DEVELOPMENT OF A DEEP-LEARNING MODEL TO IMPROVE LARGE EDDY SIMULATIONS OF TURBULENT FLOWS

Rozie Zangeneh

Department of Mechanical Engineering, Texas A&M University System, Prairie View, TX, 77446, USA

Turbulence is the most dominant characteristic of a turbulent flow. Therefore, successful modeling of turbulence can significantly improve the results of numerical simulation. Large Eddy Simulation (LES) computation of turbulent flows has been achieved a great attention recently since post-processing of LES results yields information of both mean flow and statistics of resolved fluctuations which is unique to LES and hence can model flows where persistent large-scale vortices results in flow development, e.g. flow behind bluff bodies, tumble swirl in engine combustion chambers and prediction of noise from high-speed flows (Versteeg 2009). However, this requires to address some issues to control the error sources and generate robust LES methodology for industrial applications.

In this study, a deep-learning approach is used to augment existing LES models using the relevant flow features. To this end, Random Forest Regression is developed to map relevant statistical flow-features within the LES solution to errors in calculated statistics such as sub-grid scale stresses. In this context, the exact solution is given by Direct Numerical Simulation (DNS) data. The capability of the proposed framework is examined by posteriori tests.

KEYWORDS: *Turbulence, Large Eddy Simulation, Deep learning, Sub-grid Scale stress, Random Forest Regression,*

1-INTRODUCTION

In spite of century-long efforts of fluid dynamists to develop RANS turbulence models, a comprehensive model applicable to a wide range of practical applications has so far shown to be elusive. This is due to a large extent attributable to differences in the influence of large and small eddies. A different approach to the computation of turbulent flows accepts that the larger eddies need to be simulated for each problem with a time-dependent simulation. The isotropic behavior of the smaller eddies, on the other hand, is easier to capture with a compact model. This is an essential idea behind the large eddy simulation (LES) approach to the computation of turbulence. While widely used in research and the academic community, LES has had a very limited impact on industrial simulations. The reason lies in the excessively fine grid requirements for wall boundary layers.

LES has been around since the 1960s, but computational resources to consider its application to industrially relevant problems have become commonly available recently. Since the Smagorinsky model, several sub-grid models have been proposed; most of them are categorized into the Smagorinsky model and gradient model. The main drawback of the Smagorinsky model is purely dissipative: the direction of energy flow is exclusively from resolved scales towards the sub-grid scales [1]. Leslie and Quarini [2] have shown that the gross energy flow in this direction is larger and offset by 30% backscatter – energy transfer in reverse direction from the sub-grid scale (SGS) eddies to larger, resolved scales. In contrast, the gradient model is not sufficiently dissipative since it limits of small grid size which causes instability. To overcome these issues, the dynamic version of these models was formed [3] and [4]. Although modeled SGS stresses were found to be much improved by dynamic models, they normally result in the increasing computational costs. Experience shows that LES calculation requires prior knowledge of flow so that the grid size can appropriately be specified to ensure that the turbulence energy is resolved throughout the domain [5]. This indicates that special care required to model LES and interpret the results at higher Reynolds numbers which DNS or experimental data is not available. Thus, there is a need for new LES models which perform better especially at higher Reynolds number.

Recently, there has been increasing interest in applying deep learning methods to turbulence modeling. Methods of deep learning, which are widely and successfully used in many areas, can be used to find new turbulence models. They are useful if they can automatically extract the essential grid-scale flow field required for accurate modeling. In recent years there has been a considerable effort to apply machine learning in turbulence modeling. Sarghini et al. [6] used an Artificial Neural Network (ANN) to model relationships between the grid-scale flow field and the turbulent viscosity coefficient in the mixed model. Wang et al. [7] worked on a machine-learning model to predict the discrepancies in the RANS modeled Reynolds stresses. Tracey et al. [8] used ANN to predict the source terms from the Spalart-Allmaras Reynolds Averaged Navier-Stokes (RANS) model. Moreau et al. [9] used an ANN to predict the sub-grid variance in the Cook-Riley model in isotropic turbulence for a scalar field. Ling et al. [10] demonstrated that random forest regressors based model could be used to predict the Reynolds stress anisotropy invariants more accurately than conventional RANS linear eddy viscosity models. Gamahara and Hattori [11] Developed an ANN to establish a functional relation between the grid-scale (GS) flow field and the SGS stress without any assumption of the form of function. They showed that an ANN could establish a model similar to the gradient model.

The present work aims to improve LES prediction. Random Forest Regression Model is used to develop an algorithm that maps local SG flow-field information to errors in SGS prediction from a LES simulation. Random Forest Regression (RFR) is used to predict a continuous target as a non-linear function of features. The parameters of the algorithm are learned using data where the target values are known so that the machine can make predictions about data where the target is unknown. This unique is particularly important for the improvement of LES models at higher Reynolds Number where the exact data is not available.

