Data-driven wall modeling for LES involving non-equilibrium boundary layer effects

Radhakrishnan, Sarath; Calafell, Joan; Miró, Arnau; Font, Bernat; Lehmkuhl, Oriol

DOI
10.1108/HFF-11-2023-0710

Publication date
2024

Document Version
Final published version

Published in
International Journal of Numerical Methods for Heat and Fluid Flow

Citation (APA)
Radhakrishnan, S., Calafell, J., Miró, A., Font, B., & Lehmkuhl, O. (2024). Data-driven wall modeling for LES involving non-equilibrium boundary layer effects. *International Journal of Numerical Methods for Heat and Fluid Flow*, *34*(8), 3166-3202. https://doi.org/10.1108/HFF-11-2023-0710

Data-driven wall modeling for LES involving non-equilibrium boundary layer effects

Sarath Radhakrishnan, Joan Calafell and Arnau Miró
Barcelona Supercomputing Center, Barcelona, Spain

Bernat Font
Barcelona Supercomputing Center, Barcelona, Spain and Faculty of Mechanical Engineering, Delft University of Technology, Delft, Netherlands, and

Oriol Lehmkuhl
Barcelona Supercomputing Center, Barcelona, Spain

Received 30 November 2023
Revised 23 April 2024
Accepted 30 May 2024

Abstract

Purpose – Wall-modeled large eddy simulation (LES) is a practical tool for solving wall-bounded flows with less computational cost by avoiding the explicit resolution of the near-wall region. However, its use is limited in flows that have high non-equilibrium effects like separation or transition. This study aims to present a novel methodology of using high-fidelity data and machine learning (ML) techniques to capture these non-equilibrium effects.

Design/methodology/approach – A precursor to this methodology has already been tested in Radhakrishnan *et al.* (2021) for equilibrium flows using LES of channel flow data. In the current methodology, the high-fidelity data chosen for training includes direct numerical simulation of a double diffuser that has strong non-equilibrium flow regions, and LES of a channel flow. The ultimate purpose of the model is to distinguish between equilibrium and non-equilibrium regions, and to provide the appropriate wall shear stress. The ML system used for this study is gradient-boosted regression trees.

Findings – The authors show that the model can be trained to make accurate predictions for both equilibrium and non-equilibrium boundary layers. In example, the authors find that the model is very effective for corner flows and flows that involve relaminarization, while performing rather ineffectively at recirculation regions.

Originality/value – Data from relaminarization regions help the model to better understand such phenomenon and to provide an appropriate boundary condition based on that. This motivates the authors to

---

SR acknowledges the financial support of the Ministerio de Ciencia y Innovación y Universidades, for the grant, Ayudas para contratos predoctorales para la formación de doctors (Ref: BES-2017–081982). OL has been partially supported by a Ramon y Cajal postdoctoral contract (Ref: RYC2018-025949-I). This work was partially supported by the Ministerio de Economía, Industria y Competitividad, Secretaría de Estado de Investigación, Desarrollo e Innovación, Spain (refs: PID2020-116937RB-C21 and PID2020-116937RB-C22). The authors also acknowledge the Barcelona Supercomputing Center for awarding us access to the MareNostrum IV machine based in Barcelona, Spain.

International Journal of Numerical Methods for Heat & Fluid Flow
Vol. 34 No. 8, 2024
pp. 3166-3202
Emerald Publishing Limited
0961-5539
DOI 10.1108/HFF-11-2023-0710

continue the research in this direction by adding more non-equilibrium phenomena to the training data to capture recirculation as well.

Keywords Computational fluid dynamics, Machine learning

Paper type Research paper

1. Introduction

Wall-bounded flows at high Reynolds numbers are essential because they are found in multiple fundamental industrial applications (Jiménez, 2004). Nevertheless, the use of accurate numerical methodologies such as wall-resolved large eddy simulation (LES) is often unfeasible due to their heavy computational costs. According to Choi and Moin, the computational costs in terms of spatial resolution scale with $Re^{13/7}$ (Choi and Moin, 2012). Taking into account the time integration costs, it increases to $Re^{65/21}$ (Calafell *et al.*, 2019). On the other hand, if the boundary layer (BL) is modeled instead of explicitly resolved, the scaling of the computational effort can be reduced to $Re^{4/3}$.

Among the 2030 Grand Challenges identified by NASA, the LES of a powered aircraft configuration throughout the flight envelope appears in the first position of the priority list (Slotnick *et al.*, 2014). In NASA's report, wall-modeled LES (WMLES) is proposed to circumvent the massive costs of accurately resolving the BL while benefiting from the temporal and spatial resolution of an LES computation. However, it suggests that "a longer-term high-risk effort should investigate radical new approaches to physical modeling" (Slotnick *et al.*, 2014). This work aims to make significant progress in this direction.

In line with NASA recommendations for exploring disruptive approaches, data-driven models that take advantage of recent advances in deep learning and other machine learning (ML) techniques should be considered. One of the most relevant characteristics of these approaches is that they can autonomously learn complex physical relations from the data. Eventually, this may open the way to creating innovative and general models that take advantage of this ability, leveraging physical properties that may have gone hidden in traditional equation-based models. Data-based approaches have already been used in computational fluid dynamics (CFD). The reader is directed to Brunton *et al.* (2020) for a review of ML approaches in CFD. While ML models for CFD have become very popular in recent years, their applicability is generally limited to flows of a similar nature, as evidenced by Rumsey *et al.* (2022).

Traditional wall modeling strategies take only small samples of data from the resolved outer layer. These data are then processed through a physical model to complement it, allowing more or less accurate predictions of wall shear stress depending on how far the real flow conditions are from the assumptions of the physical model (Bose and Park, 2018). On the other hand, data-driven models are intended to create a link between the outer layer and the wall shear stress through data, avoiding relying on external and non-general models. By contrast, wall modeling using data-based approaches is far more challenging because the high-fidelity data to train a model is mostly sampled from canonical cases alone. Additionally, the canonical flows must be selected so that they feature as many flow conditions as possible, thus providing a sufficiently wide physical knowledge to allow the model to generalize to any flow condition. Keeping this in mind, we intend to develop a data-driven wall model trained on different flow types that can be generalized to previously unseen cases.

Some of the first works that applied advanced deep learning techniques for the study of wall-bounded turbulence were published during 2019. Guastoni *et al.* (2020) and Srinivasan *et al.* (2019) applied deep neural networks (NNs) to directly predict mean flow and turbulent quantities of shear flows. The first data-driven wall model proposed by Yang *et al.* (2019) relied on feedforward NNs (FNNs). Later, Lozano-Durán and Bae (2023) and Zhou *et al.* (2021) proposed wall models exclusively based on an NN, while

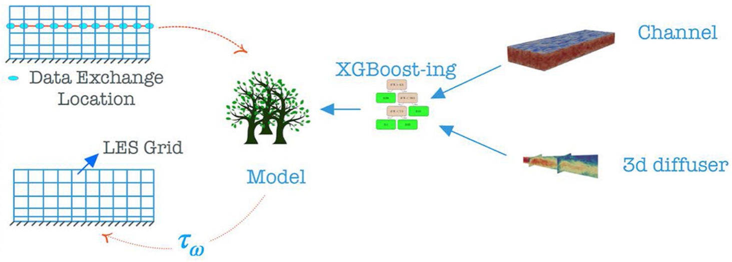
Bae and Koumoutsakos (2022) proposed a wall model based on reinforcement learning. In 2023, Zhou *et al.* (2023) proposed an improvement to their 2021 wall model, which has shown good *a priori* performance when tested on DNS data. Vadrot *et al.* (2023) provides a comprehensive literature survey of data-driven wall models and also compares the performance of the wall models of Yang *et al.* (2019), Zhou *et al.* (2023) and Bae and Koumoutsakos (2022).

According to NASA CFDVision 2030 (Slotnick *et al.*, 2014), the computational cost to resolve the laminar and transitional regions exceeds the turbulent region by two orders of magnitude in WMLES. This is characteristic of the flow near the leading edges of the wing. Due to the favorable pressure gradient, a turbulent to laminar transition develops near the vicinity of the leading edge. Although it can only be a fraction of the chord length, the downstream effects can be substantial. WMLES designed to capture the transition can be found in Drela and Giles (1987), Marques and Wang (2017), Gonzalez *et al.* (2020), Gonzalez *et al.* (2021), Dauricio and Azevedo (2023), Bodart and Larsson (2012) and Park and Moin (2014). Drela and Giles (1987) use integral methods, and Gonzalez *et al.* (2020) and Gonzalez *et al.* (2021) use Falkner–Skan similarity solutions for wedge flows to account for the transition. Marques and Wang (2017) successfully use a data-driven approach for unsteady laminar channel flows. A sensor-based approach was used in Bodart and Larsson (2012) and Park and Moin (2014).

In this work, a non-equilibrium wall model based on data is developed using gradient-boosted regression trees (GBRTs) primarily for flows involving non-equilibrium BL effects. Preliminary work on the development of a wall model that works under equilibrium conditions was published in Radhakrishnan *et al.* (2021). The current work is the second step toward developing a complete non-equilibrium wall model using GBRT. The methodology of model development is detailed in Section 2. The model is tested and validated *a priori* and *a posteriori* in Section 3, and the main findings of the paper are summarized in Section 4.

2. Methodology

The current data-driven wall modeling method consists of an iterative process that includes the following steps: data selection, model training and *a priori* and *a posteriori* testing. The procedure is repeated successively to optimize the final results. A schematic of the methodology is shown in Figure 1. As can be seen, the data on the right of the figure, i.e. channel and diffuser data, are used to create the ML-based model on XGBoost. From the solver's side, the accurate shear stress provided by the model is used as the boundary condition.



Source: Figure by authors

Figure 1. Schematic of the methodology

2.1 *Numerical method*

All the simulations presented in this work are performed using a low dissipation finite element method implemented in the code Alya (Vazquez *et al.*, 2016). Alya is a parallel multiphysics/multiscale simulation solver developed at the Barcelona Supercomputing Center to run efficiently on high-performance computing environments. A Galerkin finite element scheme is used, which conserves linear and angular momentum and kinetic energy at the discrete level, to discretize the convective term (Lehmkuhl *et al.*, 2019a). To use elements of equal order, numerical dissipation is introduced only for pressure stabilization through a fractional step scheme (Codina, 2001). A third-order Runge–Kutta explicit method combined with an eigenvalue-based time-step estimator (Trias and Lehmkuhl, 2011) is used for temporal integration. Alya is used to solve the spatially filtered non-dimensional Navier–Stokes equation given by:

$$\frac{\partial U_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial U_i}{\partial t} + \frac{\partial}{\partial x_j}(U_i U_j) = -\frac{\partial P}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 U_i}{\partial x_j^2} - \frac{\partial \tau_{ij}}{\partial x_j} \tag{2}$$

The equations are normalized using a reference length L and velocity U_0, which define the Reynolds number $Re = U_0 L/\nu$. The filtered velocity is defined as $\tilde{u}_i = U_i = u_i - u_i'$, where $U_i = (U, V, W)$, and the filtered modified pressure is $P = \tilde{p}/\rho$. Finally, $\tau_{ij} = \widetilde{u_i' u_j'} = \widetilde{u_i u_j} - U_i U_j$ is the subgrid-scale stress (SGS) tensor.

2.2 *Data sets*

To obtain accurate model predictions, the input data at inference time must be as similar as possible to the training data. Because input data at inference time will be WMLES-like in spatial or temporal resolution, ideally, training data should also have WMLES characteristics. The data sets used to train the model are derived from the WRLES of the channel flow ($Re_\tau = 950$) (Lehmkuhl *et al.*, 2019a) and the DNS of a flow through the Stanford double diffuser ($Re = 10000$) (Miró *et al.*, 2023). DNS data of the diffuser is used instead of WRLES data because: the data is already available in the literature and at the sampled Reynolds number regime, the difference in grids for DNS/WRLES/WMLES is not so significant (Choi and Moin, 2012). The details of these data sets and their selection criteria are given in the following sections.

2.2.1 Channel flow: $Re_\tau = 950$. The computational domain for the channel is $6\delta \times 2\delta \times 3\delta$ in the stream-wise, wall-normal and span-wise directions, respectively, where δ is the half-channel height. A mesh of 128^3 elements is used to discretize the domain. The mesh is uniform in the stream-wise and span-wise directions, corresponding to $\Delta x^+ \approx 45$ and $\Delta z^+ \approx 22$ in wall units, respectively. A hyperbolic tangent function is given by the following equation:

$$y(i) = \frac{\tanh\left(\gamma\left(\frac{2(i-1)}{N_y} - 1\right)\right)}{\tanh(\gamma)} \tag{3}$$

The above function is used to stretch the mesh in the wall-normal direction such that the minimum $\Delta y^+ = 1$ and the maximum $\Delta y^+ = 34$. Periodic boundary conditions on the stream-wise and span-wise boundaries and a no-slip boundary is enforced on the walls. The flow is driven by a stream-wise constant pressure gradient and the integral length-scale approximation (ILSA) (Lehmkuhl *et al.*, 2019b) SGS model is used for turbulence closure.

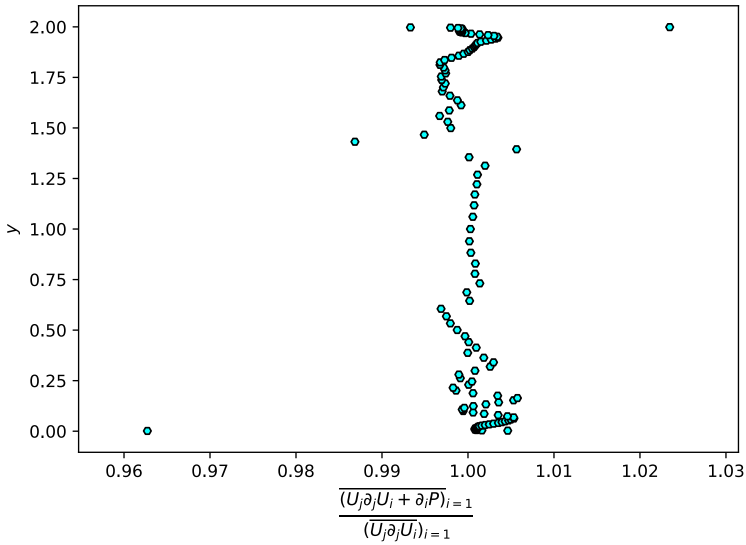
2.2.2 Stanford double diffuser: Re = 10,000. The diffuser has an inlet section, an expansion section and an outlet section. The inlet section is a rectangular duct of cross-section of 3.33×1 square units. Turbulent flow from the duct enters the diffuser at the inlet section, and a standard Dirichlet condition for pressure is applied in the outlet section. A no-slip boundary condition is applied to the walls of the diffuser. The Reynolds number based on the inlet bulk velocity and duct height is 10,000. This data set was thoroughly validated in Miró *et al.* (Miró *et al.*, 2023). Briefly, numerical results were compared with existing DNS and experimental results, resulting in very close agreement. This new data set has been made publicly available on the ERCOFTAC Wiki. The region used for model training is only the expansion section. This portion begins from the plane where the diffuser diverges from the rectangular duct cross-section of 1 × 3.33 square units to a square duct of cross-section 4×4 square units. The length of this portion is 15 units.

