Abstract

Reynolds-averaged Navier-Stokes (RANS)-based transition modeling is widely used in aerospace applications but suffers inaccuracies due to the Boussinesq turbulent viscosity hypothesis. The eigenspace perturbation method can estimate the accuracy of a RANS model by injecting perturbations to its predicted Reynolds stresses. However, there lacks a reliable method for choosing the strength of the injected perturbation, while existing machine learning models are often complex and data craving. We examined two light-weighted machine learning models to help select the strength of the injected perturbation for estimating the RANS uncertainty of flows undergoing the transition to turbulence over a Selig-Donovan 7003 airfoil. On the one hand, we examined polynomial regression to construct a marker function augmented with eigenvalue perturbations to estimate the uncertainty bound for the predicted skin friction coefficient. On the other hand, we trained a convolutional neural network (CNN) to predict high-fidelity turbulence kinetic energy. The trained CNN acts as a marker function that can be integrated into the eigenspace perturbation method to quantify the RANS uncertainty. Our findings suggest that the light-weighted machine learning models are effective in constructing an appropriate marker function that is promising to enrich the existing eigenspace perturbation method to quantify the RANS uncertainty more precisely.

1 Introduction

Turbulence is a phenomenon that manifests over a wide range of length scales. These length scales vary from micro-meters to kilo-meters. The motions at each length scale are coupled and strongly interacting. Thus, any numerical simulation needs to account for these motions. Direct numerical simulation (DNS) computes all scales of motion, large eddy simulation (LES) computes only the large scales while using turbulence models for smaller scales, Reynolds-averaged Navier-Stokes (RANS) uses turbulence models for all scales. Due to this, RANS-based simulations are computationally inexpensive and are widely used in industry and academia for analyzing turbulent flows. However, RANS models are often inaccurate when predicting complex flows due to the turbulent viscosity hypothesis [1], especially when modeling the transitional flow regime as widely encountered in turbo-machines.

Less than a decade ago, Iaccarino et al. proposed a physics-based eigenspace perturbation approach [2,3] to estimate the model-form uncertainty introduced in RANS-based models by
quantifying the model-form uncertainty via sequential perturbations to the amplitude (turbulence kinetic energy) \[4, 5\], shape (eigenvalues) \[2, 3\], and orientation (eigenvectors) \[6, 7\] of the predicted Reynolds stress tensor. Studies of RANS uncertainty quantification (UQ) are mainly focused on the eigenvalue and eigenvector perturbations, such as RANS uncertainty in flow through scramjets \[8\], aircraft nozzle jets \[7\], over streamlined bodies \[3\], supersonic axisymmetric submerged jet \[9\], canonical cases of turbulent flows over a backward-facing step \[5, 6\], and benchmark cases of complex turbulent flow \[10\].

When the eigenspace perturbation approach is used to quantify the RANS model uncertainty, it is essential to choose appropriate strength of perturbation. This is often experience-based, which usually introduces the maximum and minimum magnitude of perturbation, resulting in overly conservative perturbations \[3, 6, 11\]. Recent machine Learning models can estimate RANS model uncertainty with improved accuracy \[12–20\]; however, they are often complex and demand a large size of training data. Complex machine learning models not only require additional computational resources in training but also become less comprehensive to researchers. This hinders the understanding and shrinks the room for improvement in the existing theories. Therefore, light-weighted machine learning models are interesting and helpful to incorporate into the physics-based eigenspace perturbation framework to improve the estimates of RANS uncertainty. This study examined two light-weighted machine learning models to bridge the RANS prediction and in-house DNS. Furthermore, we focused on perturbing the turbulence kinetic energy, complementing the eigenvalue and eigenvector perturbations studied by many works.

