Bayesian model-scenario averaged predictions of compressor cascade flows under uncertain turbulence models

M. de Zordo-Banliat, X. Merle, G. Dergham, Paola Cinnella

Bayesian model-scenario averaged predictions of compressor cascade flows under uncertain turbulence models

M. de Zordo-Banliat[a,b], X. Merle[b], G. Dergham[a], P. Cinnella[b]

[a] *Safran Tech, Modelling & Simulation, Rue des Jeunes Bois, Châteaufort, 78114 Magny-Les-Hameaux, France*
[b] *DynFluid Laboratory - Arts et Métiers ParisTech - 151 boulevard de l'Hôpital, 75013 Paris, France*

Abstract

The Reynolds-Averaged Navier-Stokes (RANS) equations represent the computational workhorse for engineering design, despite their numerous flaws. Improving and quantifying the uncertainties associated with RANS models is particularly critical in view of the analysis and optimization of complex turbomachinery flows. In this work, we use Bayesian inference for assimilating data into RANS models for the following purposes: (*i*) updating the model closure coefficients for a class of turbomachinery flows, namely a compressor cascade; (*ii*) quantifying the parametric uncertainty associated with closure coefficients of RANS models and (*iii*) quantifying the uncertainty associated with the model structure and the choice of the calibration dataset based on an ensemble of concurrent models and calibration scenarios. Inference of the coefficients of three widely employed RANS models is carried out from high-fidelity LES data for the NACA65 V103 compressor cascade [1, 2]. Posterior probability distributions of the model coefficients are collected for various calibration scenarios, corresponding to different values of the flow angle at inlet. The Maximum A Posteriori estimates of the coefficients differ from the nominal values and depend on the scenario. A recently proposed Bayesian mixture approach, namely, Bayesian Model-Scenario Averaging (BMSA) [3, 4], is used to build a prediction model than takes into account uncertainties associated with alternative model forms and with sensitivity to the calibration scenario. Stochastic predictions are presented for the turbulent flow around the NACA65 V103 cascade at mildly and severe off-design conditions. The results show that BMSA generally yields more accurate solutions than the baseline RANS models and succeeds well in providing an estimate for the predictive uncertainty intervals, provided that a sufficient diversity of scenarios and models is included in the mixture.

Keywords: Compressor Flow, RANS Models, Bayesian Model Averaging, Uncertainty Quantification.

1. Introduction

The design of modern, highly-loaded axial compressors requires accurate predictions of stagnation pressure losses at the early stages of the design process. Compressor flows are characterized by high relative speeds, leading to the formation of shock waves interacting

Email addresses:
maximilien.de-zordo-banliat@safrangroup.com (M. de Zordo-Banliat), merle.xavier@ensam.eu (X. Merle), gregory.dergham@safrangroup.com (G. Dergham), paola.cinnella@ensam.eu (P. Cinnella)

with the surrounding boundary layers, as well as by the development of secondary flows (such as corner and tip vortices) at blade roots and tips, which interact with the hub and casing and have a strong impact on flow development and on the resulting efficiencies. Additional complexity is introduced by laminar-to-turbulent flow transition induced by high-intensity incoming freestream turbulence (bypass transition) or by flow separation under strongly adverse pressure gradients, for example.

Although flawed by numerous major deficiencies –especially for strongly non-equilibrium and possibly transitional flows like those of interest here–, Reynolds-Averaged Navier–Stokes (RANS) modelling remains the workhorse for turbomachinery design. The main reason is that, despite considerable advances of so-called high-fidelity simulations (namely, Direct Numerical Simulation, DNS, and Large Eddy Simulation, LES) in terms of applicability to more geometrically complex configurations and to higher Reynolds number flows (including turbomachinery flows, see [5]), the computational cost of such simulations remains hundreds of times higher than RANS.

RANS modelling uncertainties can be classified into four levels [6]: 1) uncertainties related to the validity of the averaging process itself; 2) uncertainties in representing the unclosed Reynolds stress tensor as a function of the mean field; 3) uncertainties in the constitutive law used to relate the Reynolds stresses to the mean field; 4) uncertainties in the closure parameters associated with a given model form. A review of turbulence modelling uncertainties and of methodologies for quantifying and reducing such uncertainties can be found in [7]. Several authors have recently investigated the possibility of using high-fidelity simulation data for developing improved RANS models for restricted classes of flows (see [6] for a review). For instance, in [8] the source terms of the Spalart–Allmaras were learnt from data using a single hidden layer neural network, which served as a first feasibility study. In [9], a multiplicative correction function was introduced in the turbulent kinetic energy production term of the k-equation for the $k-\omega$ model. The correction was determined via inverse modelling and served to train a Gaussian process. In [10], a deep neural network was trained to predict the anisotropic part of the Reynolds stress tensor. The network was designed to embed Galilean invariance of the predicted stresses. The above-mentioned data-driven methods all try to develop corrections of some underlying RANS model, in particular by relaxing the well-known "Boussinesq approximation" or "linear eddy viscosity" hypothesis, which is the keystone of a large majority of RANS models used in industrial applications. A particularly promising data-driven approach has been introduced in [11], based on Gene-Expression Programming (GEP). Such an approach leads to the development of Explicit Algebraic Stress Models (EARSM), a class of RANS models using nonlinear constitutive relations for the Reynolds stresses [12], directly from high-fidelity data. These models relax the linear eddy viscosity hypothesis, leading to improved and yet computationally cheap (compared to high-fidelity) RANS models. The GEP approach has been recently used to develop taylor-made EARSM for turbine flows [13]. An alternative approach for EARSM discovery, based on sparse deterministic regression, has been recently proposed in [14].

GEP, as well as other of the above-mentioned

data-driven methods, belong to the class of "non parametric" [7] methods for quantifying RANS modelling inadequacy, which try to formulate a correction to the model form (most often, the constitutive law for the Reynolds stress tensor) based on the observed data. Such non parametric methods have recently attracted considerable interest from the scientific community due to their potential for automatic learning of RANS models from data, but suffer from the following limitations: 1) they tend to lack generality, *i.e.*, they work well for flows similar to those in the training set but can be hardly extrapolated to different flows; 2) they need a significant amount of high-fidelity data (generally costly to obtain and limited to simple, low-Reynolds number configurations) and are not well suited for incomplete, noisy data such as experiments; 3) in most cases, they lead to deterministic predictions and do not provide estimates of confidence intervals due to persisting uncertainties in both model form and closure coefficients. For turbomachinery design (and for engineering design in general), confidence intervals on the predicted quantities of interest (QoI) represent as valuable information as the QoI itself, since they allow estimating uncertainties about the expected system performance early in the design phase. This is why, in this work, we focus instead on "parametric" uncertainty quantification approaches [7]. The latter use some available data for estimating and reducing uncertainties in model closure coefficients, given the model form. A natural framework for parametric approaches is Bayesian inference, whereby the model coefficients are assigned *a priori* probability distributions (based, e.g., on literature data or expert judgement) that are *a posteriori* updated by using data. Since the model coefficients are now represented as probability distributions, the model output is also a random quantity, characterized by a probability distribution. In other terms, the solution is naturally equipped with uncertainty intervals. Parametric approaches can be easily applied to small, noisy datasets and can be successively updated as soon as new or better data become available. Refs [15, 3] used Bayesian inference for calibrating the Spalart–Allmaras and $k-\varepsilon$ models, respectively, by using experimental data for turbulent flat plate boundary layers. Although parametric approaches only infer on model coefficients, they can also be used for estimating, to some extent, model-form uncertainties. One way to do that is to adopt multi-model ensemble techniques, which have been used in a variety of applications, including aerodynamics [16, 17, 18, 19, 20, 21]. Here we focus on the Bayesian Model Averaging (BMA) framework, initially proposed by Draper [22] (see also [23]). A significant extension to BMA is represented by Bayesian Model-Scenario Averaging (BMSA) [21, 3, 4]. Like BMA, BMSA combines the predictions from multiple models, thereby providing a measure for model uncertainty, using posterior distributions of the coefficients inferred from different calibration scenarios. In [3], a BMSA model was constructed by averaging five RANS models calibrated on 14 scenarios, corresponding to turbulent flat plate flows subject to various external pressure gradients. BMSA, calibrated on the scenarios of [3], was successfully applied to a transonic wing configuration in [4].

In the present work we investigate the potential of BMSA for robust predictions of turbomachinery flows under uncertain RANS models, with focus on a compressor cascade. We focus more particularly on the NACA65 V103 compressor cascade, for which high-

fidelity numerical and experimental data are available in the literature. For our study, we select three widely used RANS models, namely, the Spalart–Allmaras [24], Wilcox' $k-\omega$ [25], and Launder–Sharma $k-\varepsilon$ [26] turbulence models. The purpose of the study is manifold: 1) we investigate if BMSA calibrated on elementary external flow configurations like those of [3, 4] may still provide valuable information for the internal flow configuration of interest; 2) we set up a computationally efficient strategy for specifically calibrating BMSA for costly compressor flows; 3) finally, we apply BMSA to the NACA65 V103 compressor flow at operating conditions outside the calibration set, and we assess its capability to provide accurate predictions and the associated uncertainty intervals for new flows. The results are compared to those of BMSA based on the on-the-shelf sets of coefficients [4].

The paper is organized as follows. In Section 2, we recall the Bayesian framework, with special focus on BMSA. In Section 3, we describe the compressor flow configuration and the RANS models. In Section 4 we report BMSA results for the NACA65 V103 cascade at two off-design conditions. Finally, Section 5 summarizes the main findings and draws perspectives for future work.

2. Bayesian framework

In this section, we recall the theoretical framework for Bayesian model calibration and BMSA, following [27].

2.1. Bayesian calibration

Let us consider a physical model of the form:

$$\Delta = M(\theta) \tag{1}$$

with $\Delta = (\Delta_1, ..., \Delta_N)$ a vector of Quantities of Interest (QoI) computed by a model M given a set of parameters θ of dimension P.

In the deterministic framework, the components of θ are perfectly known and have a fixed value. In the Bayesian framework, the unknown parameters vector θ is treated as a random vector, characterised by a joint probability density function (pdf), noted f. Due to the uncertainty on θ, Δ is also a random vector.