2.3 Data sampling

The choice of channel data is motivated by the fact that it contains information on the equilibrium BL. As the channel is in equilibrium in the Reynolds average sense, the sum of the convection and pressure gradient terms should be zero (LARSSON *et al.*, 2016). This sum has a range of approximately $\left(-10^{-4}, 10^{-4}\right)$ in the sampled data, i.e.:

$$-10^{-4} < \overline{(U_j\partial_j U_i + \partial_i P)}_{i=1} < 10^{-4} \tag{4}$$

where $\overline{\cdot}$ denotes a mean flow quantity. In addition, to consider the effect of the pressure gradient, the ratio of the average sum of the convective term and the pressure gradient term to the average convective term is analyzed. The result is shown in Figure 2. The ratio ranges from 0.96 to 1.03, i.e.:



Source: Figure by authors

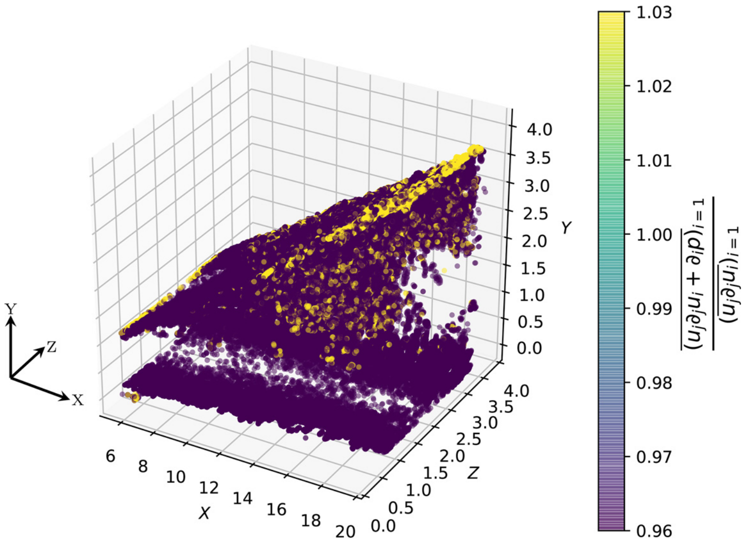
Figure 2. Ratio of the sum of the average convection and pressure gradient to the average convection for the channel in the wall-normal direction

$$0.96 < \frac{\overline{(U_j \partial_j U_i + \partial_i P)}_{i=1}}{\overline{(U_j \partial_j U_i)}_{i=1}} < 1.03 \tag{5}$$

It can also be seen that the ratio is very close to 1 in most cases, meaning the gradient of the pressure fluctuations tends to zero at those locations.

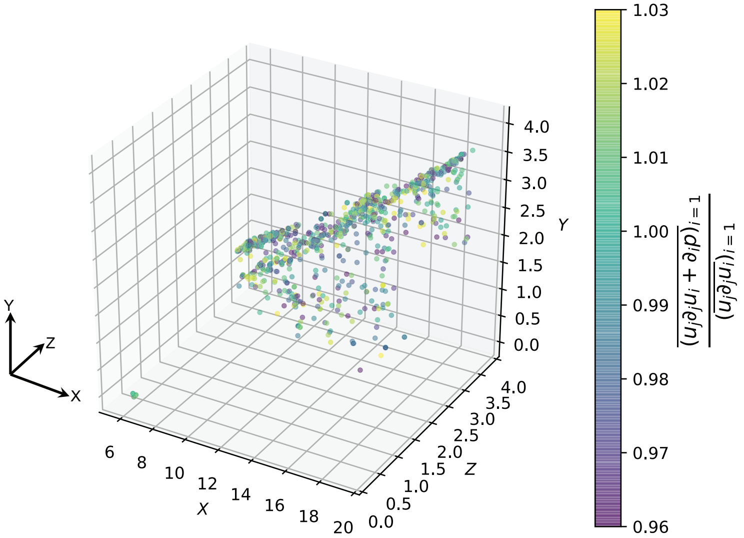
To provide information regarding non-equilibrium BL, data from the Stanford double diffuser is used. In this way, the selected data set comprises equilibrium and non-equilibrium data, such as flow around corners, three-dimensional separation or relaminarization. However, this also poses a challenge when training the model, as it is not easy to segregate the data according to the physical phenomenon, as they are mixed in a non-linear fashion. If there is an imbalance in the data with respect to the information it carries, the model will tend to be biased toward the "most seen" flow physics. The data from the simulations are studied and pre-processed so that this imbalance is reduced. The ratio of the average sum of the convection and the pressure gradient terms to the average convection is also computed for the diffuser.

From the diffuser, approximately 266×10^3 non-equilibrium points are chosen. The diffuser region from which these examples are selected is shown in Figure 3. This involves data from relaminarization and three-dimensional recirculation (Miró *et al.*, 2023). The equilibrium points in these regions are less than 1%, and it is shown in Figure 4. In the diffuser, the y^+ limit of the examples was set approximately at 150, a typical range of values for wall shear stress models (LARSSON *et al.*, 2016). A number of 120×10^3 channel flow samples are also added to this data set, accounting for 69% non-equilibrium and 31% equilibrium data. This proportion was determined by an iterative method of sampling, testing and resampling. A 90%–10% split is followed between the train data and the validation data. The evolution of the train data and the



Notes: *u* and *p* are the DNS instantaneous velocity and pressure, respectively
The sum of convection and pressure gradient is in local coordinates
Source: Figure by authors

Figure 3. Locations of non-equilibrium inside the duct plotted as a heat map in the *x* and *y* directions



Notes: *U* and *p* are the DNS instantaneous velocity and pressure, respectively. The sum of convection and pressure gradient is in local coordinates

Source: Figure by authors

Figure 4. Locations of equilibrium inside the duct plotted as a heat map in the *x* and *y* directions

validation data with respect to each iteration of the training process is shown in Figure 5. Both curves evolve in a concordant way, showing that the choice of parameters is adequate.

2.4 XGBoost

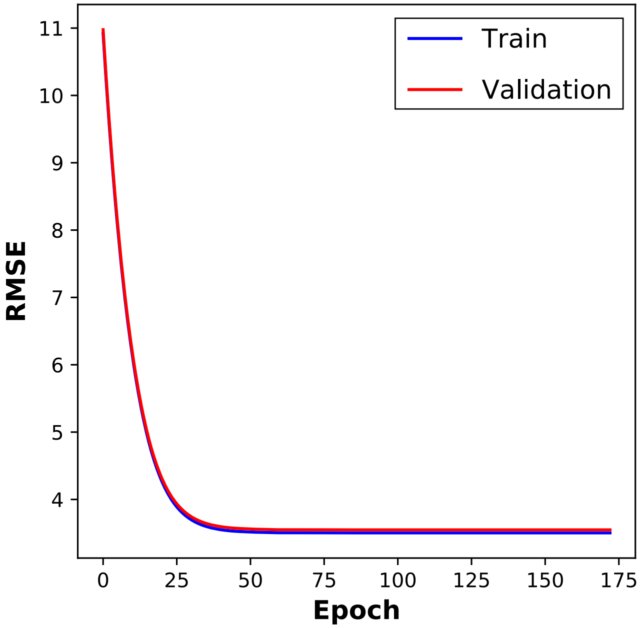
The generated data are used to build the wall model using the ML technique of gradient boosting (Friedman, 2000). The gradient boosting system used to train the model is XGBoost by Tianqi Chen and Guestrin (2016). At the time of conception of the idea of using ML for wall modeling, XGBoost was one of the popular algorithms among data scientists. XGBoost offers the *feature importance* property, which allows to assess the sensitivity of the input features with respect to the output predictions. Hence, this property helps to automatically select the best possible inputs among a list of candidates. The importance of the inputs is further discussed in the next section.

Gradient boosting uses regression trees as "weak learners" for making predictions. The model is an additive function in the form of tree structures that maps the input space to the output space. Consider the data set $\mathcal{D}$ composed of *n* samples and *m* features:

$$\mathcal{D} = \left\{(\mathbf{x}_i, y_i)\right\}\left(|\mathcal{D}| = n, \mathbf{x}_i \in \mathbb{R}^m, y_i \in \mathbb{R}\right)$$

$$\hat{y} = \phi(\mathbf{x}_i) = \sum_{k=1}^{K} f_k(\mathbf{x}_i),\ f_k \in \mathcal{F}$$

$$\text{where } \mathcal{F} = f(\mathbf{x}) = \omega_{s(\mathbf{x})}\left(s\colon \mathbb{R}^m \to \mathbf{L}, \omega \in \mathbb{R}^{\mathbf{L}}\right)$$



Source: Figure by authors

Figure 5. Evolution of the loss in terms of root mean square error (RMSE) of the train data and the validation data with respect to each epoch of the training process

$\hat{y}$ predicted from K additive functions, s represents the tree structure, ω is the output weight and L is the number of leaves. The optimizer is the gradient descent, and the loss function used for the algorithm is:

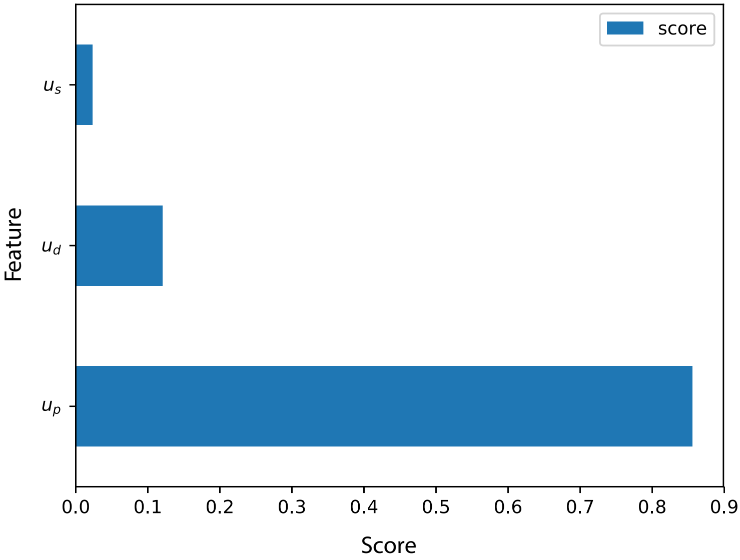
$$\mathcal{L}(\phi) = \sum_i l\left(y_i, \hat{y}_i\right) + \sum_k \Omega(f_k)$$

$$\text{where } \Omega(f) = \gamma L + \frac{1}{2}\lambda||\omega||^2$$

Here l is any differentiable convex loss function that measures the difference between y and $\hat{y}$, and λ is a regularization parameter. In XGBoost regression, the cost function used is the sum of the squared error. This function measures the squared difference between the predicted and the actual values. It is given by $l\left(y, \hat{y}\right) = \left(y - \hat{y}\right)^2$.

2.5 Model inputs and outputs

The input features processed from the flow must carry the necessary physical information to build a model that can be generalized to all types of flow configuration. After testing different input features based on velocity, strain rate and pressure, two non-dimensional reference-frame invariant features are chosen. These two input features are selected because of their high *feature importance* score. An example of this procedure is shown in Figure 6. u_s



Source: Figure by authors

Figure 6. Feature importance plot showing the importance of three inputs used for training the model. The higher the value, the higher the sensitivity of the feature in predicting the output

[equation (6)], u_d [equation (7)] and u_p [equation (8)] compose the set of candidate inputs used to train the model. The quantity u_s is a strain-based velocity scaling defined as:

$$u_s = \frac{u}{||S||\delta_n} \tag{6}$$

where u is the magnitude of the instantaneous velocity parallel to the walls, and δ_n is the normal distance from the wall. $||S||$ is the magnitude of the strain rate tensor computed as $||S|| = \sqrt{2s_{ij}s_{ji}}$, where s is the instantaneous strain rate. From the feature importance plot, it can be observed that u_s appears to be ineffective compared to u_d and u_p. u_d is called the "local Reynolds number" given by:

$$u_d = \frac{u\delta_n}{\nu} \tag{7}$$

The second, u_p, which is also a velocity scaling, is based on the magnitude of the pressure gradient. It is computed as:

$$u_p = \frac{u}{\sqrt[3]{\frac{\nu}{\rho}\partial_\theta p}} \tag{8}$$

where ρ is the density of the fluid, and $\partial_\theta p$ is the magnitude of the wall-tangential pressure gradient at a normal distance δ_n. The model output is u^+ given by:

$$u^+ = \frac{u}{\sqrt{\frac{\tau_\omega}{\rho}}} \tag{9}$$

where τ_ω is the magnitude of the wall shear stress. To teach the model to distinguish between favorable and adverse pressure gradients, the local u is set to be positive for local favorable pressure gradients, and negative for adverse pressure gradients:

$$u = \begin{cases} u, & \text{if } u \cdot \nabla p \leq 0 \\ -u, & \text{otherwise} \end{cases}$$

for the model inputs u_d and u_p. For the output of the model, regardless of the sign of the predicted u^+, the magnitude of the shear stress is calculated and the direction is chosen as the direction of the instantaneous velocity at the exchange location. This quantity was found to be more effective than the other inputs tested, as shown in Figure 6, where the strain rate-based velocity scaling has a feature importance score close to zero.

Like any other ML system, XGBoost also requires hyperparameters to define and train the model. The size of each "tree" in the model is controlled by the hyperparameter "*max_depth*." It is set to 7 so that each tree has 2^7 leaves. Beyond this value, the complexity seems to be too high, and the model gets over-fitted to the data. The leaves contain the image of the tree functions. The rate at which learning takes place is controlled by the "*learning_rate*" parameter. This hyperparameter is responsible for the extent to which the model is altered in response to loss after each epoch. Consequently, the lower the *learning_rate*, the greater the number of trees, hence increasing the forest size. The learning rate parameter is set to 0.07 for this study. As shown in Figure 5, beyond 70 epochs, there is no reduction in loss with evolution. Therefore, tree number 70 is chosen for the prediction of the output.

Finally, the model is integrated with the LES solver, as shown in the schematic Figure 1. For each time step, the stream-wise instantaneous velocity, tangential pressure gradient and wall distance sampled at approximately 10% of the thickness of the BL are used to compute u_d and u_p, the model inputs. This is consistent with the recommendations on the use of LES inputs for wall modeling, as explained in Kawai and Larsson (2012). The model computes the necessary boundary condition for the simulation, that is, the wall shear stress, which is then fed back into the LES simulation as a Neumann boundary condition.

3. Results and discussion

In this section, the performance of the model is evaluated *a priori* and *a posteriori*. In *a priori* tests, the model's performance is compared offline with a data-driven equilibrium wall model (hereafter called EQWM-ML) (Radhakrishnan *et al.*, 2021). In the *a posteriori* tests, which consist of embedding the model during the simulation, it is compared with an algebraic equilibrium wall shear stress model (hereafter called EQWM) (Owen *et al.*, 2019) and EQWM-ML.