2 Methodology

2.1 Eigenspace Perturbation

The RANS-based transition model of Langtry and Menter \[21\] is used to simulate flows over a SD7003 airfoil at \(8^\circ\) angle of attack (AoA). With the Reynolds number based on the cord length of \(Re_c = 60000\), the flow underwent transition to turbulence on the suction side of the airfoil. Langtry and Menter’s model approximates the Reynolds stresses as

\[
\langle u_i u_j \rangle = \frac{2}{3} k \delta_{ij} - 2\nu_t \langle S_{ij} \rangle ,
\]

where \(k\) is the turbulence kinetic energy, \(\delta_{ij}\) is the Kronecker delta, \(\nu_t\) is the turbulent viscosity, and \(\langle S_{ij} \rangle\) is the rate of mean strain tensor. The solution domain is a two-dimensional C-topology grid of 389 (streamwise) \(\times\) 280 (wall-normal) control volumes. At the inlet, a very low value of turbulence intensity \(Tu = 0.03\%\) was assumed. At the outlet, a zero-gradient boundary condition was used for \(\langle U_i \rangle\), \(k\), \(\omega\), and pressure. At the wall, a no-slip boundary condition was assumed. Grid dependency study showed that finer mesh in the near-wall region did not yield more accurate results.

The RANS-based transition model is implemented within the OpenFOAM framework and the transient PIMPLE solver was used for pressure-velocity coupling. The solution fields were iterated until convergence: the results of lift and drag coefficients as a function of time were used to track the instant convergence status. At the convergence state, the residuals of energy and momentum dropped more than four orders of magnitude, which required approximately 35000 iterations for all simulations.
In the eigenspace perturbation approach [2, 6], the perturbed Reynolds stresses are defined as

\[ \langle u_i u_j \rangle^* = 2k^* \left( \frac{1}{3} \delta_{ij} + v_{i\alpha}^* \hat{b}_{\alpha j}^* v_{j\alpha}^* \right), \]  

(2)

where \( k^* \) is the perturbed turbulence kinetic energy, \( \hat{b}_{\alpha j}^* \) is the perturbed eigenvalue matrix, \( v_{i\alpha}^* \) is perturbed eigenvector matrix, and \( \delta_{ij} \) is the kronecker delta.

In this study the perturbed turbulence kinetic energy can be defined as

\[ k^* = k + \Delta_k = M_k, \quad M_k \sim f(x, y), \]  

(3)

where \( M_k \) is a marker function of the \( x \) and \( y \) coordinate in a computational domain. In this study, the RANS predictions bereft of any perturbations are referred to as “baseline” solutions.

2.2 Projecting a RANS model to a DNS model

2.2.1 Polynomial regression

We applied polynomial regressions to fitting the RANS prediction and in-house DNS separately to \( n \)th order polynomials, with the \( n \)th order as a hyperparameter, as shown in Fig. 1 (a)-(f). From Figs. 1 (a) and (b), the third order polynomials in general are loosely fitted, particularly for the \( ab \) and \( cd \) zone. As the order of polynomial increases to 5, it is clear that the \( cd \) zone are better fitted. For both the \( cd \) and \( ef \) zone, the fitted polynomials almost overlap the training datasets. Figs. 1 (e) and (f), most zones are well fitted except for the \( ab \) zone shown in Fig. 1 (e). This suggests that the dynamics of turbulence kinetic energy in the \( ab \) zone, where a LSB is evolving, are more complex than polynomial.

Figure 1: Fitting polynomials of order 3, 5 and 7 to the datasets for RANS and DNS. There are 32 positions on the suction side of the airfoil. RANS\(_{actual}\) represents the baseline prediction. DNS\(_{actual}\) data are included for reference. zone \( ab \): 0.14 < \( x/c \) < 0.18; zone \( cd \): 0.18 < \( x/c \) < 0.3; zone \( ef \): 0.3 < \( x/c \) < 0.6. \( c \) is the cord length.

