the search space $\mathbf{a} \in [0.5 \mathbf{a}_{org}, 1.5 \mathbf{a}_{org}]$. This region is spanned by directly using all available DoE sample points and plotting their minimum and maximum values. It has to be noted that this coverage does not guarantee that a possible combination of parameters exists for each C_D value inside. Nevertheless, it can be seen that all experimental data points are within the covered region. The C_D characteristics using the original coefficients \mathbf{a}_{org} are also described in Fig. 9.

Results obtained by the point-estimate and the interval-estimate are shown in Fig. 10. Since the probabilistic model is exactly the same as that in the case of RAE2822 in the previous sub-section, the likelihood function used for the learning process is Eq. (10) as well. First of all, the results of the point-estimate are discussed. After the learning process, the parameters \mathbf{a} and σ in Eq. (10) are calibrated as $\hat{\mathbf{a}}$ and $\hat{\sigma}$. Figure 10 shows C_D characteristics over the nine angles of attack. The green line indicates the results computed by using the calibrated turbulence model parameter $\hat{\mathbf{a}}$. At the angle of attack of Case 4, the predicted C_D by the calibrated parameter $\hat{\mathbf{a}}$ (C_D^{calib}) corresponds well to the underlying data. For the other cases, C_D^{calib} is between the data \mathbf{Y} and C_D^{org} at each angle of attack, which is the qualitative information obtained by the point-estimate. The difference between C_D^{calib} and \mathbf{Y} is treated as the noise produced by a probability with the calibrated parameter $\hat{\sigma}$. This noise represents all of the uncertainties except the uncertainty associated with the model parameter \mathbf{a} . By following the model definition in sect. 2.2, the inaccuracy of the configuration due to aeroelasticity is also treated as the noise.

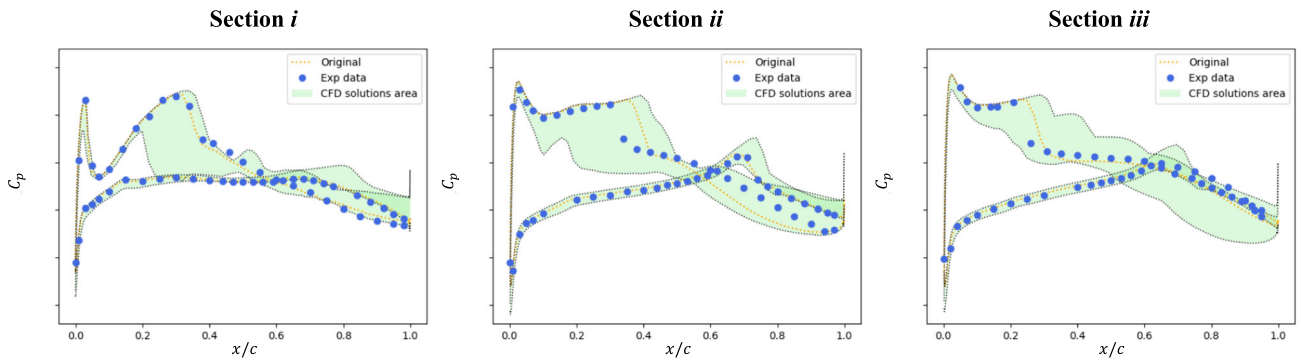


Fig. 11 Pressure distributions of experimental data and feasible CFD solutions by DoE samples at three sections

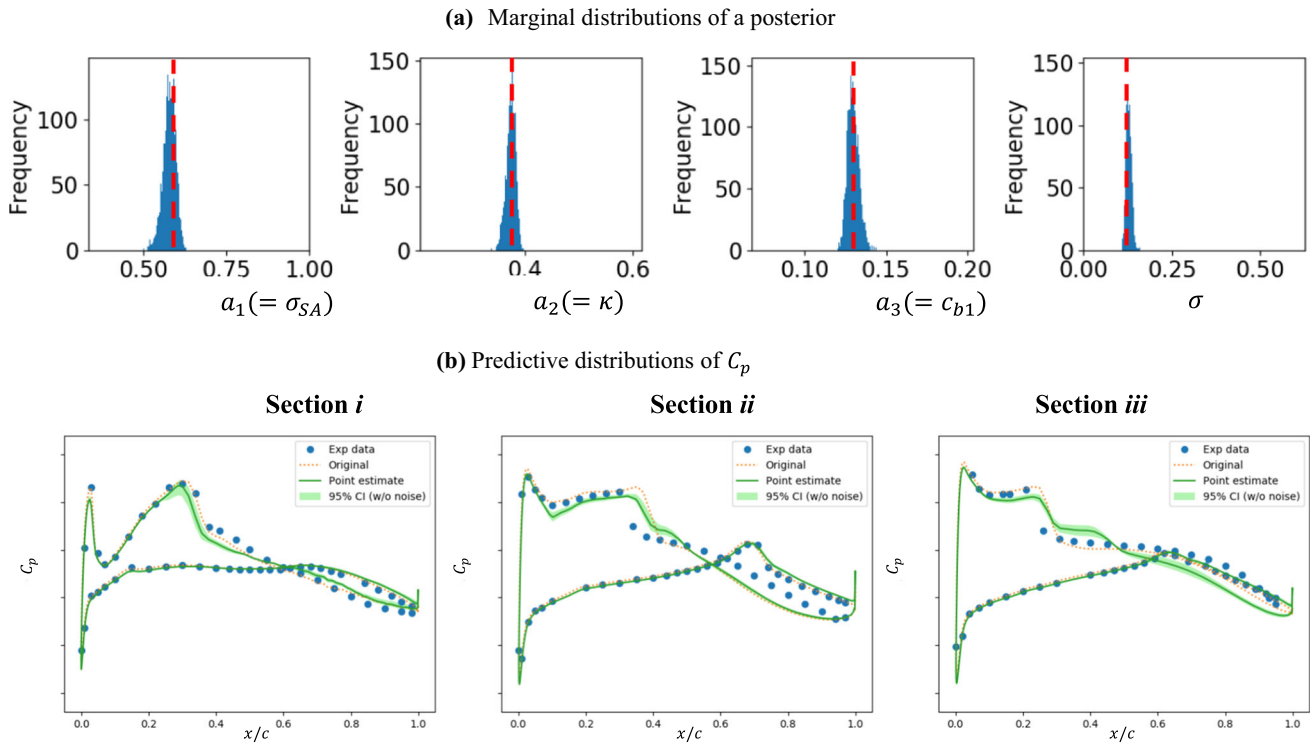


Fig. 12 Results of predictions using C_p distributions by the experimental dataset at one flow condition (Case 4) when the turbulence model parameter \mathbf{a} and the noise parameter σ are set to be common at all the three sections. **a** Marginal distributions of a posterior of the turbulence model parameter (red dotted lines indicate $\hat{\mathbf{a}}$), **b** predictions by the point-estimate and predictive distributions by the interval-estimate (showing 95% CI)

