

Evaluation of physics constrained data-driven methods for turbulence model uncertainty quantification

Marcel Matha*, Karsten Kucharczyk and Christian Morsbach

German Aerospace Center (DLR), Linder Höhe, 51147 Cologne, Germany

Abstract

In order to achieve a virtual certification process and robust designs for turbomachinery, the uncertainty bounds for Computational Fluid Dynamics have to be known. The formulation of turbulence closure models implies a major source of the overall uncertainty of Reynolds-averaged Navier-Stokes simulations. We discuss the common practice of applying a physics constrained eigenspace perturbation of the Reynolds stress tensor in order to account for the model form uncertainty of turbulence models. Since the basic methodology often leads to overly generous uncertainty estimates, we extend a recent approach of adding a machine learning strategy. The application of a data-driven method is motivated by striving for the detection of flow regions, which are prone to suffer from a lack of turbulence model prediction accuracy. In this way any user input related to choosing the degree of uncertainty is supposed to become obsolete. This work especially investigates an approach, which tries to determine an a priori estimation of prediction confidence, when there is no accurate data available to judge the prediction. The flow around the NACA 4412 airfoil at near-stall conditions demonstrates the successful application of the data-driven eigenspace perturbation framework. Furthermore, we especially highlight the objectives and limitations of the underlying methodology.

Keywords: uncertainty quantification, turbulence models, RANS, machine learning, data-driven, random forest regression

*Corresponding author

Email address: `marcel.matha@dlr.de` (Marcel Matha)

1. Introduction

In previous times, engineering design applications tried to account for a variety of uncertainties in Computational Fluid Dynamics (CFD) simulations by applying factors of safety, margins of safety, levels of redundancy, etc. Such heuristic methods need adjustments and recalibration for each new configuration. They are also highly empirical especially when applied to innovative designs or new flow configurations. Nowadays methodologies such as robust design or reliability based design have the possibility to preempt such methods. In recent years, the interest in uncertainty quantification (UQ), leading to more reliable simulation based designs, has grown significantly [1, 2, 3, 4].

As a compromise between computational time and accuracy, Reynolds-averaged Navier-Stokes (RANS) based turbulence modeling is still the prevailing tool in industrial design of turbomachinery, since the replacement of RANS by scale-resolving simulations, e.g. Direct Numerical Simulations (DNS) or Large Eddy Simulations (LES) cannot be expected for design optimization simulations in the near future. The derivation of the RANS equations reveals an unclosed term, called the Reynolds stress tensor. This tensor has to be approximated in CFD simulations by applying turbulence models. The prediction quality of the simulation is highly dependent on the accuracy of the turbulence models. However, many RANS-based models suffer from the inability to replicate fundamental turbulent processes. Throughout this paper, we consider linear eddy viscosity models (LEVM), which are widely used for complex engineering flows, by referring to RANS turbulence models. Due to simplifying assumptions used in model formulation, the turbulence model is one of the main limitations in striving for the next generation of reliable, efficient and environmentally friendly designs. These simplifications are the result of data observation, physical intuition, engineering and computational pragmatism, leading to a significant degree of epistemic uncertainty. Accounting for the uncertainties, which arise due to the structural form of the turbulence model in RANS simulation is known to be

the "greatest challenge" in CFD [5]. Nevertheless, these epistemic uncertainties could, in principle, be reduced, by increasing knowledge about turbulent processes, resulting in developing advanced models. This is contrary to aleatory uncertainties, e.g. manufacturing tolerances or operating conditions, which cannot be reduced and are not considered in the current work. Different approaches try to account for the uncertainty of the turbulence model at different modelling levels [6]. Generally, one distinguishes between parametric and non-parametric approaches. While the parametric uncertainties arise from the chosen closure coefficients and their calibration process, non-parametric methodologies directly investigate the uncertainties on modeled terms in the transport equations of the turbulence model. It is expected, that the possible solution space, with respect to the uncertainty of the turbulence model, is larger for non-parametric approaches [4].

Iaccarino and co-workers proposed an eigenspace perturbation framework, which is based on the inability of common LEVM to deal with Reynolds stress tensor anisotropy [7, 8]. This methodology belongs to the non-parametric approaches, since it tries to account for the uncertainty due to the closure model form itself. The physical rationale of the eigenspace perturbation framework is discussed in depth by Mishra and Iaccarino [9]. It enables a designer to optimize components towards an optimum design with less sensitivity to uncertainty. Successful application of design under uncertainty using the perturbation framework was already presented by Mishra et al. [10]. Moreover, this methodology was applied in multiple engineering applications such as aerospace design (aircraft nozzle [8], turbomachinery [11, 12], entire aircraft [13]), civil structural design [14] and even wind farm design [15, 16].

The increasing availability of high-fidelity simulations (such as LES and DNS) in combination with the emergence of machine learning strategies guided the path towards data-driven approaches also for the RANS turbulence modelling community [6]. Heyse et al. [17] enhanced the uncertainty estimation based on the eigenspace perturbation approach by adding a data-driven method leading to less conservative uncertainty estimates. The machine learning strategy

61 should identify flow regions, which are prone to show reduced turbulence model
62 prediction accuracy. From our point of view, their investigations suffered under
63 limited availability of data, most notably with respect to judging an appropriate
64 application of a machine learning model. We consolidate the usage of a random
65 forest regression model by investigating the credibility explicitly in the present
66 work.

67 In the current work, we present the functionality of the eigenspace perturbation
68 methodology, its data-driven extension and its implementation within DLR's
69 CFD solver suite TRACE in detail. TRACE is being developed by the Insti-
70 tute of Propulsion Technology with focus on turbomachinery flows and offers a
71 parallelized, multi-block CFD solver for the compressible RANS equations [18].
72 In order to obtain a trustworthy quantification of uncertainties for future de-
73 sign application in turbomachinery flows, we investigate our implementation of
74 the data-driven Reynolds stress tensor perturbation framework for flow config-
75 urations featuring adverse pressure gradient, flow separation and reattachment
76 with TRACE. This procedure is reasonable, as the prediction accuracy of RANS
77 turbulence models is significantly reduced in the presence of these complex flow
78 phenomena. In order to further advance the data-driven prediction capabili-
79 ties, additional data sets obtained by scale-resolving simulations of relevant test
80 cases are used to train and validate a machine learning model. These test cases
81 include complex flow physics such as adverse pressure gradient, separation and
82 reattachment. In this work, we verify the application of a trained machine learn-
83 ing model in detail. A methodology to quantify an a priori estimate of prediction
84 confidence is particularly studied as well. Finally, the data-driven perturbation
85 approach to estimate the epistemic uncertainty of turbulence models is applied
86 for the flow around the NACA 4412 airfoil at near-stall conditions featuring a
87 separation zone on the suction surface. By analyzing and comparing the un-
88 certainty estimates for certain quantities of interest (QoI) with respect to the
89 usage of the data-free and data-driven strategy, we analyze the initial intention
90 of the entire Reynolds stress tensor perturbation framework and its capabilities
91 consequently. From our point of view, this kind of subsumption was missing in

92 literature oftentimes.

93 **2. Details of the eigenspace perturbation framework**

94 *2.1. Motivation and goal*

95 RANS turbulence models are utilized in order to approximate the Reynolds
96 stress tensor $\tau_{ij} = \overline{u'_i u'_j}$ in terms of mean flow quantities, e.g. $\bar{u}_i = U_i =$
97 $u_i - u'_i$. As already described, the formulation of turbulence models brings
98 along certain assumptions. Even state-of-the-art LEVM rely for example on
99 the eddy viscosity hypothesis, also known as the Boussinesq assumption, and
100 the gradient diffusion hypothesis. This leads to the inability to account for
101 rotational effects, secondary flow, swirl and streamline curvature [19, 20, 21].
102 Besides, limited success in cases with separation and reattachment have been
103 also reported [22]. The motivation for injecting perturbations to the eigenspace
104 of the Reynolds stress tensor is the inability of LEVM to account correctly for
105 the anisotropy of Reynolds stresses. This is due to the Boussinesq assumption,
106 approximating the turbulent stresses in similar manner to the molecular viscous
107 stresses. The Boussinesq approximations reads

$$\tau_{ij} = -2\mu_t \left(S_{ij} - \frac{1}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right) + \frac{2}{3} k \delta_{ij} , \quad (1)$$

108 where the turbulent kinetic energy is defined as $k = \frac{1}{2} \tau_{kk}$ and summation over
109 recurring indices within a product is implied. The strain-rate tensor is denoted
110 as S_{ij} and the eddy viscosity, derived from the transport equations of the tur-
111 bulence model, is represented by μ_t .

112 Based on the epistemic uncertainty, which is introduced into turbulence models
113 by the choice of the actual closure model [6], the perturbation approach tries
114 to derive and quantify the effects on QoI, e.g. the pressure field, by modify-
115 ing the anisotropy of turbulence within physical limitations. The implemented
116 framework for UQ of turbulence models seeks to sample from solutions, result-
117 ing from a modified underlying structure of the turbulence model, while aiming
118 for extreme states of the Reynolds stress tensor [9]. In this manner a CFD prac-
119 titioner may get the chance to estimate the sensitivity of some QoI regarding

the uncertainty in predicting the Reynolds stress anisotropy. In the following section we explain how to obtain a perturbed state of the Reynolds stress tensor and how to apply machine learning in order to get a better-informed, less conservative uncertainty prediction.

2.2. Data-free approach

The symmetric Reynolds stress tensor can be expressed by applying an eigenspace decomposition as

$$\tau_{ij} = k \left(a_{ij} + \frac{2}{3} \delta_{ij} \right) = k \left(v_{in} \Lambda_{nl} v_{jl} + \frac{2}{3} \delta_{ij} \right) . \quad (2)$$

Equation (2) includes the split into the anisotropy tensor a_{ij} and the isotropic part of τ_{ij} . The eigenspace decomposition provides the eigenvector matrix v and the diagonal eigenvalue matrix Λ , where the eigenvalues represent the shape and the eigenvectors imply the orientation of the tensor. Emory et al. [7] propose a strategy to perturb the eigenvalues and eigenvectors in Equation (2), resulting in a perturbed state of the Reynolds stress tensor

$$\tau_{ij}^* = k \left(v_{in}^* \Lambda_{nl}^* v_{jl}^* + \frac{2}{3} \delta_{ij} \right) . \quad (3)$$

The eigenvalue perturbation (determining Λ^*) makes use of the fact, that every physical, realizable state of the Reynolds stress tensor can be mapped onto barycentric coordinates

$$\mathbf{x} = \mathbf{x}_{1C} \frac{1}{2} (\lambda_1 - \lambda_2) + \mathbf{x}_2 (\lambda_2 - \lambda_3) + \mathbf{x}_{3C} \left(\frac{3}{2} \lambda_3 + 1 \right) \quad \text{with} \quad \lambda_1 \geq \lambda_2 \geq \lambda_3 , \quad (4)$$

which is essentially a linear transformation according to $\mathbf{x} = \mathbf{B}\boldsymbol{\lambda}$ (whereby three eigenvalues λ_i are represented by the vector $\boldsymbol{\lambda}$ and $\mathbf{B}$ depends on the choice of coordinates $\mathbf{x}_{1C}$, $\mathbf{x}_{2C}$ and $\mathbf{x}_{3C}$) [23]. Figure 1(a) shows the three limiting states of the Reynolds stress tensor, represented by the corners of the triangle ($\mathbf{x}_{1C}$, $\mathbf{x}_{2C}$, $\mathbf{x}_{3C}$) standing for the one-, two- and three-component (isotropic) turbulent state (1C, 2C and 3C). Thus, Iaccarino and co-workers [7, 8] defined the eigenvalue perturbation as a shift in barycentric coordinates towards each

143 of the limiting states to location $\mathbf{x}^*$, according to

$$\mathbf{x}^* = \mathbf{x} + \Delta_B (\mathbf{x}_{(t)} - \mathbf{x}) . \quad (5)$$