2-GOVERNING EQUATIONS

The continuity and momentum equation for an incompressible viscous fluid in the absence of external forces is given by:

$$div \ (\rho u_i) = 0 \tag{1}$$

$$\rho \frac{Du_i}{Dt} = -\frac{\partial p}{\partial x_i} + div(\mu \ grad \ u_i \ ) \tag{2}$$

Using a filter function, $\bar{\emptyset}(x,t) = \int G(x',x,\Delta)\emptyset(x',t) \, dx'$, one can, by exploiting the linearity of the filtering operation, swap the order of the filtering and differentiation with respect to time, as well as the order of filtering and differentiation with respect to space coordinates. Here Δ is the grid size. Filtering of Equation (1) and (2) yields the LES continuity equation:

$$div \ (\rho \bar{u}_i) = 0 \tag{3}$$

$$\frac{D\bar{u}_i}{Dt} + div\left(\rho \overline{u_i u_j}\right) = -\frac{\partial \bar{p}}{\partial x_i} + \mu \ div(grad(\ \bar{u}_i)) \tag{4}$$

Equation set (3) and (4) should be solved to yield the filtered velocity field $\bar{u}$ and filtered pressure field $\bar{p}$. Convective terms of the form div $\left(\rho \overline{u_i u_j}\right)$ on the left-hand side need to computed; however, only the filtered velocity field and pressure field are available. The filtered Naiver-Stocks equation can be rearranged as:

$$\rho \frac{D\bar{u}_i}{Dt} + div\left(\rho \bar{u}_i \bar{u}_j\right) = -\frac{\partial \bar{p}}{\partial x_i} + \mu \ div(grad \ (\bar{u}_i)) - (div\left(\rho \overline{u_i u_j}\right) - div\left(\rho \bar{u}_i \bar{u}_j\right)) \tag{5}$$

where $\tau_{ij} = \rho \overline{u_i u_j} - \rho \bar{u}_i \bar{u}_j$ is due to the filtering operation. In recognition of the fact that a substantial portion of τ_{ij} is attributable to convective momentum transport due to interactions between the unresolved, SGS eddies; these stresses are known as SGS stresses. A turbulence model constructs a mapping from the mean velocity field $\bar{u}$ to the SGS stress field τ_{ij} aims to close the LES equations.

3-SUBGRID STRESS MODELLING

To close the equations for the filtered velocity, a model for the SGS stress tensor τ_{ij} is needed. The simplest model is that proposed by Smagorinsky [12], which also forms the basis for several of the more

advanced models. The model can be evaluated in two parts. First, the linear eddy-viscosity model. The whole stress can be modeled as a single SGS turbulence model:

$$\tau_{ij} = -2\mu_{SGS}\bar{S}_{ij} + \frac{1}{3}\tau_{ii}\delta_{ij}$$

(6)

which is used to relate the SGS stress to the filtered rate of strain. The coefficient of proportionality, μ_{SGS} is the eddy viscosity of the sub-grid motions. Second, by analogy to the mixing-length hypothesis the eddy viscosity is modeled as the SGS viscosity:

$$\mu_{SGS} = \rho(C_{SGS}\Delta)^2|\bar{S}|$$

(7)

Where C_{SGS} is the SGC coefficient. According to the eddy-viscosity model (Equation (6)), the rate of transfer of energy to the GS motions is [5]:

$$P_r = \frac{\mu_{SGS}}{\rho}|\bar{S}|^2$$

(8)

For the Smagorinsky and other eddy-viscosity models with $\mu_{SGS} > 0$, the direction of this energy transfer is everywhere, from the filtered motions to the GS motions: there is no backscatter.
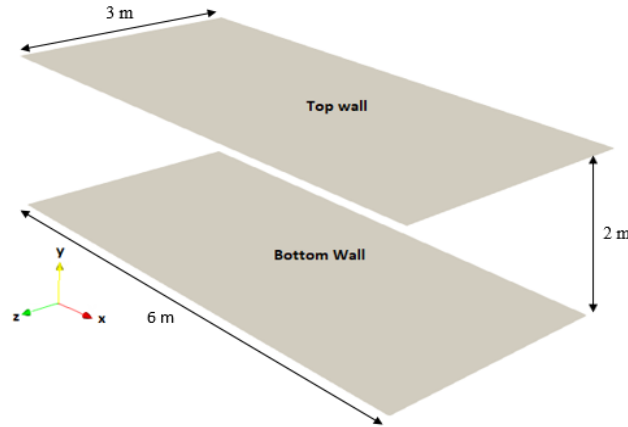

Figure 1. Computational domain and coordinate system of the Channel Flume

3-1 Simulation Setup

Fully developed channel flow, which is a classic LES model study is used here for the computational domain as shown in Figure 1. The benchmark for evaluating the accuracy of LES is the DNS of turbulent channel reported by Moin and Moser [13,14]. The bulk Reynolds number, $Re_b = U_b h/\nu$, is used where:

$$U_b = \frac{1}{h}\int_0^h <u> dy$$

(9)