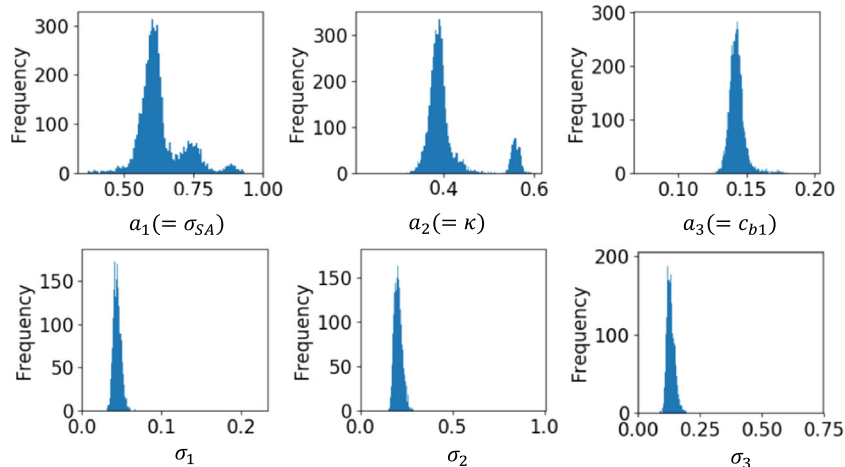
When following a stochastic approach, the marginal posterior distributions shown in the upper part of Fig. 10 have large uncertainties especially on $a_1 (= \sigma_{SA})$ since it is fully distributed in the constrained range of $[0.5 a_{org}, 1.5 a_{org}]$. The mode is around the lower bound. The noise parameter σ is also widely distributed and spreads around four times of the mode $\hat{\sigma}$. The predictive distributions $p(y|x, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ are depicted in the lower part of Fig. 10. It can be observed that the 95% CI at the highest AoA is quite wide compared to lower AoA. The predictive distribution at the second highest AoA is relatively narrow and the data point at this condition is not within the 95% CI. The noise parameter σ is actually estimated quite high even though the exact drag counts of σ

cannot be described here. Therefore, the predictive distribution considering all the uncertainties computed by Eq. (13), should contain the data points. However, in this particular application case, there exists external noise due to the geometry difference which falsifies this assumption.

3.2.2 Inference using pressure coefficient

As described in the 2D airfoil case, at each flow condition, the amount of the information when using C_p distributions is much greater than when using C_D . However, for practically relevant cases, C_p measurements are usually available only at several sections/locations. Figure 11 shows C_p distributions

(a) Marginal distributions of a posterior



(b) Predictive distributions of C_p

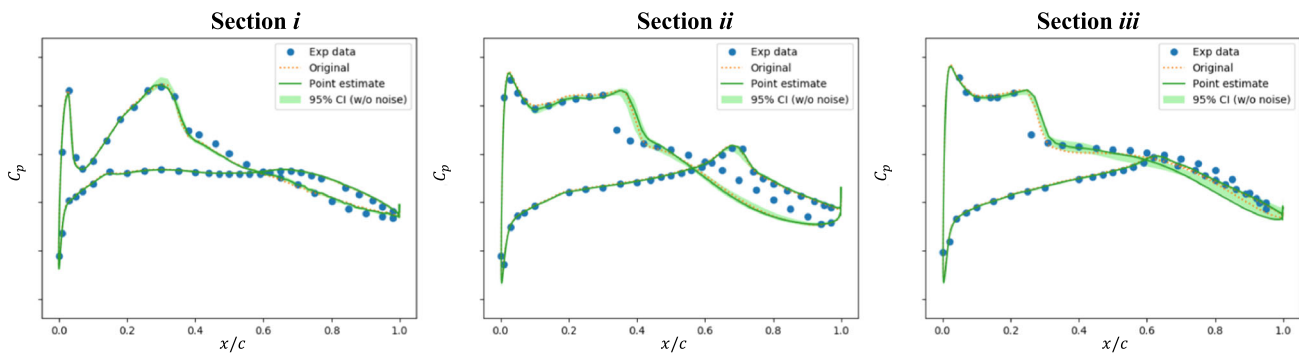


Fig. 13 Results of predictions using C_p distributions by the experimental dataset at one flow condition (Case 4) when the turbulence model parameter \mathbf{a} is set to be common at all the three sections but the noise parameter σ is set individually. Upper: marginal distributions of a posterior of the turbulence model parameter (red dotted lines indicate $\hat{\mathbf{a}}$). Lower: predictions by the point-estimate and predictive distributions by the interval-estimate (showing 95% CI)

of the experimental data and feasible CFD solutions for Case 4. Three sections are available and each of them cover the whole surface of the section where C_p data is measured, namely Section *i* to Section *iii* ordered from wing root to tip. Since the angle of attack at Case 4 is negative, shock waves are generated on the lower surface of the wing.

As the assumption in the inference process introduced in Sect. 2.2, the turbulence model parameter \mathbf{a} is stochastic. It is neither a spatio-temporal parameter like a function of the time and space nor a function of flow conditions. In definition of probabilistic models, we introduce two derivations (A) and (B) from two hypotheses:

- A. the same noise parameter at each section as one σ
- B. different noise parameters at each section as $\sigma^{(j)}$, where $j = 1, 2, 3$

The likelihood function in the hypothesis (A) is the same as Eq. (17) in the 2D RAE2822 test case as follows:

$$p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma) = \prod_{j=1}^{N_{sec}} \prod_{i=1}^{N_j} \mathcal{N}(y^{(i,j)} | f(\mathbf{x}^{(i,j)}, \mathbf{a}), \sigma^2), \quad (18)$$

where $x^{(i,j)}$ indicates in this case the spatial coordinate of i th point at the section j , and N_{sec} indicates the number of the sections. Here we use the three sections, therefore $N_{sec} = 3$. A common noise parameter denoted by σ is set for all the sections.

On the other hand, the noise parameter of the likelihood function in the hypothesis B) can be expressed by a vector. Each element of the parameter vector consists of a different noise parameter $\sigma^{(j)}$ at each section j . The likelihood function is slightly modified from Eq. (18) as:

$$p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \boldsymbol{\sigma}) = \prod_{j=1}^{N_{sec}} \prod_{i=1}^{N_j} \mathcal{N}(y^{(i,j)} | f(\mathbf{x}^{(i,j)}, \mathbf{a}), \sigma^{(j)2}), \quad (19)$$