3.1 *A priori tests*

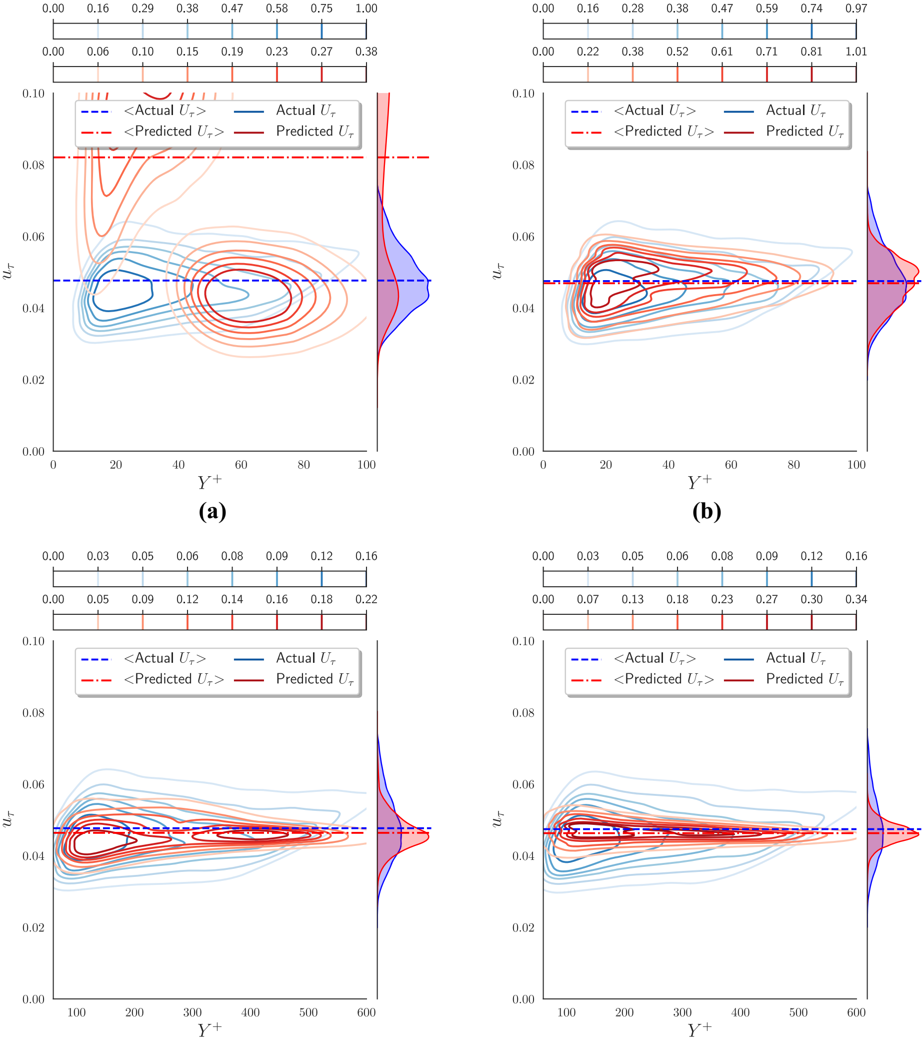
The *a priori* tests are performed by sampling unseen data from a channel case and the double diffuser and using the model to make predictions. These tests ensure that the model can distinguish between non-equilibrium and equilibrium data. This has to be confirmed, as the model is trained using a mix of both data types. The channel data used for *a priori* tests are derived from the WRLES of a channel at $Re_\tau = 550$. This data set is particularly used as it is different from the channel data distribution ($Re_\tau = 950$) that is used for training. The data derived from the diffuser for the *a priori* test is sampled from the DNS of the same Stanford double diffuser. The predicted u_τ is compared with the actual u_τ. To compare the predicted and actual result distributions, a bivariate kernel density estimate (KDE) (Simonoff, 2012) of the u_τ and y^+ quantities is computed. The KDE plot is obtained by keeping a kernel of a specific bandwidth on each point on the data-space and summing them over. This results in a smooth probability density function (PDF) even at the locations in the data-space without sampled data points. The kernel used here is Gaussian and $< \cdot >$ represents the mean value. A marginal distribution is also plotted to the right of the bivariate plot.

3.1.1 Channel $Re_\tau = 550$. The computational domain for this channel is $6\delta \times 2\delta \times 3\delta$ in the stream-wise, wall-normal and span-wise directions respectively. A mesh of $128 \times 96 \times 128$ elements is used to discretize the domain. The mesh is uniform in the stream-wise and span-wise directions, corresponding to $\Delta x^+ \approx 26$ and $\Delta z^+ \approx 13$ in wall units, respectively. The hyperbolic tangent (cf. 3.2.2) stretching with a minimum $y^+ \approx 1$ is used here as well. Periodic boundary conditions are enforced on the stream-wise and span-wise boundaries, and a no-slip boundary on the wall is maintained. The flow is driven by a stream-wise constant pressure gradient and a Vreman (Vreman, 2004) SGS model is used for turbulence closure.

In Figure 7, the plots on the left correspond to the predicted versus actual PDF of u_τ for the NEQWM-ML. The ones on the right correspond to the predicted versus actual PDF of u_τ for the EQWM-ML. The blue PDF is the actual instantaneous u_τ corresponding to y^+. The red PDF is the instantaneous predicted u_τ corresponding to y^+. A typical y^+ range that is applicable to a wall model is considered. The figures are divided according to the y^+ ranges. Figure 7(a) and 7(b), are the KDE plots of u_τ and y^+ for $y^+ < 100$. KDE plots of $y^+ > 100$ are shown in Figure 7(c) and 7(d), . From Figure 7(c), it is clear that NEQWM-ML performs very well for $y^+ > 100$ in the equilibrium domains. The peaks are well captured, and the mean is also very close. Compared to Figure 7(c), the performance of NEQWM-ML is on par with EQWM-ML. However, the behavior is different when we examine the performance of NEQWM-ML for data corresponding to $y^+ < 100$. It seems that introducing non-equilibrium data with purely equilibrium data may distort the prediction in the near-wall area where viscous effects are more relevant. Figure 7(a) shows that NEQWM-ML fails to capture u_τ while the performance of EQWM-ML is very good. Whether this affects the overall performance of the NEQWM-ML in complex flows can only be explained after *a posteriori* tests.

3.1.2 Stanford double diffuser. Unseen data from regions close to the corners of the diffuser are probed. These locations are shown in Figure 8. The four probe locations inside the cyan diffuser are chosen. With respect to the flow direction, they are labeled as lower left (LL), lower right (LR), upper left (UL) and upper right (UR).

In Figure 9, the blue PDF is the actual instantaneous u_τ corresponding to y^+. The red PDF is the predicted instantaneous u_τ corresponding to y^+. All figures on the left correspond to the predicted versus actual u_τ quantities of NEQWM-ML. The plots on the right correspond to the predicted versus actual u_τ quantities of EQWM-ML. Density values are given at the top of the figure. Figure 9(a) and 9(b), show that the actual u_τ has a peak value of approximately 0.025 at $y^+ < 100$. NEQWM-ML is able to this peak approximately, while EQWM-ML does not capture it. The average value of the predicted u_τ is not well predicted by the NEQWM-ML in this region. The prediction of EQWM-ML cannot be seen in the figure, as it is beyond the bounds of the figure, and it is therefore very inaccurate. Figure 9(c) and 9(d), show that the peak value of approximately 0.015 at $y^+ < 100$ is well predicted by NEQWM-ML. In addition, the marginal distribution is very well captured by NEQWM-ML. The average value of the predicted u_τ is also well captured by the NEQWM-ML for this region. EQWM-ML

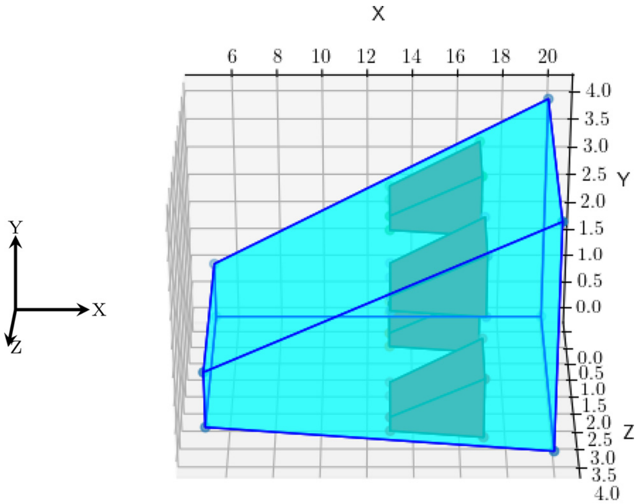


(a) **(b)** **(c)** **(d)**

Notes: (a) NEQWM-ML; (b) EQWM-ML; (c) NEQWM-ML; (d) EQWM-ML; On the left, the performance of the NEQWM-ML is shown. On the right, that of the EQWM-ML is shown. The blue color corresponds to actual values and the red color corresponds to predicted values. The dotted lines are shown for the averages. Densities are shown at the top of the figure, and marginal distributions are shown on the right of each figure

Source: Figure by authors

Figure 7. Comparison of the KDE of actual and predicted wall shear stresses for NEQWM-ML and EQWM-ML for the channel $Re_\tau = 550$



Source: Figure by authors

Figure 8. Probe locations from where the data is sampled

fails again to predict the correct PDE and average value. Similar conclusions can be drawn by comparing Figures 9(e) and 9(f), and Figure 9(g) and 9(h).

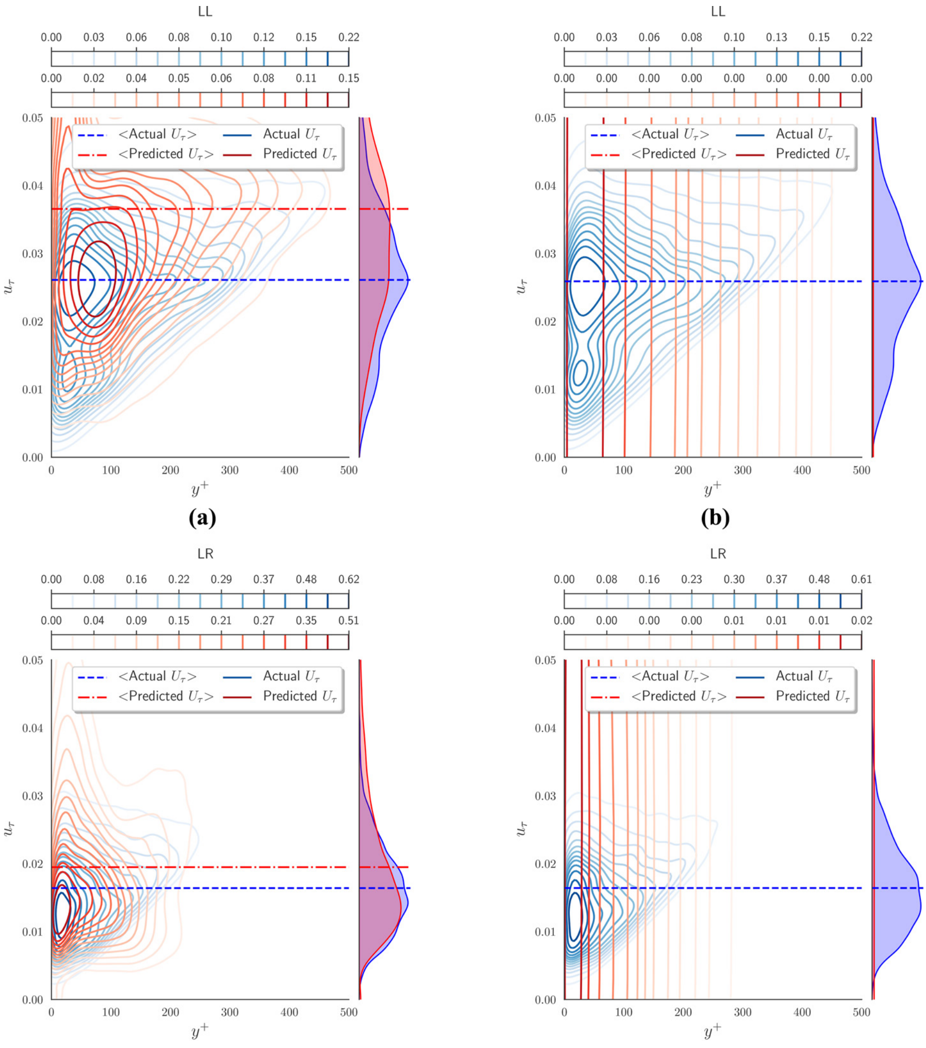
Considering the fact that NEQWM-ML is trained with only a limited amount of data, its *a priori* performance is promising. However, getting good results in *a priori* test cases does not always guarantee good results in actual simulations (Poroseva *et al.*, 2016; Thompson *et al.*, 2016; Wu *et al.*, 2019). Therefore, the performance of NEQWM-ML in actual flow simulations is evaluated in the following cases:

- flow inside a diffuser;
- turbulent channel flow of $Re_\tau = 2005$;
- turbulent channel flow of $Re_\tau \approx 4200$;
- flow over a wall-mounted hump; and
- juncture flow simulation.

The first two tests guarantee that the non-equilibrium data contained in the NEQWM-ML do not interfere with the equilibrium data when it encounters equilibrium conditions. The third test is performed to check the performance of the model when tested on the Stanford diffuser, the data from which were used to train the model. The fourth case tests the NEQWM-ML's performance when subjected to relatively smooth geometries under strong non-equilibrium conditions. The fifth case evaluates the performance of the NEQWM-ML in an actual industrial case with sharp corners. Details of the simulations are discussed below.

3.2 A posterior tests

3.2.1 Stanford double diffuser. This test is intended to check the *a posteriori* performance of the model when used in the diffuser. Geometric details are maintained as in the original set up, detailed in Miró *et al.* (2023). Three uniform meshes were used to carry out the grid convergence



(a)

(b)

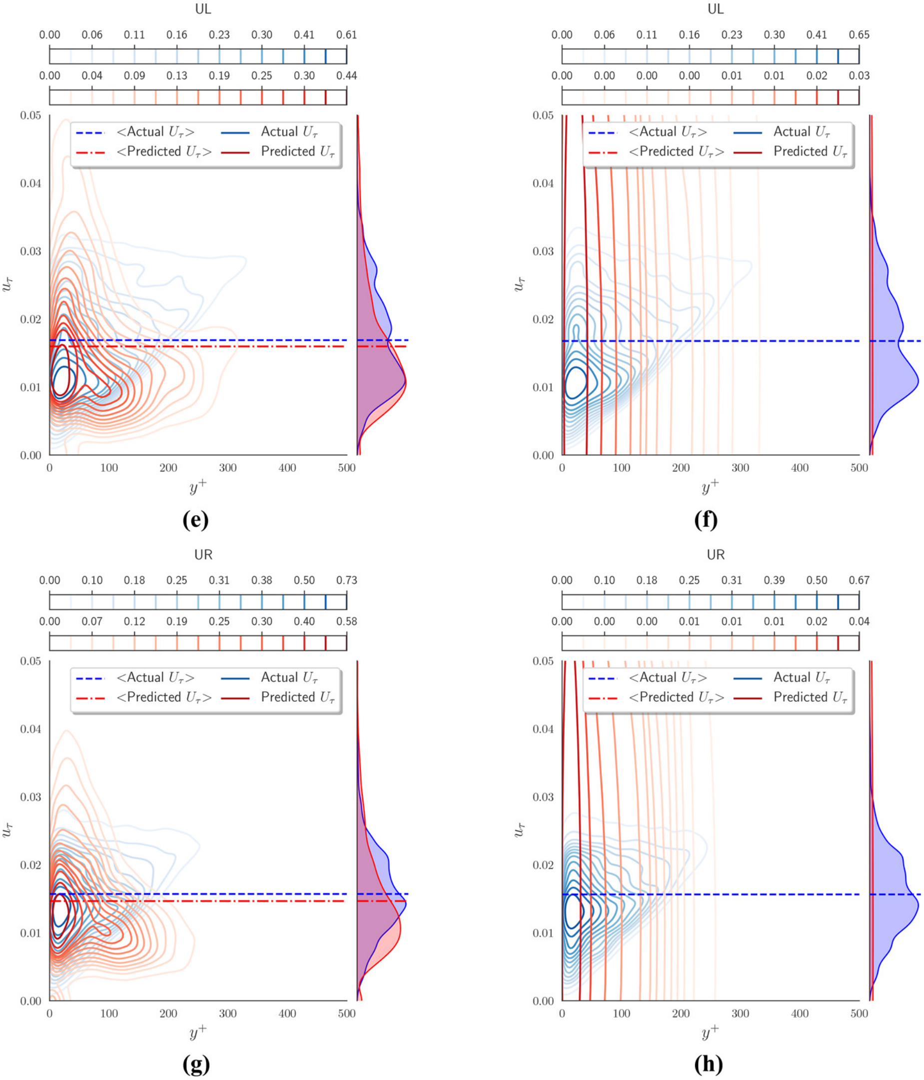
(c)

(d)

Figure 9. Comparison of the KDE of actual and predicted wall shear stresses for NEQWM-ML and EQWM-ML for data from the Stanford double diffuser

study. The first mesh has approximately 1.4 million linear elements. The second mesh, G2, has approximately 3.2 million linear elements, which is obtained by refining in the x direction. The third mesh G3 has approximately 7.1 million linear elements, which is obtained by further refinement of G2 in the x direction. The details of the meshes are shown in Table 1. Turbulent inflow is generated using a tripping mechanism as described in Miró *et al.* (2023).