The scope of Bayesian inference is to gain new knowledge about θ by constructing an improved representation of its pdf, based on prior knowledge and assimilating the observed data. For that purpose, let us note D the random vector of observed high-fidelity data. Bayes rules states that :

$$f(\theta|D = \overline{D}) = \frac{f(D = \overline{D}|\theta)}{f(D = \overline{D})} f(\theta) \tag{2}$$

Here, $f(\theta)$ is the prior pdf and represents the initial belief about θ, $f(D = \overline{D}|\theta)$ is the likelihood and corresponds to the probability to observe $\overline{D}$, a realisation of the random variable D, if θ is known exactly. The posterior pdf $f(\theta|D = \overline{D})$ represents the updated knowledge of θ given the observed data vector $\overline{D}$, of size N. In practice, calibration compares the model prediction and the observations to extract the pdf of the parameters vector θ that is the most likely to capture the data. In our case, θ is the set of closure parameters associated with a given RANS model.

From Eq. (2), it appears that the posterior distribution is entirely determined by the prior and likelihood function. Following Arnst [28] we use uninformative priors, *i.e.* uniform priors, for each component of θ (supposed independent). As RANS models have been carefully designed, we are confident in assuming that the standard values should be included in the range of the prior. We

therefore choose uniform priors that include standard values (reported in Table 2), as done in [15] [3]. Furthermore, there is no evidence that model predictions would be improved by choosing closure coefficients with significant deviations from the standard values. The prior intervals are therefore chosen to be large enough to allow a good exploration of the parameter space, while avoiding values preventing the CFD solver to converge. Also note that excessively large prior distributions may lead to overfitting problems, resulting in posterior coefficients that fit very well the calibration data, but deteriorate predictions of unobserved quantities of interest.

The likelihood function $f(D = \overline{D}|\theta)$ is a statistical model for observation errors (discrepancies between the data and their true, unobserved, values) and model inadequacies. The latter accounts for the fact that part of the physics is missed by the model due to any approximation introduced in its construction, so that the true phenomenon can never be exactly captured, even with the best possible model coefficients.

In the present calculations, the observation error is modelled as an additive noise and the model inadequacy as a multiplicative term, as also done in [15]. Specifically, the data $\overline{D}$ at a given location x_i are related to the observation error by:

$$\overline{D}(x_i) = \widehat{D}(x_i) + e_i(x_i) \tag{3}$$

with e_i the observation noise at position x_i and $\widehat{D}(x_i)$ the (unobserved) true value of the QoI vector. We choose the components of the observation noise to be independent and normally distributed, with zero mean and a standard deviation equal to 1% of the observed value.

The model-inadequacy η_i is given by:

$$\widehat{D}(x_i) = \eta_i \Delta(x_i, \theta) \tag{4}$$

with $\Delta(x_i, \theta)$ the model output at a point x_i. We choose the model errors to be independent and Gaussian, *i.e.* $\eta_i \sim \mathcal{N}(1, \sigma_\eta^2)$ where σ_η is an additional uncertain hyper-parameter that needs to be calibrated, and therefore is concatenated to the vector of parameters θ. The hyperparameter σ_η is a measure of the magnitude of the model inadequacy and thus can be taken as an indicator of the accuracy of a given model, calibrated for a given scenario. Considering a model error mitigates the influence of overfitting on the calibration, as it relaxes constraints. For more detailed discussion, see [22].

The preceding choices for η_i and e_i lead to a likelihood function of the form:

$$f(\overline{D}|\Delta, \theta) = \frac{1}{\sqrt{(2\pi)^N |K|}} exp\left[-\frac{1}{2}(\overline{D} - \Delta(\theta))^T K^{-1} (\overline{D} - \Delta(\theta))\right] \tag{5}$$

with $K = K_e + K_M$ where K_e is a diagonal matrix representing the observational error vector and $K_M = \sigma_\eta^2 I$ a diagonal matrix reflecting model inadequacy.

For complex models like those of interest in this study, the term $\Delta(\theta)$ cannot be computed analytically, and the posterior distribution for θ must be approximated numerically. Specifically, we use a Markov-Chain Monte-Carlo method to draw sample from the posterior pdf, and namely the Metropolis-Hastings algorithm [29] available in the *pymc*[1] open library. The MCMC sampling is stopped when the following criteria are satisfied: the Geweke z-score [30], the steadiness of the first two moments of the sample, and the auto-correlation of the Markov Chain. For more details concerning such creteria, we refer to [27].

Typically, $O(10^5)$ samples are needed to reach convergence, which is unacceptably high for costly RANS models. To reduce the computational effort to an

[1] https://github.com/pymc-devs/pymc

amenable level, the calibrations presented in the following are based on surrogate models, presented in section 3.4.

2.2. BMSA formulation

In this paper we call scenario, noted S, a specific flow case unambiguously described by a known and deterministic set of parameters (e.g. the geometry of the blade, boundary conditions, Reynolds Number, Mach number...). Now, consider $i = 1, ..., I$ models applicable to a set of $k = 1, ..., K$ scenarios $\mathcal{S} = \{S_1, ..., S_K\}$ for which we have K vectors of observed data $\mathcal{D} = \{\overline{D_1}, ..., \overline{D_K}\}$. Similarly, we call $\mathcal{M} = \{M_1, ..., M_I\}$ the ensemble of all available models. For each model applied to each scenario, we assume that the calibration phase resulted in a posterior for θ, *i.e.*

$$\theta_{i,k} \sim \theta | M = M_i, S = S_k, D = \overline{D_k} \tag{6}$$

Aftewards, let us consider a new scenario S' with no available data and a QoI Δ. Similarly to [22], we use the law of total probabilities to state that :

$$f(\Delta|S', \mathcal{D}, \mathcal{M}, \mathcal{S}) = \sum_{i=1}^{I}\sum_{k=1}^{K} f(\Delta|S', \overline{D_k}, M_i, S_k) p(M_i|\overline{D_k}, S_k) p(S_k) \tag{7}$$

Here, $f(\Delta|S', D = \overline{D_k}, M = M_i, S = S_k)$ represents the distribution of Δ obtained by propagating the posterior distribution $\theta_{i,k}$ for the new scenario S', p is the probability mass function of a discrete random variable and we assumed that D and S are independent, as in [21]. For the sake of conciseness, we drop the clearer but redundant formulation $M = M_i$ or $S = S_k$ to simply write M_i or S_k in the rest of the paper. We also deliberately omit $\mathcal{D}$, $\mathcal{M}$ and $\mathcal{S}$ in $f(\Delta|S')$ for the same reason.

Equation (7) leads to the following expression for the two leading moments of $f(\Delta|S')$:

$$E[\Delta|S'] = \sum_{i=1}^{I}\sum_{k=1}^{K} E\left[\Delta|S', \overline{D_k}, M_i, S_k\right] p(Mi|\overline{D_k}, S_k) p(S_k) \tag{8}$$

$$\begin{aligned} Var[\Delta|S'] = & \underbrace{\sum_{i=1}^{I}\sum_{k=1}^{K} Var\left[\Delta|S', \overline{D_k}, M_i, S_k\right] p(Mi|\overline{D_k}, S_k) p(S_k)}_{\textit{within-model, within-scenario variance}} \\ & + \underbrace{\sum_{i=1}^{I}\sum_{k=1}^{K} \left(E\left[\Delta|S', \overline{D_k}, M_i, S_k\right] - E\left[\Delta|S', \overline{D_k}, \mathcal{M}, S_k\right]\right)^2 p(Mi|\overline{D_k}, S_k) p(S_k)}_{\textit{between-model, within-scenario variance}} \\ & + \underbrace{\sum_{k=1}^{K} \left(E\left[\Delta|S', \overline{D_k}, \mathcal{M}, S_k\right] - E\left[\Delta|S', \mathcal{Z}, \mathcal{M}, \mathcal{S}\right]\right)^2 p(S_k)}_{\textit{between-scenario variance}} \end{aligned} \tag{9}$$

In Eq. (9) the variance $Var[\Delta|S']$ is decomposed in three contributions: the first one is related to the parametric uncertainty and accounts for the dispersion within the samples $\theta_{i,k}$. As such samples have been obtained in the calibration phase for given model and scenario, this first term is called *within-model, within-scenario variance*. The second term represents variance *between model, within scenario* and grows when models give contradicting predictions for the same scenario. The final term, called *between scenario variance*, reflects the fact that using different calibration scenario results in different posteriors for $\theta_{i,k}$ and in different model probabilities $p(M_i|\overline{D_k}, S_k)$. This ultimately leads to different predictions for $\Delta|S'$.

The term

$$E\left[\Delta|S', \overline{D_k}, \mathcal{M}, S_k\right]$$

in Eqs (8) and (9) represents the mean of Δ averaged

over all the models being calibrated on the same scenario. It is computed through :

$$E\left[\Delta|S', \overline{D_k}, \mathcal{M}, S_k\right] = \sum_{i=1}^{I} E\left[\Delta|S', \overline{D_k}, M_i, S_k\right] p(M_i|\overline{D_k}, S_k) p(S_k) \quad (10)$$

The posterior model probability $p(M_i, S_k|\overline{D_k})$ reflects how well the model M_i fits the data $\overline{D_k}$ for the scenario S_k . It can be computed through the Bayes rule:

$$p(M_i|\overline{D_k}, S_k) = \frac{p(\overline{D_k}|M_i, S_k) p(M_i|S_k)}{\sum_{j=1}^{J} p(\overline{D_k}|M_j, S_k) p(M_j|S_k)} \quad (11)$$

where $p(M_i|S_k)$ is a user-defined prior and $p(\overline{D_k}|M_i, S_k)$ is the model evidence :

$$p(\overline{D_k}|M_i, S_k) = \int_{\Theta} f(\overline{D_k}|\theta, M_i, S_k) f(\theta|M_i, S_k) d\theta \quad (12)$$

$p(M_i|S_k)$ is generally chosen equiproportional, i.e. $p(M_i|S_k) = 1/I$.