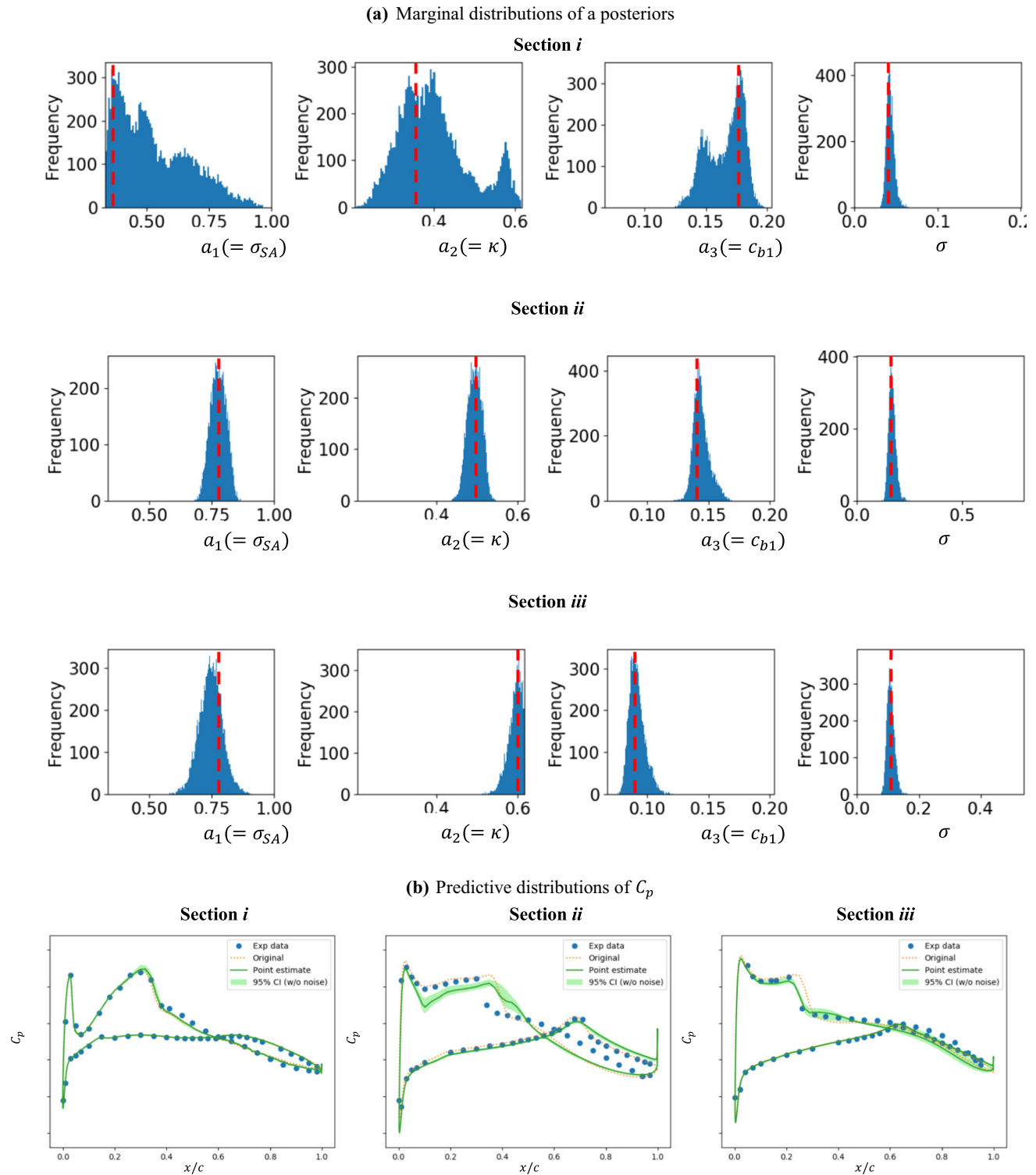


Fig. 14 Results of predictions using C_p distributions by the experimental dataset at one flow condition (Case 4) when both the turbulence model parameter \mathbf{a} and the noise parameter σ are set independently at each section. **a** Marginal distributions of a posterior of the turbulence model parameter (red dotted lines indicate $\hat{\mathbf{a}}$), **b** predictions by the point-estimate and predictive distributions by the interval-estimate (showing 95% CI)

where $\sigma = (\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(N_{sec})})$. The probabilistic models of the hypotheses (A) and (B) are summarized in Appendix 4 for further descriptions of the hypotheses behind.

The point-estimate and the interval-estimate are conducted on Eqs. (18) and (19) in order to obtain $\hat{\mathbf{a}}$, $\hat{\sigma}$ (by the point-estimate), and $p(\mathbf{a}, \sigma)$ (by the interval-estimate),

where $\hat{\sigma} = (\hat{\sigma})$ in Eq. (18), $\hat{\sigma} = (\hat{\sigma}^{(1)}, \hat{\sigma}^{(2)}, \dots, \hat{\sigma}^{(N_{sec})})$ in Eq. (19), respectively.

Figure 12 shows the marginal posterior distributions and predictive distributions $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$ when a unique noise parameter is introduced. In the predictive distributions at all three sections, the epistemic uncertainties associated with the parameter \mathbf{a} (shown as 95% CI in the figure) are slightly wider around the shock locations and downstream thereof compared to more benign flow phenomena. However, they are so small that they do not cover the experimental data. In fact, the predictive distribution at section i even diverged further from the experimental data than the CFD prediction relying on the original model parameters \mathbf{a}_{org} . These results are mainly caused by introducing the same noise parameter for all the sections.

On the hypothesis B using Eq. (19), a more flexible assumption is used in which noise at each section is expressed independently. Figure 13 shows the marginal posterior distributions and predictive distributions $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$. The overall characteristics are similar to those of the hypothesis (A). However, the predictive distribution covers the shock wave at Section i slightly better than hypothesis A). Hence, assuming different noise parameters for each section seems more promising than assuming a common noise parameter. However, the results in Fig. 13 also clearly indicate that there is no possibility for the turbulence model parameter to properly predict all flow phenomena at the three sections by adjusting just three turbulence model parameters and aiming for a global model description.

Based on the previous results, the flexibility is further increased by allowing different turbulence model parameters \mathbf{a} for each section as $\mathbf{a}^{(j)}$, where $j = 1, 2, \dots, N_{sec}$ (named hypothesis (C)). The likelihood function at each section is identical to Eq. (10). Figure 14 shows three sets of marginal posterior distributions on the three sections and the respective predictive distributions $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y})|_{\mathbf{a}}$. It can be observed in Fig. 14 that the shock wave at each section can be better detected by independently inferred parameter $\mathbf{a}^{(j)}$. As before, the uncertainties associated with the individual parameters \mathbf{a} are larger around the shock locations. This showcases that the developed framework is able to detect location of increased physical complexity, e.g. shocks, and highlights them by a larger local uncertainty.

A natural extension of the previous strategy is the assumption that the parameter is spatially varying in general. For example, Duraisamy et al. [7] proposed a model correction term and then inferred it employing a learning process. Since the spatially varying parameter has the same dimensionality than the underlying computational grid, computations of a posterior by MCMC are infeasible and they are achieved by the other approximation methods based on definition of posterior distributions such as the variational inference and the Laplace approximation (see Appendix 2).