Statistics are collected after a statistically stationary state is reached. Stream-wise velocity profiles along four different planes are considered for comparison with the DNS data. These



(e) (f) (g) (h)

Notes: (a) NEQWM-ML; (b) EQWM-ML; (c) NEQWM-ML; (d) EQWM-ML; (e) NEQWM-ML; (f) EQWM-ML; (g) NEQWM-ML; (h) EQWM-ML. The performance of the NEQWM-ML and EQWM-ML are displayed on the left and right, respectively. The four rows correspond to the probe locations. The blue color corresponds to actual values and the red color corresponds to predicted values. The dotted lines are shown for the averages. Densities are shown at the top of the figure, and marginal distributions are shown on the right of each figure

Source: Figure by authors

Figure 9. Continued

Table 1. Details of the meshes used for the simulation of the diffuser. *x*, *y* and *z* indicate the streamwise, wall-normal and spanwise directions, respectively

Mesh	No. of elements	Δx^+_{max}	Δy^+_{max}	Δz^+_{max}	y^+	z^+
G1	≈ 1.4 million	≈ 238	≈ 22	≈ 53	≈ 8.3	≈ 22.0
G2	≈ 3.2 million	≈ 143	≈ 16	≈ 38	≈ 6.9	≈ 11.5
G3	≈ 7.1 million	≈ 102	≈ 12	≈ 26	≈ 5.5	≈ 9.2

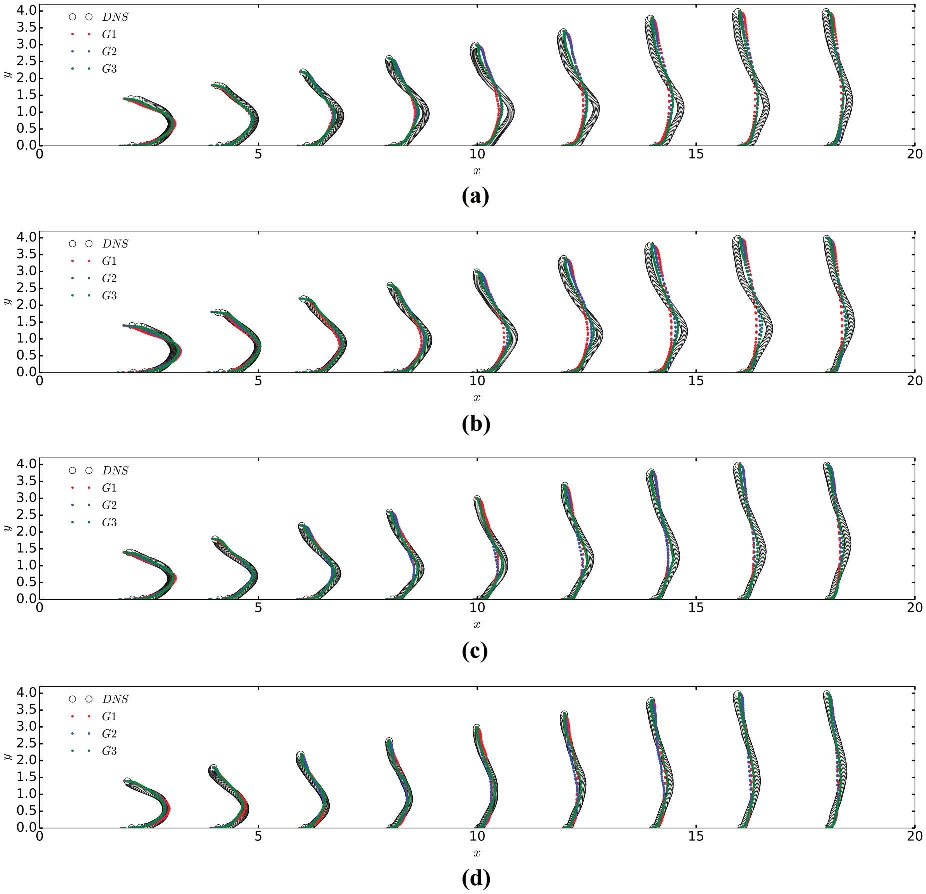
Notes: The grid sizes are non-dimensionalized in wall units. Δx_{max}, Δy_{max} and Δz_{max} are the maximum grid lengths inside the duct in the stream-wise, wall-normal and span-wise directions, respectively. y^+ and z^+ are the resolutions of the walls.
Source: Table by authors

planes are located at $z/B = 0.250$, $z/B = 0.500$, $z/B = 0.750$ and $z/B = 0.875$, where B is the width of the duct. Velocity profiles are generated at $x = 2.0$, $x = 4.0$, $x = 6.0$, $x = 8.0$, $x = 10.0$, $x = 12.0$, $x = 14.0$, $x = 16.0$, $x = 17.0$ and $x = 18.0$ for each of these planes. The comparison of these velocity profiles with DNS data for the three different meshes is shown in Figure 10. The velocity profiles show that the results are sufficiently converged. Next, the results of NEQWM-ML are compared with those of EQWM, EQWM-ML and the results of a simulation in which no wall model is used, that is, a coarse LES case (*noModel*). These are shown in Figure 11. In general, the improvement in the velocity profiles is not very significant compared to the *noModel* case. However, it can be inferred that both equilibrium wall models struggle in the recirculation region. NEQWM on the other hand, not only has an improvement with respect to the *noModel* case, but also improves with respect to the two equilibrium wall models in the recirculation regions. This can be attributed to the use of non-equilibrium data in the training process. The locations where NEQWM-ML matches the *noModel* are mostly the regions of laminar flow. Furthermore, the stream-wise velocity fluctuations predicted by the three models and the *noModel* case are compared in Figure 12. The fluctuations predicted from all simulations disagree with the DNS and with each other.

In summary, NEQWM-ML makes a notable improvement when tested *a posteriori* compared to its equilibrium counterparts, although it does not significantly improve the velocity profiles from a coarse LES (*noModel*). The key problem lies in the use of a structured mesh for the WMLES of the diffuser. Although refinement inside the duct is adequate for wall modeling, the case is not the same inside the diffuser, where the geometry diverges, hence stretching the mesh elements. This also explains the disagreement of the velocity profiles away from the walls. Anisotropic elements also affect exchange location points, causing wall models to perform inefficiently. In addition to that, the Reynolds number for this case is a notably low to perform WMLES.

3.2.2 Turbulent channel flow ($Re_\tau = 2005$). The size of the channel is $6\pi\delta \times 2\delta \times \times 2\pi\delta$ in the stream-wise, wall-normal and span-wise directions, respectively. In this study, three different mesh resolutions are considered. The meshes are uniform in the stream-wise, wall-normal and span-wise directions, details of which are given in Table. 2.

Stream-wise and span-wise directions, being homogeneous, periodic boundary conditions are applied on those boundaries. Due to the larger grid size close to the walls, the dynamically significant eddies cannot be resolved by the grid. This results in an incorrect velocity profile and, subsequently, incorrect wall shear stress. So, the no-slip condition is no longer valid at the walls, and a slip wall with the no-penetration condition is imposed on the walls. The flow is driven by a stream-wise constant pressure gradient, and the Vreman (Vreman, 2004) SGS model is used for turbulence closure. Simulations are run long enough



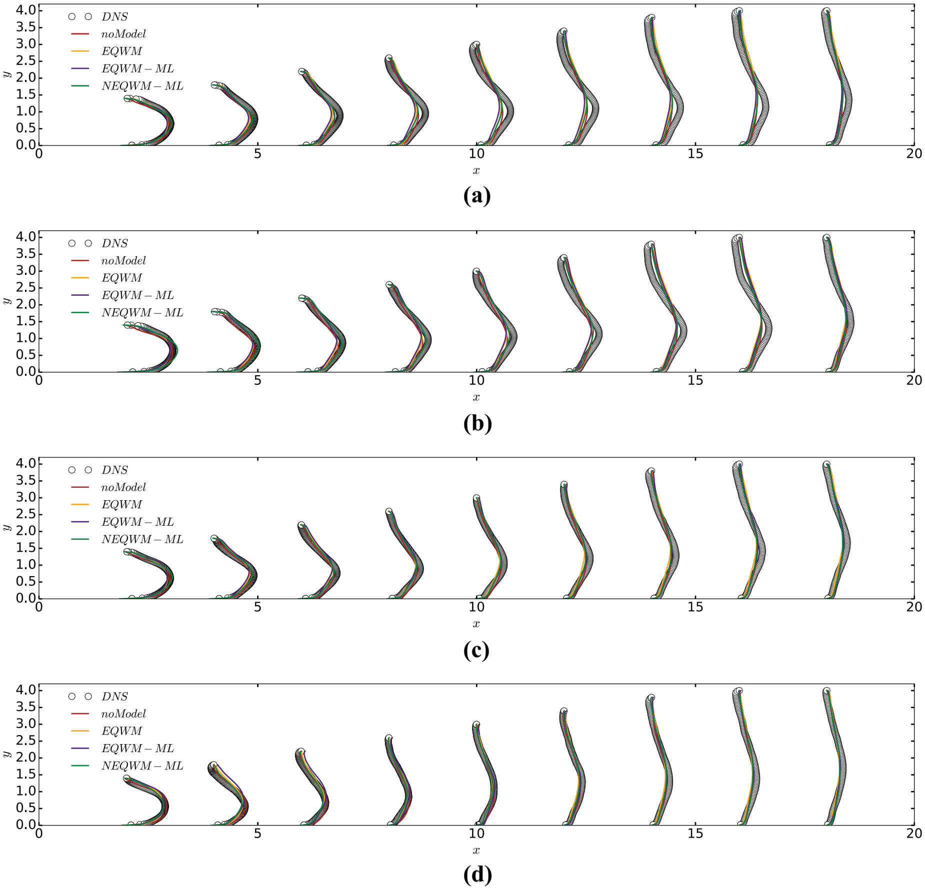
Notes: (a) *z/B* = 0.250; (b) *z/B* = 0.500; (c) *z/B* = 0.750; (d) *z/B* = 0.875, where B is the width of the duct

Source: Figure by authors

Figure 10. Comparison of the velocity profiles predicted by NEQWM-ML for meshes G1, G2 and G3 at $x = 2.0$, $x = 4.0$, $x = 6.0$, $x = 8.0$, $x = 10.0$, $x = 12.0$, $x = 14.0$, $x = 16.0$, $x = 17.0$ and $x = 18.0$ on planes

to achieve statistically stationary regimes, and statistics are collected over a period of 20 flow-through times. One flow-through time is defined as the time it takes for the center-line stream-wise velocity to cover the domain length. In addition to averaging in time, spatial averaging is also performed along the homogeneous directions, and the fields are normalized to wall units with the frictional velocity of the flow.

The mean stream-wise velocity in wall units for M1, M2 and M3 meshes compared to DNS results (Hoyas and Jiménez, 2006) is shown in Figure 13. As the mesh density increases, the velocity profiles tend to approach DNS. The error in u_τ predicted by NEQWM-ML is less than 0.5% for all simulations. This shows that the data-based NEQWM-ML is able to blend with the fluid solver like a physics-based EQWM. Furthermore, the ability of NEQWM-ML to make

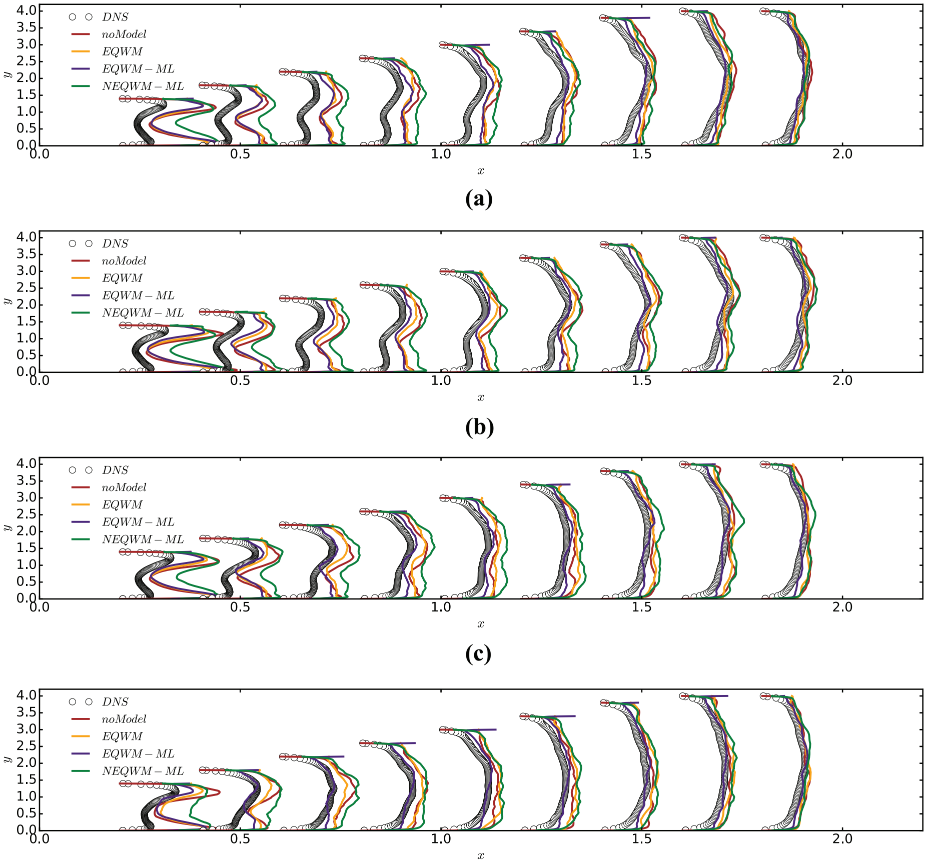


Notes: (a) *z/B* = 0.250; (b) *z/B* = 0.500; (c) *z/B* = 0.750; (d) *z/B* = 0.875, where B is the width of the duct