The BMSA formulation is completed by selecting a prior probability mass function for the scenarios, i.e. an expression for $p(S_k)$. It was shown in [3] that choosing a uniform prior for the scenario mass function brings unnecessary large variance. Following [4] and [27], we choose prior scenario based on model agreement:

$$\begin{cases} p(S = S_k) = \dfrac{\varepsilon_k^{-p}}{\sum_{k=1}^{K} \varepsilon_k^{-p}} \\ \varepsilon_k = \sum_{i=1}^{I} \left\| E\left[\Delta|S', \overline{D_k}, M_i, S_k\right] - E\left[\Delta|S', \overline{D_k}, \mathcal{M}, S_k\right] \right\|_2 \end{cases} \quad (13)$$

with $p = 2$. In this formulation, scenarios for which models give closer predictions are assigned higher probabilities.

3. Case Setup and Methodology

3.1. Compressor cascade configuration and reference data

In the following, BMSA is used to predict a compressor flow configuration. Specifically, we focus on the

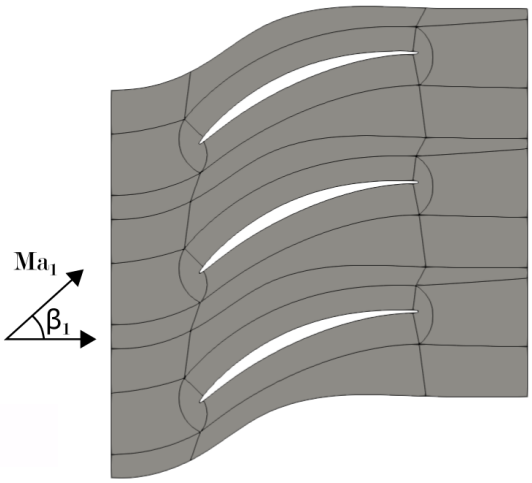


Figure 1: Sketch of the compressor Cascade V 103 adapted from [32]. Ma_1 is the inlet Mach and β_1 is the angle of attack.

NACA65 V103 cascade from Leipold [31], sketched in Figure 1, which is representative of a realistic axial compressor mid-section. For this cascade, the design conditions correspond to an inlet flow angle $\beta_1 = 42°$, an inlet Mach number of 0.67 and a Reynolds number (based on the blade axial chord and the inlet quantities) equal to 450000, respectively. This configuration has been widely studied in the past years [32, 33, 1], and the high-fidelity data available in the literature are suitable for BMSA calibration and assessment. Hereafter we consider in particular the LES data from Leggett [1, 2], who investigated the cascade at four off-design conditions, corresponding to calibration/prediction scenarios in the present Bayesian framework. The scenarios have different values of inlet angle but similar inlet Reynolds and Mach numbers and inlet turbulence intensities. Flow conditions characterizing each scenario are reported in Table 1. Previous study [1] pointed out that RANS models provide rather accurate results for near design conditions, but behave poorly at off-design.

Scenario	S_1	S_2	S_3	S_4
β_1	36.99°	39.97°	44.09°	49.2°
Ma_1	0.654	0.674	0.666	0.65
Re_1	302K	302K	298K	289K
Tu (%)	2.9	3.4	3.4	3.5

Table 1: Flow conditions for various compressor cascade scenarios.

The NACA 65 V103 cascade therefore represents a challenging configuration for assessing the BMSA methodology.

For the purpose of BMSA calibration, we extracted from LES data selected quantities of interest, namely, the tangential velocity and turbulent kinetic energy profiles in the wall-normal direction and total pressure profiles in the wake. The LES data are in good agreement with the experiments of [31]. Tangential velocity profiles at 4 streamwise positions (at $x/l = 0.56, 0.64, 0.76$ and 0.99 on the suction side, l being the chord and $x/l = 0$ the leading edge), and total pressure $(Pt_{inlet} - Pt)/(Pt_{inlet} - P_{inlet})$ profiles at 2 positions downstream of the trailing edge ($x/l = 1.02$ and 1.1) were used for the calibration/assessment of all models considered in the study. For RANS models involving a transport equation for the turbulent kinetic energy (TKE), such as the $k-\omega$ and $k-\varepsilon$ models, we also considered TKE profiles at the same positions on the suction side as the velocity profiles. For the calibrations reported in the following we used data for a small number of observation points along each profile, clustered in the near wall region and toward the wake center. The data are then concatenated to form the vector $\overline{D}$. In total, we used 82 probes for $k-\omega$ and $k-\varepsilon$ models, and 50 for Spalart–Allmaras model. As a general rule, the number of data used in the calibration is a tradeoff between the necessity of informing the model coefficients and computational cost associated with the construction and inversion of the correlation matrices involved in the likelihood function.

3.2. Flow models

The flow around the compressor cascade is modelled by the compressible RANS equations (not reported for the sake of brevity) supplemented with a turbulence model. Since it is not possible to identify *a priori* the "best" turbulence model for predicting an unseen configuration based on pure expert judgement, we adopt a multi-model ensemble constituted of three concurrent turbulence models, briefly described thereafter. The reader may refer to the original references for more details. Only linear eddy viscosity models are considered in the following since, despite the limitations intrinsic to the so-called Boussinesq hypothesis, their are robust and widespread in industrial flow solvers. For such models, *a posteriori* estimates of the closure coefficients and of the posterior model probabilities determined from model calibrations against 14 flat plate flow scenarios have been made available in [3, 4]. These posteriors have been proven useful for prediction on different flows, such as pipes and wings.

3.2.1. Launder–Sharma $k-\varepsilon$

The $k-\varepsilon$ model of Launder and Sharma [26] relies on the solution of transport equations for the turbulent kinetic energy k and the turbulent dissipation ε for computing the eddy viscosity coefficient $\nu_t = C_\mu k/\varepsilon$.The transport equations and the eddy viscosity definition involve six uncertain closure coefficients: C_μ, $C_{\varepsilon 1}$, $C_{\varepsilon 2}$, σ_ε, σ_k and κ, with κ the von Karman constant. These are not all independent since they have to satisfy the fol-

lowing relationships, derived for simple canonical flows [34] (see also [3]):

$$\kappa^2 = \sigma_\varepsilon C_\mu^{1/2}(C_{\varepsilon 2} - C_{\varepsilon 1}) \tag{14}$$

$$\frac{P}{\varepsilon} = \frac{C_{\varepsilon 2} - 1}{C_{\varepsilon 1} - 1} \tag{15}$$

Following [35] we set $\frac{P}{\varepsilon} = 2.09$ in equation (15). By enforcing the preceding relationships, we are finally left with 4 uncertain closure coefficients, namely $C_{\varepsilon 2}$, C_μ, σ_k and κ. The standard values of these coefficients for the Launder–Sharma model are given in Table 2.

3.2.2. Wilcox $k-\omega$ (2006)

The second model is Wilcox' $k-\omega$ model [25], based on transport equations for the turbulent kinetic energy k and the turbulent dissipation rate $\omega = \varepsilon/k$. This model has seven closure coefficients denoted α, β, β^*, σ, σ^*, σ_{do} and κ, whose standard values are given in Table 2. The coefficients must satisfy the relation [25]

$$\alpha = \frac{\beta}{\beta^*} - \frac{\kappa^2}{2\sqrt{\beta^*}} \tag{16}$$

so that only six independent coefficients are left. In the following, α is computed *a posteriori* once the other coefficients have been calibrated from data.

3.2.3. Spalart–Allmaras model

The Spalart–Allmaras model [24] is a single-transport-equation model for a viscosity-like quantity $\tilde{\nu}$, which merges with turbulent viscosity ν_t far from the walls. It involves 8 closure coefficients: C_{b1}, C_{b2}, σ, C_{w1}, C_{w2}, C_{w3}, C_{v1} and κ . Since the coefficient C_{w1} is related to the other coefficients by the relation:

$$C_{w1} = \frac{C_{b1}}{\kappa^2} + \frac{1 + C_{b2}}{\sigma} \tag{17}$$

only 7 independent closure coefficients are left, whose standard values are given in Table 2.

Model	Closure Coefficient	Standard value
$k-\varepsilon$	$C_{\varepsilon 2}$	1.92
	C_μ	0.09
	σ_k	1.0
	κ	0.41
$k-\omega$	κ	0.41
	σ_{do}	0.125
	σ^*	0.6
	σ	0.5
	β^*	0.09
	β	0.0708
Spalart-Allmaras	κ	0.41
	C_{w2}	0.3
	C_{w3}	2.0
	C_{v1}	7.1
	C_{b1}	0.1355
	C_{b2}	0.622
	σ	2/3

Table 2: Standard values of the closure coefficient for the $k-\varepsilon$, $k-\omega$ and Spalart–Allmaras models, according to [26] [25] [24], respectively.

3.3. RANS solver and computational setup

The simulations presented in this study are conducted by using the CFD solver *elsA*, developed by ONERA [36]. We solve the 2D steady compressible RANS equations for perfect Newtonian gases by using a cell-centered finite volume approximation on structured multi-block grids. The upwind scheme of Roe with second-order MUSCL extrapolation is used for approximating the spatial fluxes. For time stepping, we use the first-order backward Euler scheme.

The computational domain contains a single blade profile and periodic boundary conditions are applied at

the upper and lower boundaries to simulate an infinite cascade. The domain extends from 0.4 chords upstream of the leading edge to 0.5 chords downstream the trailing edge. The top and bottom boundaries are separated by a distance equal to the gap between neighboring blades, $t/l = 0.59$, with l the axial chord. In addition to the periodicity conditions at the upper and lower boundaries, non-slip adiabatic boundary condition is applied at the blade wall, and characteristic conditions are imposed at the inlet and outlet boundaries. At the inlet, the total pressure, enthalpy and angle of attack are prescribed; a constant static pressure is enforced at the outlet.The computational grid is composed by 200,000 cells distributed on 12 blocks. The near-wall grid resolution leads to an average height of the first cell closest to the wall (in wall coordinates) such that $y^+ < 1.0$ on both the suction and the pressure side of the blade. For all computations, we assume that the solution has converged to the steady state when the L_2 norm of the residuals is reduced by five order of magnitude with respect to the initial value. The simulations are run in parallel on 12 cores and the typical CPU time for obtaining a converged solution is of the order of 20 minutes. Since a very large number of numerical simulations is required for the calibration of model coefficients using MCMC, the solver output for the observed data is approximated by means of a surrogate model, described in the next section.

