20th Australasian Fluid Mechanics Conference
Perth, Australia
5-8 December 2016

Bayesian Inversion and Regression Trees for Mixing Length Model Development

P. Cheema[1], G. Mathews[2], B. Thornber[1], G. A. Vio[1]

[1]School of AMME, The University of Sydney, 2006 NSW, Australia

[2]Analytics, Data61, CSIRO, Eveleigh, 2015, NSW, Australia

Abstract

Current trends in computational fluid dynamics (CFD) sees numerical models developed primarily through parameter tuning, which is in turn based on experiments and intuition. Whilst such approaches have been traditionally useful, they are now becoming difficult to manage due to their sheer complexity, and specificity to particular flow problems. This in turn leads to CFD solutions with increased computational complexity, and little benefit in the overall accuracy or validity of the final solution. Moreover, these analytical methods do not make full use of existing direct numerical simulation (DNS) datasets. Therefore, in this paper an innovative method of improving a CFD wall model through the use of classic big-data techniques will be investigated. These techniques will make full use of current DNS by looking for important flow features, and fitting optimal regressors in order to validate and simplify current CFD wall models. This will be achieved by using a Bayesian methodology to correct a mixing length model, coupled with a regression tree to infer solutions at various friction Reynolds numbers.

Introduction

Machine learning is a science which aims to train a computer in order to find patterns in datasets. It has recently come to the forefront of research interest due to the availability of large datasets, as well as through increased computational capabilities in the modern context. Traditional uses of machine learning have been seen in fields as varied as categorising handwritten digits [5], and as a game-model to play chess [1].

Although machine learning carries a lot of potential for various research fields, it has not been used very widely in the computational fluid dynamics (CFD) area. An early example of its use in the CFD context includes the work of Shang [13], where a simple two-phase vapour problem was solved with an artificial neural network, as well as Rajkumar and Bardina [11], where a similar neural network was used to predict aerodynamic coefficients for wind tunnel data. Although these pieces of literature have claimed positive results, they have failed to show an ability to generalise such models to modern, and more complex problems.

Recently however, the CFD research groups at Stanford University, Sandia National Laboratories, and The University of Michigan have been largely focused on trying to combine these two fields in an efficient, and generalisable manner. Notable examples of this work include the works of Ling et al. [6] who have worked to encode physical invariance properties in CFD problems as the input features into a random forest regressor, Jaideep et al. [12] who have implemented a Bayesian inversion scheme in order to try and estimate the joint posterior probability of three $k-\varepsilon$ model coefficients through Markov chain Monte Carlo, Tracey et al. [14] who have trained a standard multi-layer perceptron model using a cost function based on skin friction coefficient, and the work of Parish et al. [9], where a Bayesian model inversion scheme was used to correct the production term in the $k-\varepsilon$ model.

This revitalisation of attempts to combine the two fields have no doubt come about due to the explosion of interest in machine learning in the modern context, due in large part to the variety of problems machine learning has been able to solve. Due to the clear potential that machine learning holds, this paper has been developed in an effort to use classic machine learning algorithms on direct numerical simulation (DNS) datasets to extract useful, CFD-specific information. In particular, a Bayesian inversion method will be used in this paper, which shall work to apply a corrective term to the mixing length model developed by Nikuradse [8] for the case of 1D Reynolds averaged Navier Stokes (RANS) channel flow. This method shall work by finding the maximum a posteriori (MAP) solution, which will serve as a corrective factor necessary to improve Nikuradse's formulation.

Methodology

In this section the methodology used to arrive at the corrective formulations for Nikuradse's mixing length formulation are explored. This will involve an analysis of the fluid equations, and the set-up for a machine learning scheme based-on Bayesian inverse modelling, and regression trees.

Mixing Length Model

The 1D Reynolds Averaged Navier Stokes (RANS) formulation for fluid flow is outlined in equation 1, where $\overline{U}$ is the mean flow velocity resulting from the RANS approximation, ρ is fluid density, $\overline{P}$ is average pressure, and ν is kinematic viscosity. This equation is solved computationally by first-order forward marching in time, and second-order central difference schemes for all space derivatives.

$$\frac{\partial \overline{U}}{\partial t} = -\frac{1}{\rho}\frac{\partial \overline{P}}{\partial x} + \nu\frac{\partial^2 \overline{U}}{\partial y^2} + \frac{\partial}{\partial y}\left(\nu_T \frac{\partial \overline{U}}{\partial y}\right) \tag{1}$$