And h is the channel height. For the LES simulation, Re_b is defined and Re_τ is computed. The target value is $Re_\tau = 395$. It is noted that the friction Reynolds number is defined as:

$$Re_\tau = u_\tau\delta/\nu$$

(10)

Where u_τ is the friction velocity and $\delta = h/2$. The number of grid points, $(N_x, N_y, N_z) = (120,100,90)$, and $y^+ = 0.97$. Periodic boundary conditions are applied in x and z directions. The solver pimpleFOAM, provided as part of OpenFOAM, was used to solve the equations derived in sections continuity and Navier-Stokes. The Wall-Adapting Local Eddy-Viscosity (WALE) model is used to model the eddy viscosity.
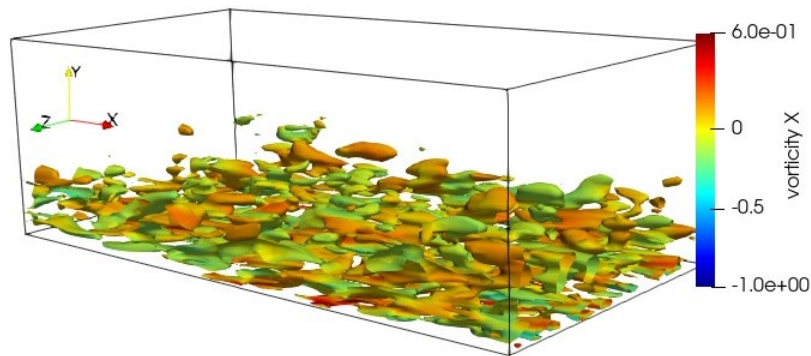

Figure 2. Isosurfaces of the Q-criterion (Q = 0.004) colored by the streamwise instantaneous vorticity at $R_\tau = 395$

3-2 LES Results

Figure 2 illustrates the instantaneous vorticity in the x-direction and is shown that coherent vortex structures built up inside the boundary layer, which is essential for momentum exchange and turbulence production in wall-bounded turbulent flows. Figure 3 depicts the mean velocity profile which is presented in global coordinates, scaled with the bulk velocity U_b. LES results underpredict the mean velocity profile especially for the region of $0.1 < \frac{y}{\delta} < 0.3$. To address this issue, it is necessary to evaluate the results in the wall coordinates and scaled with the friction velocity. The data is shown on the right plot of Figure 3. In the viscous sub-layer there are only a few points located there in the LES simulation; thus cannot resolve it adequately. In the buffer region, the results begin to diverge from the DNS leading to underprediction in the Log-Law region.
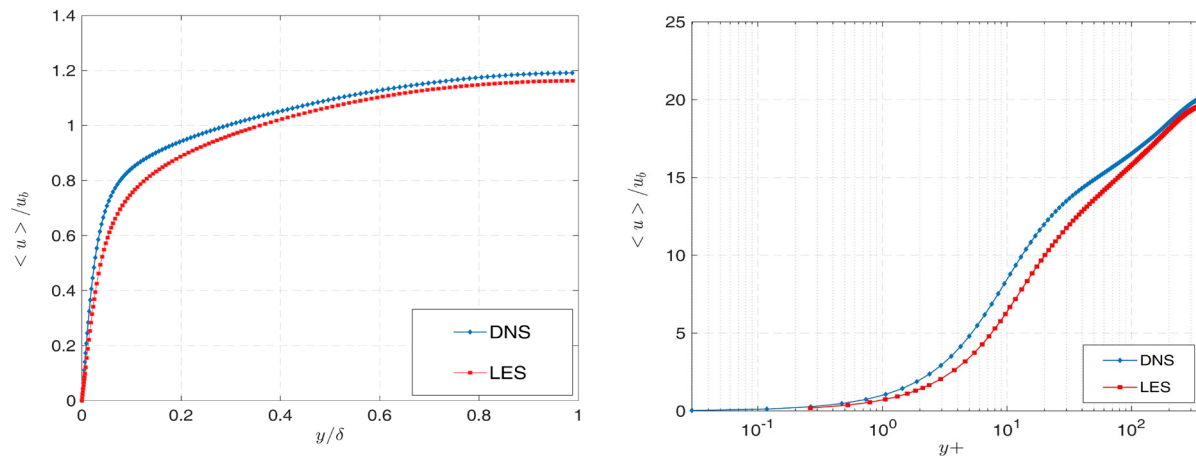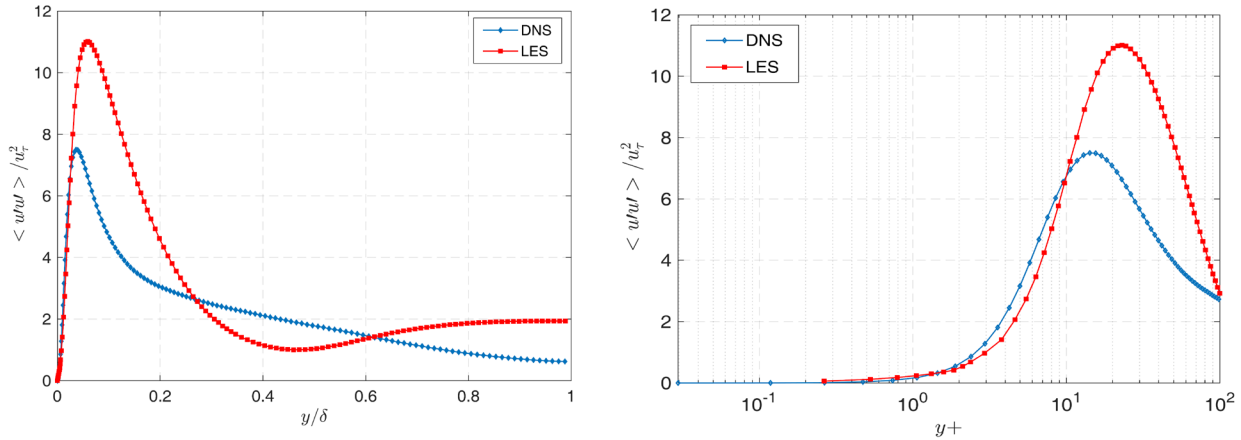