4 Conclusions

A framework for data-driven turbulence modeling under uncertainties was developed and applied to a 2D airfoil and an industrially-relevant 3D full aircraft case. The objective of this framework is to quantify the uncertainty of the closure coefficients and that of the quantities of interest associated with the closure coefficients by employing a Bayesian approach. In this paper, the closure coefficients of the Spalart Allmaras turbulence model were inferred. The partial availability of information, which is common for cases of industrial relevance, is regarded as an epistemic uncertainty that is addressed by a fully Bayesian approach, with special treatment for computing the predictive distribution. The resulting uncertainties can provide an insight into one of the model uncertainties due to the closure coefficients for predicting the quantities of interest by distinguishing it from other external uncertainties. In this work, these other external uncertainties are treated as noise. The other possible epistemic uncertainties in the CFD simulations such as the grid convergence were eliminated as much as possible in advance though they are classified into the noise if they exist. The uncertainties associated by all the parameters, which are the closure coefficients and the noise parameter, in the fully Bayesian model were rigorously evaluated by a Markov-Chain Monte Carlo (MCMC) sampling technique. To overcome the bottleneck of high computational cost in the MCMC process associated with repetitive RANS evaluation, a surrogate-based methodology has been employed.

The 2D airfoil test case with artificial data has been used to validate the methodologies integrated in the framework. Moreover, dominant closure coefficients were selected by using Sobol indices as a preprocessing step that further increased the efficiency of the surrogate-based approximation. It was shown that known model coefficients can be rediscovered by relying on data. Furthermore, it was confirmed not only that the closure coefficients and the artificial noise were replicated by the developed method but also that the epistemic uncertainty associated with the closure coefficients are large around the shock locations. What is important in this replication test case is that the information of this epistemic uncertainty was not specified in the problem setting but was derived by the developed method.

For the 3D aircraft case the information obtained from the airfoil Sobol indices analysis has been reused to keep computational cost at a feasible level. Inference of coefficient by analyzing pressure coefficients showed that the epistemic uncertainty associated by the coefficients are highly detected in regions of sensitive flow, such as the shock position and downstream thereof. It was also shown that the uncertainty associated by the noise is quite large compared with that by the coefficients in this test case.

Obtained results, and in particular the epistemic uncertainties associated by the coefficients, can be used for various applications. For example, one option could be to feedback information to the experimental measurement strategies to determine at which locations on the wing new measurement devices should be placed. Another opportunity is the propagation of uncertainties towards unobservable flow conditions with the assistance of surrogate models and the posterior of the closure coefficients to enhance the modeling of e.g. loads at the edge of the flight envelope. The developed framework is flexible in terms of modelling, i.e. other closure models can be applied as well as other probabilistic models.

Acknowledgements The DLR authors would like to thank Airbus for providing the XRF1 testcase as a mechanism for demonstration of the approaches presented in this paper. The first author would like to thank to Bernhard Eisfeld and Vamshi Togiti at DLR for discussion and assistances on the DLR-TAU code and for valuable comments, and expert knowledge on the turbulence models. The first author would also like to thank to Jan Himisch, Mohammed Abu-Zurayk, Caslav Ilic, Andrei Merle at DLR, and Thomas Engelbrecht, Cross Murray, Paul White and Alan Mann at Airbus for data offering and valuable discussion on the XRF1 test case.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

Supplementation of the omitted explanations of the methodologies and details of the theory in background in Sect. 2.2 and Sect. 3 are described here.

1. The function to be evaluated by optimizer/MCMC

In Sect. 2.2, it is explained that a likelihood function is evaluated by MCMC. Accurately to say, it has to be a posterior function by the Bayes’ theorem. It needs to be transformed into a form of a probability in terms of the parameters \mathbf{a} and σ as from $p(y|\mathbf{x}, \mathbf{a}, \sigma)$ to $p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y})$. This is the transformation from a likelihood $p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)$ to a posterior $p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y})$ and can be achieved by the Bayes’ theorem as:

$$p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y}) = \frac{p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)p(\mathbf{a}, \sigma)}{p(\mathbf{Y})}, \tag{20}$$

where $p(\mathbf{a}, \sigma)$ is a joint prior of the parameters \mathbf{a}, σ that can be defined arbitrarily. In practice, we can set independent priors for each element of the parameters, e.g. $p(\mathbf{a}, \sigma) = p(\mathbf{a})p(\sigma)$ since \mathbf{Y} is not observed yet. $p(\mathbf{Y})$ is the denominator in Bayes’ theorem as being the normalization constant required to ensure the posterior to be a probability density. This normalization constant can be represented by an integral of the numerator in the right-hand side but is normally incomputable. However, the formulation needed to be evaluated is the posterior by Eq. (20) in theory.

However, in practice, the likelihood function $p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)$ or the product of that and a prior $p(\mathbf{a}, \sigma)$ is sufficient for the formulation to be evaluated when using optimizer/MCMC. Therefore, the formulation to be evaluated is simply the product of the likelihood and a prior:

$$p(\mathbf{a}, \sigma|\mathbf{X}, \mathbf{Y}) \propto p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)p(\mathbf{a}, \sigma). \tag{21}$$

When no specific prior wants to be assumed (purely evaluated only by currently observed data), improper prior such as $p(\mathbf{a}, \sigma) = 1$ to be $p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)p(\mathbf{a}, \sigma) = p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)$ can be used. It is rational when no nonlinear transformation of the parameters and the normalize constant can be obtained [8]. Eventually the function to be evaluated by optimizer/MCMC becomes the likelihood function itself. Note that like a likelihood function in general, the improper prior either does not guarantee the integral of it to be 1 over the parameters \mathbf{a}, σ . The improper prior can be applicable since the normalization constant can be computed by the MCMC and since no nonlinear mapping of the parameter \mathbf{a} is operated in the whole procedure. Constraints of the parameter \mathbf{a} (e.g. $\mathbf{a} \in [0.5 \mathbf{a}_{org}, 1.5 \mathbf{a}_{org}]$ used in the main context) can be also treated as prior information. The prior function of these constraints is a step function extended from the improper prior as:

$$p(\mathbf{a}) = \begin{cases} 1, & 0.5\mathbf{a}_{org} \leq \mathbf{a} \leq 1.5\mathbf{a}_{org}, \\ 0, & otherwise. \end{cases} \tag{22}$$

Therefore, all of the likelihood functions (Eqs. (10), (17–19)) introduced in the main context is actually:

$$\begin{aligned} & p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma)p(\mathbf{a}, \sigma) \\ &= \begin{cases} p(\mathbf{Y}|\mathbf{X}, \mathbf{a}, \sigma), & 0.5\mathbf{a}_{org} \leq \mathbf{a} \leq 1.5\mathbf{a}_{org}, \\ 0, & otherwise, \end{cases} \end{aligned} \tag{23}$$

where $p(\mathbf{a}, \sigma) = p(\mathbf{a})p(\sigma)$, and $p(\sigma) = 1$ as the improper prior. This expression is totally omitted in the main context to avoid complexity. In practice the posterior to be a probability that the integral of posterior as a probability density function is guaranteed to be 1 by using the MCMC.