Note that the pressure differential term, $\partial\overline{P}/\partial x$, has a derivative in the x direction which renders equation 1 2D by definition. However, it is possible to interpret this term as a scalar forcing term $\mathcal{H}$, which shall work to eliminate the x direction dependency in equation 1, thus rendering it as 1D. This forcing term is defined in equation 2, where τ is the wall shear stress, and δ is the channel half height. This equation may be derived from a simple force balance between the fluid body force, and the wall shear stress, assuming steady state conditions.

$$\mathcal{H} = \frac{\tau}{\delta} \tag{2}$$

In equation 1, there is a dependency on the term ν_T, which is known as the the *turbulent viscosity* parameter. This term arises when applying the Bousinesq hypothesis to the Reynolds stress terms in the RANS equations [10]. However, this term is generally unknown and needs to be approximated in order to ensure that there exists closure for equation 1. The approximation used in this paper to model the ν_T parameter, and thus *close* equa-

tion 1, is that of the mixing length model. The mixing length model is a coarse approximation for momentum transfer in wall bounded flows. It is developed based on a naive assumption that *fluid elements* will conserve their properties under some distance of characteristic length before mixing with the surrounding fluid and dissipating [10]. This approximation is shown in equation 3.

$$\nu_T = \ell_{\text{mix}}^2 \left| \frac{\partial \overline{U}}{\partial y} \right| \tag{3}$$

Simple mixing length models assume a function for ℓ_{mix} which varies linearly with the viscous units, y^+, from the channel wall, where $y^+ = y u_\tau / \nu$, with u_τ being friction velocity as defined in [10]. However a more rigorous and experimentally determined version of the mixing length approximation, specific to pipe flow, can be obtained by using Nikuradse's formulation for the mixing length parameter [8]. This is defined in equation 4.

$$\frac{\ell_{\text{mix}_0}}{\delta} = 0.14 - 0.08\left(1 - \frac{y}{\delta}\right)^2 - 0.06\left(1 - \frac{y}{\delta}\right) \tag{4}$$

The main issue with equation 4 however, is its inability to reach a zero value at the wall. This is necessary since the turbulent viscosity parameter should be driven to zero at the wall as the flow closest to the wall is dominated by viscous effects. This condition can be incorporated into the mixing length equation by using a van Driest damping function. Such a function scales with viscous units, y^+, and decreases exponentially until it reaches a zero value at the wall. This modification is made clear in equation 5.

$$\ell_{\text{mix}_1} = \ell_{\text{mix}_0}\left(1 - \frac{y^+}{A^+}\right) \tag{5}$$

The value for A^+ is taken to be 26 according to Griffol and Giralt [4]. Equation 5 represents the *base equation* which will require a corrective factor applied. Currently this model works quite well in the near wall section, due to the behaviour of the damping function, however it is still unable to predict the flow behaviour in the outer layer. The application of the corrective factor is made specific in equation 6.

$$\ell_{\text{mix}_2} = \ell_{\text{mix}_1} \cdot \beta\left(\frac{y}{\delta}\right) \tag{6}$$

The corrective factor is defined to be a parameter, $\beta(y/\delta)$, which is a function of non-dimensional channel flow height. The value for β is required to be found at every single node point along the channel flow domain. The procedure for finding β is found using a Bayesian inversion scheme, which seeks the maximum a posterior (MAP) solution. This is clarified in the following subsection. Once the MAP solution is found for a few friction Reynolds numbers (using the definition of [10] for friction Reynolds numbers), the β parameters at different friction Reynolds numbers are determined using a regression tree. This is clarified in the following subsection.

Machine Learning

A Bayesian inversion scheme is fundamentally based on an application of Bayes rule, defined in equation 7. Bayes rule is a statement of conditional probability, which consists of a likelihood function, a prior probability, and an evidence term (defined here as $p(D|f(\beta)), p(\beta)$, and $p(D)$ respectively). In equation 7 the function, $f(\beta)$, refers to the application of the mixing length correction term, β, in the full 1D RANS model previously defined in equation 1. The D variable refers to the DNS dataset which was generated by Moser et al., which is freely available online [7].

$$p(\beta|D) = \frac{p(D|f(\beta))p(\beta)}{p(D)} \tag{7}$$

In equation 7 it is assumed that the prior probability, $p(\beta)$, follows a Gaussian distribution, and that the likelihood function, $p(D|f(\beta))$, is also Gaussian distributed. Although the definition of Gaussian distributions may seem restrictive, the optimisation method used in this paper is general enough to allow for a search of the maximum a posterior (MAP) solution to any distribution. The search for the MAP solution can be formulated as an optimisation problem [2]. This is made clear in equation 8, where the symbol of an overbar denotes a mean quantity.