Figure 3. profiles of the mean of the normalized streamwise component of velocity at $R_\tau = 395$

Figure 4. Profiles of the SGS stress, τ_{11} at $R_\tau = 395$

Evaluation of diagonal components of the SGS tensors shows that the LES simulation overpredicts the streamwise velocity fluctuations and shifts the pick away from the wall (Figure 4). The behavior is, however, different for wall-normal and spanwise velocities, and LES simulations strongly underpredict the fluctuations, as shown in Figures 5 and 6.
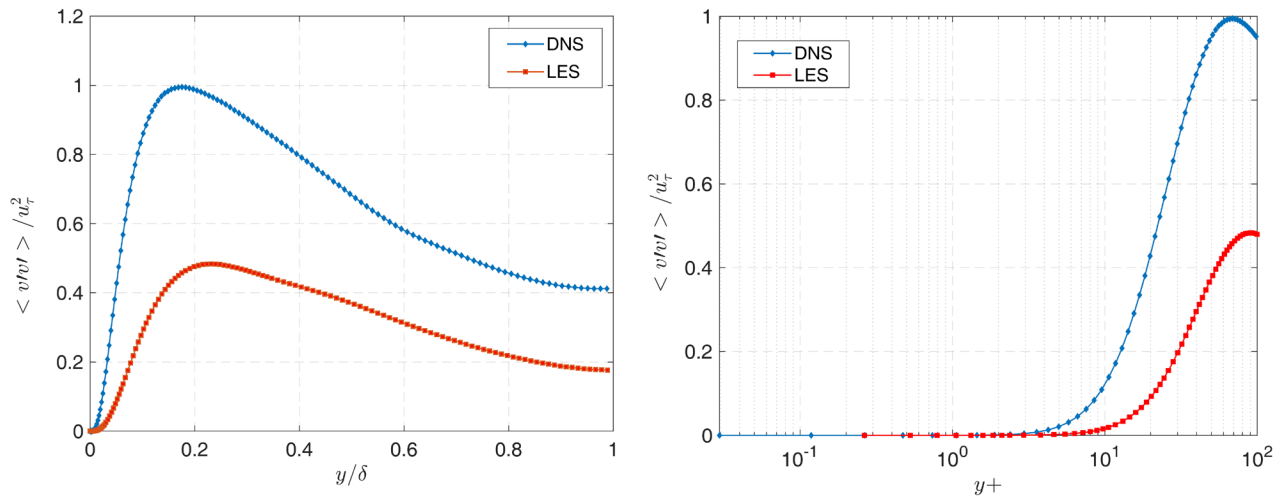
Figure 5. Profiles of the SGS stress, τ_{22} at $R_\tau = 395$

Analyzing the SGS shear stress shows the LES result predicts the location of the peak with very good accuracy, but significantly under-estimates the values of the shear stress in the viscous wall region (Figure 7). For $y^+ > 100$, LES produces a linear profile that underpredicts the DNS data.