3.4. Surrogate modelling

To reduce the number of expensive RANS simulations involved in model calibrations, we approximate the QoIs required in the argument of the likelihood function , (*i.e.* $\Delta(\theta)$) by means of surrogate models based on Gaussian process regression. For that purpose, we use the Gaussian Process Regression module available in $scikit-learn$ [37]. We select a $Matern-3/2$ kernel, whose hyperparameters are determined by optimizing the likelihood. For that purpose, we use the L-BFGS-B [38] optimizer available in the *scipy* library [39]. The initial RANS calculations required as an input to the surrogate model are distributed in the parameter space by Latin Hypercube Sampling (LHS) [40] optimized under the Maximum Projection Design criterion. More precisely, this criterion ensures optimal space filling by maximizing the minimal distance between points of the LHS, for every projection in parameter sub-spaces. We construct a separate surrogate based on 200 RANS samples for each concurrent turbulence model and each calibration scenario in Table 1, for a total of 2400 CFD calculations, run in parallel on a multi-processor computer. This is a considerable computational effort, but it is done one for all prior to the calibration phase. For a given model and a given scenario, the 200 samples are used to build a surrogate for each one of the observed QoIs involved in the likelihood function (namely, velocity, TKE and total pressure values at selected points in the flow field, as discussed in section (3.1)). We verified the accuracy of the surrogate models by Leave-One-Out cross-validation. For each model and scenario, we compute the Q_n^2 criterion for every element Δ_n of $\Delta = (\Delta_1, ..., \Delta_N)$. By definition, Q_n^2 is defined as :

$$Q_n^2 = 1 - \frac{\frac{1}{200}\sum_{i=1}^{200}\left((\Delta_n)_i^{true} - (\Delta_n)_i^{pred}\right)^2}{Var((\Delta_n)^{true})} \qquad (18)$$

For each model and scenario, we present in Table 3 the mean value of the Q_n^2 criterion, averaged on the N surrogate models.

Average Q^2	$k-\varepsilon$	$k-\omega$	Spalart–Allmaras
S_1	0.976	0.991	0.975
S_2	0.967	0.968	0.965
S_3	0.995	0.997	0.970
S_4	0.996	0.985	0.994

Table 3: Average values of Q_n^2 for models and scenarios.

4. Results

In this section, BMSA is used to predict two of the scenarios presented in section 3.1 (namely, S_2 and S_4) . In scenario S_2, the flow remains attached all over the suction side of the blade, whereas flow separation is observed in scenario S_4, which is very different from the other scenarios in the database and represents a challenging configuration for assessing BMSA predictions far outside the training set. The BMSA results reported in the following are based on different ensembles of calibration scenarios. First, a baseline BMSA model, noted BMSA1, is constructed by propagating the *maximum-a-posteriori* (MAP) estimates of model coefficients and the model posterior probabilities of [4]. Although such coefficients were obtained for flat plate flows, we may expect that the thin NACA65 V103 blades can be approximately modelled as flat plates subject to a variable (mostly adverse) pressure gradient. It is then interesting to measure the capability of BMSA to predict the present compressor cascade before having observed any data for this family of configurations. In the following, the flat-plate scenarios are noted S_{XYZW}, with $XYZW$ the four-difit code assigned to the scenarios in [3], to which we refer for more information.

Afterwards, another BMSA model, noted BMSA2, is developed by calibrating the RANS models against compressor configurations. More precisely, we calibrate the models against data available for each of the four compressor scenarios of Sec. 3.1 and we determine the corresponding model evidences. Then, we construct BMSA models of S_2 and S_4 by using the three remaining scenarios. For instance, we use models trained on scenarios S_1, S_3 and S_4 to predict scenario S_2.

Finally, a more general BMSA model, named BMSA3, is constructed by mixing together flat plate scenarios and the S_1, S_2 and S_3 NACA 65 scenarios and applied to the prediction of S_4.

In all cases the smart scenario weighting of Eq. (13) is used to assign *a priori* probabilities to the scenarios involved in the BMSA models.

Specifically, the error term ϵ_k in Eq. (13) is determined by computing the $\|.\|_2$ of local errors on the normalized velocity and total pressure profiles at the streamwise stations of section 3.

4.1. Calibration results

In order to construct a BMSA model specifically trained for compressor configurations, we apply the statistical calibration framework described in section 2.1 to infer on the parameters of the RANS models for the scenarios of section 3.1.

For each model and scenario, we assign to the closure coefficients non-informative uniform marginal distributions priors initially ranging from 10% to 250% of the standard values described in Table 2. For some models and scenarios, these large ranges are eventually restricted to values preventing the CFD solver to converge.

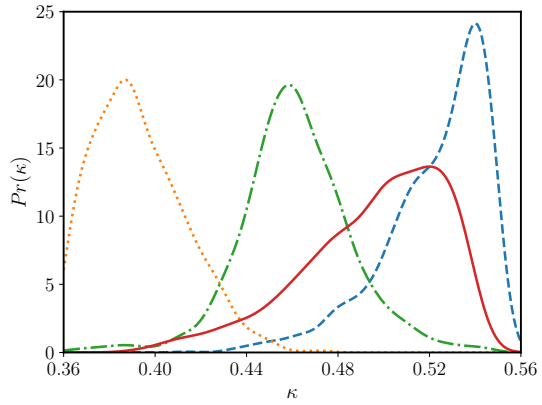
The hyper-parameter σ_η is assigned a uniform prior in the range $[0, 1]$. For illustrative purpose, we present

in Table 4 the prior ranges for the closure coefficients of Spalart–Allmaras model and scenario 3.

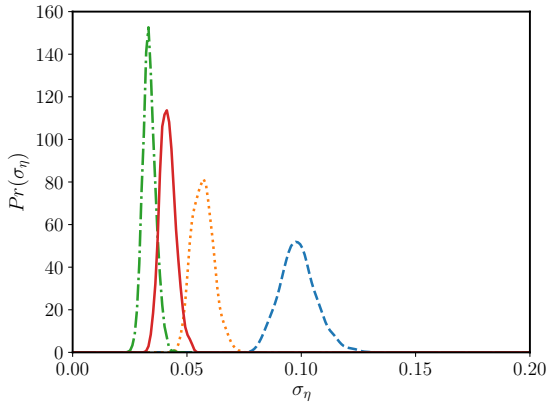
Closure Coefficient	Lower bound	Upper bound
κ	0.36	0.56
C_{w2}	1.0	2.4
C_{w3}	0.1	0.9
C_{v1}	6.5	18.0
C_{b1}	0.5	1.5
C_{b2}	0.06	0.16
σ	0.6	2.0

Table 4: Lower and upper bounds for the prior of the Spalart–Allmaras closure coefficient for the scenario 3.

Figure 2a shows typical calibration results for the κ coefficient of the Spalart–Allmaras model. As also observed in [41], the coefficient is well informed by the data but is highly sensitive to the calibration scenario. It can be noticed that calibration may assign high probabilities to values of κ that are very different from the standard value 0.41, especially for off-design scenarios farthest from the nominal conditions. Similar results are obtained for other coefficients and models, not reported for brevity. In the next Figure 2b, we present calibration results for the hyper-parameter σ_η, which is also well informed by the data. As described in Section 2.1, σ_η can be interpreted as a measure of model accuracy in the calibration scenario. We notice that for all calibrations the mean of the hyper-parameter σ_η is smaller than 10%.



(a) Coefficient κ.



(b) Hyper-parameter σ_η.

Figure 2: Posterior probabilities in case of Spalart–Allmaras model: Scenario 1 (- - -), Scenario 2 (·····), Scenario 3 (-·-·-) and Scenario 4 (——). Priors has been chosen uniform on [0.36 , 0.56] and [0 , 1] respectively.

4.2. BMSA prediction for S_2

For the rest of this paper, we present BMSA prediction in blue color, with first and second standard-deviations in degrading shades of blue. Red color is reserved for the LES reference data from [1]. Black, green and orange colors are used respectively for the baseline $k-\omega$, $k-\epsilon$ and Spalart–Allmaras models, with the nominal closure coefficients of Table 2.

In this section we first report results of BMSA of the

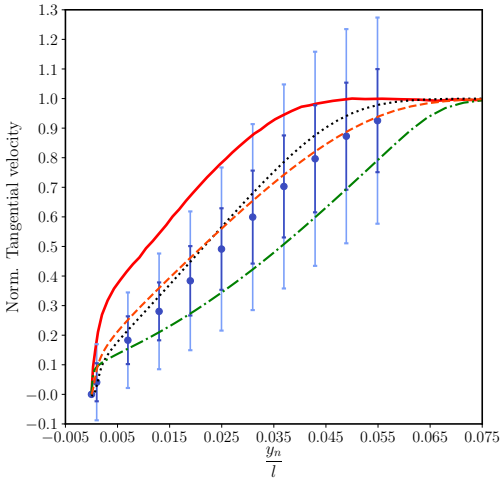
NACA65 V103 configuration at mildly off-design conditions, namely, scenario S_2. The results are discussed for selected velocity and total pressure profiles, representative of typical BMSA predictions. However, similar considerations hold for other locations in the flow.

We present in Figure 3 the tangential velocity profile at $x/l = 0.99$. The x-axis represents the normalized distance to the wall y_n/l, y_n being the distance to the blade. The BMSA results are based on the three sets of scenarios described in the above. Predictions of the baseline RANS models are also reported for comparison. These exhibit significant differences, even for the present attached 2D flow. The $k-\omega$ and Spalart–Allmaras models provide rather close predictions, in better agreement with the LES data than the $k-\epsilon$ model, which performs noticeably worse than the two other for this case.