$$\begin{aligned}\log(p(\beta|D)) \propto &-\frac{1}{2}[(f(\beta)-D)^T C_f (f(\beta)-D) + \\ &(\beta-\overline{\beta})^T C_\beta (\beta-\overline{\beta})] \\ &= \mathbb{F}(\beta)\end{aligned} \tag{8}$$

and,

$$p(\beta) = \mathcal{N}(diag(\boldsymbol{I}), \kappa_\gamma^2 \boldsymbol{I}) \propto \exp\left(-\frac{(\beta - diag(\boldsymbol{I}))^2}{\kappa_\gamma^2}\right) \tag{9}$$

$$p(D|f(\beta)) = \mathcal{N}(D, \kappa_\delta^2 \boldsymbol{I}) \propto \exp\left(-\frac{(D - f(\beta))^2}{\kappa_\delta^2}\right) \tag{10}$$

The value for κ_γ was chosen to be 10^{-2} in order to ensure a well regularised MAP solution. For a discussion on the similarities behind the selection of a Gaussian prior and regularisation performed in linear regression problems, the reader is referred to [2]. In addition to this, the value for κ_δ is taken as 10^{-1} since it is assumed that the DNS solution is known to be reasonably accurate. Moreover, the value for the mean β prior is taken to be unity, in order to ensure that optimisation begins without a modified base model (since β is multiplied into the equation defined in equation 6). This value is generally chosen to reflect an initial belief in the value of β. Note that all β variables stated in this paper are vectors with length equal to the number of nodes in the domain.

Therefore the overall optimisation problem may be stated succinctly as in equation 11.

$$\beta_{\text{MAP}} = \underset{\beta}{\text{argmin}}(\mathbb{F}(\beta)) \tag{11}$$

In order to search for the MAP solution, an optimisation scheme is carried out on equation 8. As mentioned before, this effectively finds the mode of the Bayesian distribution, given the covariance, and mean β values defined in the prior and likelihood functions. The optimisation scheme used in this paper is the Nelder-Mead simplex method, which works to explore a surface in a derivative-free manner. Such a method is robust against undefined function values, which is important in this context since

the perturbation of the mixing length model can lead to unrealistic, and undefined solution instances. Although the formulation in this paper only seeks the MAP solution, it is possible to obtain the Gaussian uncertainty bounds about the MAP by using methods which approximate the posterior distribution.

In addition to performing the Bayesian inversion scheme, a regression tree is used to estimate the value of β_{MAP} at different friction Reynolds numbers. In particular, since the DNS dataset from Moser et al. consists of high resolution mean flow solutions at three different friction Reynolds numbers (see [7]), the proposed method will learn the β_{MAP} inversion solution at two of these Reynolds numbers, and then predict what the solution should be at the third friction Reynolds number. Regression trees are used in this analysis, since they are incredibly fast to train. Further information on regression trees may be seen in the seminal work of Olshen et al [3].

Results

Through the Bayesian inversion scheme discussed in the previous section, it is possible to modify the base mixing length model (equation 6) by finding the MAP solution for the β parameter. This results in the plots shown in figure 1. From this figure, it can be seen that the although the Nikuradse formulation is generally followed by the corrected solutions, at certain sections along the Nikuradse base model curve, the corrective curves deviate as necessary, to ensure that the DNS datasets are reproducible. In particular, for the lower friction Reynolds number case, Nikuradse's model appears to need additional damping near the wall. However as the friction Reynolds number increases, the magnitude of this required damping appears to decrease. Additional friction Reynolds DNS cases would be required to confirm that this observation is true as a general statement.

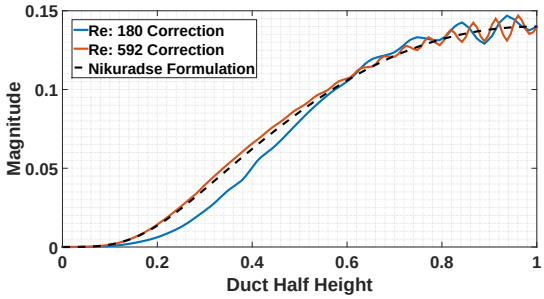

Figure 1: A comparison between the base model mixing length formulation (defined in equation 5), and those learned from a Bayesian inversion scheme using the DNS data of Moser et al.