Given the LES simulation discrepancies with DNS, it is useful to review the physical basis of the two components of the model. The first component is the basic Smagorinsky model assumption that the SGS stress τ_{ij}, is aligned with the filtered rate of strain $\bar{S}_{ij}$, Equation (6). However, Pope [5] showed that there is a much weaker correlation between τ_{ij} and $\bar{S}_{ij}$ than the near-perfect correlation implied by the Smagorinsky model. Hence no value of the Smagorinsky coefficient, C_{sgs}, can yield the correct levels both

of C_{sgs} and the rate of energy transfer of $p_r = \frac{\mu_{SGS}}{\rho}|\bar{S}|^2$. The second component of the model - the dynamic aspect - is based on the assumption that the Smagorinsky coefficient is independent of the filter width. This assumption is well-founded for the case of the filter being within the inertial subrange of high-Reynolds-number turbulence. However, the success of the model is attributed to its behavior for flow regimes for which the assumption lacks justification - laminar flow, transitional flow, and flow in the viscous wall region.
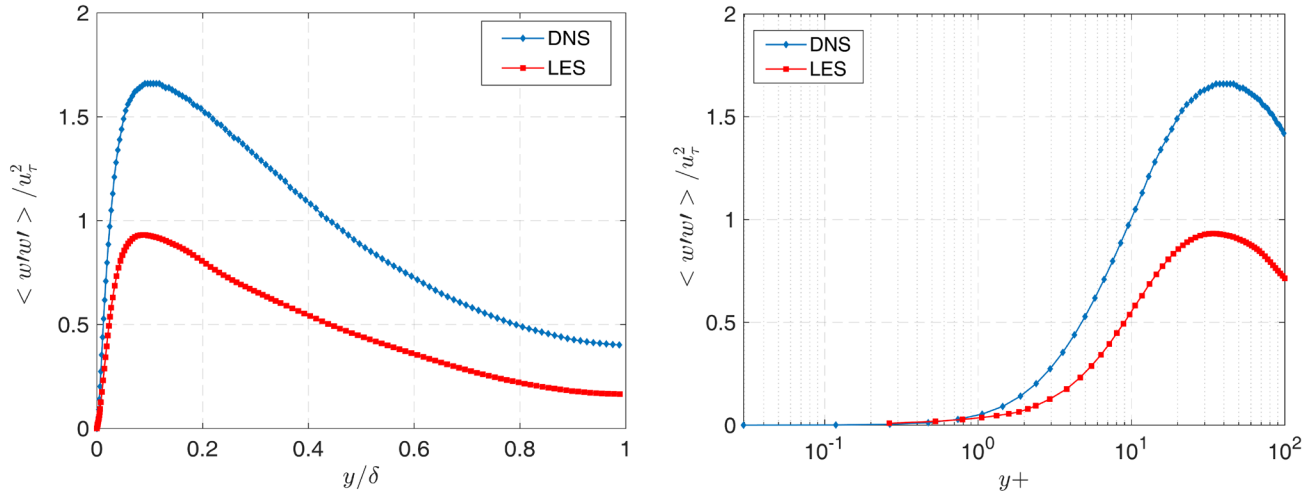

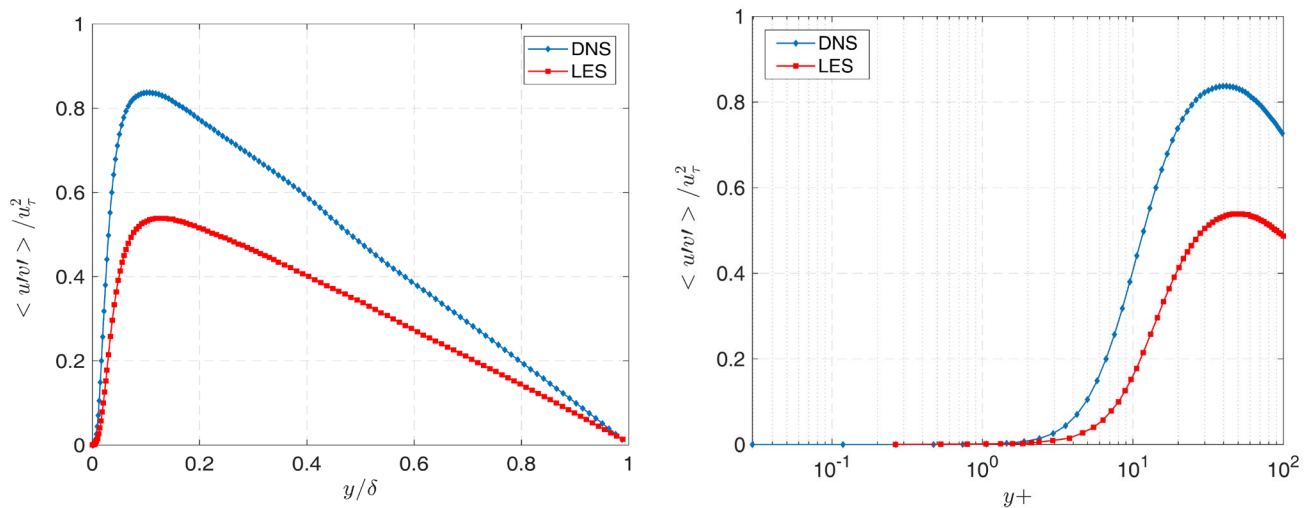Figure 6. Profiles of the SGS stress, τ_33   at R_τ =395