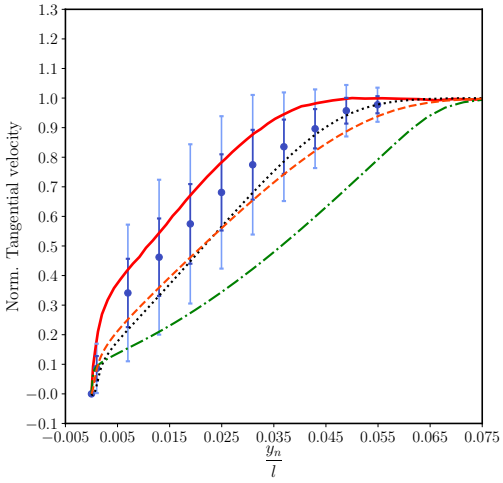
Figure 3a displays results for the BMSA1 model, *i.e.* using on-the-shelf MAP estimates of model coefficients calibrated on flat plates from [4]. The prediction expectancy for this model does not yield better results than the best baseline model but performs much better than the worst one. Moreover, the prediction error bars, corresponding to ± 2 standard deviations, encompass rather well the reference data, except in the region closest to the wall.

In Figure 3b we report the results for BMSA2, calibrated on compressor scenarios $\{S_1, S_3, S_4\}$ and applied to S_2. In this figure, the complete posterior distributions are propagated through the models based on kriging surrogates of the output QoI. Propagation of the full posterior distributions is based on surrogate models for each RANS model in the mixture and each QoI, as discussed in the above. The predictive accuracy of BMSA improves significantly when we consider closeby scenarios for model calibrations. In particular, the mean prediction $E[\Delta|S']$ is significantly better than the best RANS model, and the reference data are now captured within only one standard deviation.

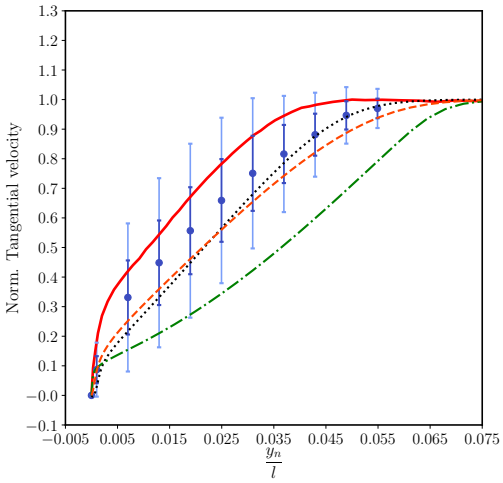
The cost of the kriging surrogate increases with the size of the parameter space but can be roughly estimated to $O(10\,P)$, with P the number of uncertain parameters. A way of reducing the computational overcost associated with the propagation step is to approximate the full posterior distributions with MAP estimates of the coefficients. Using this approximation neglects the posterior parametric uncertainty but involves only $K \times I$ RANS calculations using the MAP estimates for the coefficients. Furthermore, since the parameters are no longer considered as random variables in the propagation step, the BMSA formula can be applied to any QoI in the output solution and not only to selected QoI and flow locations for which a surrogate is available. Figure 3c reports results for BMSA2 based on the propagation of MAP estimates of the coefficients for scenarios $\{S_1, S_3, S_4\}$ through scenario S_2. It turns out that the BMSA prediction using MAP estimates is very close to the one using the full posteriors, both in terms of expectancy and of standard deviation.



(a) MAP estimates calibrated on flat-plate [4].



(b) Complete distributions obtained on $\mathcal{S} = \{S_1, S_3, S_4\}$.



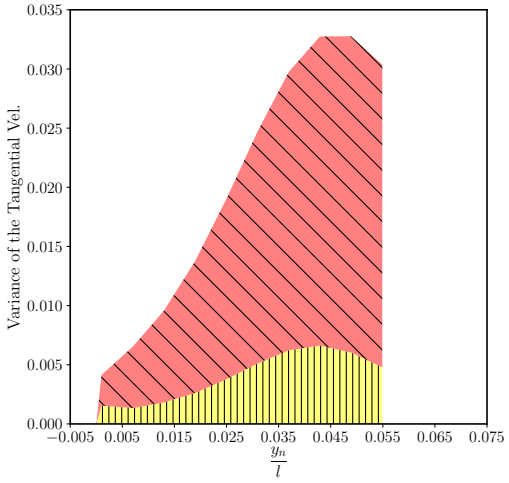
(c) MAP estimates obtained on $\mathcal{S} = \{S_1, S_3, S_4\}$.

Figure 3: Prediction of the normalized tangential velocity profile at $x/l = 0.99$ on the suction side for scenario 2. LES data from *Leggett et al.* [1] (—), $E[\Delta|S'] \pm \sqrt{Var[\Delta|S']}$ (—●—), $E[\Delta|S'] \pm 2\sqrt{Var[\Delta|S']}$ (—●—), Baseline $k-\omega$ (·····), Baseline Spalart–Allmaras (- - -) and Baseline $k-\varepsilon$ (-·-·-).

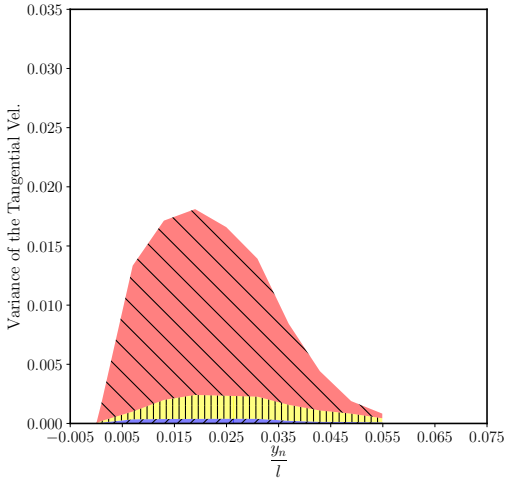
Figure 4 shows the variance decomposition according to equation 9, for each prediction of Figure 3. The total variance for BMSA1 is larger than for the other two cases, due to the greater diversity of scenarios included in the model. The wall-normal locations associated with the largest variance is close to the boundary layer edge in this case, whereas it is located in the near-wall region for BMSA2 predictions, either using full posteriors and MAP estimates. A possible explanation is that the flat-plate scenarios used in BMSA1 mostly differ in the wake region. As a consequence, the calibration mostly adjusts the coefficients to fit velocity profiles in the outer part of the boundary layer. On the contrary, for NACA65 scenarios the near wall region is found to be the most sensitive to the RANS model.

As expected, the *within-model, within scenario* variance is strictly equal to zero for the MAP-based BMSA models. However, inspection of Figure 4b shows that that this term is also very small when propagating the full posteriors. The reason is that the latter are rather peaked (*i.e.* not too far from a Dirac function), since the model coefficients are well informed from the data. The residual parametric uncertainty is then negligible compared to the *between-model, within scenario*. On the other hand, the total variance of the MAP-based BMSA2 model (Figure 4c) is comparable to the one of the full BMSA2 or slightly larger. The discrepancy is due to the different probabilities assigned to the scenarios in the two cases, which are discussed below. Overall, these results further support the choice of MAP estimates for BMSA predictions. To complete the discussion of this figure, we also observe that the larger contribution to the variance is due to the *between scenarios* component. This indicates that the uncertainty associated with the calibration of the closure coefficients

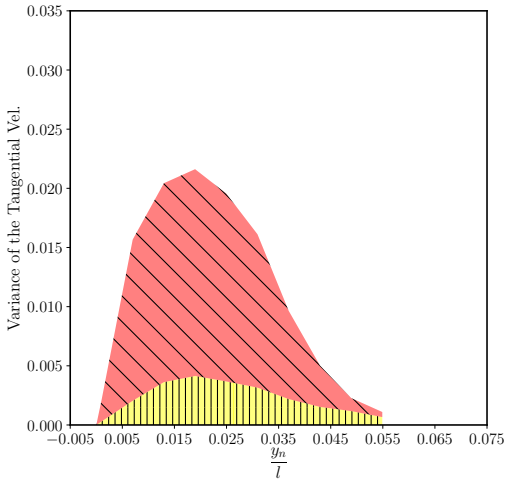
against different datasets is larger than the uncertainty about the more suitable model form.



(a) MAP coefficients calibrated on flat-plate [4].



(b) Complete distributions obtained on $\mathcal{S} = \{S_1, S_3, S_4\}$.

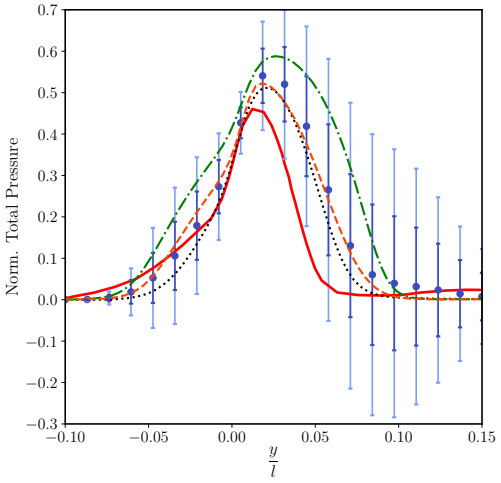


(c) MAP coefficients obtained on $\mathcal{S} = \{S_1, S_3, S_4\}$.

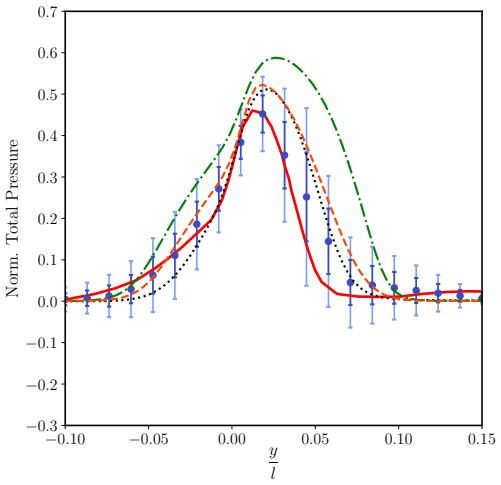
Figure 4: Variance decomposition of prediction for the normalized tangential velocity profile on the suction side at $x/l = 0.99$ for scenario 2. *within-model, within scenario variance*, *between models, within scenario variance* and *between scenario variance*

BMSA predictions of a normalized total pressure profile in the compressor wake are presented in Figure 5. Results are reported again for BMSA1 and for BMSA2 based on full posterior distributions and MAP estimates of the coefficients. The quantity on the $x-axis$ (namely y/l) represents the normalized crossflow position, with the origin aligned with the trailing edge.