144 The relative distance $\Delta_B \in [0, 1]$ controls the magnitude of eigenvalue perturba-
 145 tion towards the corner state $\mathbf{x}_{(t)} \in \{\mathbf{x}_{1C}, \mathbf{x}_{2C}, \mathbf{x}_{3C}\}$. The perturbed eigenvalues
 146 λ_i^* can be remapped by

$$\boldsymbol{\lambda}^* = \mathbf{B}^{-1} \mathbf{x}^* . \quad (6)$$

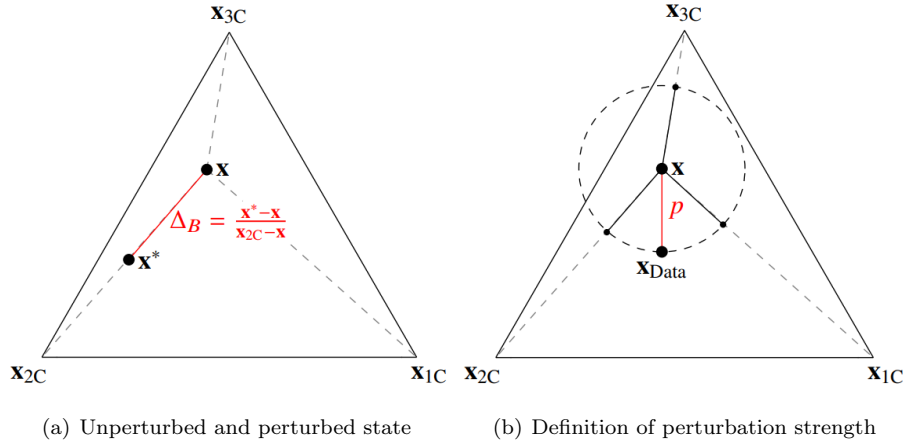


Figure 1: Schematic representation of the eigenvalue perturbation approach

147 The creation of the perturbed eigenvector matrix v^* is purely motivated by ma-
 148 nipulating the turbulent production term $P_k = -\tau_{ij} \frac{\partial u_i}{\partial x_j}$. Changing the align-
 149 ment of the eigenvectors of the Reynolds stress tensor and the strain rate tensor
 150 S_{ij} limits the production term to a maximum and minimum value [8]. Maxi-
 151 mum turbulent production is obtained by not changing the eigenvectors of the
 152 Reynolds stress tensor, meaning that they are identical to the eigenvectors of
 153 the strain rate tensor $v_{kS_{ij}}$ due to the Boussinesq assumption in Equation (1).
 154 Commuting the first and the last eigenvector of the Reynolds stress tensor leads
 155 to minimum turbulent production:

$$\begin{aligned} v_{\max} &= \begin{pmatrix} v_{1S_{ij}} & v_{2S_{ij}} & v_{3S_{ij}} \end{pmatrix} \rightarrow P_{k_{\max}} \\ v_{\min} &= \begin{pmatrix} v_{3S_{ij}} & v_{2S_{ij}} & v_{1S_{ij}} \end{pmatrix} \rightarrow P_{k_{\min}} . \end{aligned} \quad (7)$$

When combining the eigenvalue and eigenvector perturbation, not only the shape of the Reynolds stress ellipsoid is modified but also the relative alignment with the principle axes of the mean rate of the strain rate tensor is changed (orientation). It should be noted, that targeting the 3C turbulent state with $\Delta_B = 1$ results in identical eigenvalues and consequently the eigenvector matrix cancels out with its inverse. That is the reason, why there will be no distinction between minimized and maximized turbulent production.

To sum up, the data-free perturbation framework promises to only need five distinct simulations $\in \{(1C, P_{k_{\max}}), (1C, P_{k_{\min}}), (2C, P_{k_{\max}}), (2C, P_{k_{\min}}) \text{ and } 3C\}$, in order to get the entire information with reference to the epistemic uncertainty of the underlying turbulence model, if $\Delta_B = 1$ is chosen.

2.3. Data-driven approach

The eigenspace perturbation approach is a purely physics-based methodology, aiming for understandable uncertainty bounds for the turbulence modelling community. The data-free procedure applies a uniform perturbation to the entire flow domain. But the perturbation amplitude is a reflection of the inability of the turbulence model to reflect the underlying physics with high fidelity. This discrepancy between the turbulence model's dynamics and those of the turbulence physics are not uniform, but differ between different turbulent flows and even across different regions of the flow domain in the same flow. Thus, enabling variation in the magnitude of the perturbations is a better reflection of the actual model form uncertainty. As an additional advantage, if executed correctly, such a varying perturbation approach would enable more precise and less conservative uncertainty bounds on the QoIs. Moreover, a user has to choose the degree of uncertainty by selecting Δ_B before each investigation, which is another major drawback of the proposed method. This might be especially unfavorable in the design phase of turbomachinery components, when even a CFD practitioner experienced in turbulence modeling does not know how to set the degree of uncertainty.

Consequently, the user defined bounds on the eigenspace perturbation procedure

186 need to be replaced. Data-driven modeling approaches can be very beneficial for
 187 such surrogate modeling. Such machine learning surrogate models have found
 188 wide application in turbulence modeling [6, 24, 25]. As the amount of high-
 189 fidelity simulations (such as LES and DNS) increases, this data can be used
 190 to estimate the perturbation magnitudes. Heyse et al. propose a strategy to
 191 obtain a locally varying perturbation strength by using a random forest model
 192 [17]. Physical flow features are extracted to train a machine learning model in
 193 order to predict the local perturbation strength

$$p = |\mathbf{x}_{\text{Data}} - \mathbf{x}_{\text{RANS}}| = |\mathbf{x}_{\text{RANS}}^* - \mathbf{x}_{\text{RANS}}|, \quad (8)$$

194 as illustrated in Figure 1(b). Forward propagating CFD simulations follow train-
 195 ing the model, where the predicted perturbation strength is used to modify the
 196 Reynolds stress towards the same three limiting states as in the data-free ap-
 197 proach.

198 In this work, we go several steps beyond the initial proposed strategy. First
 199 of all, Heyse et al. [17] applied pure eigenvalue perturbation of the anisotropy
 200 tensor. We combine the data-driven eigenvalue perturbation with the data-free
 201 manipulation of the eigenvectors of the Reynolds stress tensor, as already de-
 202 scribed above. From our point of view, this procedure is inevitable, since the
 203 assumption that the Reynolds stress tensor and the mean rate of strain tensor
 204 share their eigenvectors (see Equation (1)), is known to be invalid in flow sit-
 205 uations, where streamline curvature, rotational effects, flow separation or reat-
 206 tachment, etc play a role [26]. We decided against utilizing machine learning to
 207 adjust the perturbations for the eigenvectors of the Reynolds stress tensor and
 208 the turbulent kinetic energy directly, as this would lead to a data-augmented,
 209 corrected turbulence model instead of obtaining uncertainty estimates. Since
 210 no justifiable bounds for the turbulent kinetic energy budget exist, the present
 211 methodology perturbs it indirectly by applying the described eigenvector ma-
 212 nipulation. Furthermore, we evaluate the adequate usage of a machine learning
 213 model for the desired application by extending the training data set and per-
 214 forming certain verification checks. The latter also includes the question, how to

215 build trust in such a trained machine learning model, when there is no accurate
216 data available anymore to estimate the prediction error. And finally, we assess
217 the limitations and capabilities of this method.

218 *2.3.1. Choice of machine learning model*

219 The general concept of machine learning is to approximate the relationship be-
220 tween input quantities (features) and output quantities (targets) to make pre-
221 diction under similar conditions. There are multiple approaches to approximate
222 these relationships. For the sake of interpretability and usability, decision trees
223 are chosen to be the machine learning model in the present work. Decision
224 trees (also called regression trees for solving regression problems) learn binary
225 rules (if/else decision rules) to predict target values based on given features [27].
226 Decision trees are prone to overfitting, which means that the model is not able
227 to generalize. A machine learning model is able to generalize, if it performs
228 adequate predictions based on a feature space, that is different than the fea-
229 ture space of the training data. Machine learning models, which are prone to
230 overfitting reveal a high variance. The potential accuracy of a machine learning
231 model is also dependent on its bias, which is characterized by the difference be-
232 tween the averages of the predictions and the true values. An inflexible model is
233 not capable to fit the total number of data sufficiently, which is determined as a
234 high bias of the model. The fact, that an increasing flexibility (lower bias) comes
235 along with worse generalization (high variance), is known as the bias-variance
236 trade-off. This trade-off describes the aim to choose a machine learning model,
237 that has low variance and low bias simultaneously [28, 29]. Random forests are
238 based on a number of uncorrelated regression trees and offer the possibility to
239 handle the bias-variance trade-off, while enabling powerful predictions [29]. For
240 this reason, we have chosen to use this ensemble learning technique. Instead
241 of just averaging the prediction of individual regression trees, a random forest
242 makes use of two essential key concepts:

- 243 • bootstrapping: Random sampling (with replacement) of the training data
244 for each individual tree, i.e. each tree is trained on a different data set

245 with equal size.

- 246 • feature bagging: Random subsampling of features at each decision point
247 (also known as split) for each tree, i.e. every tree uses a different feature
248 space at each binary decision.

249 In combination with bootstrap aggregation (bagging), which implies averaging
250 the prediction of a number of bootstrapped regression trees, the variance is
251 reduced and overfitting is avoided [30]. In this work, the python library *scikit-*
252 *learn* [31] is used to train the random forests and evaluate their predictions.

253 2.3.2. Choice of flow features

254 The selection of input features, which are relevant for predicting more accurate
255 perturbation magnitude p , is critical for turbulence modelling purposes. It has
256 to be ensured, that the chosen features represents physical significance with
257 respect to desired target quantity. Wang et al. [32] identified four raw quantities

$$Q = (\mathbf{S}, \mathbf{\Omega}, \nabla p, \nabla k) \quad (9)$$

258 to be a reasonable choice as input data for conducting machine learning based
259 on LEVM. The two raw input tensors $\mathbf{S}$, $\mathbf{\Omega}$ represent strain rate and rotation
260 rate, while ∇p and ∇k are the gradients of pressure and turbulent kinetic energy.
261 In our work, we agree on the usage of Q , and make use of the normalization
262 scheme, derived by Ling et al. [33]. A normalization by a factor β and the
263 absolute value of each element α of Q according to

$$\hat{\alpha} = \frac{\alpha}{|\alpha| + |\beta|} , \quad (10)$$

264 lead to the determination of non-dimensional raw flow features, which are pre-
265 sented in Table 1.

266 Since we are aiming for Galilean invariant features, which means, that they
267 should stay the same in different inertial frames of reference, it is essential to
268 embed invariance properties into a machine learning model (if the model should
269 not learn these properties during training). One method to do this is by formu-
270 lating all inputs and output quantities of the model such that they are invariant.

Table 1: Raw flow features for constructing the invariant basis

Description	Normalized input $\hat{\alpha}$	raw input α	normalization factor β
Strain rate	$\hat{\mathbf{S}}$	$\mathbf{S}$	ω
Rotation rate	$\hat{\mathbf{\Omega}}$	$\mathbf{\Omega}$	$ \mathbf{\Omega} $
Pressure gradient	$\widehat{\nabla} p$	∇p	$\rho \mathbf{U} \cdot \nabla \mathbf{U} $
Turbulent kinetic energy gradient	$\widehat{\nabla} k$	∇k	$\omega \sqrt{k}$

271 The velocity gradients (its symmetric part $\hat{\mathbf{S}}$ and antisymmetric part $\hat{\omega}$) and the
 272 gradients of pressure and turbulent kinetic energy are Galilean invariant. In or-
 273 der to determine the invariant feature basis of the raw flow features, Wang et al.
 274 [32] make use of the Hilbert basis theorem. This theorem states, that a finite
 275 number of invariants belongs to each minimal integrity basis for a finite tensorial
 276 set [34]. A minimal integrity basis is the minimal set of invariants, that represent
 277 all polynomial invariants associated with a tensorial set under transformation.
 278 The integrity basis constructed here is supposed to be rotational invariant [35].
 279 In this manner the minimal integrity basis amount to 47 invariants, which are
 280 in the following used as input features for training and evaluating the random
 281 forest. We add additional physical meaningful flow features to this exhaustive
 282 list of features based on domain knowledge and physical intuition. The addi-
 283 tional raw input features, which are presented in Table 2, can be computed by
 284 providing the turbulent kinetic energy k , the specific turbulent dissipation rate
 285 ω , the molecular viscosity μ , the eddy viscosity μ_t , the distance to the nearest
 286 wall d , the local Mach number Ma , the mean velocity U_i and its gradient tensor
 287 and the mean pressure p and its gradient vector.

288 The normalization procedure is retained in accordance to Equation (10). It
 289 should be noted, that the raw inputs and the normalization factors are Galilean
 290 invariant as well. Thus, a total number of 56 input features is used for train-
 291 ing and evaluating the random forests. Lastly, each feature is standardized by
 292 removing the mean and scaling to unit variance by applying a standard scaler
 293 preprocessing functionality of *scikit-learn* [31].

294 Although Wang et al. [36] reported significant improvements in prediction ac-

Table 2: Physical flow features

Numbering	Description	raw input α	normalization factor β
q_1	Q-criterion	$\frac{1}{2} (\mathbf{\Omega} ^2 - \mathbf{S} ^2)$	$ \mathbf{S} ^2$
q_2	Turbulent kinetic energy	k	$\frac{1}{2} U_i U_i$
q_3	Wall-distance based Reynolds Number	$\min\left(\frac{\sqrt{k}d}{50\nu}, 2\right)$	-
q_4	Pressure gradient along streamline	$U_k \frac{\partial p}{\partial x_k}$	$\sqrt{\frac{\partial p}{\partial x_j} \frac{\partial p}{\partial x_j} U_i U_i}$
q_5	Turbulent time scale	$\frac{1}{\omega}$	$\frac{1}{ \mathbf{S} }$
q_6	Production term	P_k	$k\omega$
q_7	Mach number	Ma	-
q_8	Eddy viscosity	μ_t	μ
q_9	Norm of Reynolds stresses	$ \overline{u'_i u'_j} $	k

295 curacy for using the invariant feature basis over a smaller number of features
 296 (e.g. physical motivated scalars) in a comparable study, we cannot confirm such
 297 observations in our work. We rather think, that due to the limited number of
 298 data (for training and testing purposes) a smaller number of input quantities
 299 performs excellently as well. As soon as the diversity of the data starts to in-
 300 crease (significantly different geometries featuring various flow situations) the
 301 need for a large set of features may occur. Being aware of the fact, that the
 302 present feature list may be needlessly large, we assure, that this exhaustive list
 303 of features does not involve any disadvantages in terms of accuracy. For this
 304 reason, we use the total amount of 56 input features in the present work.