Towards the end of the Bayesian inversion models in figure 1 (near a duct half height of unity), there is a collection of spurious oscillations which normally implies that the obtained solution is a high variance, overfit solution. Such solutions are difficult to generalise to novel scenarios [2]. However, this phenomenon is conjectured to be a result of the $\partial\overline{U}/\partial y$ term in the mixing length model, being driven towards a zero value at the channel half height. That is, the turbulent viscosity parameter value near the duct half height can take almost any value since it is being multiplied by a near zero value.

The application of these corrected mixing length functions applied to the 1D RANS flow equation (that is, applying the mixing length curves of figure 1, to equation 1) results in figures 2 and 3. From these figures it can be seen that when the corrected mixing length expression is used, the resulting mean channel flow matches the DNS dataset significantly better than when the base mixing length model is used. This can be seen in the portions of the image which are closer to the channel half height, where it can be seen that the base mixing length model misses the DNS model, running almost parallel to it.

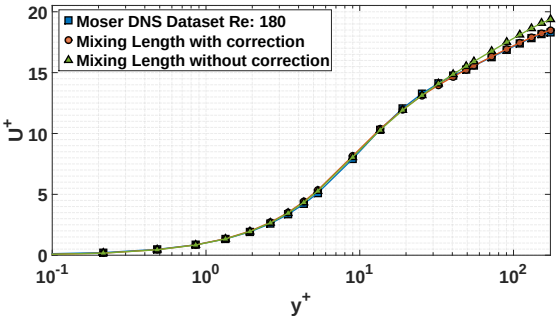

Figure 2: Comparison netween Moser DNS data, and the base and corrected models, at a friction Reynolds number of 180.

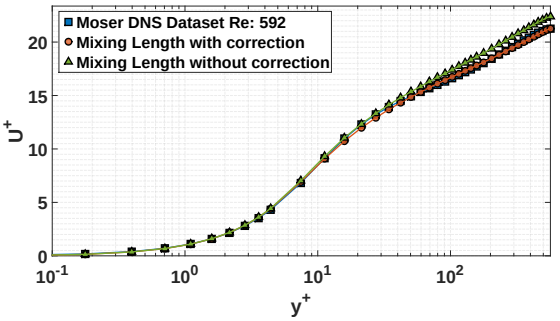

Figure 3: Comparison netween Moser DNS data, and the base and corrected models, at a friction Reynolds number of 592.

Once the Bayesian inverse models have been learned from the DNS datasets (figure 1), the β parameter at various other friction Reynolds numbers can be determined by a regression tree algorithm, learned over the inverse mixing length models in figure 1. During training of the tree, the minimum entropy splits were determined by using a 10-fold cross validation scheme, and the minimum number of branches per non-leaf node in the tree was determined to be five by manual tuning. This ensures that the tree is inexpensive to train and run, and aids in its ability to generalise well to new scenarios. The ability for the regression tree model to predict a mixing length function at the friction Reynolds number of 395, using the mixing length curves in figure 1 as the training dataset, is shown in figure 4.

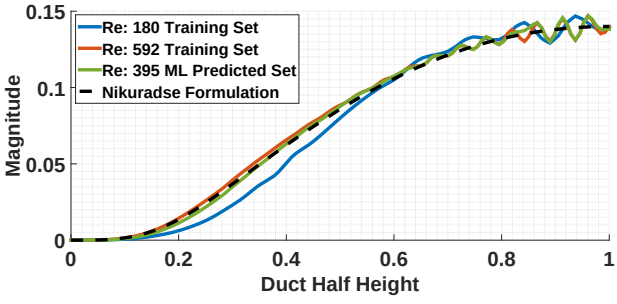

Figure 4: The mixing length curve predicted at a friction Reynolds number of 395 by a regression tree.

As can be seen, this new curve at the friction Reynolds number of 395, shares characteristics of the curves at the other friction Reynolds numbers. There is an initial dip, underneath the base solution, and then oscillatory motion towards the channel half height. Naturally, it is difficult to quantify and comment on how well this estimated mixing length curve works without, using it in the mixing length model. Hence, the results of implemented this learned curve is revealed in figure 5 (and its zoomed in version, figure 6).