2.2.2 Convolutional neural network

We employed a one-dimensional convolutional neural network (1D-CNN) to learn the projection from RANS estimated function \( f^{RANS}(x, y) \) to DNS estimated function \( f^{DNS}(x, y) \). For a given \( x \), we can rewrite the estimated function as \( g_x(k, y) \). Assuming that there exists a morphism \( F \) from \( g_x^{RANS}(k, y) \) to \( g_x^{DNS}(k, y) \), then every \( x \) and \( g_x(k, y) \) is smooth. Our CNN is trained to depict \( F \) in
the hope of projecting $g^{\text{RANS}}_x(k,y)$ to $g^{\text{DNS}}_x(k,y)$. This is conducted by training the paired RANS- and DNS-estimated functions at the selected $x$ coordinates. Taking advantage of the smoothness assumption of $g_x(k,y)$, our 1D-CNN is trained to predict DNS estimated function at $(x,y_{\text{target}})$ given a series of RANS estimated function at $(x,y_i)$, where $y_i \in [y - \epsilon, y + \epsilon], \epsilon > 0$ belongs to the neighbor of $y_{\text{target}}$. Our 1D-CNN has four-layers and in total 86 parameters: a single model for all zones at any $x$ to project RANS to DNS. We trained our 1D-CNN with normalized pairs of $(g^{\text{RANS}}_x(k,y), g^{\text{DNS}}_x(k,y))$ at only three positions $x = 0.4, 0.56, 0.58$ with mean squared error as the loss function and a 80%–20% split as training–testing dataset. We validated our trained 1D-CNN by comparing the L1 loss of RANS, denoted as $L^1_{\text{ran}} = |CF^{\text{RANS}}_k - CF^{\text{DNS}}_k|$ with the L1 loss of 1D-CNN projected RANS, denoted as $L^1_{\text{pred}} = |CF^{\text{CNN}}_k - CF^{\text{DNS}}_k|$. 

![Figure 2: CNN projected DNS (CNN_DNS) compared with ground truth (DNS_actual). There are 32 positions on the suction side of the airfoil.](image)

### 3 Results

#### 3.1 Polynomial regression

In Figures 3 (a) - (c), the mean of the seventh-order regression-based normalized turbulence kinetic energy profile as the representative for each zone shows the discrepancy between RANS and in-house DNS. The discrepancy marks the degree of untrustworthiness in the $y/c$ direction. Note that all profiles are shifted down to the origin of $y/c$, denoted $y/c_o$. The corresponding marker for each zone is shown in Figs. 3 (d) - (f). Marker functions can be constructed by fitting appropriate models to the discrepancy data for each zone, i.e., a seventh-order polynomial for the $ab$ and $ef$ zone, and a Fourier series for the $cd$ zone.

Augmenting the eigenvalue perturbation with the marker function ($1c_{-}M_k$, $2c_{-}M_k$ and $3c_{-}M_k$) using Eqs. 2 and 3, the estimated model-form uncertainty (red envelope) for the predicted skin friction coefficient is constructed and shown in Fig. 4. The $1c$ and $3c$ eigenvalue perturbations are included for reference. It is clear that the uncertainty bound successfully encompasses the

---

1 The strength of eigenvalue perturbation is denoted $\Delta_B$, which varies from 0 to 1.
Figure 3: Mean of seventh order polynomials for normalized turbulence kinetic energy (a) - (c), and the corresponding marker function (d) - (f). (a) and (d) zone ab; (b) and (e) zone cd; (c) and (f) zone ef.

ILES/LES data of Galbraith et al. [22] and Garmann et al. [23] for 0.25 < x/c < 0.45. This region falls into the cd and part of the ef zone, where the LSB is forming and the flow is re-attaching on the wall surface. In comparison to the eigenvalue perturbations, the red envelop exhibits a significant increase in the magnitude of $C_f$, showing a tendency to retain the shape of the reference data. This marks a significant improvement in the RANS model prediction for $C_f$. The shape of the red envelope is not as smooth as the eigenvalue perturbations, reflecting the effect of spatial variability in $M_k$.

Figure 4: Skin friction coefficient. Displayed are uncertainty bounds for 1c_Mk, 2c_Mk, and 3c_Mk perturbations (red envelope). $\Delta B1$ stands for $\Delta B = 1$. Profile of the baseline prediction and eigenvalue perturbations (1c and 3c) are provided for reference.
3.2 CNN

The 1D-CNN can predict DNS at any zone given RANS, thus acting as the marker function $M_k$ in Eq. 3. From Figs. 2 (a) and (b), the CNN predicted DNS profile for $k$ shows overall good resemblance to the DNS dataset, although an over-prediction exists at the beginning of the $ab$ zone. From the Fig. 3 the series of CNN predicted DNS profiles in the first row are then smoothed with the moving average with a window size of six consecutive estimations. Our CNN predicted DNS profiles resemble the ground truth DNS despite being trained with only a few pairs of RANS and DNS results. From Fig. 5, the discrepancy in general reduces as the flow proceeds further downstream. Consequently, the CNN predicted DNS given the RANS estimated function acts as the marker function $M_k$ in Eq. 3. From the Fig. 5, the second row shows the computed error of the baseline solution and the CNN predicted DNS, and it is clear that the error for CNN predicted DNS is significantly reduced in magnitude compared to that for the baseline solution.