305 2.4. Integration of UQ computation in CFD solver suite TRACE

306 2.4.1. Implementation

307 The aim of running a CFD simulation, propagating a perturbed Reynolds stress
 308 tensor, is to obtain a sensitivity with respect to the solution. For smooth and
 309 time-efficient simulations, it is advisable to start the perturbation from a suffi-
 310 ciently converged baseline RANS simulation (baseline means standard unmod-
 311 ified turbulence model). Mishra et al. [37] apply a factor to march the solution

312 based on the perturbed Reynolds stress tensor to a fully converged state. In our
 313 implementation, we use a factor f for the reconstruction of Reynolds stresses
 314 in order to be able to achieve fully converged perturbed solutions as well. We
 315 discuss the necessity and the effect of this factor in Section 2.4.2. The pertur-
 316 bation of the Reynolds stress tensor was implemented to the existing C code of
 317 TRACE and can be subdivided in several steps within each pseudo-time step of
 318 steady simulations:

- 319 1. calculate Reynolds stresses based on Boussinesq approximation in Equa-
 320 tion (1)
- 321 2. determine anisotropy tensor (see Equation (2))
- 322 3. perturb anisotropy tensor within physical realizable limits by selecting Δ_B
 323 (see Equation (5)) and whether the turbulent production term should be
 324 minimized or maximized (see Equation (7))
- 325 4. reconstruct perturbed Reynolds stress tensor according to

$$\tau_{ij_f}^* = \tau_{ij} + f \left[k \left(a^* + \frac{2}{3} \delta_{ij} \right) - \tau_{ij} \right], \quad (11)$$

326 where $f \in [0, 1]$ is the introduced moderation factor, adjusting the total
 327 amount of newly perturbed anisotropy tensor to be considered

- 328 5. update of the viscous fluxes using perturbed Reynolds stresses explicitly
- 329 6. update of the turbulent production term $P_k = -\tau_{ij} \frac{\partial u_i}{\partial x_j}$ using the per-
 330 turbed Reynolds stresses explicitly

331 TRACE features a python interface, called pyTRACE [38], which can be used
 332 to conduct a full set of perturbed simulations and sample the results for some
 333 QoI. In case of applying a data-driven perturbation of the Reynolds stresses,
 334 the python script takes also charge of evaluating a previously in preprocessing
 335 trained machine learning model based on extracted mean flow quantities. The
 336 high-level python script takes input parameters, containing information regard-
 337 ing the geometry, mesh resolution and additional solver settings. Furthermore,
 338 the set of intended perturbed simulations is set up, including selecting the tur-
 339 bulence limit state ($\mathbf{x}_{(t)} \in \{\mathbf{x}_{1C}, \mathbf{x}_{2C}, \mathbf{x}_{3C}\}$), the relative distance $\Delta_B \in [0, 1]$,

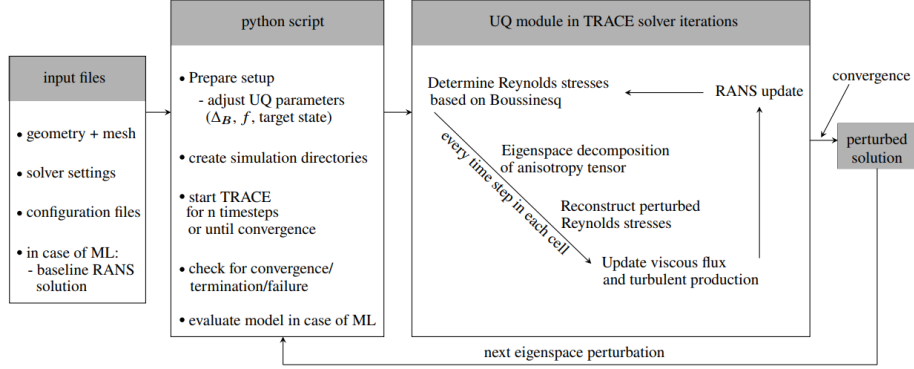


Figure 2: Implementation of the UQ framework within CFD solver suite TRACE

the alignment of the Reynolds stress eigenvectors with the strain rate tensor ($v_{\min}$ or $v_{\max}$) and the moderation factor $f \in [0, 1]$, as described earlier. As illustrated in Figure 2 the integration of the UQ module in the TRACE simulation run is conducted every time step (steady simulation) in each cell of the computational domain. When the converged perturbed solution is reached, the python script takes charge of setting up the next desired perturbation.

2.4.2. Discussion about restrictions of the Reynolds stress tensor perturbations

The eigenspace perturbation methodology, presented in Section 2, is solely motivated on quantifying the epistemic uncertainties of LEVM due to the inaccurate account for anisotropic flow phenomena. Iaccarino and co-worker [7, 8, 37] designed this method based on the mathematical derivations of Lumley [39] and Banerjee [23]. These derivations map the states of the Reynolds stress tensor, when it features one, two or three non-zero eigenvalues, onto corners of a constructed triangle, called the barycentric triangle. These states are described to be the extreme states of the Reynolds stress tensor, since the turbulence is only present in one, two or three directions - the corresponding directions of the eigenvectors.

A CFD practitioner is interested in ascertain the effect of the turbulence model's uncertainty on certain QoI, which are relevant for design. However, the re-

359 lation between the one-, two-, and three-component corner of the Reynolds
 360 stress tensor and some QoI is anything but linear. As a consequence, mod-
 361 ifying/perturbing the turbulent state of the Reynolds stress tensor seeks to
 362 estimate the uncertainty bounds rather than create extreme state of QoI.
 363 Nonetheless, we analyzed the relation of barycentric coordinates and QoI for
 364 selected flow cases by sampling points inside the barycentric triangle and prop-
 365 agating the perturbed Reynolds stress tensor in an earlier prior investigation
 366 [40]. Therefore, assessed against currently available data, we agree on the fact,
 367 that the corners of the barycentric coordinate produce adequate estimate of the
 368 uncertainty bounds in most of the flow regions. Although there might be areas
 369 of the flow solution, where the extreme state of turbulence is not corresponding
 370 to the extreme state of some QoI. This observation will be also discussed in
 371 Section 5.

372 Additionally, we would like to discuss the effect of the moderation factor f ,
 373 which was initially mentioned by Mishra et al. [37]. The main goal of ap-
 374 plying f is to reach a converged solution based on the perturbation approach.
 375 We agree on the fact, that this factor is needed for convergence issues, since
 376 some perturbed states tend to be unstable. Understandably, this is especially
 377 the case for perturbations seeking to decrease the turbulent production term
 378 ($P_{k_{\min}}$ and/or $3C$). Nevertheless, it is shown in the Appendix, that the effect of
 379 applying the moderation factor is actually identical to reducing Δ_B in case of
 380 pure eigenvalue perturbation. The need for moderating the effect of Reynolds
 381 stress tensor perturbation by an additional factor according to Equation (11)
 382 emerges, when combining eigenvector and eigenvalue perturbation. Generally
 383 speaking, using $f \leq 1$ stabilizes the CFD-simulation by weakening the impact
 384 of perturbation. Accordingly, users are encouraged to not only state the pre-
 385 scribed Δ_B but also the factor f . It has to be stated, though, that damping
 386 the effects of the actual perturbation (eigenvalues and eigenvectors) weakens the
 387 interpretability of the limiting states of turbulence, represented by the corners
 388 of the barycentric triangle.

389 Last but not least, the data-driven extension of the eigenspace perturbation

framework is build on Reynolds stress reference data, whereas other machine learning approaches in the field of turbulence modeling utilize indirect mean flow quantities like velocity and pressure. Hence, appropriate training data is mainly limited to well-resolved DNS/LES and cannot make use of experimental measurement data containing no second-moment statistics.

2.5. Data sets for training, testing and applying the machine learning model

As already described in Section 1, the final well-trained machine learning model should be sensitive to flow phenomena such as adverse-pressure gradient, separation and reattachment due to the known shortcomings of the LEVM. Consequently, it is reasonable to use data sets for training, which include these flow situations. We are continuously striving for extending our database, which contains various flow cases for machine learning. For the present study, we use the following flow cases:

- DNS of turbulent channel flow at $Re_\tau \in \{180, 550, 1000, 2000, 5200\}$ based on Lee and Moser [41]
- DNS at $Re_H \in \{2800, 5600\}$ and LES at $Re_H = 10595$ of periodic hill flow based on Breuer et al. [42]
- DNS of wavy wall flow at $Re_H = 6850$ based on Rossi [43]
- DNS of converging-diverging channel flow at $Re_\tau = 617$ based on Laval and Marquillie [44].

All the DNS and LES data of the described test cases are generated using incompressible solvers. In order to simulate these incompressible flows using the compressible solver TRACE without low-Mach preconditioning, the simulations are scaled (adapting dimensions of the geometry and/or molecular viscosity) to an incompressible Mach number of approximately 0.1, while preserving the intended Reynolds numbers. The two-equation, linear eddy viscosity Menter SST $k - \omega$ turbulence model is selected as the baseline model for all conducted RANS simulations [45]. In order to obtain steady state solutions, an implicit

time marching algorithm is applied. A flux difference splitting approach is employed to discretize the convection terms making use of a second order accurate Roe scheme in combination with MUSCL extrapolation.

To evaluate proper features as input parameters based on the RANS simulations, we conducted a mesh convergence study for each of the listed flow cases. Although the mesh convergence studies are not presented here due to the scope of the paper, we affirm, that we only use RANS simulation data, which shows sufficient grid convergence using a low-Reynolds resolution ($y^+ \leq 1$) at solid walls. The perturbation magnitude p can be determined accordingly to Equation (8) by comparing scale-resolving and RANS solutions. In order to compute the intended target quantity, the scale-resolving data has to be interpolated onto the RANS data points for every test case first. Due to numerical issues some RANS data samples may be located outside the barycentric triangle in terms of anisotropy tensor (see Figure 1). Therefore, we included the opportunity to remove these samples from the training or testing sets. The final application of the UQ perturbation approach is presented for the airfoil test case NACA 4412 at $Re_c = 1.52 \cdot 10^6$.

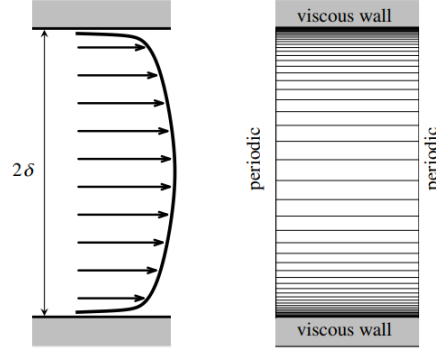
2.5.1. Turbulent channel

Although we are interested in more complex cases, the channel flow data serves as one of the key properties, what the model should be able to recognize and predict: turbulent boundary layer with inaccurate anisotropy represented by the LEVM close to the wall. The configuration for simulating the turbulent boundary layer is sketched in Figure 3. The characteristic Reynolds number is defined by

$$Re_\tau = \frac{\rho u_\tau \delta}{\mu}, \quad (12)$$

where δ is the channel half-height and the friction velocity is known as $u_\tau = \sqrt{(\tau_w/\rho)}$ with $\tau_w = \mu \frac{\partial U}{\partial y}|_{y=\text{wall}}$. The turbulent channel flow is homogeneous in the streamwise direction x and the spanwise direction z . A constant pressure gradient $\partial P/\partial x$ is applied to balance the skin friction at the wall. We use the available RANS grid cells in one half of the channel at the five different Reynolds

447 numbers as subsequent data points for training the random forest.



(a) Schematic sketch of a fully developed turbulent boundary layer
(b) $Re_\tau = 1000$ (every fourth line shown) mesh and boundary conditions; symmetry is enforced in spanwise direction

Figure 3: Turbulent channel flow simulation

448 2.5.2. Periodic hill

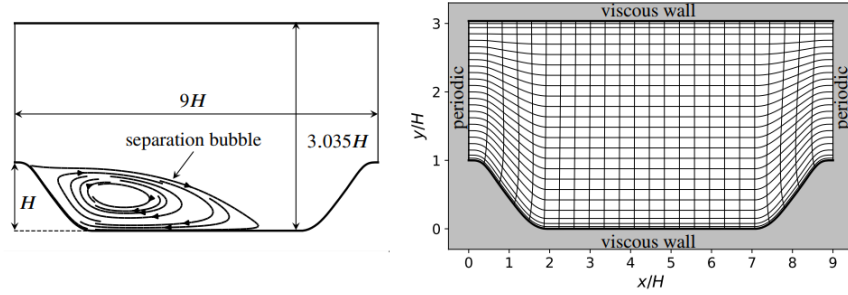
449 The flow over periodic hills features flow separation from curved surfaces, re-
450 circulation and a subsequent reattachment on the flat bottom of the channel.
451 Since the Reynolds number has a strong impact on the actual size of the sepa-
452 ration bubble, it is worthwhile to add three different Reynolds number flows to
453 our training set.