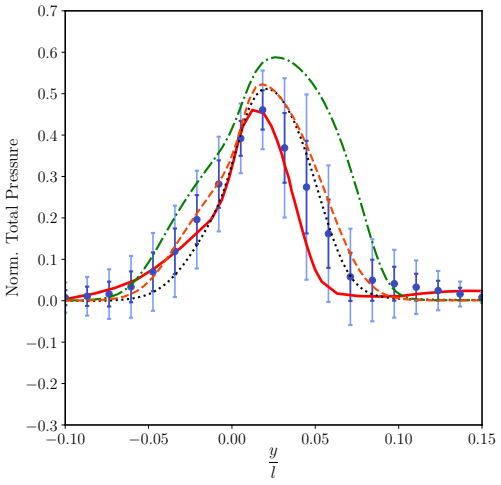
For this QoI, the BMSA models exhibit a behavior similar to the velocity profiles. Specifically, the $k-\epsilon$ baseline model predicts a wake profile farther from the LES reference compared to the two other baselines. Second, the prediction using the BMSA1 model predicts a wake profile relatively close to the best performing nominal RANS model, with LES reference data falling within two standard deviations from the mean prediction. As for the velocity profile, the BMSA2 model provides results in very good agreement with the reference data (Figure 5b), especially for the peak and the left-hand side of the profile. For the right-hand side, corresponding to flow coming from the suction side (characterized by a more challenging physics), BMSA still improves over the nominal models but with higher standard deviations than for the rest of the profile. The results do not change much when using MAP estimates instead of full posteriors. In fact, Figure 6 shows that, once again, the contribution of parametric uncertainty to the total variance is very small, which justifies the use of MAPs.



(a) MAP coefficients calibrated on flat-plate [4].



(b) Complete distributions obtained on $\mathcal{S} = \{S_1, S_3, S_4\}$.



(c) MAP coefficients obtained on $\mathcal{S} = \{S_1, S_3, S_4\}$.

Figure 5: Prediction of the normalized pressure wake profile at $x/l = 1.10$ for scenario 2. LES data from *Leggett et al.* [1] (—), $E[\Delta|S'] \pm \sqrt{Var[\Delta|S']}$ (—●—), $E[\Delta|S'] \pm 2\sqrt{Var[\Delta|S']}$ (—●—),Baseline $k-\omega$ (·····), Baseline Spalart–Allmaras (- - -) and Baseline $k-\varepsilon$ (-·-·-).

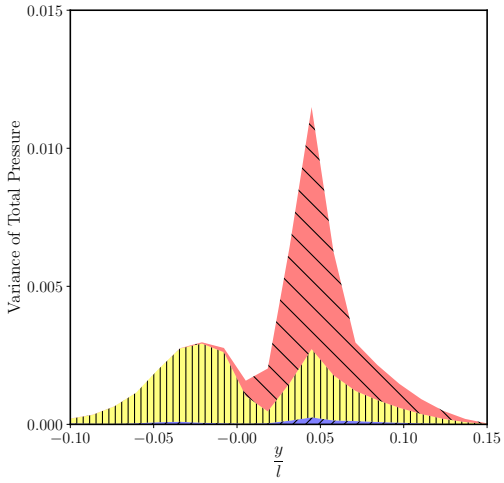
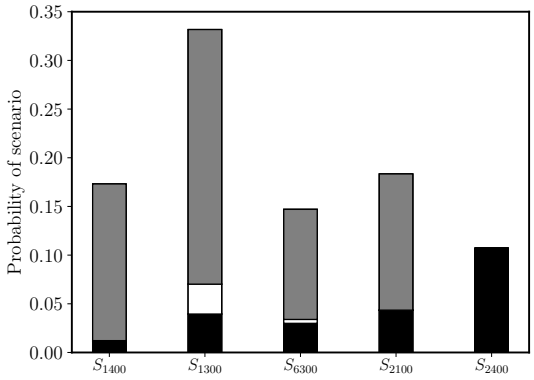
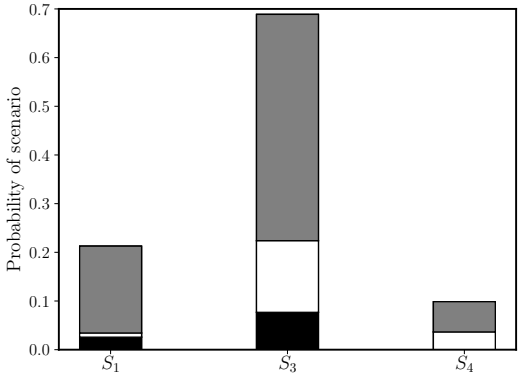


Figure 6: Variance decomposition of prediction for the normalized pressure wake profile at $x/l = 1.10$ for scenario 2 for the complete distributions obtained on $\mathcal{S} = \{S_1, S_3, S_4\}$. *within-model, within scenario variance*, *between models, within scenario variance* and *between scenario variance*

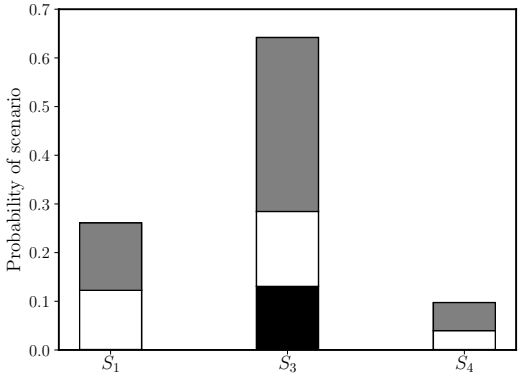
Finally, in Figure 7 we compare the scenario weighting for the various scenarios. Only scenarios that are assigned a probability of 5% or more are shown. For each scenario, we also report the fraction assigned to each RANS model in the mixture, *i.e.* $P(M_i|\overline{D_k}, S_k)\,P(S_k)$ For BMSA1, the calibration scenarios are labelled as in [3]. The scenario weighting criterion automatically assigns higher probabilities to scenarios corresponding to mixed pressure gradients (airfoil-like cases like S_{2100}) or to zero-gradient (S_{1400}) and mildly favorable cases (S_{6300}), which is a bit counter-intuitive. This is probably due to the fact that model agreement for the prediction scenarios is better for regions of favorable pressure gradient (the left part of the blade), leading to lower errors and the higher weighting of such scenarios. For BMSA2, scenario weighting is little affected by the MAP approximation. In both cases, the scenarios are assigned similar probabilities, with scenarios S_1 and S_3 being preferentially weighted with respect to S_4. This can be explained by the proximity of the inlet flow angle of S_2, S_3 and S_1. For the first two scenarios, the flow is qualitatively similar to S_2, which is not the case for S_4, as discussed in the next section. In all BMSA, the Spalart–Allmaras model is generally assigned the highest probability, and $k-\epsilon$ the lowest. Using the MAP approximation changes slightly the model evidences, and subsequently model weighting within each scenario, but the results are overall very close to the BMSA2 using the full posterior distributions.



(a) MAP coefficients calibrated on flat-plate [4]

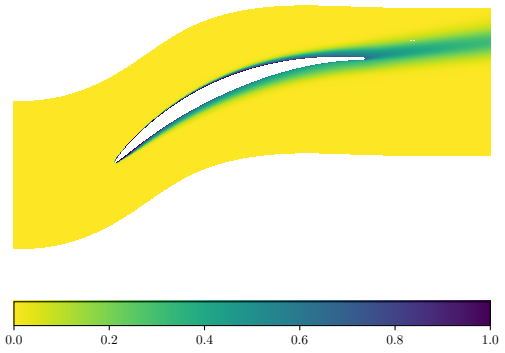


(b) Obtained with complete distributions on $\mathcal{S} = \{S_1, S_3, S_4\}$.

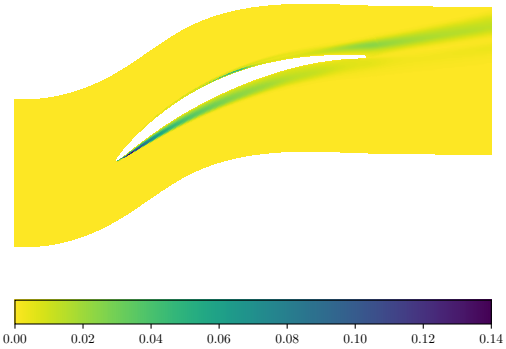


(c) Obtained with MAP estimates of the distributions on $\mathcal{S} = \{S_1, S_3, S_4\}$.

Figure 7: Distribution of $p(S_k)$ and $p(M_i|\overline{D_k}, S_k)$ in case of scenario 2. Only scenarios with probability superior to 5% are shown on Figure 7a. Each bar sums to the probability of the scenario. Each probability of scenario is then decomposed into probabilities of models, given this scenario. $k-\varepsilon$ (■), $k-\omega$ (□) and Spalart–Allmaras (▩).



(a) $E\left[P_t(x_i)|S_2\right]$



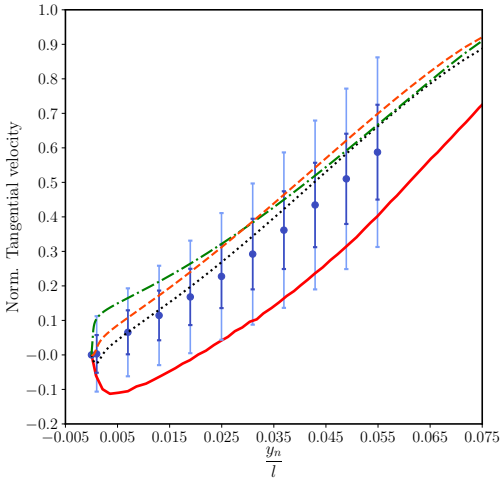
(b) $Var\left[P_t(x_i)|S_2\right]$

Figure 8: 2-D contour of first two moments of the BMSA prediction for normalized total pressure for scenario 2. In this case, we considered MAP estimates on scenarios $\mathcal{S} = \{S_1, S_3, S_4\}$

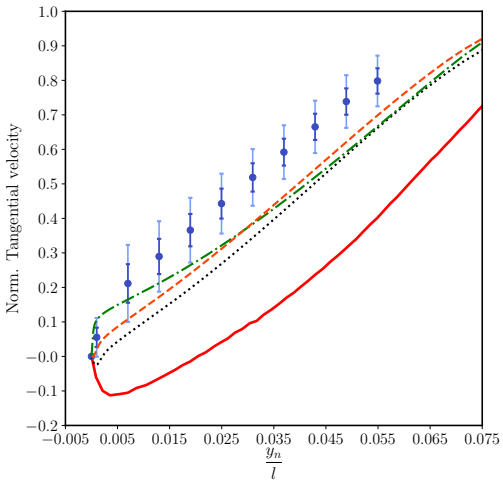
A clear advantage of the MAP approximation of the posteriors, in addition to speeding up the prediction phase, is that it allows constructing a BMSA prediction for any QoI in the flow, compared to the surrogate-based propagation. For instance, Figure 8 shows the iso-contours of the mean and standard deviation of the total pressure field around the blade. The latter provides a global view of flow regions that are the most sensitive to the turbulence model. Based on the preceding discussion, only MAP-based BMSA models are considered in the following.