Source: Figure by authors

Figure 11. Comparison of the velocity profiles predicted by the wall models at $x = 2.0$, $x = 4.0$, $x = 6.0$, $x = 8.0$, $x = 10.0$, $x = 12.0$, $x = 14.0$, $x = 16.0$, $x = 17.0$ and $x = 18.0$ on planes

accurate predictions in a higher Re_τ channel flow shows that NEQWM-ML has learned the law of the wall without being explicitly given any information about it during training. This is further confirmed in Figure 14 where the results are compared with those of a simulation using EQWM. The EQWM (Owen *et al.*, 2019) is used for this simulation, and the mesh used is M3. Everything else remains the same, except for the input to the EQWM, which is the average velocity at the point of exchange. The mean stream-wise velocity profiles and the mean fluctuations predicted from both simulations are in very good agreement. It is interesting to note that even with a model that works instantaneously, the mean fluctuations predictions could not be improved. The algebraic model relies on the law of the wall, which is a mean velocity profile, while the NEQWM-ML does not rely on any temporal averaging procedure. One reason for the



(a)

(b)

(c)

(d)

Notes: (a) *z/B* = 0.250; (b) *z/B* = 0.500; (c) *z/B* = 0.750; (d) *z/B* = 0.875, where B is the width of the duct

Source: Figure by authors

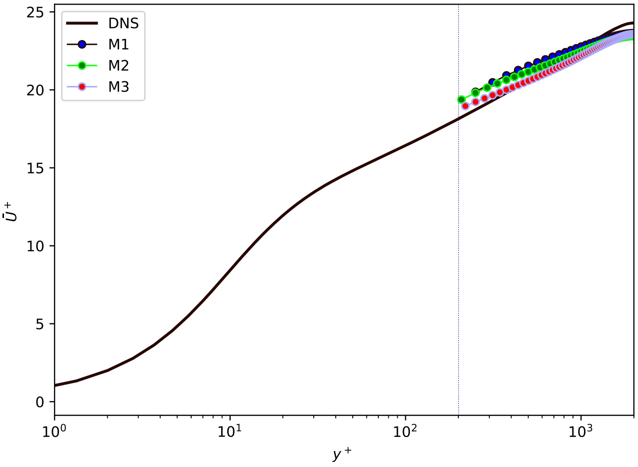
Figure 12. Comparison of the stream-wise velocity fluctuations predicted by the wall models at $x = 2.0$, $x = 4.0$, $x = 6.0$, $x = 8.0$, $x = 10.0$, $x = 12.0$, $x = 14.0$, $x = 16.0$, $x = 17.0$ and $x = 18.0$ on planes

Table 2. Details of the meshes used for the simulation of the turbulent channel flow at $Re_\tau = 2005$

Mesh	size ($N_x \times N_y \times N_z$)	Δx^+	Δy^+	Δz^+
M1	$64 \times 64 \times 64$	≈ 591	≈ 63	≈ 196
M2	$128 \times 96 \times 96$	≈ 295	≈ 42	≈ 131
M3	$256 \times 128 \times 128$	≈ 147	≈ 31	≈ 98

Note: N_x, N_y and N_z represent the number of elements in the stream-wise, wall-normal and span-wise directions, respectively

Source: Table by authors



Source: Figure by authors

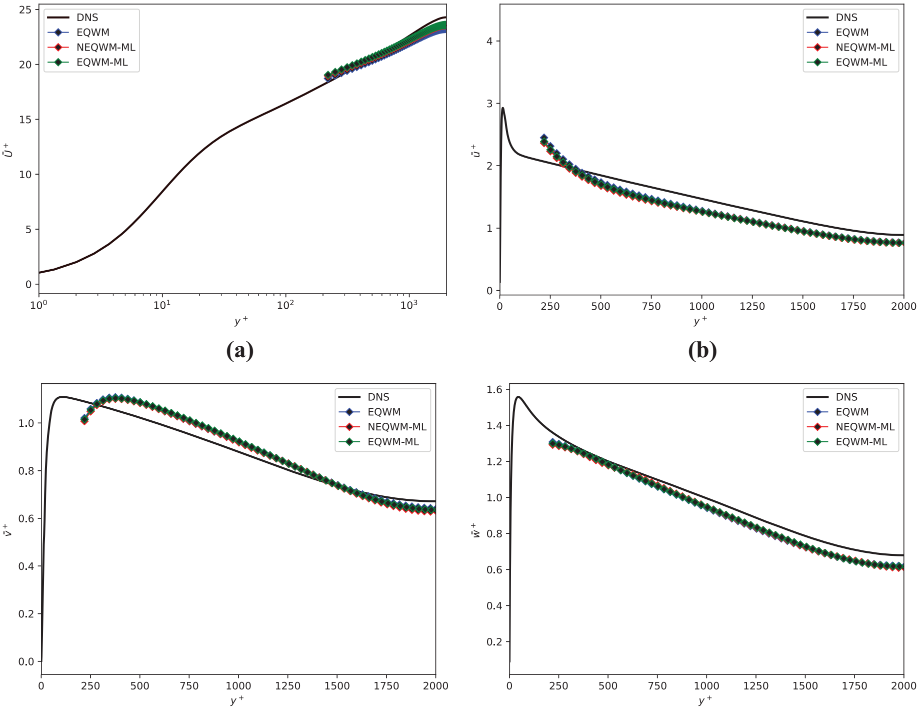
Figure 13. Mean stream-wise velocity profiles in wall units predicted by the NEQWM-ML for M1, M2 and M3 compare with DNS (Hoyas and Jiménez, 2008) for channel flow at $Re_\tau = 2005$. The blue vertical dotted line indicates the location of data exchange

inability to capture fluctuations is that the time steps and the grid size are still too large for the simulation to correctly resolve the near-wall dynamics. Second, 10% of the BL always is above the primary production peak and, too often, the secondary production peak. The figures also show the comparison of the performance of NEQWM-ML with that of EQWM-ML. The results show that the NEQWM-ML can perform as an equilibrium wall model when subjected to equilibrium flow conditions, and its performance is unaltered even though it is partially trained with non-equilibrium data.

3.2.3 Turbulent channel flow ($Re_\tau = 4200$). The size of this channel is maintained as in the previous case. Three uniform meshes are used to test for convergence. Details of the meshes are given in Table 3. Δy^+ is maintained constant for all three meshes, and refinements are made in the x and z directions. Periodic boundary conditions are applied in the stream-wise and span-wise directions, and a no-penetration condition is imposed on the walls. Like in the previous case, the flow is driven by a streamwise constant pressure gradient. Vreman (Vreman, 2004) SGS model is used to close the turbulence. Once the statistically stationary regime is reached, statistics are collected to study the performance of the model.

Figure 15 shows the mean stream-wise velocity profiles in wall units predicted by NEQWM-ML for the three meshes M1, M2 and M3 compared to DNS (Lozano-Durán and Jiménez, 2014). Velocity profiles tend to approach DNS values as the mesh density increases. As in the previous example, the error in the wall shear stress computed by NEQWM-ML is less than 0.5%. Furthermore, the performance of NEQWM-ML is compared with that of EQWM and EQWM-ML in Figure 16, showing that all models provide similar results and are in good agreement with the DNS reference data.

3.2.4 Wall-mounted hump. In this section, the performance of the NEQWM-ML in the simulation of the flow over a wall-mounted hump is studied. This flow involves turbulent to laminar transition, separation, reattachment and recovery of the BL, which are hallmarks of many industrial flows, and it is therefore considered a benchmark for testing turbulence



(a) **(b)**

(c) **(d)**

Notes: (a) y^+ vs *U, shows the mean stream-wise velocity and*; (b) y^+ vs u^+_{rms}; (c) y^+ vs v^+_{rms}; (d) y^+ vs $w^+_{rms,}$ show the mean fluctuations

Source: Figure by authors

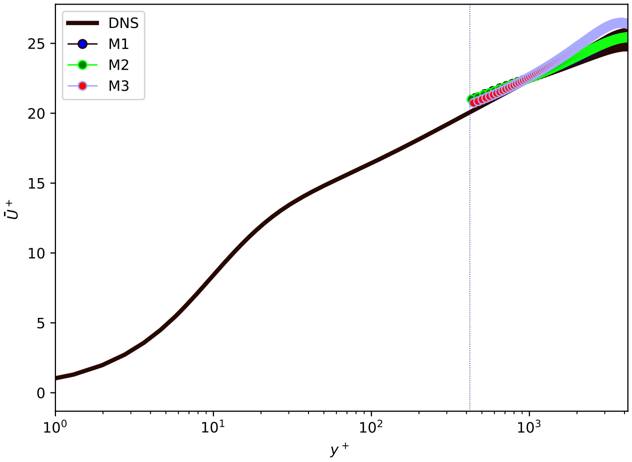
Figure 14. Comparison of the performance of the NEQWM-ML with that of the EQWM and EQWM-ML for channel flow at $Re_\tau = 2005$

Table 3. Details of the channel mesh $Re_\tau \approx 4200$. Δx^+, Δy^+ and Δz^+ are refinements in wall units in the streamwise, wall-normal and spanwise directions, respectively

Mesh	No. of elements	Δx^+	Δy^+	Δz^+
M1	≈ 2.1 million	≈ 615	≈ 65	≈ 205
M2	≈ 9.4 million	≈ 309	≈ 44	≈ 137
M3	≈ 33.5 million	≈ 154	≈ 33	≈ 103

Source: Table by authors

models. The geometry of the hump is defined following the guidelines of the NASA CFDVAL2004 workshop (Rumsey *et al.*, 2006). The results of the numerical simulation are compared with the experimental data of Naughton *et al.* (2006). The NEQWM-ML performance is also compared with that of EQWM-ML.



Source: Figure by authors

Figure 15. Mean stream-wise velocity profiles in wall units predicted by the NEQWM-ML for M1, M2 and M3 compare with DNS (Lozano-Durán and Jiménez, 2014) for channel flow at $Re_\tau = 4200$

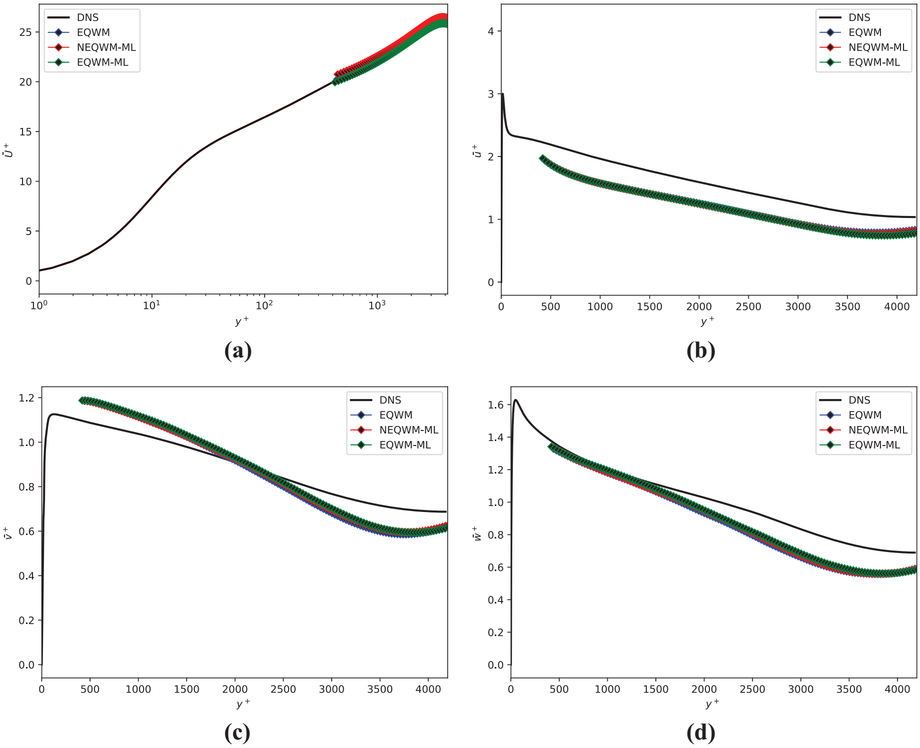
The computational domain for this case is $4.64c$, $0.909c$, and $0.3c$ in the stream-wise, normal and span-wise directions, respectively, where c is the chord length of the hump. The span-wise length is determined so that it does not restrict any turbulent structures in the span-wise direction. To ensure this, the span-wise two-point correlations (R) of the velocity components over a period of 32 time units in the vicinity of the separation bubble were calculated for the fine mesh (mesh details explained below). R is computed as:

$$R(dz) = \frac{< \phi(z)\phi(z+dz) >}{< \phi^2(z) >} \tag{10}$$

where $\phi = u', v', w'$, respectively, for stream-wise, wall-normal and span-wise velocity fluctuations. $< \cdot >$ denotes the mean. The results are shown in Figure 17. It can be seen that the span-wise length is sufficient to contain the large eddies in the vicinity of the bubble. The inlet and outlet planes of the domain are at $x/c = -2.14$ and $x/c = 2.5$, respectively, so the leading edge of the hump can be placed at $x/c = 0$. Because end plates were used in the experiment, to account for the effects of these, the top wall of the domain has a contour from $x/c = -0.5$ to $x/c = 1.5$, as shown in Figure 18.

Three different grids are used for simulations. The coarse mesh (G1) has approximately 3.1 million linear elements, with 742 × 70 × 60 elements in the stream-wise, normal and span-wise directions, respectively. The fine mesh (G2) has approximately eight million linear elements with 900 × 110 × 80 elements. G2 is generated by refining the tangential directions and reducing the mesh growth rate in the normal direction from 1.06 to 1.03. The third grid G3 is only used to ensure grid convergence. The mesh is obtained by refining G2 in the stream-wise direction. G3 has approximately 15 million linear elements with a refinement of 1654 × 110 × 80 in the stream-wise, wall-normal and span-wise directions, respectively.