454 The Reynolds number based on the bulk velocity U_B , evaluated at the crest of
455 the hill, and the hill height H is defined as

$$Re = \frac{\rho U_B H}{\mu} . \quad (13)$$

456 For simulating the periodic hill configuration, periodic boundary conditions
457 are applied as illustrated in Figure 4. A constant pressure gradient $\partial P/\partial x$
458 is applied to move the fluid through the configuration. The available scale-

459 resolving data sets of the periodic hill only contain data at certain slices ($x/H \in$
 460 $\{0.05, 0.5, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0\}$). Consequently, the RANS solution
 461 is sliced at these locations accordingly. The scale-resolving data is interpolated
 462 onto the wall-normal RANS data positions, in order to generate the desired
 463 target quantity for the machine learning model.

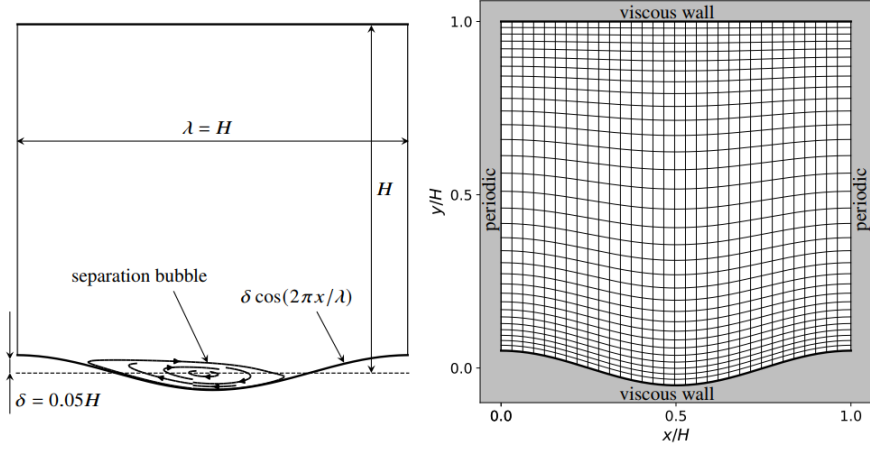


(a) Relative dimensions and sketch of the flow
 (b) Mesh ($Re_H = 10595$, every tenth line shown) and boundary conditions; slip conditions/inviscid walls are applied in spanwise direction

Figure 4: Schematic periodic hill setup

464 2.5.3. Wavy wall

465 The wavy wall test case is confined by a plane wall and a wavy surface, which is
 466 sketched in Figure 5. In former experimental settings, the desired flow situation
 467 was generated by stringing together multiple hills and valleys, described by a
 468 cosine function. For the CFD simulations (DNS and RANS) periodic bound-
 469 ary conditions in streamwise direction can be applied. In order to adjust the
 470 intended Reynolds number of $Re_H = 6850$, based on the bulk velocity and the
 471 mean channel height evaluated on the hill crest (accordingly to Equation (13)),
 472 a constant pressure gradient $\partial P/\partial x$ is used. Since the available DNS data set is
 473 two-dimensional and covers the entire domain size, we use all available RANS
 474 grid cells as subsequent data points for training the random forest.



(a) Relative dimensions and sketch of the flow (b) Mesh (every fourth line shown) and boundary conditions; slip conditions/inviscid walls are applied in spanwise direction

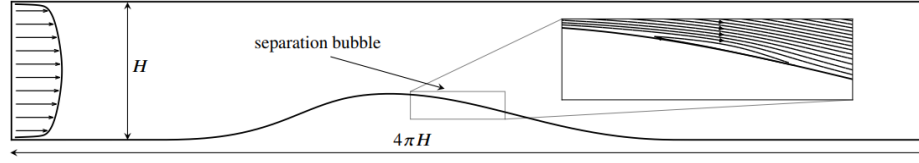
Figure 5: Schematic wavy wall setup

2.5.4. Converging-diverging channel

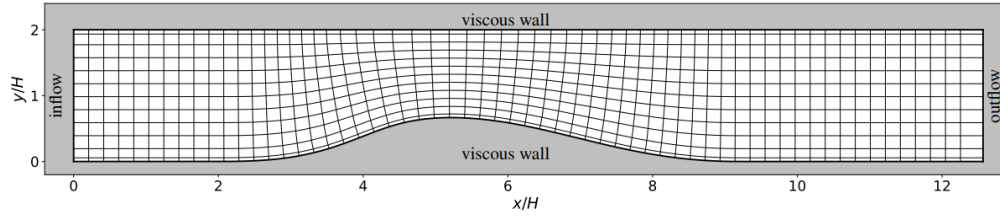
The configuration of a converging followed by a diverging section is an ideal test case to investigate the effect of an adverse pressure gradient with and without curvature. The flow separates slightly at the diverging part at the lower wall, but not on the flat top wall, as shown in Figure 6. Similar to the DNS, the inflow boundary conditions are derived from a fully developed turbulent boundary layer at $Re_\tau = 617$ (RANS predicted). A constant mass flow rate is prescribed at the outflow of the domain, which was derived based on the domain size and bulk quantities of the inflow profile (streamwise velocity and density). Since the available DNS data set is two-dimensional as well, we are able to provide all available RANS grid solution points as subsequent data points for training the machine learning model.

2.5.5. NACA 4412 airfoil

To demonstrate the application of the UQ framework with and without a machine learning model, the near-stall NACA 4412 airfoil is chosen in the presented

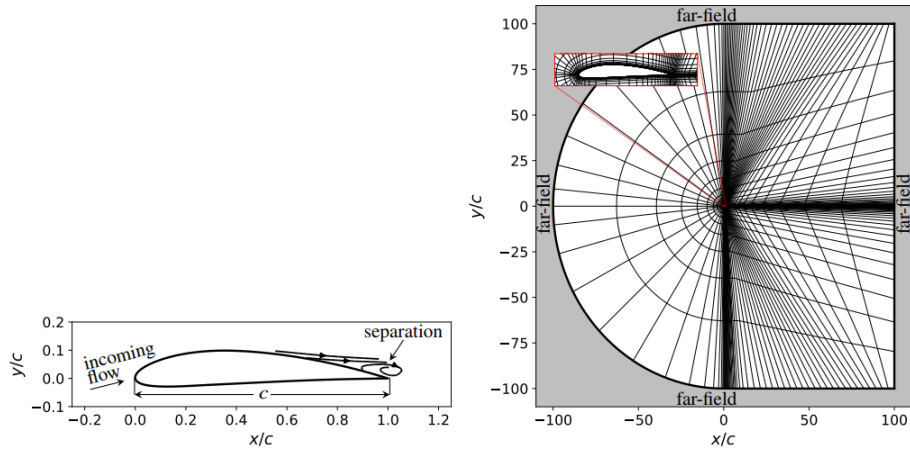


(a) Relative dimensions and sketch of the flow



(b) Mesh (every fourth line in streamwise direction and every twentieth line in wall normal direction shown) and boundary conditions; slip conditions/inviscid walls are applied in spanwise direction

Figure 6: Schematic converging-diverging setup



(a) Relative dimensions and sketch of the flow (b) Mesh (every eighth line shown) and boundary conditions; slip conditions/inviscid walls are applied in spanwise direction

Figure 7: Schematic NACA 4412 setup

work. This test case is a NASA benchmark case for turbulence models, featuring boundary layer separation close to the trailing edge. This airfoil is operated at a Reynolds number of $Re_c = 1.52 \cdot 10^6$ (based on the freestream velocity U_{inf} and the chord length c) and a Mach number of $Ma = 0.09$ (based on U_{inf}). The angle of attack is 13.87° as sketched in Figure 7(a). The CFD results are compared against experimental measurements of Coles and Wadcock [46]. In order to minimize the effect of boundary conditions on the CFD simulation, far-field conditions are applied to prescribe the specified flow conditions (see Figure 7(b)). A turbulence intensity of 0.086% and an eddy viscosity ratio μ_t/μ of 0.009 is prescribed in accordance with the description of NASA [47]. The mesh topology is the so-called C-grid featuring a grid cell resolution of $n_x, n_y, n_z = 896, 256, 1$, which can be downloaded from NASA’s turbulence database [47].

3. Hyperparameter selection based on generalization study

Before the training of the final random forest regression model was conducted, the impact of four different hyperparameters:

- maximum tree depth: maximum number of decision nodes from the root down to the furthest node allowed
- minimum sample count: minimum number of data samples required at a decision node allowed
- maximum number of features: maximum number of features randomly chosen at each decision node allowed
- number of trees: total number of individual decision trees used

on the accuracy is evaluated. Since the final trained model should be able to generalize for different geometries and flow conditions, it seems to be reasonable to evaluate these hyperparameters with focus on generalization capabilities of the random forest. Therefore, we apply a leave-one-out-cross validation, which is an appropriate procedure for small data sets. This means, that three out of

517 four available training data sets (see Section 2.5) are used for training, while
518 the remaining flow case is used to verify the model (see Table 3). Data sam-
519 ples featuring non-physical Reynolds stress tensors (barycentric coordinates are
520 located outside the barycentric triangle) are removed from each data set.

Table 3: Scenarios for hyperparameter study: x means part of training data, $\circ$ means testing data

Flow cases	Scenario			
	I	II	III	III
turbulent channel	x	x	x	$\circ$
periodic hill	x	x	$\circ$	x
wavy wall	x	$\circ$	x	x
converging-diverging channel	$\circ$	x	x	x

521 For each of the first three hyperparameters several different values were studied
522 over a range of the total amount of individual regression trees, while the other
523 two hyperparameters were set to default values (see *scikit-learn* documentation
524 for further information). As an example, Figure 8 presents the effect of the con-
525 sidered hyperparameters on the accuracy of the model prediction in scenario I,
526 where the accuracy is expressed in terms of the root mean square error (RMSE).

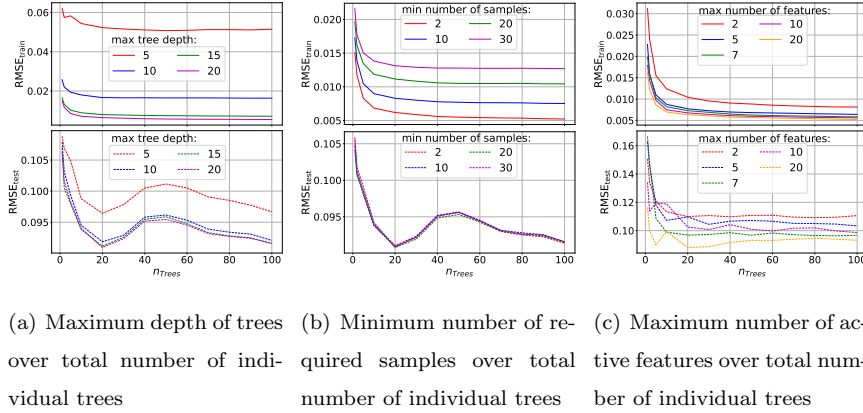


Figure 8: Training accuracy (solid / top) and testing accuracy (dashed / bottom) based on RMSE for selection of hyperparameters in scenario I

527 The subsequent statements made coincide with all evaluated scenarios as well.
 528 With increasing maximum depth of each individual tree, the model prediction
 529 error on the training data is reduced at the expense of increasing complexity.
 530 Increasing complexity of the tree may result in reduced generalizability capa-
 531 bilities (reduced test data accuracy) of the model. Because a maximum tree
 532 depth of 15 showed excellent performance for the training as well as for testing
 533 data, it is preferred compared to a higher value of 20, which increases compu-
 534 tational costs (see Figure 8(a)). The smaller the number of data samples for
 535 each decision node, the more accurate the performance on the training data
 536 (see Figure 8(b)). Since the RMSE based on evaluated testing data is not sig-
 537 nificantly affected by this hyperparameter, a minimum sample count of 10 is
 538 chosen. This enables the model to generalize to a greater extend, than selecting
 539 a smaller value. A larger number of selected features for each decision node
 540 lowers the training error and increases the risk of overfitting. Since a maximum
 541 number of 7 features produces accurate prediction performance for the test data
 542 as well (see Figure 8(c)), it is selected as inferred hyperparameter. In terms of
 543 total number of individual trees, one can observe a steep drop in RMSE for small
 544 numbers followed by a constant level of accuracy. The computational costs scale
 545 linearly with the number of individual trees. Although computational costs do
 546 not really play a relevant role for our application, as we only evaluate the model
 547 once before each simulation run, we sought for the minimum number of trees for
 548 maximized performance of the model. Therefore, we concluded to use a total
 549 number of 30 individual trees for the random forest by evaluating all described
 550 scenarios in Table 3.