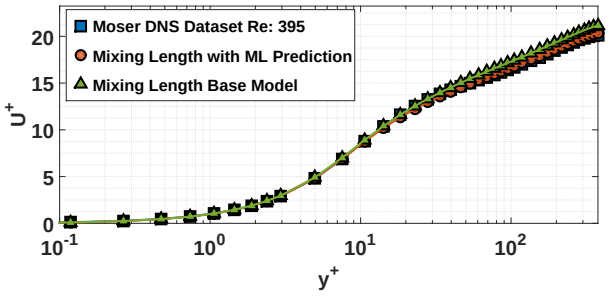

Figure 5: Comparison betwen the mean flow solution obtained with a predicted correction parameter, against the base mixing length model, and the DNS dataset, at a friction Reynolds number of 395.

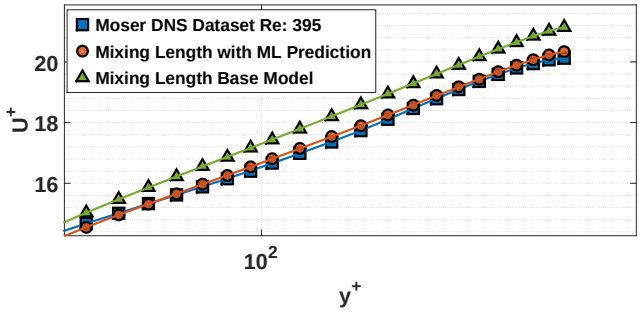

Figure 6: The zoomed-in version of figure 5, at the channel outer layer.

Figure 5 makes clear that the mixing length function predicted by the regression helps to ensure that the mean flow field matches the DNS dataset. This is emphasised in the zoomed in plot of figure 6, were it can be seen that the base mixing length model curve, results in a mean flow field which does not match the DNS solution, and in fact runs parallel to it for y^+ values in the channel outer layer. This trend has been seen previously in figures 2, and 3.

Conclusion

This paper has explored a methodology of applying a Bayesian inversion scheme to correct for deficiencies in Nikuradse's 1D mixing length equation. The inversion scheme was found to work well, when compared against the DNS dataset of Moser et al. In addition to this, it was proven possible to infer the value of the corrective parameter at a different friction Reynolds number by using a simple, and efficient regression tree architecture. Such a methodology is typical in the machine learning field and is a primary reason of why such methods show great promise in any field where datasets exist in the form of a ground truth.

Acknowledgements

The principal author of this paper would like to thank Vanja Zecevic for his help in formulating some of the CFD equations used in this paper.

References

[1] Baxter, J., Tridgell, A. and Weaver, L., Learning to play chess using temporal differences, *Machine Learning*, **40**, 2000, 243–263.

[2] Bishop, C. M., *Pattern recognition and machine learning*, Springer, New York, 2006.

[3] Breiman, L., Friedman, J., Stone, C. J. and Olshen, R. A., *Classification and regression trees*, CRC press, 1984.

[4] Grifoll, J. and Giralt, F., The near wall mixing length formulation revisited, *International Journal of Heat and Mass Transfer*, **43**, 2000, 3743–3746.

[5] Hinton, G. E., Osindero, S. and Teh, Y.-W., A fast learning algorithm for deep belief nets, *Neural Computation*, **18**, 2006, 1527–1554.

[6] Ling, J., Jones, R. and Templeton, J., Machine learning strategies for systems with invariance properties, *Journal of Computational Physics*, **318**, 2016, 22–35.

[7] Moser, R. D., Kim, J. and Mansour, N. N., Direct numerical simulation of turbulent channel flow up to re =590, *Physics of Fluids*, **11**, 1999, 943–945.

[8] Nikuradse, J., Laws of flow in rough pipes, in *VDI Forschungsheft*, Citeseer, 1933.

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

[10] Pope, S. B., *Turbulent flows*, IOP Publishing, 2001.

[11] Rajkumar, T. and Bardina, J. E., Prediction of aerodynamic coefficients using neural networks for sparse data, in *FLAIRS Conference*, 2002.

[12] Ray, J., Lefantzi, S., Arunajatesan, S. and Dechant, L., Bayesian parameter estimation of a k-ε model for accurate jet-in-crossflow simulations, *AIAA Journal*, 1–17.

[13] Shang, Z., Application of artificial intelligence cfd based on neural network in vapor–water two-phase flow, *Engineering Applications of Artificial Intelligence*, **18**, 2005, 663–671.

[14] Tracey, B., Duraisamy, K. and Alonso, J. J., A machine learning strategy to assist turbulence model development, *AIAA Paper*, **1287**, 2015, 2015.