Figure 7. Profiles of the SGS stress, τ_{12} at R_τ =395

4-METHODOLOGY OF THE DEEP LEARNING MODEL

The overall objective of this work is to create, using deep learning, a model that maps the GS flow field predicted by LES linear eddy viscosity model to errors in the prediction of SGS stress. A functional mapping can be obtained from the flow features obtained from LES models to τ^E by using deep learning:

$$\{F\} \rightarrow \tau^E$$

And then:

$$\tau^R = \tau^{LES} + \tau^E \tag{11}$$

τ^R is the actual SGS stress calculated from DNS and τ^E is the SGS stress discrepancies. τ^E can be calculated by training on DNS data from similar flows.
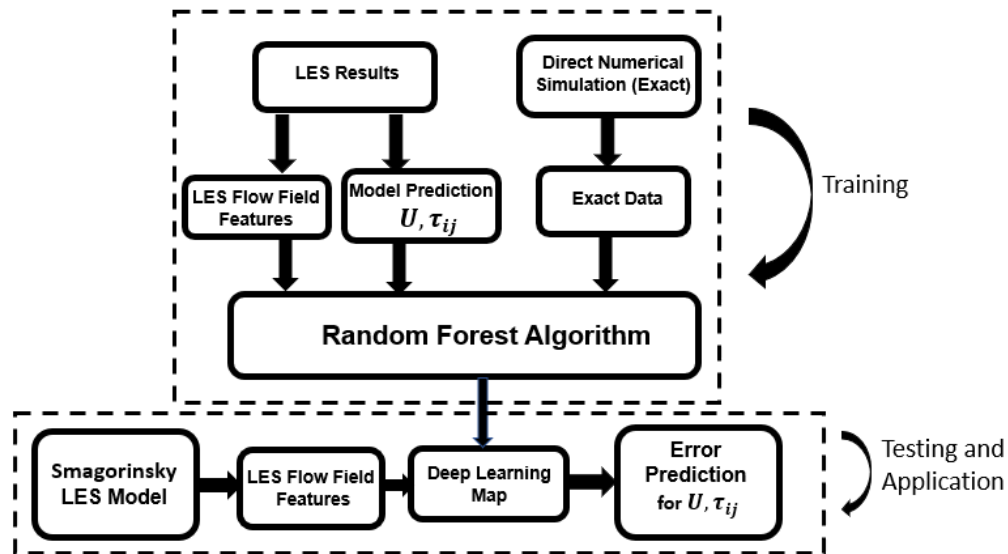

Figure 8. Thee method of creating an algorithm that maps LES flow-field features to model error

Figure 8 illustrates the proposed framework for creating this model. The τ^R, predictions τ^{LES}, and feature set, $\{F\}$, are used to train a deep learning map that takes as inputs a subset of $\{F\}$ and outputs a model error, a measure of the discrepancy between prediction and actual.

4-1-Random Forest Regression Scheme

Random Forest Regression is the problem of learning a functional relationship between input features and an output target using training data where the specific functional form learned depends on the choice of model. The parameters of the function are learned using data where the target values are known so that the algorithm can make predictions about data where the target is unknown. Therefore, the goal of a regression model is to learn to predict an output based on an input set of features. More formally, Random Forest is a collection of tree predictors, $h(\boldsymbol{x}, \theta_k)$, where $\boldsymbol{x}$ represents the observed input vector and θ_k are the independent random vectors. For regression, the random forest prediction is the unweighted average over the collection: $h(x) = \left(\frac{1}{k}\right) \sum_1^k h(x, \theta_k)$ [15]. An advantage of the Random Forest Regression is that it can provide importance scores for inputs after training, which can be further used to assist the modelers in

improving the existing turbulence models. Besides, Random forest has robust performances with only a small set of training parameters, which is in contrast to the commonly used Neural Networks.

4-2-CHOICE OF INPUT VARIABLES

A set of velocity gradient and distance from the wall $\{\nabla u, y\}$, was suggested by Gamahara et al. [11] for successful learning based on correlation coefficients. Therefore, in the rest of the paper, the set of input variables is fixed to $\{\nabla u, y\}$. The RFR algorithm is trained by the DNS data of the channel flow at $Re_\tau = 180$ and then the data at $Re_\tau = 395$ is used as the test data set. Finally, the error between prediction and actual SGS tensor, τ^E, obtained from the test is implemented to pimpleFOAM and the LES simulation repeated for $Re_\tau = 395$ to investigate whether the error predicted by RFR would improve the LES model.