More details are shown in Figure 20 where the grid spacing in wall units between $x/c = -0.5; x/c = 2.0$ for all the meshes are depicted. The Reynolds number of the flow



(a) (b) (c) (d)

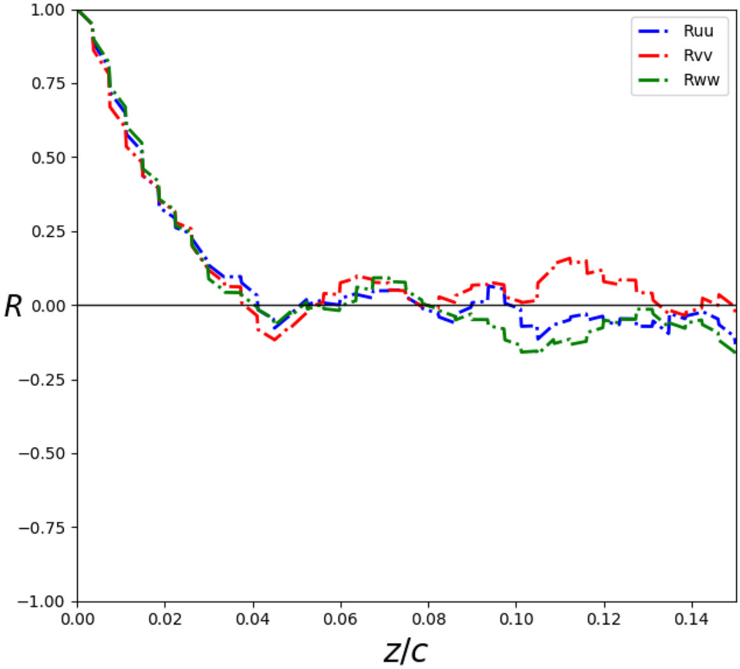
Notes: (a) y^+ vs U^+; (b) y^+ vs u^+_{rms}; (c) y^+ vs v^+_{rms}; (d) y^+ vs w^+_{rms}. a shows the mean stream-wise velocity and b, c, d show the mean fluctuations

Source: Figure by authors

Figure 16. Comparison of the performance of the NEQWM-ML with that of the EQWM and EQWM-ML for channel flow at $Re_\tau = 4200$

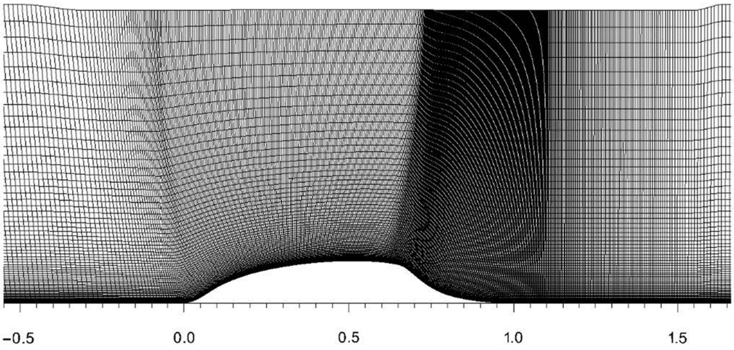
based on the length of the hump chord, c, and the free stream velocity, U_∞, at the inlet is $Re_c = 936{,}000$. Periodic boundary conditions are imposed in the span-wise direction. A slip boundary condition is applied at the top boundary of the domain, and at the bottom wall where the wall stress is predicted, a no-penetration condition is imposed. For turbulence closure, Vreman (2004) SGS model is used. Synthetic turbulence as inflow data is generated based on the technique described in Kempf *et al.* (2005). Due to missing experimental data at the inflow plane, the missing Reynolds stresses are specified to match those used by Park (2015) (left panels of Figure 19). The evolution of realistic turbulence is ensured before the flow reaches the hump by comparing the mean velocity and Reynolds stresses of the present simulation with that of Avdis *et al.* (2009) at a downstream location ($x/c = -0.81$), shown in the right panels of Figure 19 for G2. After achieving a statistically stationary state, statistics are collected and results are averaged over a period of 20 flow-through times. Subsequently, spatial averaging is also done.

The average values of the coefficients of skin friction (C_f) and pressure (C_p) on the hump for the G1 and G2 meshes are shown in Figure 22. The coefficients for both simulations are close to



Source: Figure by authors

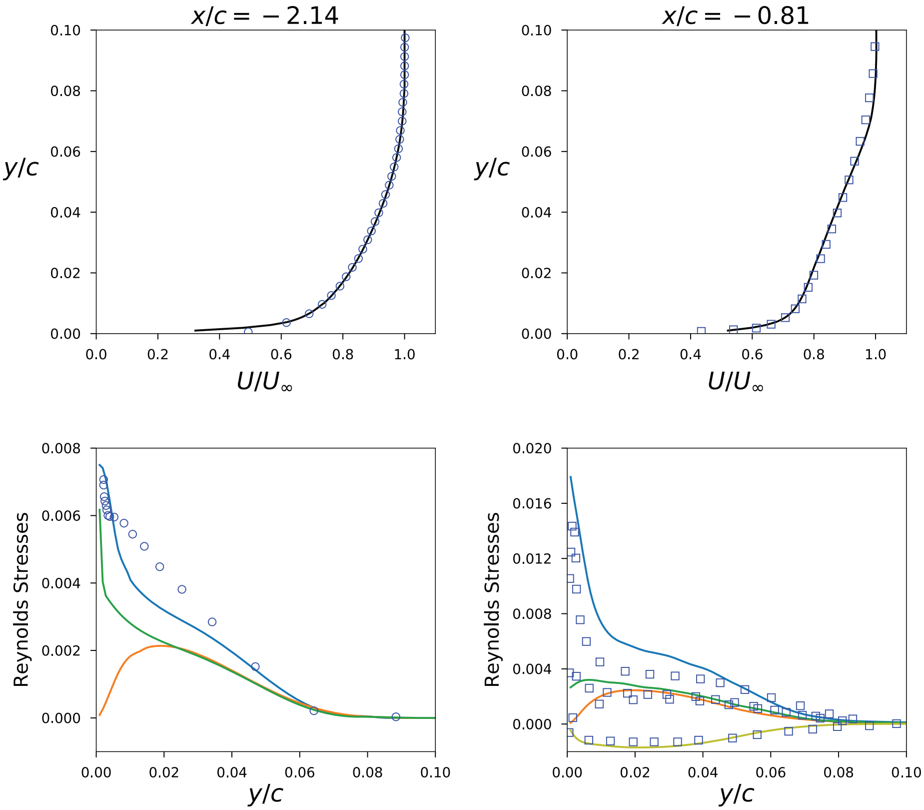
Figure 17. Span-wise auto-correlations (*R*) of the velocity components in the vicinity of the separation bubble. *Ruu*, *Rvv* and *Rww* are, respectively, the auto-correlations of the stream-wise, wall-normal and span-wise fluctuations of velocity



Notes: The contour between *x/c* = −0.5 and *x/c* = 1.5 is visible. Figure courtesy of Owen *et al*. (2019) 2018

Source: Figure by authors

Figure 18. Computational mesh around the hump

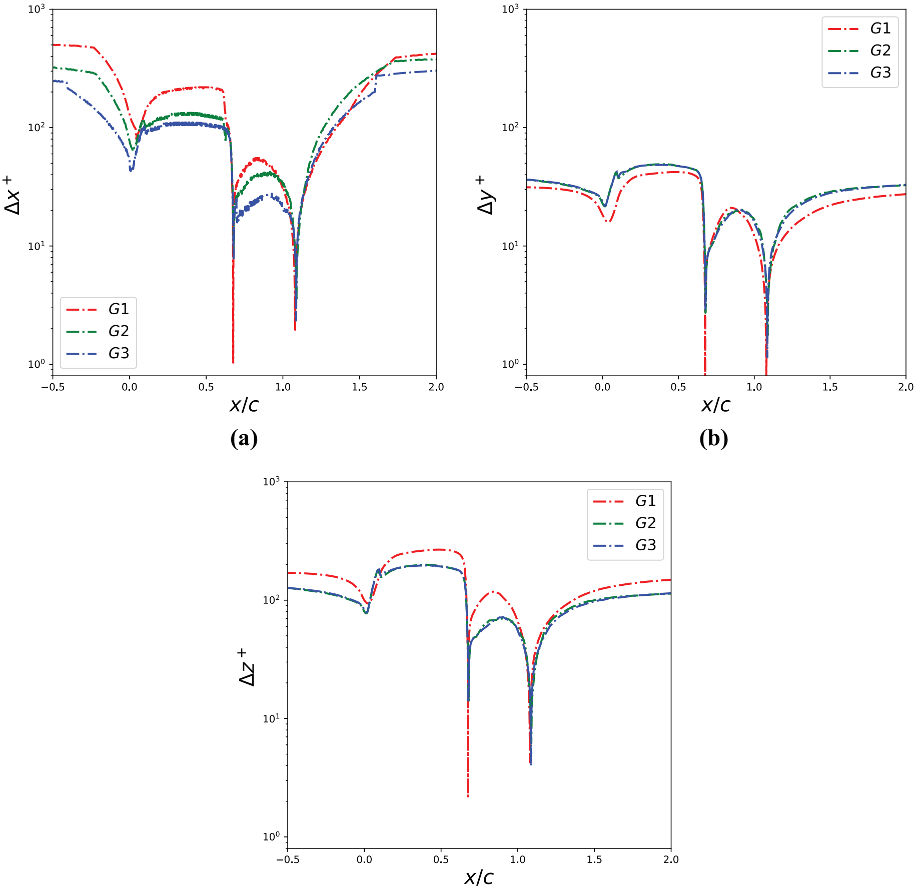


Note: Blue lines: $u'u'$, orange lines: $v'v'$, green lines: $w'w'$, yellow line: $u'v'$, circles: experiment (Naughton *et al.*, 2006), squares: WMLES (Avdis *et al.*, 2009)
Source: Figure by authors

Figure 19. Mean stream-wise velocity and Reynolds stresses at the inflow ($x/c = -2.14$) and at a downstream location ($x/c = -0.81$) for G2

each other, except in the recirculation region. Both simulations do not capture the primary suction peak before the recirculation bubble and the secondary suction peak within the bubble, although the quality of C_p improves slightly with mesh refinement. In addition to that, no noticeable improvements can be observed in the figures. However, a significant improvement is observed in the estimation of the bubble length (see Table 4). The percentage error in the position of the reattachment compared to the experimental values is reduced from 4.5% to 0.9%, upon refining the mesh. To check for convergence, the results from the simulation run on G3 is compared with G2 and G1. The results are also shown in Figure 22. Observing the plot of C_f, it can be seen that there is no improvement in the C_f or C_p profiles after refinement. Therefore, for further analysis, the results of the simulation performed on G2 are considered.

Subsequently, the performance of NEQWM-ML is compared to that of the EQWM. The skin friction and pressure coefficients computed from the simulations performed on the G2 mesh using



(a)

(b)

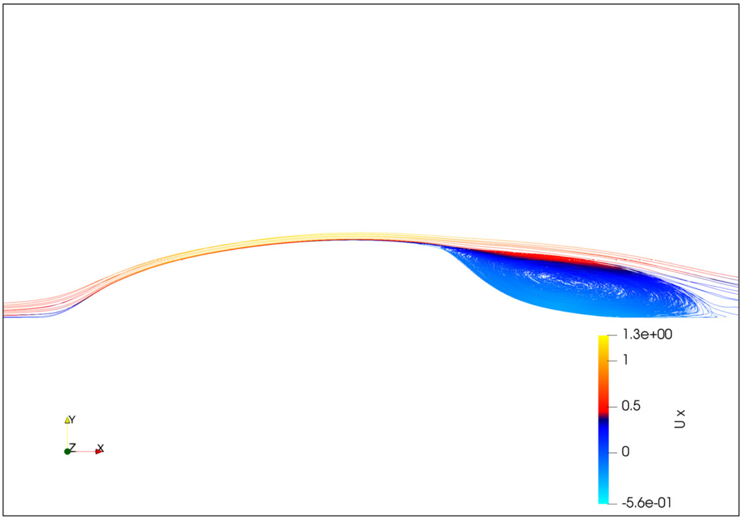
(c)

Notes: (a) Δx^+, (b) Δy^+, and (c) Δz^+ between $x/c = -0.5$ and $x/c = 2.0$
Source: Figure by authors

Figure 20. Grid spacing in wall units for G1, G2 and G3

the two different models are compared in Figure 24(a) and 24(b). There is no difference in the prediction of C_p. However, the differences in the prediction of C_f are noteworthy, especially in the region of relaminarization (Figure 21 shows the location of the relaminarization on the hump). Figure 23 shows the C_f profile, which also includes the relaminarization parameter, K (Bourassa and Thomas, 2009; Uzun and Malik, 2017). It describes the condition that influences the return of the turbulent regime to the laminar regime. It is computed as:

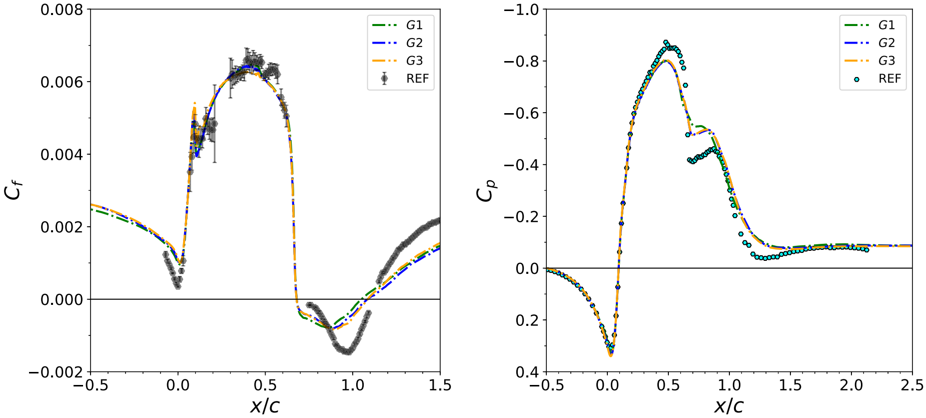
$$K = -\frac{1}{2}\frac{1}{Re_c}\left(\frac{1}{1-C_p}\right)^{\frac{3}{2}}\frac{\partial C_p}{\partial s^*} \tag{11}$$



Notes: The recirculation predicted by the NEQWM-ML is shown. The colormap is according to the average stream-wise velocity. In addition to the recirculation the relaminarization can also be seen as the region of the yellow-colored contours

Source: Figure by authors

Figure 21. Flow over the hump



(a) **(b)**

Notes: (a) Skin friction coefficient; (b) pressure coefficient

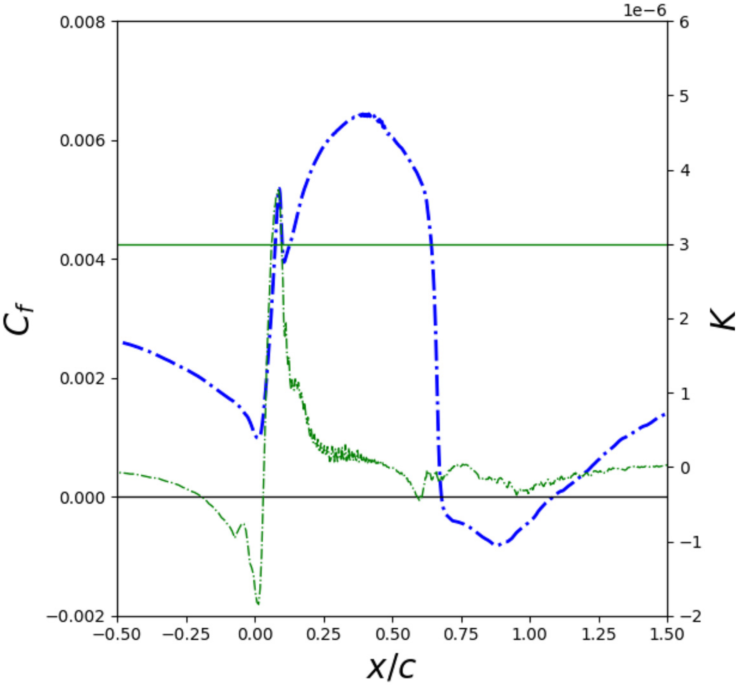
Source: Figure by authors

Figure 22. Skin friction and pressure coefficients for G1, G2 and G3 predicted by the NEQWM-ML compared with experimental results

Table 4. Details of the recirculation bubble for NEQWM-ML and EQWM on G1

Case	Mesh	$x/c\|_{sep}$	$x/c\|_{reatt}$	Bubble length	$error\|x_{reatt}$
Experiment (Naughton *et al.*, 2006)	–	∼0.665	∼1.1	∼0.435	–
EQWM	G1	0.67	1.05	0.38	4.5%
NEQWM-ML	G1	0.68	1.05	0.37	4.5%
EQWM	G2	0.68	1.09	0.41	0.9%
NEQWM-ML	G2	0.68	1.09	0.41	0.9%

