Workshop on Frontiers of Uncertainty Quantification in Fluid Dynamics
11–13 September 2019, Pisa, Italy

# Using Machine learning to predict and understand turbulence modelling uncertainties

***Ashley Duncan Scillitoe***[1]

[1]The Alan Turing Institute, London, UK – ascillitoe@turing.ac.uk

*Key words:* Epistemic uncertainty, Turbulence modelling, Sensitivity analysis, Machine learning.

Many industrial flows are turbulent in nature, and Reynolds-Averaged Navier-Stokes (RANS) based turbulence models play an important role in the prediction of such flows. Many simplifying assumptions are made in the development of a RANS model. These assumptions can lead to epistemic uncertainties which are still a major obstacle for the predictive capability of RANS models. If experimental measurements or higher fidelity simulation results are not available for comparison, it can be difficult to evaluate the accuracy of RANS predictions. Recently, data-driven strategies have been proposed to either quantify model uncertainty [2], or improve model accuracy [3]. Many of these strategies involve training supervised machine learning (ML) models on high fidelity CFD data, such as that from large eddy simulations [1].

Ling and Templeton [4] proposed a technique whereby a machine learning classifier is trained to detect whether a number of RANS modelling assumptions are broken. The classifier can then be used to predict regions of high and low uncertainty in RANS simulations where corresponding high fidelity data is not available. In the present work, the original technique is further refined, and applied to a number of flow configurations such as a recent family of bumps dataset [5]. The non-dimensional “features" used by Ling and Templeton to describe each flow to the ML model are revisited, with improvements made to ensure the ML classifier is more generalisable to other flows. New error metrics are also proposed to allow for other areas of RANS modelling uncertainty to be explored. Figure 1 shows a prediction made for one of the bump cases, with the region in red being a region where the Boussinesq hypothesis used by many RANS models would be invalid due to the high turbulent anisotropy.

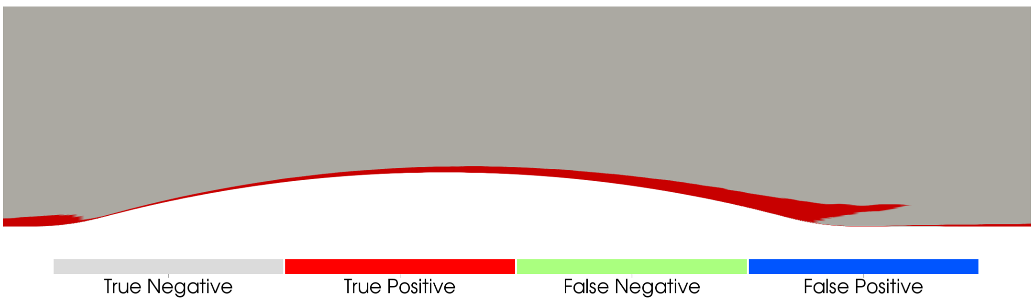


Figure 1: Regions in the flow over a bump [5] where the Boussinesq hypothesis is predicted to be invalid due to significant turbulent anisotropy, predicted by a random forest classifier.

Within the ML community there is an increasing interest in interpretability. Indeed, a common criticism of ML augmented RANS models is that they are a “black box", with the machine able to make accurate predictions but not explain why it has made them. The proposed RANS error classifier is examined using recently proposed ML interpretation methods such as individual conditional expectation [6] and Shapley additive explanations (SHAP) [7] plots. These novel techniques provide global and local explanations of model predictions. A SHAP summary plot, which summarises the global effect of each flow feature (the model inputs) on the turbulent anisotropy metric (a model output), is shown in Figure 2. More detailed SHAP plots are discussed, such as SHAP dependence and local force plots. Now, the proposed classifier can not only be used to predict regions of uncertainty, but also to explain what physical flow features have caused this uncertainty. This brings with it the possibility of using the classifier to aid in the further understanding and development of turbulence models, in addition to its use as a tool in predicting “trust" regions in RANS simulations.

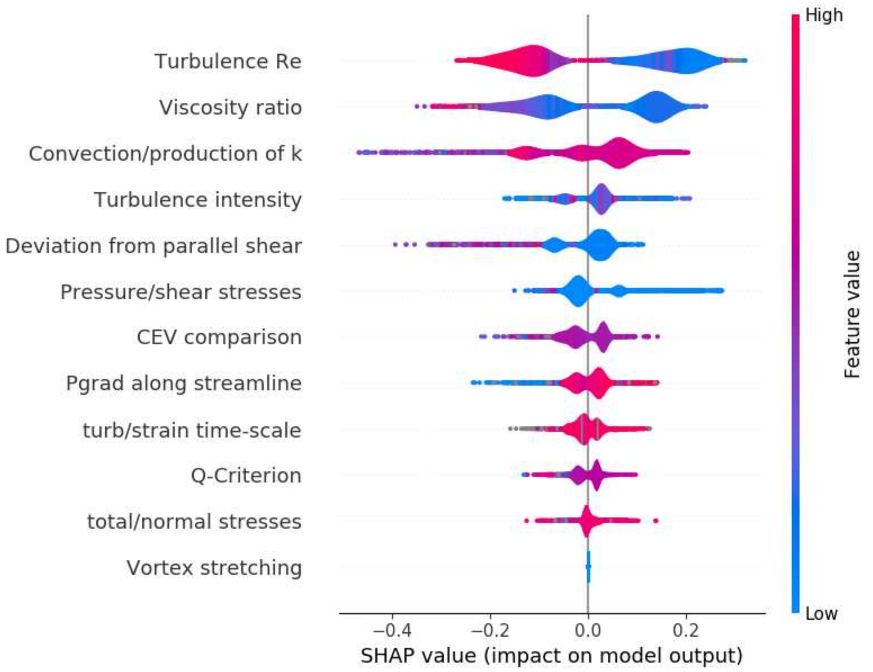


Figure 2: A SHAP summary plot showing the impact of flow features on the turbulent anisotropy error metric predicted in Figure 1.

[1] A. D. Scillitoe, P. G. Tucker and P. Adami, “Large eddy simulation of boundary layer transition mechanisms in a gas-turbine compressor cascade" Journal of Turbomachinery **141**, 6 (2019).

[2] W. Edeling, G. Iaccarino and P. Cinnella, “Data-Free and Data-Driven RANS Predictions with Quantified Uncertainty" Flow, Turbulence and Combustion **100**, 3 (2018).

[3] J. Wu, H. Xiao and E. Paterson, “Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework" Physical Review Fluids **7**, 3 (2018).

[4] J. Ling and J. Templeton, “Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty" Physics of Fluids **27**, 8 (2015).

[5] R. Matai and P. Durbin, “Large-eddy simulation of turbulent flow over a parametric set of bumps" Journal of Fluid Mechanics **866** (2019).

[6] A. Goldstein et al., “Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation" Journal of computational and graphical statistics **24**, 1 (2015).

[7] S. M. Lundberg and S. I. Lee, “A Unified approach to interpreting model predictions" Advances in neural information processing systems **30** (2017).