5- POSTERIORI TESTS

LES is a high-fidelity model to be used for high Reynolds number flows which are not accessible by DNS. Thus, when applying a deep learning model to LES of high Reynolds number flows, no training data are available from DNS. In this regard, it is of importance to check whether an RFR algorithm trained at low Reynolds numbers can make predictions at high Reynolds numbers. The proposed framework is examined by posteriori tests to determine whether these improved SGS tensor values would translate to improved mean velocity predictions.
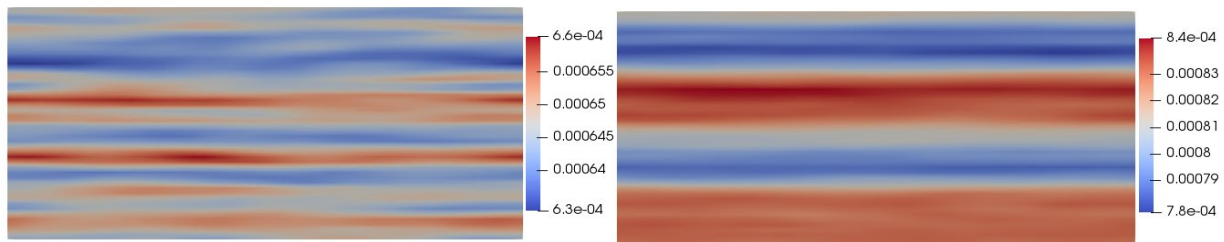

Figure 9. Comparison between τ_{11} predicted by the RFR trained at $Re_\tau = 395$ with LES model . y = 0.1.

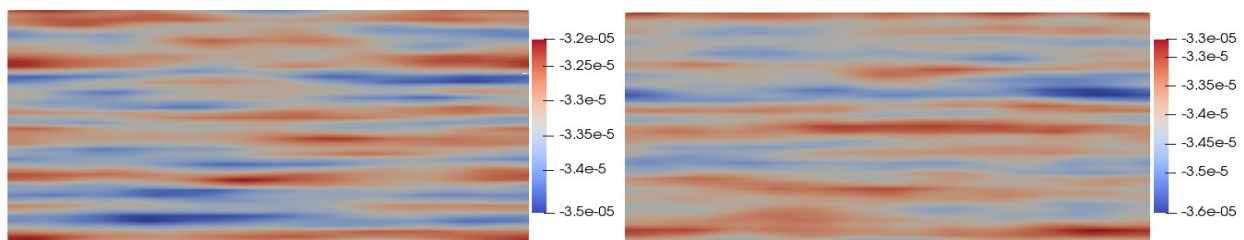

Figure 10. Comparison between τ_{12} predicted by the RFR trained at $Re_\tau = 395$ with LES model. y = 0.1.

First, details of the learning results for $Re_\tau = 395$ are investigated. Figures 9 and 10 compare the distributions of the SGS stress obtained by LES simulations, and that predicted by trained RFR. The plane of y = 0.1 is chosen where the rms amplitudes of all components are nearly largest. It is clear from Figure 8 that RFR improves the prediction of the first normal component of SGS tensor. However, there is a slight change in predicting SGS shear stress at this plane as shown in Figure 10. Next, these two SGS components investigated along with the channel height, as shown in Figure 11.
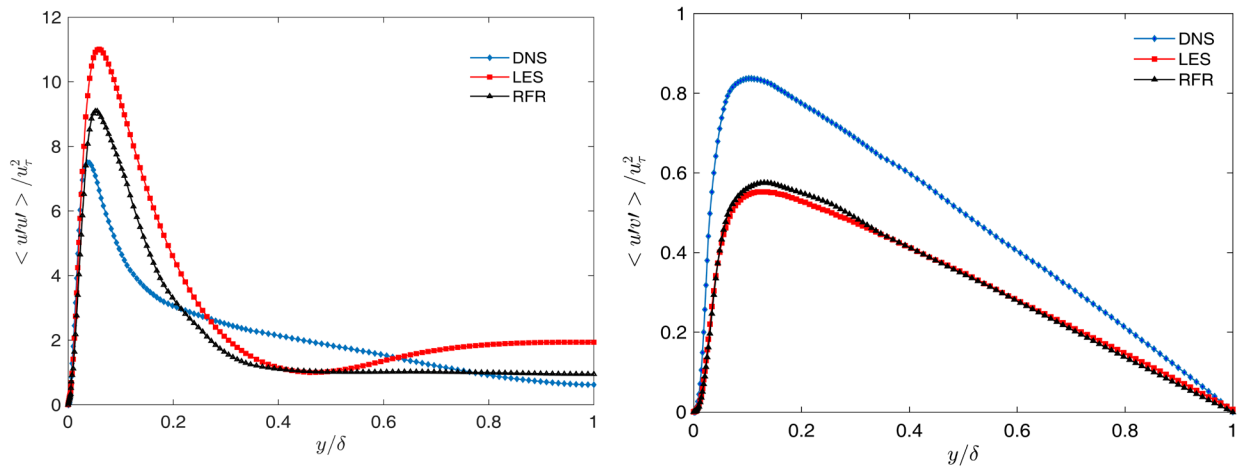
Figure 11. Comparison of the SGS stress predicted by LES with Smagorinsky model and RFR model; Left: τ_{11}, Right: τ_{12}