2. How to compute the predictive distributions

First of all, it has to be reminded that cases where the regression model defined by Eq. (7) is a linear regression model, and the probabilistic model defined by Eq. (9) is a Gaussian distribution, can analytically represent the predictive distribution in general [8]. The posterior can be analytically obtained by assist of a conjugate prior, which can share the same probability distribution between the prior and the posterior, and leads to efficient update of the posterior. These cases are often used in practical applications. In fact, some properties in this process to obtain analytical solutions are applied to the approximation methods for their efficiency. For example, one of the most widely used analytical representations is the case where the noise parameter (of a Gaussian probabilistic model) is deterministic and the prior is a Gaussian distribution. This modeling can be categorized in one of the empirical Bayes approaches shown in Appendix 3. In this case, the posterior becomes also a Gaussian distribution. This approach would be rather used for rapid approximations in iterative processes. On the other hand, when the noise parameter is treated as stochastic such as the case in this article, the conjugate prior is not a Gaussian distribution but a Gauss-Gamma distribution and more generally a Gaussian-Wishart distribution for multiple input parameters.

In our applications, since the regression model is a nonlinear regression model, the predictive distribution by Eq. (13) does not have analytical solutions even in the empirical Bayesian approaches (see Appendix 3). As an example of other cases, the neural networks in general can be also regarded as nonlinear regression models in general, so as it to the same.

There are several approximation methods in computing the predictive distributions [8]

- MCMC,
- Laplace approximation,
- variational inference.

The first one is a method with a help of the MCMC. The integral in Eq. (13) is replaced by the samples. The aim of the latter two methods is basically to approximate the posterior distribution by a parametric distribution. These can be rather used in terms of efficiency or intractability due to high dimensionality of the parameter to be inferred. These require other alternative computations such as optimization (used for the point-estimate) and Hessian predictions of the regression model $f(\mathbf{x}, \mathbf{a})$, the second derivatives with respect to \mathbf{a} . If computational resources are sufficient, the MCMC strategy is expected to be more accurate in general than the others.

When computing the predictive distributions with all mixed uncertainties represented by Eq. (13), the computation is a mixture of Gaussian distributions, which consists

of equally weighted N_{mcmc} Gaussian distributions, each of which is expressed by a known mean and variance as:

$$\begin{aligned} p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}) &\approx \frac{1}{N_{mcmc}} \sum_{i=1}^{N_{mcmc}} p(y|\mathbf{x}, (\mathbf{a}, \sigma)^{(i)}) \\ &= \frac{1}{N_{mcmc}} \sum_{i=1}^{N_{mcmc}} \mathcal{N}(y|f(\mathbf{x}, \mathbf{a}^{(i)}), \sigma^{(i)^2}). \end{aligned} \quad (24)$$

The predictive distribution is not guaranteed to be a Gaussian distribution but its mean and variance can be analytically obtained by the generated MCMC samples in hand whereas it is computationally expensive to estimate the distribution itself by nonparametric approaches.

Thus, the predictive distribution at arbitrary output value y represented by Eq. (13) can be computed by the mixture of the distributions $p(y|\mathbf{x}, (\mathbf{a}, \sigma)^{(i)})$, where $i = 1, 2, \dots, N_{mcmc}$. Each distribution $p(y|\mathbf{x}, (\mathbf{a}, \sigma)^{(i)})$ is properly weighted by dividing the sum of them by the MCMC sample size N_{mcmc} . However, although any complex distributions can be represented by the MCMC sample points, the probability $p(y)$ is not parametric. Therefore, the probability needs to be computed at arbitrary y . The direct computation of that is expensive so that approximation methods are recommended to be used.

An approximation example when the empirical Bayes approach is described here. The empirical Bayes approach is to treat the hyperparameter (the noise parameter σ in the main context) as deterministic (e.g. as $\hat{\sigma}$). The mixture of Gaussian expressed by Eq. (24) is adjusted to as follows in the empirical Bayes approach:

$$\begin{aligned} p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma}) &= \int p(y|\mathbf{x}, \mathbf{a}, \hat{\sigma}) p(\mathbf{a}|\mathbf{X}, \mathbf{Y}, \hat{\sigma}) d\mathbf{a} \\ &\approx \frac{1}{N_{mcmc}} \sum_{i=1}^{N_{mcmc}} p(y|\mathbf{x}, \mathbf{a}^{(i)}, \hat{\sigma}) \\ &= \frac{1}{N_{mcmc}} \sum_{i=1}^{N_{mcmc}} \mathcal{N}(y|f(\mathbf{x}, \mathbf{a}^{(i)}), \hat{\sigma}^2), \end{aligned} \quad (25)$$

where $\hat{\sigma}$ is a fixed noise parameter. The concept of this approximation is the same as the Laplace approximation and the variational inference shown in the above. It is to approximate the posterior distribution $p(\mathbf{a})$ as a Gaussian distribution. Then only two parameters are needed to completely describe the distribution. Therefore, the mean $\mu_{\mathbf{a}, \sigma}$ and variance $\sigma_{\mathbf{a}}^2$ are computed from the MCMC samples describing $p(\mathbf{a})$. The variance $\sigma_{\mathbf{a}}^2$ is used later to approximate the predictive distribution.

Here let us take the analytical form of the predictive distribution of Eq. (13) when the regression model defined in Eq. (7) is a linear regression and the probabilistic model defined in Eq. (9) is a Gaussian distribution.

$$p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}) \approx \mathcal{N}(y|f(\mathbf{x}, \hat{\mathbf{a}}), \sigma_{pre}^2), \tag{26}$$

where σ_{pre}^2 is the variance of the predictive distribution. For example, when the regression model is a linear regression as $f(\mathbf{x}, \hat{\mathbf{a}}) = \hat{\mathbf{a}}^T \boldsymbol{\varphi}(\mathbf{x})$ can be analytically solved as the analytical solution of the least-squares as $\hat{\mathbf{a}} = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T \mathbf{Y}$ where $\boldsymbol{\Phi}$ is a matrix whose elements consist of $\varphi_j(\mathbf{x}_i)$ (so-called design matrix).

The variance σ_{pre}^2 is now easily computed by the following two facts. One is that the two distributions, the probabilistic model and the posterior distribution, both of which compose the predictive distribution, are independent of each other. This is actually general characteristics of predictive distributions. The other one is that both of the two distributions are now Gaussian distributions. Based on these two facts, the variance σ_{pre}^2 is computed simply the sum of the variance of each distribution as:

$$\sigma_{pre}^2 = \hat{\sigma}^2 + \sigma_a^2, \tag{27}$$

where $\hat{\sigma}^2$ is the variance of the probabilistic model the noise defined by Eq. (9) [and it was obtained by Eq. (11)]. σ_a^2 is the variance of the posterior distribution. It is regarded as the uncertainty of the model parameter \mathbf{a} . Both $\hat{\sigma}^2$ and σ_a^2 are the variances obtained after the learning process. The mixed predictive distributions stemmed from all the uncertainties, which are not valued as important in this article, can be then efficiently computed.