551 **4. Verification of trained machine learning model**

552 Based on the choice of hyperparameters, which was discussed in the previous
 553 section, the prediction accuracy of the random forest should be evaluated on the
 554 available data (see Section 2.5). Ten different scenarios based on combinations
 555 of the flow data cases, listed in Table 4, serve as verification of functionality

Table 4: Prediction accuracy of random forest: x means part of training data, $\circ$ means not part of training data, **red** means data set for evaluation of RMSE

Flow cases	Scenario									
	I	Ia	II	IIa	III	IIIa	IIIb	III	IIIa	V
turbulent channel	x	x	x	x	x	x	x	$\circ$	x	x
periodic hill										
· $Re_H \in \{2800, 5600, 10595\}$	x	x	x	x	$\circ$	x		x	x	x
· $Re_H \in \{2800, 10595\}$							x			
· $Re_H = 5600$							$\circ$			
wavy wall	x	x	$\circ$	x	x	x	x	x	x	x
converging-diverging channel	$\circ$	x	x	x	x	x	x	x	x	x
RMSE ($p_{\text{pred}}, p_{\text{true}}$)	0.098	0.010	0.133	0.029	0.095	0.028	0.041	0.051	0.014	0.013

and present the accomplishment of the intended generalization of every model. While data samples featuring non-physical Reynolds stress tensors (barycentric coordinates are located outside the barycentric triangle) are removed from the data sets, the RMSE classifies the prediction accuracy of different scenarios. As the target quantity p can vary between zero and one (based on the construction of the equilateral triangle with edge length equals to one), the resulting RMSE indicate less than 10% absolute prediction error except for scenario II. As soon as a trained machine learning model should make predictions on flow cases, for which accurate data does not exist, judging the model's prediction in terms of accuracy becomes difficult. Comparing the input feature spaces of training and testing data (previously unseen case) resulting in extrapolation metrics, in order to build confidence in a machine learning model is a reasonable idea. Extrapolation metric measures the distance between a test point $\tilde{m}$ and the training data feature set $m^{(i)}$ for $i = 1, \dots, n$ with n as the number of training data points. In this paper, we use the Kernel Density Estimation (KDE) to compute the distance by estimating the probability density

$$f_{\text{KDE}} = \frac{1}{n\sigma^d} \sum_{i=1}^n \prod_{j=1}^d K\left(\frac{\tilde{m}_j - m_j^{(i)}}{\sigma}\right), \quad (14)$$

with the number of features d and the bandwidth σ , determined by Scott's rule [48]. According to the work of Wu et al. [35], we use a Gaussian kernel $K(t) = 1/\sqrt{(2\pi)} \exp(-t^2/2)$ and the distance is computed as follows:

$$d_{\text{KDE}} = 1 - \frac{f_{\text{KDE}}}{f_{\text{KDE}} + 1/A} \quad \text{with} \quad (15)$$

$$A = \prod_i^n \left(\max_j \left(m_j^{(i)} \right) - \min_j \left(m_j^{(i)} \right) \right) \quad \text{for } j = 1, \dots, d$$

The applied Gaussian kernel K ensures, that the smaller the difference of $\tilde{m}_j$ and $m_j^{(i)}$, the larger the output of K . In other words, Equation (14) increases, if $\tilde{m}$ becomes close to a concentrated feature space of the training data points and vice versa. The quantity A , which is only dependent on the features of the training set, can be interpreted as the volume of a cuboid with d dimensions. Thus $1/A$ is the probability density with respect to a uniform distribution inside such a cuboid. Due to the normalization of the distance in Equation (15), the metric is able to measure the distance of $\tilde{m}$ to the training data with respect to a uniform distribution. Consequently, on the one hand, if $\tilde{m}$ is close to a concentrated feature space, $f_{\text{KDE}} \gg 1/A$ implies $d_{\text{KDE}} \rightarrow 0$. On the other hand, $d_{\text{KDE}} \rightarrow 1$ follows from $f_{\text{KDE}} \ll 1/A$. This enables a user to interpret the rate of extrapolation needed based on the training data set. Thus, two extreme scenarios are represented according to:

- $d_{\text{KDE}} = 0$: no extrapolation is required; the features of the training data set cover the feature of the test point $\tilde{m}$
- $d_{\text{KDE}} = 1$: high extrapolation is required; the features of the test point $\tilde{m}$ are far off the features of the training data.

Since the extrapolation metric only assesses the closeness of the features between training and test data sets, Wu et al. [35] demonstrate, that the KDE extrapolation metric can be used to estimate the prediction confidence by quantifying the correlation between the degree of extrapolation and the prediction accuracy. In our work, the flow case of the converging-diverging channel serves to present the application of the extrapolation metric. The remaining data (turbulent channel

598 flow, flow over periodic hills and wavy walls) is used for training individual ran-
 599 dom forests, while each random forest is evaluated on the converging-diverging
 600 channel.
 601 Selected input features attributed with significant feature importance are consid-
 602 ered for computing the KDE distance d_{KDE} . The individual feature importance
 603 for each of the 56 input features, is determined after training the final random
 604 forest (scenario V in Table 4) using the chosen hyperparameters (see Section 3).
 605 The utility of each feature is determined by the permutation feature importance
 606 approach, accounting for the reduction in the model accuracy, when the values
 607 of this feature are randomly shuffled. Consequently, the selected five most im-
 608 portant features to be considered for determining d_{KDE} are the eddy viscosity
 609 q_8 , the normalized wall-distance q_3 , the Mach number q_7 , the turbulent kinetic
 610 energy q_2 and the Q-criterion q_1 . Contrary to the work of Wu et al. [35], we
 611 cannot confirm a strong correlation between the accuracy of the model, evalu-
 612 ated by predicting the perturbation magnitude p for the converging-diverging
 613 channel, and the mean of the KDE distance d_{KDE} for different training data
 614 sets, as illustrated in Figure 9.

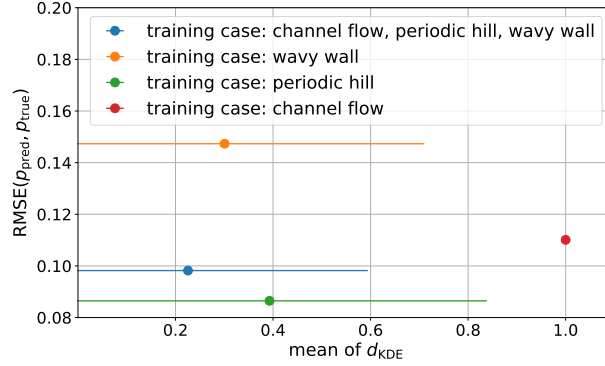
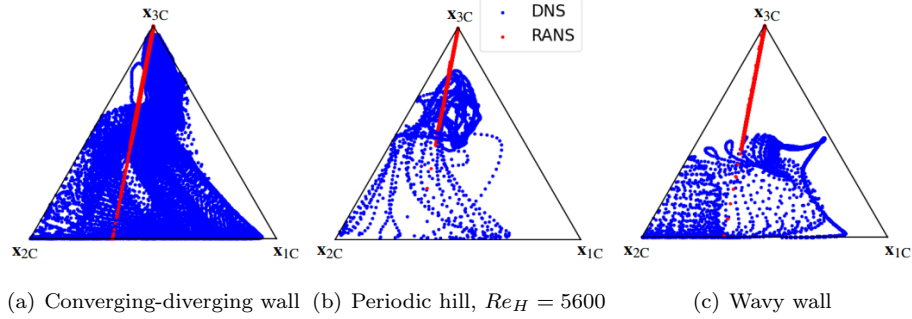
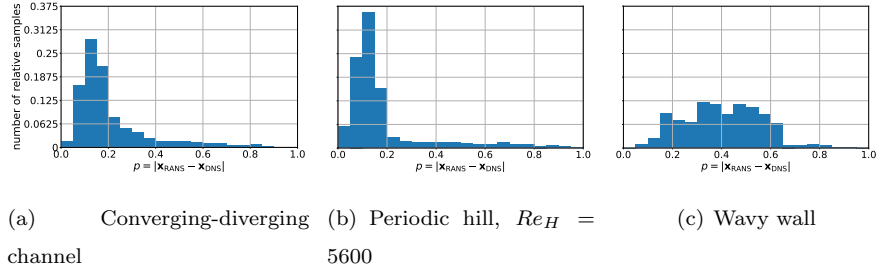


Figure 9: Relationship between the RMSE of the prediction for the converging-diverging chan-
 nel and the mean value of KDE extrapolation metric (standard deviation of the extrapolation
 metric is shown as the horizontal bars). All 56 input features are considered for the prediction
 and training of the random forest models, while only q_1 , q_2 , q_3 , q_7 and q_8 are used to compute
 d_{KDE} .



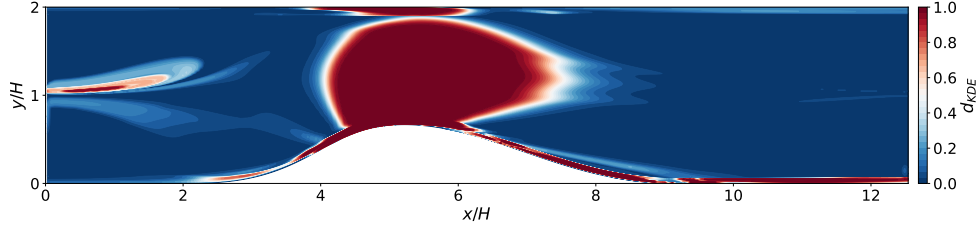
(a) Converging-diverging wall (b) Periodic hill, $Re_H = 5600$ (c) Wavy wall
Figure 10: Barycentric coordinates for selected flow cases; legend of (b) corresponds to (a) and (c) as well



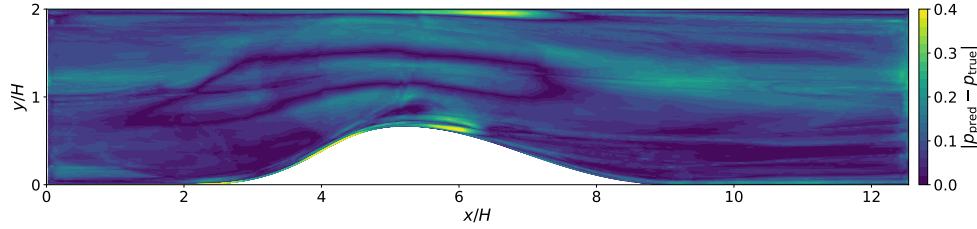
(a) Converging-diverging channel (b) Periodic hill, $Re_H = 5600$ (c) Wavy wall
Figure 11: Frequency of target quantity p for selected flow cases; vertical axis correspond to (a), (b) and (c)

615 However, Figure 10 presents a possible explanation for reduced prediction error,
616 when training on the periodic hill compared to the wavy wall. The DNS
617 data based barycentric coordinates of the converging-diverging channel and the
618 periodic hill cover similar areas in the barycentric triangle, while true values of
619 barycentric coordinates for the wavy wall test case are only located in the lower
620 range of the triangle.

621 Thus, the target quantity, which is the distance in barycentric coordinates,
622 becomes more frequent in a similar range of absolute values for the converging-
623 diverging channel and the periodic hill (see Figure 11). Even Wu et al. [35]
624 mention, that the correlation between accuracy and extrapolation metric is less
625 correlated, if the training set is very similar or very different from the test set,
626 what we might be facing here as well.



(a) Evaluated extrapolation metric



(b) Model prediction p_{pred} presented as absolute deviation based on true values

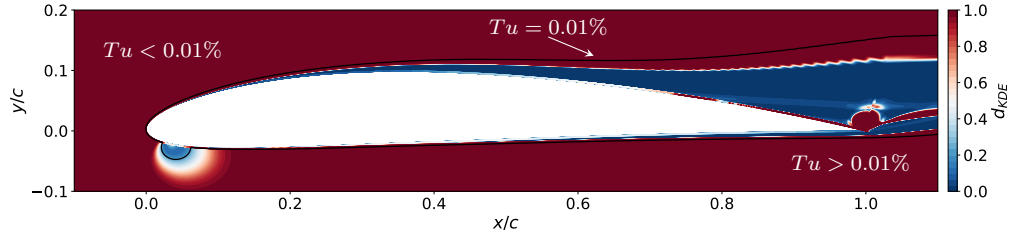
Figure 12: Verification of metric for converging-diverging channel

Nevertheless, the result of the extrapolation metric is highly dependent on the set of considered features. Thus, it seems reasonable to limit the evaluation of the metric to certain important features for the random forest.