It is evident from Figure 11 that RFR model drastically improves the streamwise velocity fluctuations, especially in the vicinity of the pick region as well as at $\frac{y}{\delta} > 0.6$, which the Smagorinsky model declines to predict accurate fluctuations. Figure 11 shows with an exception near the pick region, the trained algorithm predicts the SGS shear stress values, τ_{12}, close to the Smagorinsky model. Finally, Figure 12 compares the mean flow obtained by LES simulations. LES results underpredict the mean velocity profile especially for the region of 0.1<y/δ<0.4. As was explained before, in the buffer region, results begin to diverge from the DNS leading to underprediction in the Log-Law region. However, the mean flow obtained by the Smagorinsky model trained by RFR algorithm is in reasonable agreement with DNS except that it slightly underestimated for y/δ >0.8. Therefore, it can be concluded that the RFR model has an advantage over the Smagorinsky model for prediction of the mean flow.
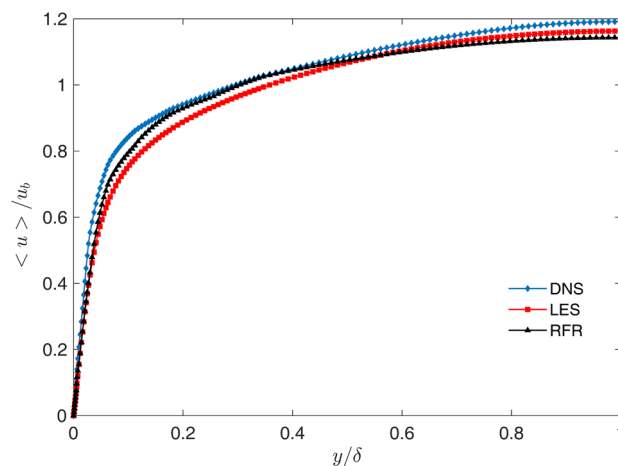

Figure 12. Comparison of mean flow obtained by LES with Smagorinsky model and RFR model

6- CONCLUSION

A new deep-learning-based approach was proposed to improve Large Eddy Simulations of Turbulent flows. The new framework is based on a trained Random Forest Regression algorithm using Direct Numerical

Simulation to make predictions when the exact data is not available due to high CPU cost. A set of flow features obtained from LES was used to train the RFR algorithm and then obtained results were implemented to PimpleFOAM, which is an OpenFOAM solver for LES simulations. A posteriori evaluation showed that the RFR was able to predict the mean flow in a channel with a considerable improvement in contrast to the Smagorinsky model. Although the deep-learning approach showed an advantage over the LES model, for this novel approach to become a practical turbulence model, RFR model needs to be trained and tested across a much broader set of flows.

ACKNOWLEDGMENT

REFERENCES

[1] Versteeg, H., Malalasekra W. *An Introduction to Computational Fluid Dynamics*, 2rd edition, Pearson, 2009.

[2]. Leslie D. C., Quarini G. L*., The application of turbulence theory to the formulation of subgrid modeling procedures*, J. Fluid Mech. , 91(1), 65-91, 1979.

[3] Germano M., *Turbulence: The filtering approach*, J. Fluid Mech. 238, 325,1992.

[4] Vreman B., Geurts B., Kuerten H., *Large eddy simulation of the temporal mixing layer using the Clark model*, Theor. Comput. Fluid Dyn. 8, 309,1996.

[5] Pope S.P, *Turbulent Flows*, 3rd printing, Cambridge University Press, 2018

[6] Sarghini F., de Felice G., Santini S., *Neural networks based subgrid scale modeling in large eddy simulations*, Comput. Fluids, 32, 97, 2003.

[7] Wang J.X., Wu J.L., Xiao H., *Physics-informed machine learning approach for recon-*

structing Reynolds stress modeling discrepancies based on DNS data, Physical Review Fluids 2 (3), 2017.

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

[9] Moreau A., Teytaud O., Bertoglio J. P., *Optimal estimation for large-eddy simulation of turbulence and application to the analysis of subgrid models*, Phys. Fluids 18, 2006.

[10] Ling J. L, H. Xiao, Paterson E., Physics-Informed Machine Learning Approach for Augmenting Turbulence Models: A Comprehensive Framework, Phys. Rev. Fluids 3, 074602 2018.

[11] Gamahara M., Hattori Y., *Searching for turbulence models by artificial neural network*, Phys. Rev. Fluids 2, 054604, 2017.

[12] Smagorinsky J., *General circulation experiments with the primitive equations*, Mon. Weather Rev. 91, 99, 1963.

[13] Moin P., Kim J., *Numerical Investigation of Turbulent Channel Flow*, J. Fluid Mech., 118, 341-377, 1982.

[14] Moser, R., Kim, J., Mansour, N. Direct Numerical Simulation of Turbulent Channel Flow Up to Re= 590, Physics of Fluids, 11(4), 943-945, 1999.

[15] Breiman L., *Random forests*, Machine learning 45 (1), 2001.