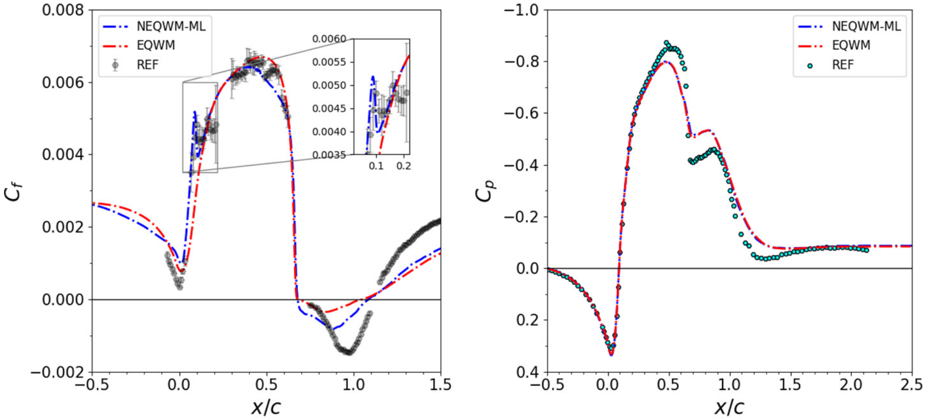
Notes: $x/c|_{sep}$ is the location of separation, and $x/c|_{reatt}$ is the location of reattachment. $error|x_{reatt}$ is the percentage error in the reattachment location.
Source: Table by authors



Note: The green horizontal line shows the threshold of relaminarisation
Source: Figure by authors

Figure 23. Skin friction coefficient for the G2 mesh predicted by the NEQWM-ML shown along with the relaminarization parameter K

where $s^* = s/c$, and s is the distance from the solid wall. The threshold value of K above which the relaminarization occurs is around 3×10^{-6}, indicated as the green horizontal line in the figure. This region is well captured by NEQWM-ML as shown in the inset of Figure 24(a). None of the previous RANS models or WMs has been able to capture this phenomenon because of their inherent inability to capture relaminarization. This is attributed to the ability



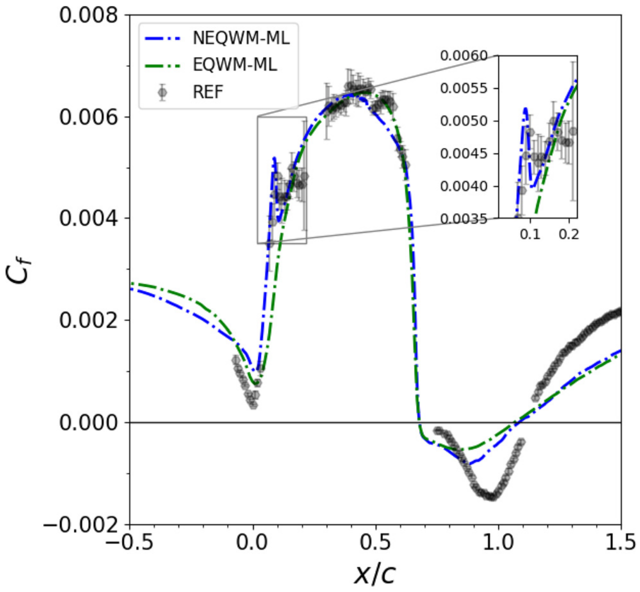
(a) **(b)**

Notes: (a) Skin friction coefficient; (b) pressure coefficient
Source: Figure by authors

Figure 24. Skin friction and pressure coefficient using NEQWM-ML and EQWM compared with experimental results

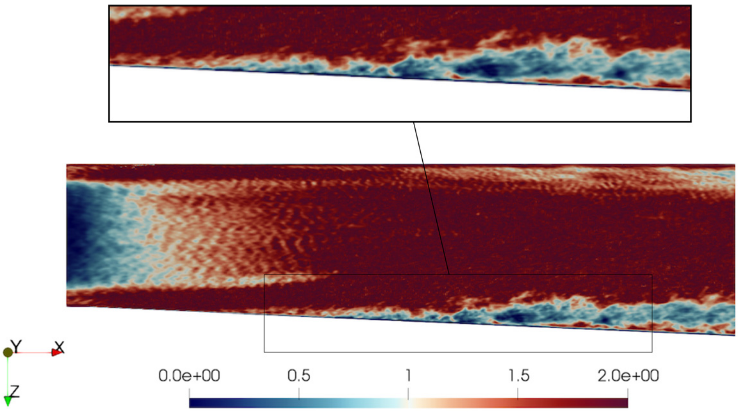
of the model to learn from the data it is trained on. Figure 26 shows the region within the diffuser where the relaminarization occurs. The inset shows one of the regions from the diffuser from where the data is sampled, which is also in a state of relaminarization. This state is estimated by computing the ratio of turbulent production to dissipation. When this ratio is less than 1, it can be considered as the onset of laminarization according to Narasimha and Sreenivasan (Narasimha and Sreenivasan, 1979). In addition to predicting relaminarization, the performance in recirculation is improved compared to EQWM, although not as much as expected. Data sampled for model training also have information on recirculation (Miró *et al.*, 2023). The recirculation in the hump is mainly driven by strong adverse pressure gradients, while in the diffuser, it is rather geometry-induced. Recovery of C_f is also under-predicted by both models. This is possibly due to the inability of the simulation to reproduce the corner vortices with a simple top-wall contour. These vortices are not present in the simulation with the periodic boundary conditions. The absence of corner vortices also explains why there is no improvement in bubble length (see Table 4). The WRLES study of the hump by Uzun and Malik (Uzun and Malik, 2017) also confirms that the bubble length is sensitive to the details of the top-wall contour. In addition, NEQWM-ML is also compared with EQWM-ML in Figure 25. It can be observed that the improvement in the prediction of C_f by NEQWM-ML compared to that of EQWM-ML is similar to what is observed in the case of EQWM. This confirms that the addition of non-equilibrium data has remarkable effects. A comparison of the average mean-stream velocity profiles and Reynolds stresses in the recirculation region, predicted by the two models, is shown in Figure 27. The inability to predict suction peaks has affected the longitudinal curvature of the velocity profiles and the behavior of the stress profiles within the recirculation bubble for both models. The momentum recovery predicted by NEQWM-ML is in agreement with the predictions of EQWM. NEQWM-ML also agrees with EQWM in the prediction of Reynolds stresses inside the recirculation bubble.

3.2.5 Juncture flow. The juncture flow configuration based on the experiments conducted by Lee *et al.* (2018) in the NASA Langley 14 × 22 ft. subsonic wind tunnel is investigated next. These



Source: Figure by authors

Figure 25. Skin friction using NEQWM-ML and EQWM-ML compared with experimental results

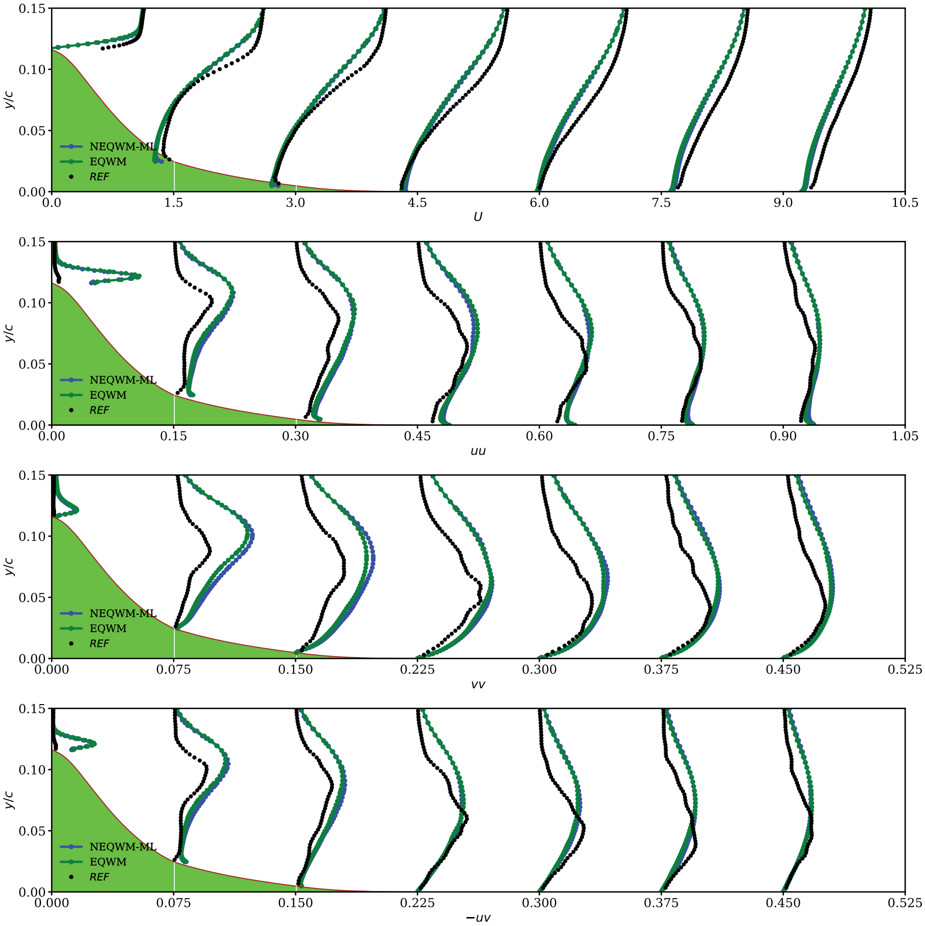


$$Ratio = \frac{Production}{Dissipation}$$

Notes: The color map is based on the ratio of turbulence production to turbulence dissipation. The inset contains the region from which data are taken to train the model

Source: Figure by authors

Figure 26. Location inside the diffuser where the relaminarization takes place (gradients of blue)

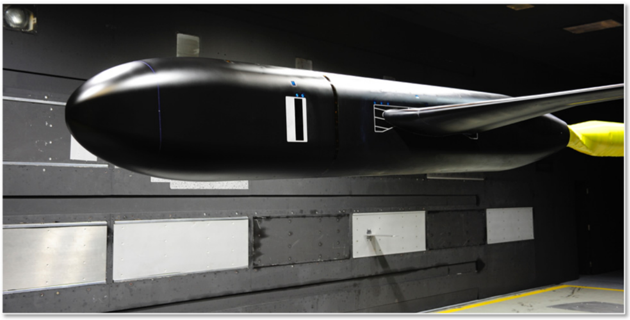


Notes: $x/c = \{0.65, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3\}$. The profiles are respectively shifted by $\Delta = 1.5$, $\Delta = 0.15$, $\Delta = 0.075$ and $\Delta = 0.075$ units

Source: Figure by authors

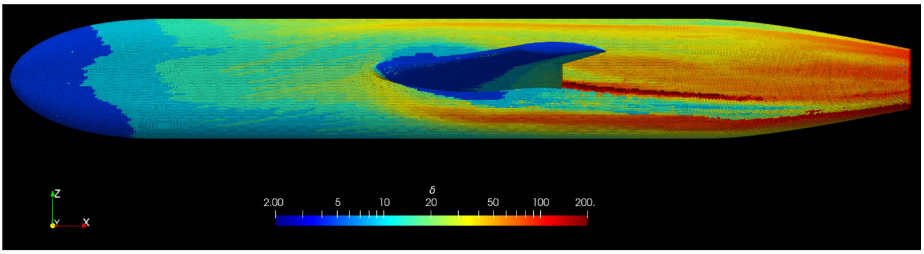
Figure 27. Mean stream-wise velocity and Reynolds stress at stream-wise positions

experiments provide information to validate turbulence models for separated corner flows. Experiments were carried out in a full-span wing-body configuration at an angle of attack of 5° with a focus on the upper junction. The flow separation is observed near the junction of the trailing edge. The experimental setup is shown in Figure 28. The wingspan of the model is nominally 3,397.2 mm, the fuselage length is 4,839.2 mm and the crank chord (chord length in the Yehudi break) is $c =$ 557.1 mm. An 87 million-element mesh is used to simulate this flow. The mesh is adapted according to the BL thickness estimation method proposed by Griffin *et al.* (2021). The estimated thickness of the BL is shown in Figure 29. The generated mesh follows industrial standards and the reader is referred to Iyer and Malik (2020) for meshes of similar dimensions of the juncture flow case. The



Notes: Figure available in Langley Research Center Turbulence Modeling Resource, Exp: Juncture Flow
Source: Figure by authors

Figure 28. Juncture flow experimental setup in the wind tunnel

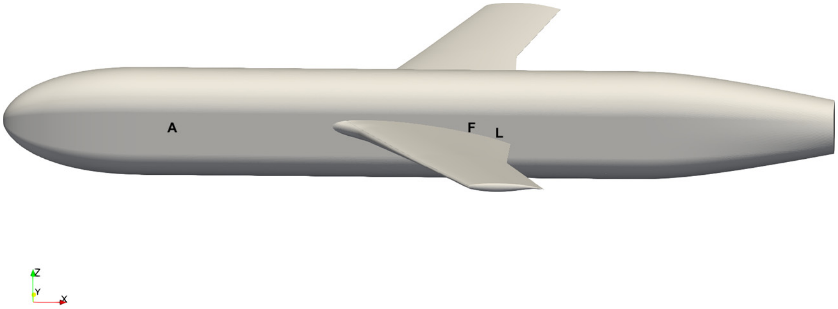


Source: Figure by authors

Figure 29. The estimated BL thickness over the aircraft

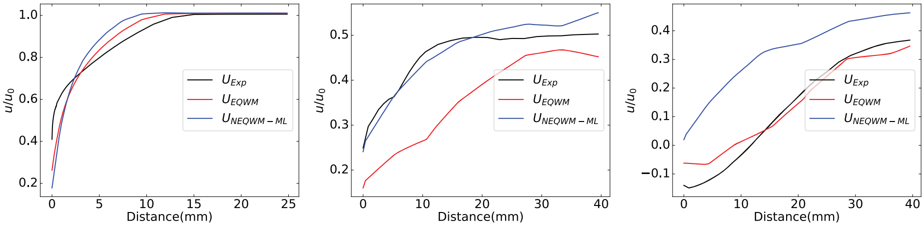
angle of attack considered for this simulation is 5°. The Reynolds number based on the crank chord is 2.4 million. The turbulence closure used is Vreman. A spherical domain of diameter 5×10^4 is used to solve the case. As an inflow boundary condition, a bulk velocity is imposed on the upstream half of the spherical domain, and on the downstream half, a null traction is applied as an outflow condition. The model was tripped on the upper and lower surfaces of both wings, as well as close to the front of the fuselage during the experiment. Preliminary numerical calculations indicated that tripping was also required to initiate the change from laminar to turbulent air above the wings, and hence, the surface mesh was modified accordingly.