3. Comparison with the empirical Bayes approach and the predictive distribution composed of all the mixed uncertainties

The objective of this paper is UQ of QoIs derived from the turbulence closure coefficients \mathbf{a} taking the uncertainty of the noise into account. This modeling can cover all the uncertainties that are present under the probabilistic model on our definition. Here influences of the modeling that the noise parameter is also treated as stochastic are demonstrated by comparing the results of the empirical Bayes approach in which the noise parameter is treated as deterministic as $\hat{\sigma}$ (e.g. obtained by Eq. (11)). The Eqs. (13) and (15a) are adjusted to Eq. (25) and the following equation, respectively:

$$p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})|_a = \int p(y|\mathbf{x}, \mathbf{a})p(\mathbf{a}|\mathbf{X}, \mathbf{Y}, \hat{\sigma})d\mathbf{a} \tag{28}$$

$$\approx \frac{1}{N_{mcmc}} \sum_{i=1}^{N_{mcmc}} p(y|\mathbf{x}, \mathbf{a}_i)$$

with $p(y|\mathbf{x}, \mathbf{a}) = \delta(y - f(\mathbf{x}, \mathbf{a}))$.

The only difference between Eqs. (15a) and (28) is the posterior distributions $p(\mathbf{a}|\mathbf{X}, \mathbf{Y}, \hat{\sigma})$ in Eq. (28) and $p(\mathbf{a}|\mathbf{X}, \mathbf{Y})$ in Eq. (15a), whose distributions are computed by MCMC. The difference between these posterior distributions can be shown in the next paragraph with Fig. 15. The predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})$ by Eq. (25) reflects the contribution of all the uncertainties (see Fig. 16 as an example). The results of reflecting all the uncertainties are shown in the paragraph after the next.

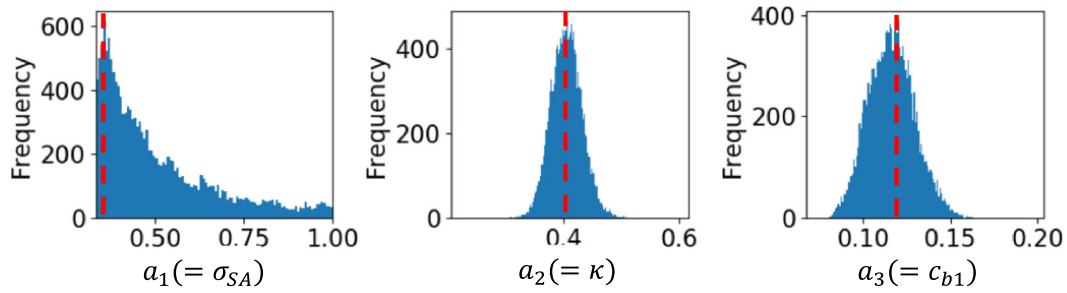
The difference of the predictive distributions between the full Bayes approach and the empirical Bayes approach are expected to be remarkable when the posterior distribution of σ is wide, i.e. when a large uncertainty on σ . Figure 15 shows marginal posterior distributions and predictive distributions for the inference using drag coefficient, on comparison of Fig. 10, which has relative wide distribution in terms of σ . Overall the uncertainties of the marginal posterior distributions (of \mathbf{a}) and the accompanying predictive distributions are estimated to be smaller than the cases when the full Bayes approach in the main context is used. Note that the mode $\hat{\mathbf{a}}$ is common for the both approaches since it can be evaluated just by the point-estimate.

Next, an example of the predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})$ when the contribution of all the uncertainties is accounted is presented. Figure 16 shows comparison between the predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})|_a$ computed by Eq. (28) and the predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})$ computed by Eqs. (25–27)). The approach commonly used here is the empirical Bayes. Therefore, the noise parameter σ is fixed at $\hat{\sigma}$. The result obtained by the fully Bayesian approach is shown in Fig. 14-Section *ii* in the main context. It can be interesting to compare Fig. 16a and Fig. 14-Section *ii* like comparison between Figs. 10 and 15 for C_D , which shows that there is fewer differences in the cases of C_p than those of the case of C_D . However, the main focus here is to confirm the differences of the uncertainties between $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})|_a$ and $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})$. The predictive distribution $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})$ (see 95% CI in the figures) is not useful anymore to observe the uncertainty associated with the parameter \mathbf{a} which is mixed to be lost in $p(y|\mathbf{x}, \mathbf{X}, \mathbf{Y}, \hat{\sigma})$.

4. Probabilistic models

The probabilistic models of hypotheses (A), (B) and (C) in Sect. 3.2, i.e. how the likelihoods expressed by Eqs. (17),

(a) Marginal distributions of a posterior



(b) Predictive distributions of C_D

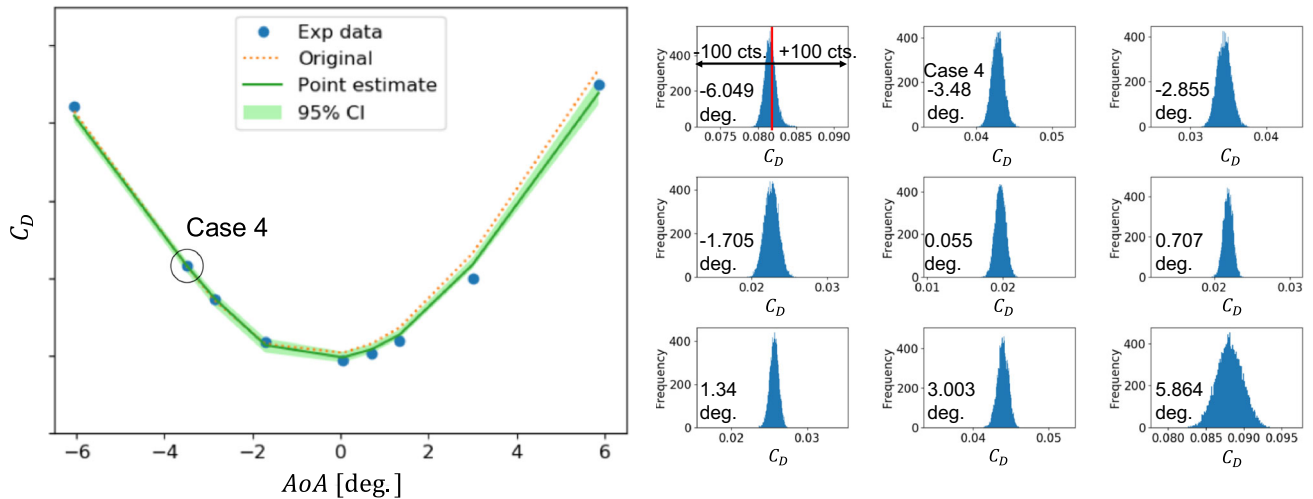


Fig. 15 Results of predictions by the empirical Bayes approach, using C_D by the experimental dataset at nine flow conditions. **a** Marginal distributions of a posterior of the turbulence model parameter (red dotted lines indicate \hat{a}), **b** predicted C_D distributions by the point-estimate (green line), and the predictive distributions of C_D by the interval-estimate (light green fill as 95% CI), which are also shown in the right side of the figure as distributions. The dotted orange line is C_D distributions by the original model