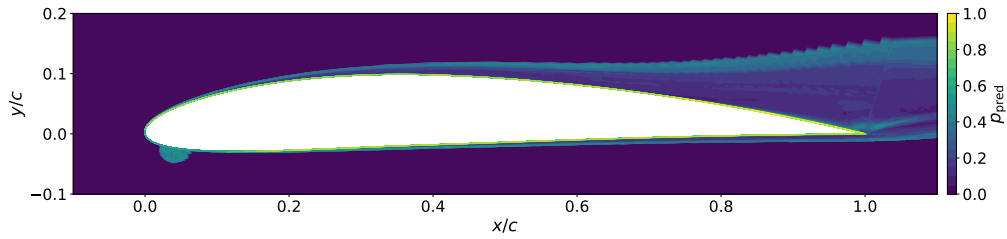
Figure 12(a) presents the two-dimensional distribution of the KDE metric, evaluated based on the five most important features q_8, q_3, q_7, q_2 and q_1 (corresponding to the blue data point in Figure 9 and scenario I in Table 4). Although, some spatial correlated regions between KDE distances d_{KDE} and model errors $|p_{\text{pred}} - p_{\text{true}}|$ in Figure 12(b) can be recognized, their overall correlation is not strong (Pearson correlation coefficient of approx. 0.2). Nonetheless, similar observations based on comparable test cases as in the work of Wu et al. [35] (not shown here) reinforce the trust in the presented KDE distance, when applied in an adequate manner.

Before the actual application of the data-driven UQ perturbation framework on the flow around NACA 4412 can be conducted, the prediction of the random

forest model for this case should be discussed. Figure 13 shows the evaluated
 extrapolation metric and the corresponding random forest predicted p . The
 estimated distribution of the KDE distance based on the five most important
 flow features q_8, q_3, q_7, q_2 and q_1 is shown in Figure 13(a). Solely regions close
 around the airfoil contain $d_{\text{KDE}} \leq 1$. This is due to the widely differing ge-
 ometry, Reynolds number and flow situation compared to the training data.
 The feature space of NACA 4412 contains especially higher Mach numbers q_7
 in regions far off the boundary layer with less turbulent kinetic energy q_2 (re-
 spective turbulent eddy viscosity q_8) and limited wall-distance based Reynolds
 Number q_3 . This fact also manifests, when taking a look at the level of turbu-
 lence intensity for the baseline solution. Because far-field boundary conditions
 are used, laminar flow is present almost everywhere in the CFD domain. An
 area with $Tu \geq 0.01\%$ can only be identified in the boundary layer around the



(a) Evaluated extrapolation metric, black line divides areas with higher and lower turbulence intensity Tu with a threshold of 0.01%



(b) Model prediction in the area of $Tu \geq 0.01\%$ (p_{pred} is set to zero for area with $Tu < 0.01\%$)

Figure 13: Evaluated metric and perturbation magnitude for NACA 4412 profile

654 profile, in the wake flow and around the stagnation point on the pressure side
 655 at $x/c \approx 0.035$. Consequently, the chosen training data does not sufficiently
 656 cover the feature space of NACA 4412 in the outer regions. But since the en-
 657 tire mapping of the Reynolds stress tensor onto barycentric coordinates makes
 658 only sense for relevant turbulent stresses, which are significantly larger than
 659 machine precision, the random forest is restricted to predict values in the area
 660 at $Tu \geq 0.01\%$ (see Figure 13(b)). This decision is purely based on our observa-
 661 tions and physical intuition. The verification of this procedure was numerically
 662 justified by comparing forward data-driven UQ computations based on model
 663 predicted p . Some simulations contained p , determined by the random forest,
 664 everywhere in the domain, others only in the area $Tu \geq 0.01\%$. Although the
 665 random forest is able to predict certain values greater than zero in the region
 666 featuring $Tu \leq 0.01\%$, the evaluated flow quantities around the NACA 4412
 667 airfoil did not show any significant difference (not shown here). Even the KDE
 668 distance in Figure 13(a) confirms the chosen procedure, by revealing reduced
 669 extrapolation distance in areas with increased turbulence intensity. Thus, re-
 670 stricting the model prediction to certain area closely around the airfoil can be
 671 justified from a physical and a machine learning perspective. However, the pre-
 672 sented extrapolation metric and model prediction close to the separation zone
 673 ($x/c \geq 0.8$ based on RANS simulation) reveal, that predicting the influence of a
 674 separated region in terms of anisotropy discrepancy is a challenging task. This
 675 issue might be only overcome with an increasing number of training data sets
 676 involving varying flow conditions and geometries.

677 **5. Application of UQ perturbation framework**

678 The flow around NACA 4412 at $Re_c = 1.52 \cdot 10^6$, $Ma = 0.09$ and an angle of
 679 attack of 13.87° demonstrates the general framework of the UQ perturbation
 680 approach presented in Section 2. Before analyzing the actual perturbed solu-
 681 tions and derived UQ estimates, the general performance of the Menter SST
 682 $k - \omega$ turbulence model is discussed briefly. The baseline simulation is in accor-

683 dance with the presented RANS solutions using the identical turbulence model
 684 provided by NASA’s turbulence modeling resource site [47]. Similar to NASA’s
 685 observations, when conducting steady simulation with TRACE on the given
 686 grid resolution, $CFL = 1$ has to be used in order to reach a fully converged
 687 steady-state solution. The main difference in comparison with the experimen-
 688 tal surface pressure measurements conducted by Coles and Wadcock [46] can
 689 be observed at the trailing edge of the suction side. The pressure coefficient,
 690 shown for example in Figure 14(a), is defined as

$$c_p = \frac{p - p_{\text{inf}}}{\frac{1}{2}\rho_{\text{inf}}U_{\text{inf}}^2}, \quad (16)$$

691 while the reference quantities are the ones at infinity $U_{\text{inf}} = 31.2$ m/s, $p_{\text{inf}} =$
 692 76914.1 Pa and $\rho_{\text{inf}} = 0.9$ kg/m³ (far-field boundary condition). Although the
 693 reference velocity is evaluated at different locations in the experiment, we apply
 694 the far-field freestream velocity instead based on best practice guidelines and in
 695 order to retain similar CFD results compared to the NASA findings [47].

696 The data-free uncertainty estimates, presented in Figure 14 are the result of
 697 perturbed turbulence model simulations using a relative perturbation magnitude
 698 $\Delta_B = 1.0$, since there is no justifiable physical reason to reduce the amount
 699 for targeting the extreme states of turbulence [7]. Unfortunately, as already
 700 discussed in Section 2.4.2, the perturbed simulations, trying to minimize the
 701 turbulent production, come along with stability issues in terms of convergence or
 702 even completely diverge from steady-state solutions. An appropriate moderation
 703 factor f has to be adjusted to retain acceptable, converged simulations. Besides
 704 examining the overall residuals of each simulation, we evaluate the mean blade
 705 force in y -direction to distinguish between an unacceptable unstable and an
 706 acceptable converged solution. Based on physical experience and consequent
 707 intuition, a threshold of 2% relative standard deviation with respect to the
 708 mean of the overall blade force in y -direction is chosen.

709 Put simply, as one is increasing the moderation factor f for simulations min-
 710 imizing the turbulent production term ($P_{k_{\text{min}}}$), the standard deviation of the
 711 mean blade force rises. In our investigations, a high-level python script (see

Table 5: Moderation factor f for every perturbed simulation of NACA 4412 UQ estimation. Application of $\Delta_B \leq 1$ in the data-driven approach necessitates a distinction between $P_{k_{\min}}$ and $P_{k_{\max}}$ for 3C.

Data-free						
Perturbed simulation:	$(1C, P_{k_{\max}})$	$1C, P_{k_{\min}}$	$(2C, P_{k_{\max}})$	$(2C, P_{k_{\min}})$	3C	
moderation factor	1.0	0.03	1.0	0.05	0.1	
Data-driven						
Perturbed simulation:	$(1C, P_{k_{\max}})$	$(1C, P_{k_{\min}})$	$(2C, P_{k_{\max}})$	$(2C, P_{k_{\min}})$	$(3C, P_{k_{\max}})$	$(3C, P_{k_{\min}})$
moderation factor	1.0	0.2	1.0	0.2	1.0	0.2

Figure 2) is applied to march the moderation factor as high as possible (by steps of 0.1 for $f \in [0.1, 1.0)$ and by steps of 0.01 for $f \in [0, 0.1)$). As the designated solutions may still contain small variations, we instrument probes on the airfoil surface and average the solution in order to get the mean for QoI. Due to the described convergence issues only a fraction of the perturbed Reynolds stress tensor can be utilized to update the Navier-Stokes equations for the NACA 4412 simulations (see moderation factors in Table 5). To the authors knowledge these low values are in accordance with the implementation of the eigenspace perturbation in the solver suite SU2, as their default value for the moderation factor is 0.1 [49].

Figure 14 presents the uncertainty estimates based on the perturbed target states, using the moderation factors presented in Table 5. As discussed above, the baseline Menter SST $k - \omega$ simulation shows significant deviation for the prediction of the separation zone close to the trailing edge. The results of (1C, $P_{k_{\max}}$) and (2C, $P_{k_{\max}}$) minimize this gap on the suction side for $x/c < 0.7$, whereas perturbed simulations minimizing the turbulent production term by modifying the eigenvectors of the Reynolds stress tensor predict an increased static pressure on the suction side for $x/c > 0.7$. Targeting for the 3C turbulent state with $\Delta_B = 1$ results in minimizing the turbulent production term as well [50], which can be seen on the presented surface quantities. Reduced turbu-

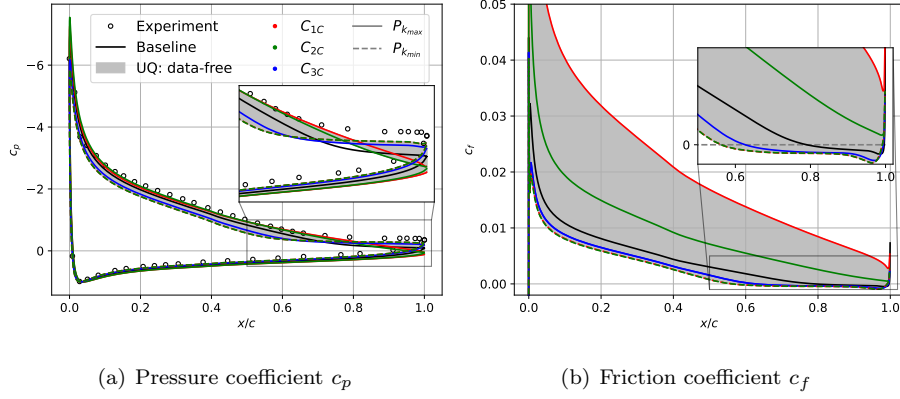


Figure 14: Surface quantities for flow around NACA 4412 including data-free evaluation of uncertainty estimates for Menter SST $k - \omega$ turbulence model; legend in (a) applies also for (b)

lent kinetic energy moves the separation zone towards the front of the airfoil,
indicated by the friction coefficient

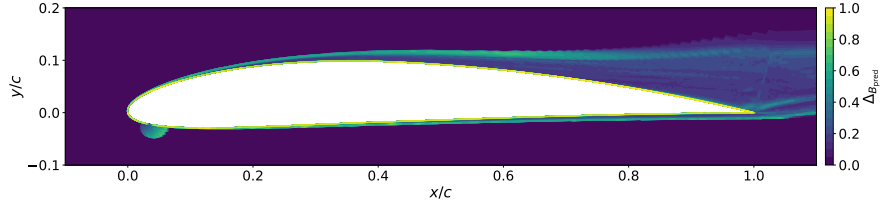
$$c_f = \frac{\tau_w}{\frac{1}{2}\rho_{\text{inf}}U_{\text{inf}}^2} \quad (17)$$

in Figure 14(b).

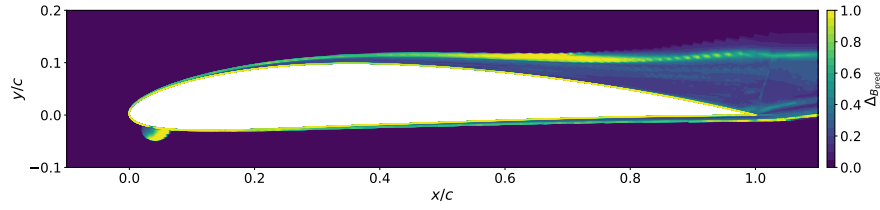
In contrast, the boundary layers of the perturbed solutions ($1C$, $P_{k_{\text{max}}}$) and ($2C$, $P_{k_{\text{max}}}$) reveal significant increased momentum transfer into the viscous sublayers, inducing complete suppression of the separation bubble.

As already mentioned in Section 2.4.2, the estimated uncertainty bounds by the perturbation framework are only aiming for the extreme state of turbulence in terms of Reynolds stress tensor's shape and orientation, which may not necessarily need to coincide with extreme state of some QoI. In the range of $0.72 < x/c < 0.82$ the baseline solution lies outside of the determined, grey shaded, UQ estimate for the pressure coefficient.