![Figure 5](image-url)

Figure 5: First row: CNN projected DNS (pred) compared with ground truth (dns). Second row: Validation of 1D-CNN by comparing L1 loss between $L_{1,c}^{\text{rans}}$ and $L_{1,c}^{\text{pred}}$.

4 Conclusion

In this study, the flow being considered is over a SD7003 airfoil at 8° angle of attack and the Reynolds number based on the cord length of $Re_c = 60000$. A laminar separation bubble evolves within the separated boundary layer whereby the flow undergoes transition to turbulence. The goal of this study was to assess the effect of the polynomial regression and the CNN approaches on the RANS model uncertainty quantification. The learning algorithms need to be coupled to the eigenspace perturbation approach of Emory et al. [2] that is implemented within the OpenFOAM framework to construct a marker function for the turbulence kinetic energy perturbation.
Seventh-order polynomial regressions overall captures the discrepancy in the predicted turbulence kinetic energy between RANS and in-house DNS. Correspondingly, the the marker function is augmented with the eigenvalue perturbation to significantly increase the uncertainty bound for $C_f$. Around the peak of the $C_f$ curve, the uncertainty bound successfully encompassed the reference data in the aft portion of the LSB and the re-attachment point.

To the best of our knowledge, we are among the first to examine the projection from RANS to DNS using the CNN approach. Our experiment results suggest that the CNN approach can help us project the RANS estimated marker function to the in-house DNS data. A projection that can approximate the in-house DNS reasonably well from RANS might exist independent of $x$. Our methodology can be easily extended to analyze flows over different types of airfoils.

Our findings are subject to following limitations. The results of our machine learning based uncertainty quantification are subject to the RANS-based transition model of Langtry and Menter when simulating flows over the SD7003 airfoil at 8° angle of attack (AoA). With the Reynolds number based on the cord length of $Re_c = 60000$. Future work may include evaluating more machine learning models in generating marker functions with different types of airfoils, as well as integrating the CNN approach into the eigenspace perturbation framework to result in accurate perturbations and hence accurate RANS UQ.

5 Impact statement

Transitional flow is important because the transitional flow regime is frequently encountered in aerospace applications, while this flow regime is hard to model. LES and DNS can provide high fidelity solution, but the calculations are often expensive. Therefore, the estimation of the RANS uncertainty is a valuable alternative tool for improving the RANS predictive capability in engineering applications. More expensive LES or DNS would only be adopted if the RANS uncertainty is too large. A recent physics-based eigenspace perturbation method evaluates the accuracy of RANS model for practical usage based on experts selected perturbations, which is less reliable in terms of giving inaccurate strength of perturbation. We proposed a machine-learning based eigenspace perturbation method that can effectively increase the precision of the estimates of RANS uncertainty and build confidence in aeronautical simulations by which risky decision-making is always accompanied. Our method is highly generalizable to a variety of flow scenarios. In addition, the method can also be employed to shed lights on predicting deviations of RANS from DNS, promising a correction to RANS to improve its accuracy.

References

[1] TJ Craft, BE Launder, and K Suga. Development and application of a cubic eddy-viscosity model of turbulence. *International Journal of Heat and Fluid Flow*, 17(2):108–115, 1996.

[2] Michael Emory, Johan Larsson, and Gianluca Iaccarino. Modeling of structural uncertainties in reynolds-averaged navier-stokes closures. *Physics of Fluids*, 25(11):110822, 2013.

[3] C Gorlé, S Zeoli, M Emory, J Larsson, and G Iaccarino. Epistemic uncertainty quantification for reynolds-averaged navier-stokes modeling of separated flows over streamlined surfaces. *Physics of Fluids*, 31(3):035101, 2019.
[4] C Gorlé and G Iaccarino. A framework for epistemic uncertainty quantification of turbulent scalar flux models for reynolds-averaged navier-stokes simulations. *Physics of Fluids*, 25(5):055105, 2013.

[5] Luis F Cremades Rey, Denis F Hinz, and Mahdi Abkar. Reynolds stress perturbation for epistemic uncertainty quantification of rans models implemented in openfoam. *Fluids*, 4(2):113, 2019.