(a) $p(y|x, \mathbf{X}, \mathbf{Y}, \hat{\sigma})|_a$

(b) $p(y|x, \mathbf{X}, \mathbf{Y}, \hat{\sigma})$

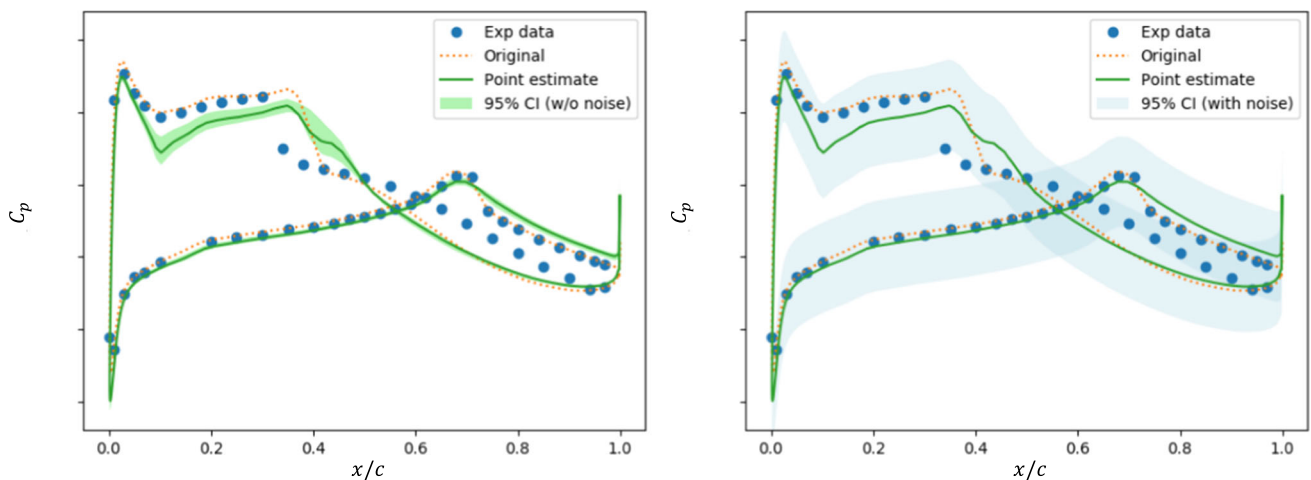


Fig. 16 Predictive distributions at Section *ii* by the empirical Bayes approach to compare the left figure to the result obtained by the full Bayes approach shown in Fig. 14-Section *ii*. Another comparison from the plots in this figure as two kinds of predictive distributions. **a**: $p(y|x, \mathbf{X}, \mathbf{Y}, \hat{\sigma})|_a$ as discussed in the main context, **b** $p(y|x, \mathbf{X}, \mathbf{Y}, \hat{\sigma})$

where $\lambda^{(j)}$ can be a regularization parameter defined as a user input parameter instead of defining $\sigma_{pri}^{(j)}$. M is the dimensionality of the model parameter \mathbf{a} . e.g. $M = 12$ when using the original turbulence model parameter. Equation (31) is a kind of ridge regressions. Therefore, the point-estimate using a Gaussian prior distribution in Sect. 3.1 is a ridge regression itself. It is also well-known that the equation corresponding to Eq. (31) becomes a lasso regression when a Laplace distribution is used for the prior $p(\mathbf{a})$ in Eq. (16).

Our case is to take into account the case where both the model parameter and the noise parameter are learned by data. The error function is represented by the following formulation to execute an optimization in the parameter space of (\mathbf{a}, σ) :

$$\hat{\mathbf{a}}, \hat{\sigma} = \operatorname{argmin}_{\mathbf{a}, \sigma} \left(N \ln(\sigma^2) + \frac{1}{\sigma^2} \sum_{i=1}^N \{y^{(i)} - f(\mathbf{x}^{(i)}, \mathbf{a})\}^2 + \sum_{j=1}^M \left\{ \frac{a^{(j)} - a_{org}^{(j)}}{\sigma_{pri}^{(j)}} \right\}^2 \right). \tag{32}$$

Since the noise parameter is a scalar, ordinary optimization process is expected to work better than the case of Eq. (30) because of the dimensionality of the search space.

6. Generalized perspective of statistical modelling

The statistical modelling described in Sect. 2 can be summarized by a generalized perspective which can be more practical when the probabilistic model that we define is complicated. The process of the modelling is then generalized with assistance of the graphical model (the Bayesian network) as follows:

1. describe a graphical model (a Bayesian network),
2. describe the joint probability distribution of all the stochastic variables as $p(y, \mathbf{Y})$,
3. describe the conditional probability distribution of the prediction output conditioned by the observed variables as $p(y|\mathbf{Y})$ with marginalizing all the latent variables,

where \mathbf{Y} and y specify a dataset and a (arbitrary) prediction of the output, respectively. The graphical model in the step 1 is described by your defined probabilistic model. Figure 17 shows the graphical model of the fully Bayesian approach defined in Sect. 2. Please refer some books [8] to follow the rules of describing graphical models.

In the second step, the joint probability distribution composed of all the stochastic variables $(y, \mathbf{Y}, \mathbf{a}, \sigma)$ according to the graphical model:

$$p(y, \mathbf{Y}, \mathbf{a}, \sigma | \mathbf{x}, \mathbf{X}) = p(y | \mathbf{x}, \mathbf{a}, \sigma) p(\mathbf{Y} | \mathbf{X}, \mathbf{a}, \sigma) p(\mathbf{a}) p(\sigma) \tag{33}$$

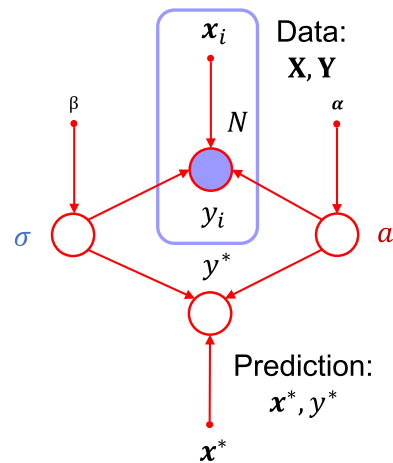


Fig. 17 Graphical model (the Bayesian network) of the probabilistic model showing also a new input \mathbf{x} together with the corresponding model prediction y

Deterministic hyperparameters on the independent priors $p(\mathbf{a})$ and $p(\sigma)$ are omitted for simplicity (e.g. $p(\mathbf{a}|\alpha)$ and $p(\sigma|\beta)$).