4.3. BMSA prediction for S_4

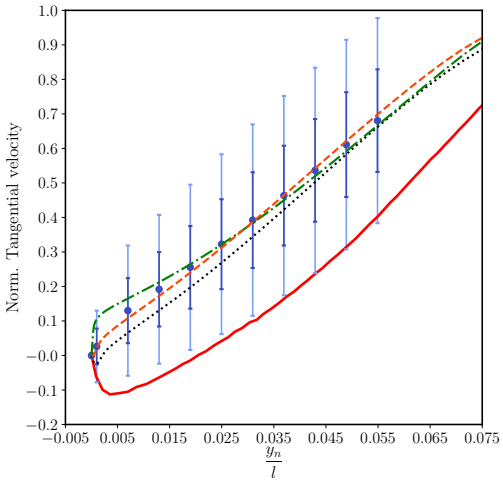
The BMSA mixture models are then applied to the prediction of a more challenging off-design conditions, *i.e.* the separated flow scenario S_4. We show in Figure 9 the predictions for the tangential velocity profile at $x/l = 0.99$ on the suction side in this case. The BMSA1 solution is reported in Figure 9a. The solution clearly under-estimates the size of the backward flow region. Nevertheless, the predicted velocity profile exhibits incipient separation and the $\pm 2\sigma$ error bars encompass reasonably well the reference LES solution. Figure 9b reports results for BMSA2 calibrated on scenarios S_1, S_2 and S_3. In this case, the mean solution compares poorly with the reference LES. Since BMSA2 has been calibrated on attached scenarios, the posterior coefficients tend to provide even fuller velocity profiles than the baseline models, which already fail to predict flow separation, except for the baseline $k-\omega$ that under-estimates the size of the reversed flow. We also observe that, in this case, the error bars are small and do not encompass the reference data. This is due to the fact that the models in the mixture strongly agree on the wrong solution. This result shows the importance of including sufficientlu diverse scenarios in BMSA models. In the present BMSA, predictions are based on models with similar characteristics (linear eddy viscosity), furthermore calibrated on similar attached flow scenarios. As a consequence, the resulting BMSA model is very good at predicting flow scenarios similar to the calibration ones but generalizes badly to a different flow, leading to less accurate results than BMSA1. In Figure 9c we present results for BMSA3, which aggregates together the flat plate scenarios and the NACA65 scenarios. Increasing the diversity of scenarios in the model mixture has a beneficial effect on the solution, which is not worst than the baseline RANS models in the average, but provides an estimate for the error bars. The reference data are captured within approximately two standard deviations from the average.



(a) MAP estimates calibrated on flat-plate [4].



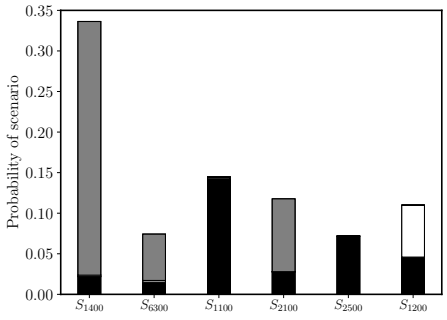
(b) MAP estimates obtained on $\mathcal{S} = \{S_1, S_2, S_3\}$.



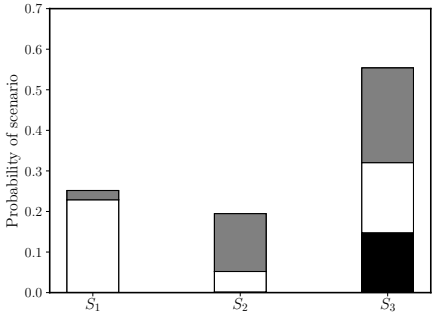
(c) MAP estimates obtained on $\mathcal{S} = \{S_{1400}, ...S_{2134}, S_1, S_2, S_3\}$.

Figure 9: Prediction of the tangential velocity profile at $x/l = 0.99$ on the suction side for scenario 4. LES data from *Leggett et al.* [1] (——), $E[\Delta|S'] \pm \sqrt{Var[\Delta|S']}$ (—●—), $E[\Delta|S'] \pm 2\sqrt{Var[\Delta|S']}$ (—●—), Baseline $k-\omega$ (·····), Baseline Spalart–Allmaras (- - -) and Baseline $k-\varepsilon$ (-·-·-).

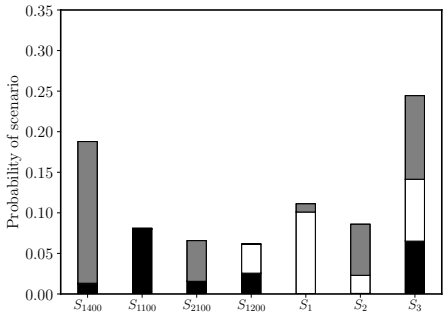
The scenario probabilities are reported in Figure 10 for the three BMSA. Once again we focus only on scenarios with a probability of 5% or higher. For BMSA1 (10a), the most influential scenario is S_{1400}, *i.e.* the zero pressure gradient flat plate, probably due to strong model agreement in the upstream portion of the flow. Interestingly, the BMSA now also assigns significant weights to S_{1100} and S_{2500}, characterized by mildly adverse pressure gradients, and S_{1200} which is representative of a "diverging channel, with eventual separation". Such scenarios were not assigned any significant probability in the S_2 solution. For BMSA2, the highest probability is assigned to S_3, followed by S_1 and finally S_2. This shows that the scenario weighting criterion tends to promote scenarios with inlet angles closest to the one of the prediction scenario. Finally, BMSA3 (Figure 10c) assigns the higher weights to the mixed and adverse pressure gradient scenarios from BMSA1 and to the NACA65 scenarios.



(a) MAP coefficients calibrated on flat-plate [4].



(b) MAP estimates obtained on $\mathcal{S} = \{S_1, S_2, S_3\}$.



(c) MAP estimates obtained on $\mathcal{S} = \{S_{1400}, ...S_{2134}, S_1, S_2, S_3\}$.

Figure 10: Distribution of $p(S_k)$ and $p(M_i|\overline{D_k}, S_k)$ in case of scenario 4. Only scenarios with probability superior to 5% are shown on Figures 10a and 10c. Each bar sums to the probability of the scenario. Each probability of scenario is then decomposed into probabilities of models, given this scenario. $k-\varepsilon$ (■), $k-\omega$ (□) and Spalart–Allmaras (▒).

	Scenario 2		Scenario 4	
	$U_{t,RMS}$	$P_{t,RMS}$	$U_{t,RMS}$	$P_{t,RMS}$
Baseline	0.728	0.492	0.797	0.345
Flat plate	0.710	0.448	0.567	0.232
NACA65 configuration	0.275	0.183	1.199	0.554
NACA65 and flat plate	0.561	0.356	0.843	0.372

Table 5: Root-Mean Square values for the baseline models (averaged value for the 3 considered models), BMSA with models calibrated on flat plates [4], BMSA with models calibrated on NACA65 configurations and BMSA with models calibrated on NACA65 and flat plates together.

5. Conclusions

A recently developed Bayesian framework is assessed for the quantification and reduction of modelling uncertainties in RANS-based simulations of turbomachinery flows. In this framework, modelling uncertainties are treated in terms of probabilities. Specifically, the closure coefficients associated with RANS models are treated as random variables, which are assigned an a priori probability distribution based on their nominal values and expert judgement. Bayesian inference from observed data for selected Quantities of Interest (QoI) is used to reduce the uncertainty ranges of the coefficients, leading to a posteriori distributions. The latter can be propagated through the model by means of an Uncertainty Quantification (UQ) method to obtain predictions with quantified uncertainty of a new flow. Additionally, the proposed framework leverages information from a set of concurrent RANS models and a set of concurrent calibration scenarios to build a mixture model based on Bayesian-Model-Scenario-Averaging (BMSA). BMSA

allows to account to some extent for uncertainties associated with the mathematical form of the RANS model and for uncertainties associated with the choice of flow scenarios for model calibration in view of the stochastic prediction of a new flow not included in the calibration set.

BMSA models were constructed by averaging three linear-eddy viscosity models widely used for industrial applications, namely, Spalart–Allmaras, Wilcox' $k-\omega$ and Launder–Sharma $k-\varepsilon$. A baseline mixture model, named BMSA1, was constructed by using on-the-shelf sets of model coefficients calibrated for fourteen turbulent flat-plate flow scenarios corresponding to different external pressure gradients [3, 4]. A second model, named BMSA2, was specifically tailored for the targeted flow configuration, *i.e.* the compressor cascade NACA65 V103. In this case, each RANS model in the mixture was calibrated against reference LES data [1, 2] available for 3 off-design scenarios and validated against data available for a fourth scenario, not included in the calibration set. The main parameter differentiating the scenarios is the flow inlet angle. The resulting posterior distributions of the model coefficients assign high probability to radically different values of the coefficients than the nominal values. The results show that, even if BMSA1 was not calibrated for the flows of interest, the results obtained for a mildly off-design and a highly off-design scenario are globally not worst than the nominal models and the estimated error bars encompass rather well the reference solution. On the other hand, the compressor-specific BMSA2 model significantly improves the predictions compared to the baseline RANS models when it is used to predict scenario characterized by an intermediary inlet angle with respect to those included in the BMSA. Additionally, the predicted error bars encompass the reference data. However, this strategy may leads to overfitting problems. When applied to a scenario with operating conditions leading to radically different flow features compared to the training scenarios, BMSA provides less accurate predictions than the baseline models. In addition, the error bars are strongly underestimated due to the insufficient diversity of models and scenarios included in the mixture.