Results from three different zones are chosen for comparison of the NEQWM-ML's performance. Zones A, F and L are shown in Figure 30. Zone A is a location of equilibrium flow, Zone F is a location of corner flow and Zone L is a location of recirculation. First, the performance of NEQWM-ML is compared with that of EQWM. The results of the simulation are shown in Figure 31 by comparing velocity profiles at different locations. In the region of equilibrium, the performance of both models is similar, although the EQWM gives a lesser error. At the corner, NEQWM-ML yields more accurate profiles as a result of learning corner flows physics from the diffuser data. However, in the recirculation region, the NEQWM-ML has problems predicting the velocity profile correctly. On the other hand, the EQWM



Source: Figure by authors

Figure 30. Zones A, F and L on the test model



(a) **(b)** **(c)**

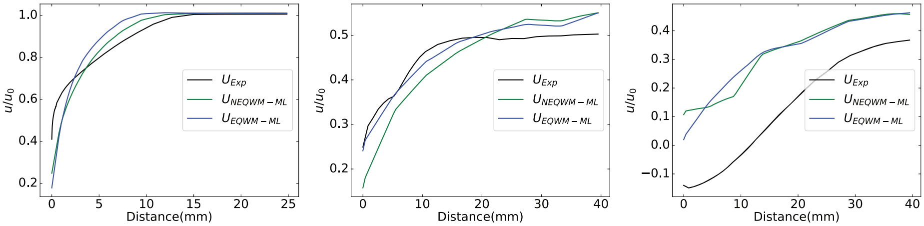
Notes: The black lines are from the experimental results. Blue lines are predicted by NEQWM-ML, and red lines by EQWM
Source: Figure by authors

Figure 31. Mean velocity profiles at locations (a) $x = 1{,}168.4$ mm and $z = -30$ mm (Zone A), (b) $x = 2{,}822.6$ mm and $y = 237.1$ mm (Zone F) and (c) $x = 2{,}922.6$ mm and $y = 237.1$ mm (Zone L)

performs far better than the NEQWM-ML. To better understand this, the same simulation is run with the EQWM-ML. The results of this simulation are shown in Figure 32. The EQWM-ML performs as well as the EQWM at equilibrium zone A. In the corner, the EQWM-ML performs better than the EQWM, but not as well as the NEQWM-ML. At the recirculation region, both ML-based models perform almost equivalently. This implies that the NEQWM-ML has failed to learn the recirculation from the non-equilibrium data. Similar to the observations derived from the hump case, the performance of the NEQWM-ML at the recirculation region is not as good as expected. A possible reason is the mismatch of the recirculation type between the learned data and the physics of the current case. On the other hand, the great agreement of the EQWM is probably coincidental.

4. Conclusion

One of NASA’s 2030 Grand Challenges is to numerically analyze the performance of a powered aircraft configuration across the entire flight envelope through high-fidelity models. The computational resources needed for LES are still beyond current capabilities, so WMLES is the best alternative. In this study, a data-driven wall-modeling methodology is proposed.



(a) **(b)** **(c)**

Notes: The black lines are from the experimental results. Blue lines are predicted by NEQWM-ML, and red lines by EQWM-ML

Source: Figure by authors

Figure 32. Mean velocity profiles at locations (a) $x = 1{,}168.4$ mm and $z = -30$ mm (Zone A), (b) $x = 2{,}822.6$ mm and $y = 237.1$ mm (Zone F) and (c) $x = 2{,}922.6$ mm and $y = 237.1$ mm (Zone L)

The ML system used to build the model is based on GBRTs. The model is trained using high-fidelity data from the Stanford double diffuser DNS and the LES of a channel at $Re_\tau = 950$. After carefully adjusting the hyperparameters and balancing the train data, a model is obtained. The model is first tested *a priori* on unseen data. After that, the model is tested *a posteriori* on the Stanford double diffuser. A notable improvement was observed in the prediction of the velocity profiles when NEQWM-ML is used. Then the model is tested on two channel cases at $Re_\tau = 2005$ and $Re_\tau = 4200$, respectively, to prove the model's ability to switch between equilibrium and non-equilibrium conditions. The model performs similarly to an algebraic wall shear stress model when tested on channel flow, demonstrating that the non-equilibrium predictor of the domain does not interfere with the equilibrium part. The percentage errors in the mean wall shear stress is less than 0.5% for both $Re_\tau = 2005$ and $Re_\tau = 4200$ channel flows. Finally, the model is tested on a wall-mounted hump and the juncture flow. For the wall-mounted hump, the model performs better than the algebraic wall shear stress model and produces noteworthy performance in relaminarization region. The performance of NEQWM-ML at adverse pressure-induced recirculations is slightly better than the other two models. The percentage error in C_f is less than 21.5% for NEQWM-ML, while it is 25.4% and 27.7% for EQWM and EQWM-ML, respectively. For the juncture flow, while the model performs adequately at the corners, its accuracy at the recirculation region is mediocre, showing the limitations of the model and, therefore, potential aspects to improve.

The improved performance of the model in the relaminarization regions is highly unlikely to be captured by traditional wall models. The model has learned the turbulent-to-laminar transition phenomenon from the data it is supplied with. It shows that, given different sets of non-equilibrium data and trained appropriately, the model can be used in all types of non-equilibrium flows. Future work in this direction will focus on training the model with varied sets of balanced non-equilibrium data.

References

Avdis, A., Lardeau, S. and Leschziner, M. (2009), "Large eddy simulation of separated flow over a two-dimensional hump with and without control by means of a synthetic slot-jet", *Flow, Turbulence and Combustion*, Vol. 83 No. 3, pp. 343-370.

Bae, H.J. and Koumoutsakos, P. (2022), “Scientific multi-agent reinforcement learning for wall-models of turbulent flows”, *Nature Communications*, Vol. 13 No. 1, p. 1443.

Bodart, J. and Larsson, J. (2012), “Sensor-based computation of transitional flows using wall-modeled large eddy simulation”, *Center for Turbulence Research Annual Briefs*, Vol. 2012, pp. 229-240.

Bose, S.T. and Park, G.I. (2018), “Wall-modeled large-eddy simulation for complex turbulent flows”, *Annual Review of Fluid Mechanics*, Vol. 50 No. 1, pp. 535-561.

Bourassa, C. and Thomas, F.O. (2009), “An experimental investigation of a highly accelerated turbulent boundary layer”, *Journal of Fluid Mechanics*, Vol. 634, pp. 359-404.

Brunton, S.L., Noack, B.R. and Koumoutsakos, P. (2020), “Machine learning for fluid mechanics”, *Annual Review of Fluid Mechanics*, Vol. 52 No. 1, pp. 477-508.

Calafell, J., Trias, F., Lehmkuhl, O. and Oliva, A. (2019), “A time-average filtering technique to improve the efficiency of two-layer wall models for large eddy simulation in complex geometries”, *Computers and Fluids*, Vol. 188, pp. 44-59.

Chen, T. and Guestrin, C. (2016), “XGBoost”, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Choi, H. and Moin, P. (2012), “Grid-point requirements for large eddy simulation: Chapman’s estimates revisited”, *Physics of Fluids*, Vol. 24 No. 1, p. 11702.

Codina, R. (2001), “Pressure stability in fractional step finite element methods for incompressible flows”, *Journal of Computational Physics*, Vol. 170 No. 1, pp. 112-140.

Dauricio, E.T. and Azevedo, J.L.F. (2023), “A wall model for external laminar boundary layer flows applied to the wall-modeled LES framework”, *Journal of Computational Physics*, Vol. 484, p. 112087.

Drela, M. and Giles, M.B. (1987), “Viscous-inviscid analysis of transonic and low Reynolds number airfoils”, *AIAA Journal*, Vol. 25 No. 10, pp. 1347-1355.

Friedman, J.H. (2000), “Greedy function approximation: a gradient boosting machine”, *Annals of Statistics*, Vol. 29, pp. 1189-1232.

Gonzalez, C., Harris, S. and Moin, P. (2021), “Falkner-Skan wall model baseflow generation for the parabolized stability equations”, *Annual Research Briefs, Center for Turbulence Research*, Stanford University, pp. 73-81.

Gonzalez, C., Karp, M. and Moin, P. (2020), “Wall-stress modeling for laminar boundary layers in coarse grids”, *Annual Research Briefs, Center for Turbulence Research*, Stanford University, pp. 85-95.

Griffin, K.P., Fu, L. and Moin, P. (2021), “General method for determining the boundary layer thickness in nonequilibrium flows”, *Physical Review Fluids*, Vol. 6 No. 2, p. 24608.

Guastoni, L., Encinar, M.P., Schlatter, P., Azizpour, H. and Vinuesa, R. (2020), “Prediction of wall-bounded turbulence from wall quantities using convolutional neural networks”, *In: Journal of Physics: Conference Series*, Vol. 1522 No. 1, p. 12022.

Hoyas, S. and Jiménez, J. (2008), “Reynolds number effects on the Reynolds-stress budgets in turbulent channels”, *Physics of Fluids*, Vol. 20 No. 10.

Hoyas, S. and Jiménez, J. (2006), “Scaling of the velocity fluctuations in turbulent channels up to Re $\tau=$ 2003”, *Physics of Fluids*, Vol. 18 No. 1, p. 11702.

Iyer, P.S. and Malik, M.R. (2020), “Wall-modeled LES of the NASA juncture flow experiment”, *AIAA Scitech 2020 Forum*. 1307.

Jiménez, J. (2004), “Turbulent flows over rough walls”, *Annual Review of Fluid Mechanics*, Vol. 36 No. 1, pp. 173-196.

Kawai, S. and Larsson, J. (2012), “Wall-modeling in large eddy simulation: length scales, grid resolution, and accuracy”, *Physics of Fluids*, Vol. 24 No. 1, p. 15105.

Kempf, A., Klein, M. and Janicka, J. (2005), “Efficient generation of initial-and inflow-conditions for transient turbulent flows in arbitrary geometries”, *Flow, Turbulence and Combustion Formerly: Applied Scientific Research*, Vol. 74 No. 1, pp. 67-84.

Larsson, J., Kawai, S., Bodart, J. and Bermejo-Moreno, I. (2016), "Large eddy simulation with modeled wall-stress: recent progress and future directions", *Mechanical Engineering Reviews*, Vol. 3 No. 1, pp. 15-00418.

Lee, H.C., Pulliam, T.H., Rumsey, C.L. and Carlson, J.R. (2018), "Simulations of the NASA Langley 14- by 22-Foot subsonic tunnel for the juncture flow experiment", available at: https://api.semanticscholar.org/CorpusID:201289296

Lehmkuhl, O., Houzeaux, G., Owen, H., Chrysokentis, G. and Rodríguez, I. (2019a), "A low-dissipation finite element scheme for scale resolving simulations of turbulent flows", *Journal of Computational Physics*, Vol. 390, pp. 51-65.

Lehmkuhl, O., Piomelli, U. and Houzeaux, G. (2019b), "On the extension of the integral length-scale approximation model to complex geometries", *International Journal of Heat and Fluid Flow*, Vol. 78, p. 108422.

Lozano-Durán, A. and Bae, H.J. (2023), "Machine learning building-block-flow wall model for large-eddy simulation", *Journal of Fluid Mechanics*, Vol. 963, p. A35.

Lozano-Durán, A. and Jiménez, J. (2014), "Effect of the computational domain on direct simulations of turbulent channels up to Re τ= 4200", *Physics of Fluids*, Vol. 26 No. 1.

Marques, A.N. and Wang, Q. (2017), "Integral formulation of under-resolved near-wall scales for coarse-grid LES", *55th AIAA Aerospace Sciences Meeting*. 980.

Miró, A., Eiximeno, B., Rodríguez, I. and Lehmkuhl, O. (2023), "Self-Induced large-scale motions in a three-dimensional diffuser", *Flow, Turbulence and Combustion*, Vol. 112 No. 1.

Narasimha, R. and Sreenivasan, K. (1979), "Relaminarization of fluid flows", *Advances in Applied Mechanics*, Vol. 19, pp. 221-309.

Naughton, J.W., Viken, S. and Greenblatt, D. (2006), "Skin friction measurements on the NASA hump model", *AIAA Journal*, Vol. 44 No. 6, pp. 1255-1265.

Owen, H., Chrysokentis, G., Avila, M., *et al*. (2019), "Wall-modeled large-eddy simulation in a finite element framework", *International Journal for Numerical Methods in Fluids*, Vol. 92 No. 1.

Park, G.I. and Moin, P. (2014), "An improved dynamic non-equilibrium wall-model for large eddy simulation", *Physics of Fluids*, Vol. 26 No. 1.

Park, G.I. (2015), "Wall-modeled large-eddy simulation of a separated flow over the NASA wall-mounted hump By".

Poroseva, S.V., Colmenares F, J.D. and Murman, S.M. (2016), "On the accuracy of RANS simulations with DNS data", *Physics of Fluids*, Vol. 28 No. 11, p. 115102.

Radhakrishnan, S., Gyamfi, L.A., Miró, A., Font, B., Calafell, J. and Lehmkuhl, O. (2021), "A data-driven wall-shear stress model for LES using gradient boosted decision trees", International Conference on High Performance Computing, *Springer*, p. 105-21.

Rumsey, C.L., Coleman, G.N. and Wang, L. (2022), "In search of data-driven improvements to RANS models applied to separated flows", *AIAA Scitech 2022 Forum*, p. 937.

Rumsey, C.L., Gatski, T.B., Sellers, W.L., Vasta, V.N. and Viken, S.A. (2006), "Summary of the 2004 computational fluid dynamics validation workshop on synthetic jets", *AIAA Journal*, Vol. 44 No. 2, pp. 194-207.

Simonoff, J.S. (2012), *Smoothing Methods in Statistics*, Springer Science and Business Media, Cham,

Slotnick, J., Khodadoust, A., Alonso, J., *et al*. (2014), "CFD vision 2030 study: a path to revolutionary computational aerosciences".

Srinivasan, P.A., Guastoni, L., Azizpour, H., Schlatter, P. and Vinuesa, R. (2019), "Predictions of turbulent shear flows using deep neural networks", *Physical Review Fluids*, Vol. 4 No. 5, p. 54603.

Thompson, R.L., Sampaio, L.E.B., deBragança Alves, F.A., Thais, L. and Mompean, G. (2016), "A methodology to evaluate statistical errors in DNS data of plane channel flows", *Computers and Fluids*, Vol. 130, pp. 1-7.

Trias, F. and Lehmkuhl, O. (2011), "A self-adaptive strategy for the time integration of Navier-Stokes equations", *Numerical Heat Transfer, Part B: Fundamentals*, Vol. 60 No. 2, pp. 116-134.

Uzun, A. and Malik, M.R. (2017), "Wall-resolved large-eddy simulation of flow separation over NASA wall-mounted hump", *55th AIAA Aerospace Sciences Meeting*, p. 538.

Vadrot, A., Yang, X.I. and Abkar, M. (2023), "Survey of machine-learning wall models for large-eddy simulation", *Physical Review Fluids*, Vol. 8 No. 6, p. 64603.

Vazquez, M., Houzeaux, G., Koric, S., *et al*. (2016), "Alya: multiphysics engineering simulation towards exascale", *Journal of Computational Science*, Vol. 14, pp. 15-27.

Vreman, A.W. (2004), "An eddy-viscosity subgrid-scale model for turbulent shear flow: Algebraic theory and applications", *Physics of Fluids*, Vol. 16 No. 10, pp. 3670-3681.

Wu, J., Xiao, H., Sun, R. and Wang, Q. (2019), "Reynolds-averaged Navier–stokes equations with explicit data-driven Reynolds stress closure can be ill-conditioned", *Journal of Fluid Mechanics*, Vol. 869, pp. 553-586.

Yang, X.I.A., Zafar, S., Wang, J.X. and Xiao, H. (2019), "Predictive large-eddy-simulation wall modeling via physics-informed neural networks", *Phys. Rev. Fluids*, Vol. 3 No. 4, p. 34602.

Zhou, Z., He, G. and Yang, X. (2021), "Wall model based on neural networks for LES of turbulent flows over periodic hills", *Physical Review Fluids*, Vol. 6 No. 5, p. 54610.

Zhou, Z., Yang, X.I., Zhang, F. and Yang, X. (2023), "A wall model learned from the periodic hill data and the law of the wall", *Physics of Fluids*, Vol. 35 No. 5.

Corresponding author
Bernat Font can be contacted at: b.font@tudelft.nl