The third step is composed by two small steps. As indicated in the graphical model, the stochastic variable \mathbf{Y} is an observed variable. Therefore, the joint distribution by Eq. (33) is modified as conditioned by \mathbf{Y} as follows by using the product rule:

$$\begin{aligned} p(y, \mathbf{a}, \sigma | \mathbf{x}, \mathbf{X}, \mathbf{Y}) &= p(y | \mathbf{x}, \mathbf{a}, \sigma) \underbrace{\frac{p(\mathbf{Y} | \mathbf{X}, \mathbf{a}, \sigma) p(\mathbf{a}) p(\sigma)}{p(\mathbf{Y})}}_{\text{posterior}} \\ &= p(y | \mathbf{x}, \mathbf{a}, \sigma) p(\mathbf{a}, \sigma | \mathbf{X}, \mathbf{Y}). \end{aligned} \tag{34}$$

Here, the Bayes' theorem and the property of *head-to-head* among \mathbf{a}, σ and \mathbf{Y} as an observed variable visualized in the graphical model are used for the deformation of the formula, resulting in a joint probability of the posterior distribution. All the latent variables (here they are \mathbf{a} and σ) in Eq. (34) are finally marginalized (integrated out) by using the sum rule of the probability:

$$\begin{aligned} &\iint p(y, \mathbf{a}, \sigma | \mathbf{x}, \mathbf{X}, \mathbf{Y}) d\mathbf{a} d\sigma \\ &\left(= \iint p(y | \mathbf{x}, \mathbf{a}, \sigma) p(\mathbf{a}, \sigma | \mathbf{X}, \mathbf{Y}) d\mathbf{a} d\sigma \right) \\ &= p(y | \mathbf{x}, \mathbf{X}, \mathbf{Y}). \end{aligned} \tag{35}$$

Equation (35) is exactly equal to Eq. (13).

Thus, by specifying the stochastic variables in the model, and by classifying them into the observed variables and the latent variables, a general process of obtaining the predictive

distribution is introduced in any complex models with visualization assist by graphical modeling and useful properties equipped in them.

References

- Brunton, S.L., Noack, B.R., Koumoutsakos, P.: Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **52**, 477–508 (2020)
- Tinoco, E.N., Brodersen, O.P., Keye, S., et al.: Summary of Data from the Sixth AIAA CFD Drag Prediction Workshop: CRM Cases 2 to 5. *Journal of Aircraft* (2018), vol. 55, no. 4 (2018)
- Da Ronch, A., Panzeri, M., Drofelnik, J., et al.: Sensitivity and calibration of turbulence model in the presence of epistemic uncertainties. *CEAS Aeronaut. J.* **11**, 33–47 (2020)
- He, X., Zhao, F., Vahdati, M.: Machine Learning Uncertainty Quantification of Spalart-Allmaras Turbulence Model for Compressors. GPPS Paper No. 2019-BJ-0050 (2019)
- Singh, A.P., Duraisamy, K.: Using field inversion to quantify functional errors in turbulence closures. *Phys. Fluids* **28**(4), 045110 (2016)
- Parish, E.J., Duraisamy, K.A.: Paradigm for data-driven predictive modeling using field inversion and machine learning. *J. Comput. Phys.* **305**, 758–774 (2016)
- Duraisamy, K., Singh, A.P., Pan, S.: Augmentation of turbulence models using field inversion and machine learning. In: *Proceedings of 55th AIAA Aerospace Sciences Meeting, AIAA 2017–0993*, (2017)
- Bishop, C.M.: *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, New York (2007)
- Robert, C.P., Casella, G.: *Monte Carlo Statistical Methods*. Springer, New York (2004)
- Ray, J., Lefantzi, S., Arunajatesan, S. et al.: Bayesian calibration of a k - ϵ turbulence model for predictive jet-in-crossflow simulations. In: *Proceedings of 44th AIAA Fluid Dynamics Conference, AIAA 2014–2085* (2014)
- Subbian, G., Botelho e Souza, A.C., Radespiel, R., et al.: Calibration of an extended eddy viscosity turbulence model using uncertainty quantification. In: *Proceedings of AIAA Scitech 2020 Forum, AIAA 2020–1031* (2020)
- Edeling, W.N., Cinnella, P., Dwight, R.P., et al.: Bayesian estimates of parameter variability in the k - ϵ turbulence model. *J. Comput. Phys.* **258**, 73–94 (2014)
- Schaefer, J., Hosder, S., West, T., et al.: Uncertainty quantification of turbulence model closure coefficients for transonic wall-bounded flows. *AIAA J.* **55**(1), 195–213 (2017)
- Xiao, H., Cinnella, P.: Quantification of model uncertainty in RANS simulations: a review. *Prog. Aerosp. Sci.* **108**, 1–31 (2019)
- Galle, M., Gerhold, T., Evans, J.: Parallel computation of turbulent flows around complex geometries on hybrid grids with the DLR-TAU code. In: A. Ecer, D.R. Emerson (Eds.), *Proc. 11th Parallel CFD Conf.* (1999)
- Gerhold, T., Hannemann, V., Schwamborn, D.: On the validation of the DLR-TAU code. In: W. Nitsche, H.-J. Heinemann, R. Hilbig (Eds.), *New Results in Numerical and Experimental Fluid Mechanics, Notes on Numerical Fluid Mechanics*, vol. 72, Vieweg, ISBN 3-528-03122-0, pp. 426–433 (1999)
- Schwamborn, D., Gerhold, T., Heinrich, R.: The DLR TAU-code: Recent applications in research and industry, invited lecture. In: Wesseling, P., Oate, E., Priaux, J (Eds.), *Proceedings of the European Conference on Computational Fluid Dynamics (ECOMAS CFD 2006)* (2006)
- Allmaras, S.R., Johnson, F.T., Spalart, P.R.: Modifications and clarifications for the implementation of the Spalart-Allmaras turbulence model. In: *Seventh International Conference on Computational Fluid Dynamics (ICCFD7), ICCFD7-1902* (2012)
- Kennedy, M.C., O'Hagan, A.: Bayesian calibration of computer models. *J. R. Stat. Soc. Ser. B* **63**(3), 425–464 (2001)
- Bekemeyer, P., Matteo, R., Heinrich, R., et al.: Nonlinear unsteady reduced-order modeling for gust-load predictions. *AIAA J.* **57**(5), 1839–1850 (2019)
- Stradtner, M., Liersch, C.M., Bekemeyer, P.: An aerodynamic variable-fidelity modelling framework for a low-observable UCAV. *Aerospace Sci. Technol.* **107**, 106232 (2020)
- Sobol, I.M.: Distribution of points in a cube and approximate evaluation of integrals. *Zh. Vychisl. Mat. Mat. Fiz.* **7**(4), 784–802 (1967)
- Joe, S., Kuo, F.Y.: Remark on Algorithm 659: implementing Sobol's quasirandom sequence generator. *ACM Trans. Math. Softw.* **29**, 49–57 (2003)
- Herman, J., Usher, W.: An open-source Python library for sensitivity analysis. *J. Open Source Softw.* **2**(9), 97 (2017). <https://doi.org/10.21105/joss.00097>
- Iooss, B., Lemaître, P.A.: *Review on Global Sensitivity Analysis Methods. Uncertainty Management in Simulation-Optimization of Complex Systems*, pp. 101–112. Springer, New York (2004)
- Papadimitriou, D.I., Papadimitriou, C.: Bayesian uncertainty quantification of turbulence models based on high-order adjoint. *Comput. Fluids* **120**, 82–97 (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.