Since it is difficult to select a priori the most suitable scenarios to be included in the BMSA based on pure expert judgement (for instance, one could argued that flat plate scenarios are *a priori* less suitable than the NACA65 scenarios to predict another NACA65 flow condition), it is very important to include in the mixture sufficiently diverse flow scenarios and RANS model to mitigate overfitting and avoid underestimation of variance. For instance, predictions of the strongly off-design scenario based on a mixture of the flat plat and NACA65 scenarios preserved or improved the average prediction with respect to the baseline RANS models and delivered sufficiently large error bars to encompass the reference data.

A serious limitation to the number of models and scenarios in a BMSA is the computational cost of the stochastic prediction. In fact BMSA combines stochastic predictions of a new flow scenarios from K several models using posterior pdf of the closure coefficients calibrated for I flow scenarios. Each stochastic prediction involves an UQ calculation, corresponding to a high number ($O(100)$ or more, according to the UQ method in use and to the number of uncertain coefficients), leading to an unacceptably high number of costly deterministic RANS simulations, $O(100\,K{\times}I)$. A first method for drastically reducing the required number of determinis-

tic simulations, first proposed in [4] and further assessed in this work, is to approximate the posterior pdf by Dirac pdf based on Maximum A Posteriori (MAP) estimates of the closure coefficients. This approximation is shown to affect weakly the quality of the BMSA predictions, both in terms of mean and variance, while reducing the number of deterministic simulations to only $K \times I$, *i.e.* nine deterministic simulations in the present application. All the required simulations are independent and can be run in parallel and the BMSA intervenes as a post-processing step. Since MAP-based BMSA does not rely on any surrogate model for the UQ propagation step, it can be use to extract potentially any QoI at any point in the flow, provided that such QoI is computable with the baseline models in the mixture (for instance, a BMSA prediction of a QoI like the turbulent kinetic energy k or the turbulent dissipation ε can be obtained only if all the models in the mixture are at least two-transport equation models).

Further reduction of the computational cost can be achieved by using alternative criteria to assign weights to the BMSA scenarios. In this work, we used a criterion based on model agreement in the prediction scenario derived in [3, 4]. This criterion has proved to be effective in assigning high weights to scenarios in the mixture that are more similar to the prediction case. However, this criterion requires computing the new flow with the K models using the coefficient from all of the I scenarios, even when many of them are in the end assigned a very low probability. Alternative criteria have been proposed in the literature (e.g. [27]) that allow selecting the more suitable scenarios *a priori*, thus excluding from the beginning scenarios that are affected a probability below a given threshold and finally reducing the number of deterministic calculations required for the prediction of a new flow. The development and assessment of smarter and computationally efficient scenario-selection criteria will make the object of further research.

Acknowledgements

The authors acknowledge the french National Agency for Research and Technology (ANRT) for providing support through the CIFRE PhD grant number 2018-1370.

References

[1] J. Leggett, S. Priebe, R. Sandberg, V. Michelassi, and A. Shabbir. Detailed investigation of rans and les predictions of loss generation in an axial compressor cascade at off design incidences. In *ASME Turbo Expo 2016: Turbomachinery Technical Conference and Exposition*. American Society of Mechanical Engineers Digital Collection, 2016.

[2] J. Leggett. *Detailed investigation of loss prediction of an axial compressor cascade at off-design conditions in the presence of incident free-stream disturbances using large eddy simulations*. PhD thesis, University of Southampton, 2018.

[3] WN Edeling, P. Cinnella, and R. Dwight. Predictive rans simulations via bayesian model-scenario averaging. *Journal of Computational Physics*, 275:65–91, 2014.

[4] W. Edeling, M. Schmelzer, R. P Dwight, and P. Cinnella. Bayesian predictions of reynolds-averaged navier–stokes uncertainties using maximum a posteriori estimates. *AIAA Journal*, 56(5):2018–2029, 2018.

[5] R. Sandberg and V. Michelassi. The current state of high-fidelity simulations for main gas path turbomachinery components and their industrial impact. *Flow, Turbulence and Combustion*, 102(2), 2019.

[6] K. Duraisamy, G. Iaccarino, and H. Xiao. Turbulence modeling in the age of data. *Annual Review of Fluid Mechanics*, 51, 2019. doi: 10.1146/annurev-fluid-010518-040547.

[7] X. Heng and P. Cinnella. Quantification of model uncertainty in rans simulations: A review. *Progress in Aerospace Sciences*, 108:1–31, 2019.

[8] B. Tracey, K. Duraisamy, and J.J. Alonso. Application of supervised learning to quantify uncertainties in turbulence and combustion modeling. In *51st AIAA Aerospace Sciences Meeting*, 2013. Dallas, TX, paper 2013-0259.

[9] E.J. Parish and K. Duraisamy. A paradigm for data-driven predictive modeling using field inversion and machine learning. *Journal of Computational Physics*, 305:758–774, 2016.

[10] J. Ling, A. Kurzawski, and J. Templeton. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, 807:155–166, 2016.

[11] J. Weatheritt and R. Sandberg. A novel evolutionary algorithm applied to algebraic modifications of the RANS stress–strain relationship. *Journal of Computational Physics*, 325:22–37, 2016.

[12] S.B. Pope. A more general effective-viscosity hypothesis. *Journal of Fluid Mechanics*, 72:331—340, 1975.

[13] H. Akolekar, R. Sandberg, N. Hutchins, V. Michelassi, and G. Laskowski. Machine-learnt turbulence closures for low pressure turbines with unsteady inflow conditions. *Journal of Turbomachinery*, 05 2019.

[14] M. Schmelzer, R. P. Dwight, and P. Cinnella. Machine learning of algebraic stress models using deterministic symbolic regression, 2019.

[15] S. Cheung, T. A. Oliver, E. E. Prudencio, and R. Prudhomme, S.and D. Moser. Bayesian uncertainty analysis with applications to turbulence modeling. *Reliability Engineering & System Safety - RELIAB ENG SYST SAFETY*, 96:1137–1149, 09 2011.

[16] Svetlana V Poroseva, M Yousuff Hussaini, and Stephen L Woodruff. Improving the predictive capability of turbulence models using evidence theory. *AIAA Journal*, 44(6):1220–1228, 2006.

[17] P.D. Meyer, M. Ye, M. L. Rockhold, S. P Neuman, and K. J. Cantrell. Combined estimation of hydrogeologic conceptual model, parameter, and scenario uncertainty with application to uranium transport at the hanford site 300 area. Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2007.

[18] Q Duan, N K Ajami, X Gao, and S Sorooshian. Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources*, 30(5):1371–1386, 2007.

[19] C Tebaldi and R Knutti. The use of the multimodel ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1857):2053–2075, 2007.

[20] T Diomede, S Davolio, C Marsigli, MM Miglietta, A Moscatello, P Papetti, T Paccagnella, A Buzzi, and P Malguzzi. Discharge prediction based on

multi-model precipitation forecasts. *Meteorology and Atmospheric Physics*, 101(3-4):245–265, 2008.

[21] Rojas R., Kahunde S., Peeters L., Batelaan O., Feyen L., and Dassargues A. Application of a multimodel approach to account for conceptual model and scenario uncertainties in groundwater modelling. *Journal of Hydrology*, 394(3-4):416–435, 2010.

[22] D. Draper. Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):45–70, 1995.

[23] J A Hoeting, D Madigan, A E Raftery, and C T Volinsky. Bayesian model averaging: a tutorial. *Statistical science*, 14(4):382–401, 1999.

[24] P. R. Spalart and S. R. Allmaras. One-equation turbulence model for aerodynamic flows. *Recherche aerospatiale*, (1):5–21, 1994.

[25] David C Wilcox. *Turbulence modeling for CFD*. DCW Industries , La Canada, CA, 2006.

[26] B. E. Launder and BI Sharma. Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc. *Letters in heat and mass transfer*, 1(2):131–137, 1974.

[27] X. Merle and P. Cinnella. Robust prediction of dense gas flows under uncertain thermodynamic models. *Reliability Engineering & System Safety*, 183:400–421, 2019.

[28] M. Arnst, R. Ghanem, and C. Soize. Identification of bayesian posteriors for coefficientsof chaos expansions. *Journal of Computational Physics*, 229(9):3134–3154, 2010.

[29] W K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[30] J. Geweke et al. *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, volume 196. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, 1991.

[31] Leipold R., Boese M., and Fottner L. The influence of technical surface roughness caused by precision forging on the flow around a highly loaded compressor cascade. *Journal of turbomachinery*, 122(3):416–424, 2000.

[32] L. Hilgenfeld and M. Pfitzner. Unsteady boundary layer development due to wake passing effects on a highly loaded linear compressor cascade. *Journal of Turbomachinery*, 126(4):493–500, 2004.

[33] T. A Zaki, J.G Wissink, and P. A Rodi, W.and Durbin. Direct numerical simulations of transition in a compressor cascade: the influence of free-stream turbulence. *Journal of Fluid Mechanics*, 665:57–98, 2010.

[34] Stephen B Pope. Turbulent flows, 2001.

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

[36] L. Cambier, S. Heib, and S. Plot. The onera elsa cfd software: input from research and feedback from industry. *Mechanics & Industry*, 14(3):159–174, 2013.

[37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine*

Learning Research, 12:2825–2830, 2011.

[38] C. Zhu, R.H Byrd, P. Lu, and J. Nocedal. Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)*, 23(4):550–560, 1997.

[39] E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python, 2001–.

[40] V R. Joseph, E. Gul, and S. Ba. Maximum projection designs for computer experiments. *Biometrika*, 102(2):371–380, 2015.

[41] WN Edeling, Pasquale Cinnella, Richard P Dwight, and Hester Bijl. Bayesian estimates of parameter variability in the k–ε turbulence model. *Journal of Computational Physics*, 258:73–94, 2014.