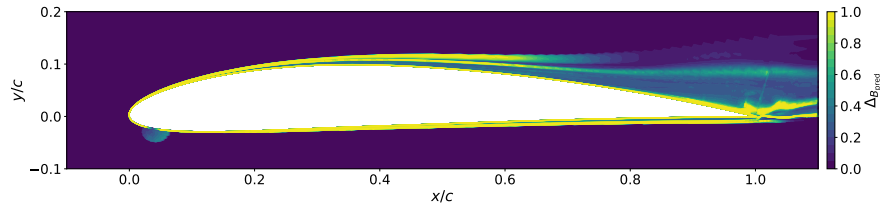
The random forest predicted perturbation magnitude p (see Figure 13(b)), is forward propagated towards the same three limiting states as in the data-free approach, as described in Section 2.3. This two-dimensional distribution of p is used to determine the respective Δ_B for each target state (see Figure 15 for



(a) Data-driven relative perturbation magnitude evaluated for 1C target state



(b) Data-driven relative perturbation magnitude evaluated for 2C target state



(c) Data-driven relative perturbation magnitude evaluated for 3C target state

Figure 15: Comparison of the effect of identical model predicted perturbation magnitude p on the relative perturbation magnitude Δ_B

748 1C, 2C and 3C). Due to the fact, that the unperturbed RANS solution data
 749 points are distributed along the plane strain line, the spatially averaged rela-
 750 tive distance Δ_B is highest for the the simulations targeting the isotropic corner
 751 (3C), followed by the two-component corner (2C) and the one-component corner
 752 (1C). In order to reach an acceptable steady-state solution for each perturbed
 753 simulation the moderation factor f is adjusted in the same manner as discussed
 754 above for the data-free procedure. Since the overall perturbation is weaker than
 755 using $\Delta_B = 1$, the moderation factor could be increased (see Table 5).

756 The estimated uncertainty bands for the surface quantities, shown in Figure 16,
 757 become narrower. Especially the uncertainty estimates for the pressure coeffi-
 758 cient based on $(1C, P_{k_{\max}})$ and $(2C, P_{k_{\max}})$ are very close to the baseline solution.
 759 As the overproduction of turbulent kinetic energy for the data-free 1C and 2C
 760 cases disappears, all data-driven perturbed simulations feature a separation zone
 761 on the suction surface.

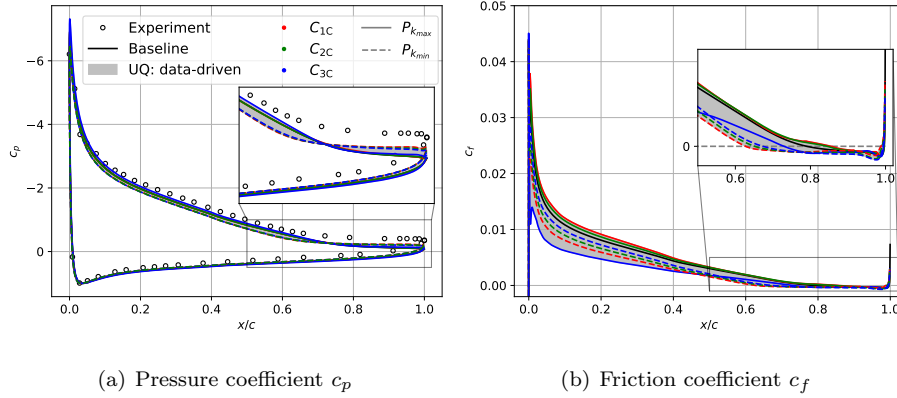


Figure 16: Surface quantities for flow around NACA 4412 including data-driven evaluation of
 uncertainty estimates for Menter SST $k - \omega$ turbulence model; legend in (a) applies also for
 (b)

762 Last but not least, the fact, that none of the presented UQ estimates envelopes
 763 the experimental surface pressure measurements, needs to be discussed. To start
 764 with, as discussed already in the beginning of this section, the CFD setup seems
 765 to come along with certain weaknesses. Moreover, the underlying intention of
 766 applying the UQ perturbation framework is not to include certain high-fidelity
 767 data, whether it originates from experiments or scale-resolving simulation, into
 768 its resulting envelope. The methodology seeks to produce limiting states of the
 769 Reynolds stress tensor, propagates these states and results in modified QoI. But
 770 there is no need that accurate data points for some QoI have to be covered
 771 by the uncertainty estimates resulting from the Reynolds stress tensor pertur-
 772 bation framework. Therefore, we disagree with some recent publications (e.g.

[51]), where the quality of the eigenspace perturbation framework is judged by covering certain QoI. As the uncertainty envelopes do not represent confidence or strict intervals at all [10], we encourage to validate the uncertainty estimates with respect to the underlying physical concept of the eigenspace perturbation framework. Analyzing simulation results related to the perturbed state of the Reynolds stress tensor should be chosen over evaluating the coverage of certain reference data. The main reason is that the perturbation framework is only able to account for structural uncertainties limited to the chosen RANS turbulence model in the CFD-solver. However, other sources of uncertainties related to RANS simulations are not considered by the eigenspace perturbations, such as:

- approximation of 3D geometries with 2D CFD setups
- neglecting of geometry and manufacturing tolerances
- choice of boundary conditions, which must not necessarily coincide with reference conditions
- assuming steady-state flow conditions, when flow might be already unsteady in reality.

Consequently, it cannot be expected, that the perturbed simulations envelope experimental or reference data.

The described machine learning procedure only accounts for the spatially varying deviation in Reynolds stress anisotropy, since the turbulence model's uncertainties are not uniform across the computational domain. The impact of the discrepancy in terms of anisotropy between RANS and scale-resolving data on certain QoI, was not part of the machine learning process. Thus, with reference to the disregarded sources of uncertainty mentioned above, even the data-driven perturbed turbulence model simulations cannot account for adequate entirely enveloping bounds for selected QoI.

6. Conclusion and Outlook

The present work aims to consolidate an arisen method in the field of turbulence model UQ. We demonstrate the possibility to estimate uncertainty bounds for turbulence models with state-of-the-art methods in DLR's CFD solver suite TRACE. The UQ perturbation framework is described extensively, presenting its underlying idea while mentioning its limitations and what it is not able to do for industrial applications. Additionally, since TRACE is eminent in turbomachinery industry, designing and implementing a framework to easily conduct uncertainty estimation for turbulence models using TRACE was a major goal of this work.

Moreover, we applied a proposed machine learning strategy to further enhance the interpretability of generated uncertainty estimates by being less conservative and nonetheless physics constrained simultaneously. Our extension of this data-driven eigenvalue perturbation approach is the enlargement of flow cases featuring separated flows, adverse pressure gradient and reattachment for training and testing purposes on the one hand. This enables us to check and verify the appropriate application of trained random forest regression model in-depth. On the other hand, considering eigenvector perturbation of the Reynolds stress tensor on top of the data-driven eigenvalue perturbation was the plausible next step in this specific research field.

In our investigations, we outline tools and methodologies for assessing and analyzing data-driven models, especially in the context of turbulent flows. We address key points in the field of machine learning such as selection of input features, tuning of hyperparameters, judging the model's accuracy in an a posteriori and an a priori way.

In order to predict the desired target quantity for the selected flow cases by the random forest, we admit, that we might not have to use this abundant number of input features, as described above. This is due to the fact, that the considered cases show certain similarities in terms of input and target quantities. However, if the amount of training data sets increases, covering a wider range of flow

phenomena, the machine learning model will likely take advantage of a larger number of input features.

We confirm, that the perturbation approach to account for turbulence model uncertainties, is a purely physics-based, comprehensible framework. Nevertheless, it suffers from reduced convergence or even divergence of steady state solutions. The necessity to moderate certain perturbations by an arbitrary factor, seems unsatisfactory for such an advanced approach. Currently, we are not aware of any other remedy for convergence issues as well, as even the machine learning does not help to overcome this particular issue. Moreover, we also agree on the underlying idea to account for spatially varying of turbulence model uncertainties by using data-driven methodologies to determine certain areas of high deviations from accurate Reynolds stress anisotropy states. Training a machine learning model to predict an appropriate model-form uncertainty will always help CFD users to get an indication, in which regions the LEVM assumptions might be violated. However, we believe, based on our experiences with the data-driven Reynolds stress tensor perturbation framework (which are not only limited to the NACA 4412 flow case), that an enhanced applicability can only be achieved, if the stability issues in terms of convergence for steady state simulation were solved.

Appendix

Mathematical effect of the moderation factor f in case of pure eigenvalue perturbation

By applying the moderation factor $f \in [0, 1]$ the perturbed Reynolds stress tensor can be expressed as

$$\begin{aligned}\tau_{ij_f}^* &= \tau_{ij} + f \left[k \left(a_{ij}^* + \frac{2}{3} \delta_{ij} \right) - \tau_{ij} \right] \\ &= (1 - f) \tau_{ij} + f \tau_{ij}^*,\end{aligned}\tag{18}$$

where τ_{ij} is the Reynolds stress tensor, which was calculated based on Boussinesq assumption in step 1 (see Section 2.4.1). Based on the perturbed anisotropy

855 tensor, the reconstructed Reynolds stress tensor is indicated by τ_{ij}^* . The anisotropy
 856 tensor related to this perturbed Reynolds stress tensor can be written as

$$\begin{aligned}
 a_{ij_f}^* &= \frac{\tau_{ij_f}^*}{k} - \frac{2}{3}\delta_{ij} \\
 &= \frac{(1-f)\tau_{ij} + f\tau_{ij}^*}{k} + \frac{2}{3}\delta_{ij} \\
 &= \frac{(1-f)\left[k\left(a_{ij} + \frac{2}{3}\delta_{ij}\right)\right] + f\left[k\left(a_{ij}^* + \frac{2}{3}\delta_{ij}\right)\right]}{k} - \frac{2}{3}\delta_{ij} \quad (19) \\
 &= (1-f)\left(a_{ij} + \frac{2}{3}\delta_{ij}\right) + f\left(a_{ij}^* + \frac{2}{3}\delta_{ij}\right) - \frac{2}{3}\delta_{ij} \\
 &= (1-f)a_{ij} + fa_{ij}^* .
 \end{aligned}$$

857 When just applying eigenvalue perturbation of the anisotropy tensor, a_{ij} and
 858 a_{ij}^* share identical eigenvectors. Thus, the eigenvalues of $a_{ij_f}^*$ are

$$\lambda_{i_f}^* = (1-f)\lambda_i + f\lambda_i^* . \quad (20)$$

859 The barycentric weights C_{iC} , which are used to calculate the barycentric coor-
 860 dinates in Equation (4), can be expressed in terms of moderation factor:

$$\begin{aligned}
 C_{1C_f}^* &= \frac{1}{2} \left[\lambda_{1_f}^* - \lambda_{2_f}^* \right] \\
 &= \frac{1}{2} [(1-f)\lambda_1 + f\lambda_1^* - (1-f)\lambda_2 + f\lambda_2^*] \\
 &= (1-f)\frac{1}{2}(\lambda_1 - \lambda_2) + f\frac{1}{2}(\lambda_1^* - \lambda_2^*) \quad (21) \\
 &= (1-f)C_{1C} + fC_{1C}^*
 \end{aligned}$$

861

$$\begin{aligned}
 C_{2C_f}^* &= \lambda_{2_f}^* - \lambda_{3_f}^* \\
 &= (1-f)\lambda_2 + f\lambda_2^* - (1-f)\lambda_3 + f\lambda_3^* \\
 &= (1-f)(\lambda_2 - \lambda_3) + f(\lambda_2^* - \lambda_3^*) \quad (22) \\
 &= (1-f)C_{2C} + fC_{2C}^*
 \end{aligned}$$

$$\begin{aligned}
C_{3C_f}^* &= \frac{3}{2}\lambda_{3_f}^* + 1 \\
&= \frac{3}{2}(1-f)\lambda_3 + \frac{3}{2}f\lambda_3^* + (1-f) + f \\
&= (1-f)\left(\frac{3}{2}\lambda_3 + 1\right) + f\left(\frac{3}{2}\lambda_3^* + 1\right) \\
&= (1-f)C_{3C} + fC_{3C}^*
\end{aligned} \tag{23}$$

863 The perturbed barycentric coordinates $\mathbf{x}_f^*$ (modified by the moderation factor
864 f) can be written using Equations (21), (22) and (23) as

$$\begin{aligned}
\mathbf{x}_f^* &= \mathbf{x}_{1C}C_{1C_f}^* + \mathbf{x}_{2C_f}^* + \mathbf{x}_{3C}C_{3C_f}^* \\
&= \mathbf{x}_{1c}[(1-f)C_{1C} + fC_{1C}^*] + \mathbf{x}_{2c}[(1-f)C_{2C} + fC_{2C}^*] + \mathbf{x}_{3c}[(1-f)C_{3C} + fC_{3C}^*] \\
&= (1-f)(\mathbf{x}_{1c}C_{1C} + \mathbf{x}_{2c}C_{2C} + \mathbf{x}_{3c}C_{3C}) + f(\mathbf{x}_{1c}C_{1C}^* + \mathbf{x}_{2c}C_{2C}^* + \mathbf{x}_{3c}C_{3C}^*) \\
&= (1-f)\mathbf{x} + f\mathbf{x}^* .
\end{aligned} \tag{24}$$

865 Remembering Equation (5) and rearranging leads to

$$\mathbf{x}^* = (1 - \Delta_B)\mathbf{x} + \Delta_B\mathbf{x}_{(t)} . \tag{25}$$

866 The analogy of Equation (24) and (25) reveals the similar effect of adjusting Δ_B
867 or f in case of only perturbing the eigenvalues of the anisotropy tensor. Thus,
868 one can rewrite the actual intended location inside the barycentric triangle as a
869 relative distance towards the corners

$$\mathbf{x}_f^* = (1 - \Delta_B f)\mathbf{x} + \Delta_B f\mathbf{x}_{(t)} . \tag{26}$$

870 Acknowledgments

871 The project on which this paper is based was funded by the German Fed-
872 eral Ministry for Economic Affairs and Climate Action under the funding code
873 03EE5041A. The authors are responsible for the content of this publication.
874 Additionally, the authors thank the financial funding of the prescribed work by
875 the internal DLR project SuperCOOL.