[6] Gianluca Iaccarino, Aashwin Ananda Mishra, and Saman Ghili. Eigenspace perturbations for uncertainty estimation of single-point turbulence closures. *Physical Review Fluids*, 2(2):024605, 2017.

[7] Aashwin Ananda Mishra and Gianluca Iaccarino. Uncertainty estimation for reynolds-averaged navier–stokes predictions of high-speed aircraft nozzle jets. *AIAA Journal*, 55(11):3999–4004, 2017.

[8] Michael Emory, Vincent Terrapon, Rene Pecnik, and Gianluca Iaccarino. Characterizing the operability limits of the hyshot ii scramjet through rans simulations. In *17th AIAA international space planes and hypersonic systems and technologies conference*, page 2282, 2011.

[9] AA Mishra and G Iaccarino. Rans predictions for high-speed flows using enveloping models. *arXiv preprint arXiv:1704.01699*, 2017.

[10] Roney L Thompson, Aashwin Ananda Mishra, Gianluca Iaccarino, Wouter Edeling, and Luiz Sampaio. Eigenvector perturbation methodology for uncertainty quantification of turbulence models. *Physical Review Fluids*, 4(4):044603, 2019.

[11] Catherine Gorlé, Michael Emory, Johan Larsson, and Gianluca Iaccarino. Epistemic uncertainty quantification for rans modeling of the flow over a wavy wall. *Center for turbulence research, annual research briefs*, pages 81–91, 2012.

[12] Heng Xiao, J-L Wu, J-X Wang, Rui Sun, and CJ Roy. Quantifying and reducing model-form uncertainties in reynolds-averaged navier–stokes simulations: A data-driven, physics-informed bayesian approach. *Journal of Computational Physics*, 324:115–136, 2016.

[13] Jin-Long Wu, Jian-Xun Wang, and Heng Xiao. A bayesian calibration–prediction method for reducing model-form uncertainties with application in rans simulations. *Flow, Turbulence and Combustion*, 97(3):761–786, 2016.

[14] Heng Xiao, Jian-Xun Wang, and Roger G Ghanem. A random matrix approach for quantifying model-form uncertainties in turbulence modeling. *Computer Methods in Applied Mechanics and Engineering*, 313:941–965, 2017.

[15] Jian-Xun Wang, Jin-Long Wu, and Heng Xiao. Physics-informed machine learning approach for reconstructing reynolds stress modeling discrepancies based on dns data. *Physical Review Fluids*, 2(3):034603, 2017.

[16] Jian-Xun Wang, Jinlong Wu, Julia Ling, Gianluca Iaccarino, and Heng Xiao. A comprehensive physics-informed machine learning framework for predictive turbulence modeling. *arXiv preprint arXiv:1701.07102*, 2017.
[17] Jin-Long Wu, Heng Xiao, and Eric Paterson. Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework. *Physical Review Fluids*, 3(7):074602, 2018.

[18] Jan Felix Heyse, Aashwin A Mishra, and Gianluca Iaccarino. Estimating rans model uncertainty using machine learning. *Journal of the Global Power and Propulsion Society*, 2021(May):1–14, 2021.

[19] Jan Felix Heyse, Aashwin Ananda Mishra, and Gianluca Iaccarino. Data driven physics constrained perturbations for turbulence model uncertainty estimation. In *AAAI Spring Symposium: MLPS*, 2021.

[20] Fanzhi Zeng, Wei Zhang, Jinping Li, Tianxin Zhang, and Chao Yan. Adaptive model refinement approach for bayesian uncertainty quantification in turbulence model. *AIAA Journal*, pages 1–15, 2022.

[21] Robin B Langtry and Florian R Menter. Correlation-based transition modeling for unstructured parallelized computational fluid dynamics codes. *AIAA journal*, 47(12):2894–2906, 2009.

[22] Marshall Galbraith and Miguel Visbal. Implicit large eddy simulation of low-reynolds-number transitional flow past the sd7003 airfoil. In *40th fluid dynamics conference and exhibit*, page 4737, 2010.

[23] Daniel J Garmann, Miguel R Visbal, and Paul D Orkwis. Comparative study of implicit and subgrid-scale model large-eddy simulation techniques for low-reynolds number airfoil applications. *International Journal for Numerical Methods in Fluids*, 71(12):1546–1565, 2013.