876 References

- 877 [1] Karniadakis G. Quantifying uncertainty in CFD. *Journal of Fluids*
878 *Engineering-transactions of the ASME* 2002;124(1):2–3. doi:10.1115/1.
879 1447925.
- 880 [2] Oberkampf W, Roy C. *Verification and Validation in Scientific Computing*.
881 2010. doi:10.1017/CB09780511760396.
- 882 [3] Emory MA. Estimating model-form uncertainty in Reynolds-averaged
883 Navier-Stokes closures. Ph.D. thesis; Stanford University, Department of
884 Mechanical Engineering; 2014.
- 885 [4] Xiao H, Cinnella P. Quantification of model uncertainty in RANS
886 simulations: A review. *Progress in Aerospace Sciences* 2019;108:1–31.
887 doi:10.1016/j.paerosci.2018.10.001.
- 888 [5] Zang T, Hensch M, Hilburger M, Kenny S, Luckring J, Maghami P,
889 et al. Needs and opportunities for uncertainty-based multidisciplinary de-
890 sign methods for aerospace vehicles 2002;.
- 891 [6] Duraisamy K, Iaccarino G, Xiao H. Turbulence modeling in the age of
892 data. *Annual Review of Fluid Mechanics* 2019;51(1):357–77. doi:10.1146/
893 annurev-fluid-010518-040547.
- 894 [7] Emory M, Larsson J, Iaccarino G. Modeling of structural uncertain-
895 ties in Reynolds-averaged Navier-Stokes closures. *Physics of Fluids*
896 2013;25(11):110822. doi:10.1063/1.4824659.
- 897 [8] Iaccarino G, Mishra A, Ghili S. Eigenspace perturbations for uncertainty
898 estimation of single-point turbulence closures. *Physical Review Fluids*
899 2017;2:024605. doi:10.1103/PhysRevFluids.2.024605.
- 900 [9] Mishra A, Iaccarino G. Theoretical analysis of tensor perturbations for
901 uncertainty quantification of Reynolds-averaged and subgrid scale closures.
902 *Physics of Fluids* 2019;31(7):075101. doi:10.1063/1.5099176.

- 903 [10] Mishra A, Mukhopadhyaya J, Alonso J, Iaccarino G. Design exploration
904 and optimization under uncertainty. *Physics of Fluids* 2020;32(8):085106.
905 doi:10.1063/5.0020858.
- 906 [11] Emory M, Iaccarino G, Laskowski GM. Uncertainty quantification in tur-
907 bomachinery simulations. *Turbo Expo: Power for Land, Sea, and Air*
908 2016;Volume 2C: Turbomachinery. doi:10.1115/GT2016-56798.
- 909 [12] Razaaly N, Gori G, Iaccarino G, Congedo PM. Optimization of an ORC
910 supersonic nozzle under epistemic uncertainties due to turbulence models.
911 In: *GPPS 2019 - Global Power and Propulsion Society*. 2019,.
- 912 [13] Mukhopadhyaya J, Whitehead BT, Quindlen JF, Alonso JJ, Cary AW.
913 Multi-fidelity modeling of probabilistic aerodynamic databases for use in
914 aerospace engineering. *International Journal for Uncertainty Quantification*
915 2020;10(5):425–47.
- 916 [14] Lamberti G, Gori C. Uncertainty quantification for RANS predictions
917 of wind loads on buildings. In: *Proceedings of the XV Conference of the*
918 *Italian Association for Wind Engineering*. Cham: Springer International
919 Publishing; 2019, p. 402–12.
- 920 [15] Eidi A, Ghiassi R, Yang X, Abkar M. Model-form uncertainty quantifica-
921 tion in RANS simulations of wakes and power losses in wind farms. *Re-*
922 *newable Energy* 2021;179:2212–23. doi:10.1016/j.renene.2021.08.012.
- 923 [16] Hornshøj-Møller SD, Nielsen PD, Forooghi P, Abkar M. Quantifying
924 structural uncertainties in Reynolds-averaged Navier–Stokes simulations of
925 wind turbine wakes. *Renewable Energy* 2021;164:1550–8. doi:10.1016/j.
926 *renene*.2020.10.148.
- 927 [17] Heyse J, Mishra A, Iaccarino G. Estimating RANS model uncertainty using
928 machine learning. *Journal of the Global Power and Propulsion Society*
929 2021;:1–14doi:10.33737/jgpps/134643.

- 930 [18] Morsbach C. Reynolds stress modelling for turbomachinery flow applica-
 931 tions. Ph.D. thesis; Technische Universität Darmstadt, Darmstadt, Ger-
 932 many; 2017. URL: elib.dlr.de/113258/.
- 933 [19] Speziale CG. Analytical methods for the development of Reynolds-Stress
 934 closures in turbulence. *Annual Review of Fluid Mechanics* 1991;23(1):107–
 935 57. doi:10.1146/annurev.fl.23.010191.000543.
- 936 [20] Mompean G, Gavrilakis S, Machiels L, Deville M. On predicting the
 937 turbulence-induced secondary flows using nonlinear k - ϵ models. *Physics*
 938 *of Fluids* 1996;8:1856–68. doi:10.1063/1.868968.
- 939 [21] Craft T, Launder B, Suga K. Development and application of a cubic eddy-
 940 viscosity model of turbulence. *International Journal of Heat and Fluid Flow*
 941 1996;17(2):108–15. doi:10.1016/0142-727X(95)00079-6.
- 942 [22] Lien F, Leschziner M. Assessment of turbulence-transport models including
 943 non-linear rng eddy-viscosity formulation and second-moment closure for
 944 flow over a backward-facing step. *Computers & Fluids* 1994;23(8):983–1004.
 945 doi:10.1016/0045-7930(94)90001-9.
- 946 [23] Banerjee S, Krahel R, Durst F, Zenger C. Presentation of anisotropy prop-
 947 erties of turbulence, invariants versus eigenvalue approaches. *Journal of*
 948 *Turbulence* 2007;8:N32. doi:10.1080/14685240701506896.
- 949 [24] Ihme M, Chung WT, Mishra A. Combustion machine learning: Princi-
 950 ples, progress and prospects. *Progress in Energy and Combustion Science*
 951 2022;91:101010.
- 952 [25] Brunton SL, Noack BR, Koumoutsakos P. Machine learning for fluid me-
 953 chanics. *Annual Review of Fluid Mechanics* 2020;52:477–508.
- 954 [26] Launder BE, Tselepidakis DP, Younis BA. A second-moment closure
 955 study of rotating channel flow. *Journal of Fluid Mechanics* 1987;183:63–75.
 956 doi:10.1017/S0022112087002520.

- 957 [27] Breiman L. Random forests. *Machine Learning* 2001;45:5–32. doi:10.1023/
958 A:1010950718922.
- 959 [28] Geman S, Bienenstock E, Doursat R. Neural networks and the
960 bias/variance dilemma. *Neural Computation* 1992;4:1–58. doi:10.1162/
961 **neco.1992.4.1.1.**
- 962 [29] Louppe G. Understanding random forests: From theory to practice. Ph.D.
963 thesis; 2014. doi:10.13140/2.1.1570.5928.
- 964 [30] Breiman L. Bagging predictors. *Machine Learning* 2004;24:123–40. doi:10.
965 1007/BF00058655.
- 966 [31] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O,
967 et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learn-*
968 *ing Research* 2011;12:2825–30.
- 969 [32] Wang JX, Wu JL, Xiao H, Paterson E. Physics-informed machine learning
970 approach for augmenting turbulence models: A comprehensive framework.
971 *Physical Review Fluids* 2018;3(7). doi:10.1103/physrevfluids.3.074602.
- 972 [33] Ling J, Jones RE, Templeton J. Machine learning strategies for systems
973 with invariance properties. *Journal of Computational Physics* 2016;318:22–
974 35.
- 975 [34] Spencer A, Rivlin R. Isotropic integrity bases for vectors and second-order
976 tensors. *Archive for Rational Mechanics and Analysis* 1962;9:45 – 63.
- 977 [35] Wu JL, Wang JX, Xiao H, Ling J. A priori assessment of prediction con-
978 fidence for data-driven turbulence modeling. *Flow, Turbulence and Com-*
979 *bustion* 2017;99(1):25–46. doi:10.1007/s10494-017-9807-0.
- 980 [36] Wang JX, Wu J, Ling J, Iaccarino G, Xiao H. A comprehensive physics-
981 informed machine learning framework for predictive turbulence modeling
982 2018;.

- [37] Mishra A, Mukhopadhyaya J, Iaccarino G, Alonso J. An uncertainty estimation module for turbulence model predictions in SU2. *AIAA Journal* 2018;57. doi:10.2514/1.J057187.
- [38] DLR, German Aerospace Center, Institute of Propulsion Technology . Trace user guide. <http://www.trace-portal.de/userguide/trace/index.html>; accessed April 2022.
- [39] Lumley JL. Computational modeling of turbulent flows. In: *Advances in Applied Mechanics*; vol. 18. Elsevier; 1979, p. 123 – 176. doi:10.1016/S0065-2156(08)70266-7.
- [40] Matha M, Kucharczyk K. Applicability of machine learning in uncertainty quantification of turbulence models. 2022. doi:10.48550/ARXIV.2210.16358.
- [41] Lee M, Moser RD. Direct numerical simulation of turbulent channel flow up to $Re_\tau \approx 5200$. *Journal of Fluid Mechanics* 2015;774:395–415. doi:10.1017/jfm.2015.268.
- [42] Breuer M, Peller N, Rapp C, Manhart M. Flow over periodic hills – numerical and experimental study in a wide range of reynolds numbers. *Computers & Fluids* 2009;38(2):433 – 457. doi:10.1016/j.compfluid.2008.05.002.
- [43] Rossi R. Passive scalar transport in turbulent flows over a wavy wall. Ph.D. thesis; Università degli Studi di Bologna, Bologna, Italy; 2006.
- [44] Laval J, Marquillie M. Direct numerical simulations of converging–diverging channel flow. In: *Progress in Wall Turbulence: Understanding and Modeling*. 2011, p. 203 – 209. doi:10.1007/978-90-481-9603-6_21.
- [45] Menter F, Kuntz M, Langtry R. Ten years of industrial experience with the SST turbulence model. *Turbulence, Heat and Mass Transfer* 2003;4:625 – 632.

- 1009 [46] Coles D, Wadcock AJ. Flying-hot-wire study of flow past an NACA 4412
1010 airfoil at maximum lift. AIAA Journal 1979;17(4):321 – 329. doi:10.2514/
1011 3.61127.
- 1012 [47] NASA, National Aeronautics and Space Administration . Turbulence mod-
1013 eling resource database. [https://turbmodels.larc.nasa.gov/index.](https://turbmodels.larc.nasa.gov/index.html)
1014 [html](https://turbmodels.larc.nasa.gov/index.html); accessed April 2022.
- 1015 [48] Scott D. Multivariate density estimation: Theory, practice, and visualiza-
1016 tion: Second edition. 2015. doi:10.1002/9781118575574.
- 1017 [49] Stanford University, Aerospace Design Lab of the Department of Aero-
1018 nautics and Astronautics . SU2 code (version 7.4.0 "blackbird"). [https:](https://github.com/su2code/SU2)
1019 [/github.com/su2code/SU2](https://github.com/su2code/SU2); accessed December 2022.
- 1020 [50] Cook L, Mishra A, Jarrett J, Willcox K, Iaccarino G. Optimization un-
1021 der turbulence model uncertainty for aerospace design. Physics of Fluids
1022 2019;31:105111. doi:10.1063/1.5118785.
- 1023 [51] Eidi A, Zehtabiyani-Rezaie N, Ghiassi R, Yang X, Abkar M. Data-driven
1024 quantification of model-form uncertainty in reynolds-averaged simulations
1025 of wind farms. Physics of Fluids 2022;34:085135. doi:10.1063/5.0100076.