



HAL
open science

Data-driven wall modeling for turbulent separated flows

Dorian Dupuy, Nicolas Odier, Corentin Lapeyre

► **To cite this version:**

Dorian Dupuy, Nicolas Odier, Corentin Lapeyre. Data-driven wall modeling for turbulent separated flows. *Journal of Computational Physics*, 2023, 487, pp.112173. 10.1016/j.jcp.2023.112173 . hal-04255181

HAL Id: hal-04255181

<https://hal.science/hal-04255181v1>

Submitted on 23 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data-driven wall modeling for turbulent separated flows

Dorian Dupuy¹, Nicolas Odier¹, and Corentin Lapeyre¹

¹*European Centre for Research and Advanced Training in Scientific Computing, Toulouse F-31057 Cedex 1, France*

*(Published version: Journal of Computational Physics 487, 112173 (2023);
<https://doi.org/10.1016/j.jcp.2023.112173>)*

Abstract

The large-eddy simulation of wall-bounded turbulent flows at high Reynolds numbers is made more efficient by the use of wall models that predict the wall shear stress, allowing coarser cell sizes at the wall. In this paper, a data-driven approach for the modeling of the wall shear stress is examined using filtered high-fidelity numerical data from two fully developed turbulent channel flows and two turbulent flows with separated regions: a three-dimensional diffuser and a backward-facing step. The model is a multilayer perceptron based on the flow information in the vicinity, given by the distance to the wall and the velocity components at a given number of grid points above the wall. The model is Mach number equivariant at the quasi-incompressible limit, Galilean invariant, statistically rotational invariant and can extrapolate to flow conditions unseen in the training dataset. The relevance of the machine-learning procedure is verified *a priori* using the filtered numerical data and *a posteriori* by performing wall-modeled large-eddy simulations implementing the model. The results show that the model is able to leverage the local spatial information to discriminate developed wall turbulence and separated regions in a flow configurations not included in the training dataset.

1 Introduction

The direct simulation of wall-bounded turbulent flows becomes increasingly intractable as the Reynolds number of the flow increases. Indeed, the size of the near-wall turbulent structures becomes increasingly small compared to the boundary-layer thickness, which demands an increasingly fine computational grid [87, 17]. The computational requirements may be reduced significantly by simulating only the larger scales of the fluid motion. As part of this approach, known as large-eddy simulation (LES), the large-scale motions are typically defined using a low-pass filter of the order of the computational grid [53]. However, the analysis of complex engineering flows at practical Reynolds number requires grid size and thus filter size so large, that none of the turbulent structures in the viscous sublayer would be resolved by the large-eddy simulation. In this context, the no-slip boundary condition may no longer be used at the walls. It is instead replaced by a wall model accounting for the effect of the non-resolved part of the boundary layer [13, 74, 48]. Numerical simulations implementing such wall models are referred to as wall-modeled large-eddy simulations (WMLES). Most commonly, the wall model provides a condition on the wall shear stress. A non-zero Dirichlet boundary condition for the velocity can also be provided, either because the filtered velocity does not tend to zero at the wall with the selected filter [10, 2], or because the boundary of the computational domain is at a finite distance from the wall [3, 43, 67]. In any case, the wall solution may be computed assuming the shape and scaling of the velocity profile [21, 84, 75], for instance using a statistical equilibrium assumption to enforce the law of the wall instantaneously. Hybrid LES/Reynolds-averaged Navier–Stokes (RANS) strategies can also be used [73]. This includes both zonal methods and seamless methods [5, 20, 93]. Detached-eddy simulation [90] and related approaches can also be included in that category, including delayed detached-eddy simulation [91], improved delayed detached-eddy simulation [86, 36, 55] or zonal detached-eddy simulation [22]. Hybrid LES/RANS methods are cost effective since they can provide an overall quite satisfactory prediction of turbulence dynamics (e.g. separation point, non-equilibrium turbulence...) with a coarser mesh than would be used in a wall-resolved large-eddy simulation [80, 94]. An alternative to the resolution of the full RANS equation is to solve the thin-boundary-layer

equations (TBLE) on an embedded grid, as first proposed by Balaras *et al.* [4]. In addition, the resolution of vertically integrated RANS boundary layer equations has been proposed as a less expensive hybrid method capable of handling some non-equilibrium effects [18, 42, 100, 14].

There are still challenges associated with the application of current wall models to complex turbulent flows, involving phenomena such as shock-boundary layer interaction, laminar-turbulent transition and flow detachment [6, 7, 30, 27, 31]. In particular, flow separation has long proved difficult to handle by wall-modeled LES since, as the equilibrium assumption gets violated, it is difficult to assume a shape for the velocity profile in algebraic wall stress models. Data-driven approaches have recently emerged as a promising novel approach to wall modeling which could in principle handle non-equilibrium flow phenomena, including flow separation, for scenarios where high-fidelity data is available and sufficient. Indeed, the capability of data-driven approaches to produce models beyond the capability of models based on human expertise and manual feature engineering has been demonstrated for various complex tasks, such as image classification, speech recognition or natural language processing [51]. The development of a wall model capable of handling the immense variety of possible flow behavior might, similarly, involve a high degree of complexity that could be learned by a machine-learning (ML) algorithm provided with a sufficient number of high-fidelity experimental or numerical data. Machine-learning methods have been applied to the modeling of turbulence in a number of contexts [28, 12], including LES modeling [82, 33, 95, 63, 64, 104, 99, 69] and RANS modeling [56, 57, 96, 98]. In particular, the development of wall models using machine-learning techniques, to the author's knowledge, has been assessed in a turbulent channel flow [102, 65], a spanwise rotating channel [38] the turbulent flow over periodic hills [105], a boundary layer with shocks [8], a wing-fuselage juncture flow [58, 59] and supersonic turbulent flows [103]. In each case, a different arbitrarily engineered set of input features is used, typically involving the scaling of the flow variables and the preliminary computation of the wall-normal distance logarithm and velocity or pressure derivatives.

This paper investigates the modeling of the wall shear stress for turbulent flows with separation using a machine-learning approach. The procedure is validated *a priori*, by filtering the instantaneous three-dimensional fields of high-fidelity numerical simulations, and *a posteriori*, by performing simulations implementing the models. The machine-learning model is a feed-forward artificial neural network trained using a database composed of two channel flow simulations and the simulation of the flow in a three-dimensional diffuser, involving mean separated regions. The model assumes a locally structured and uniform grid and uses the velocity components at one or several grid points above the wall as input, scaled to ensure the Mach number equivariance of the model at the quasi-incompressible limit but with minimal feature engineering. The training procedure ensures in addition that the model is statistically invariant to reflections or rotations of the coordinate system. The number of input grid points is varied to assess the relevance of an increasing amount of spatial information on the ability of the model to discriminate attached and detached flow regions, that is to make a different prediction in the two regions. Care is taken to ensure that this discrimination is achieved while maintaining at least parity performance with the law of the wall for fully developed wall turbulence. The resulting model is implemented in a flow solver and assessed *a posteriori* for the wall-modeled large-eddy simulation of a channel flow and of the flow over a backward-facing step, demonstrating the generalization performance of the model. The results are compared to corresponding large-eddy simulations using an algebraic wall stress model.

The training dataset is described in section 2, as well as the procedures used to filter and prepare the data for the machine-learning models. The architecture of the artificial neural networks is given in section 3. The *a priori* and *a posteriori* results are discussed and analyzed in section 4.

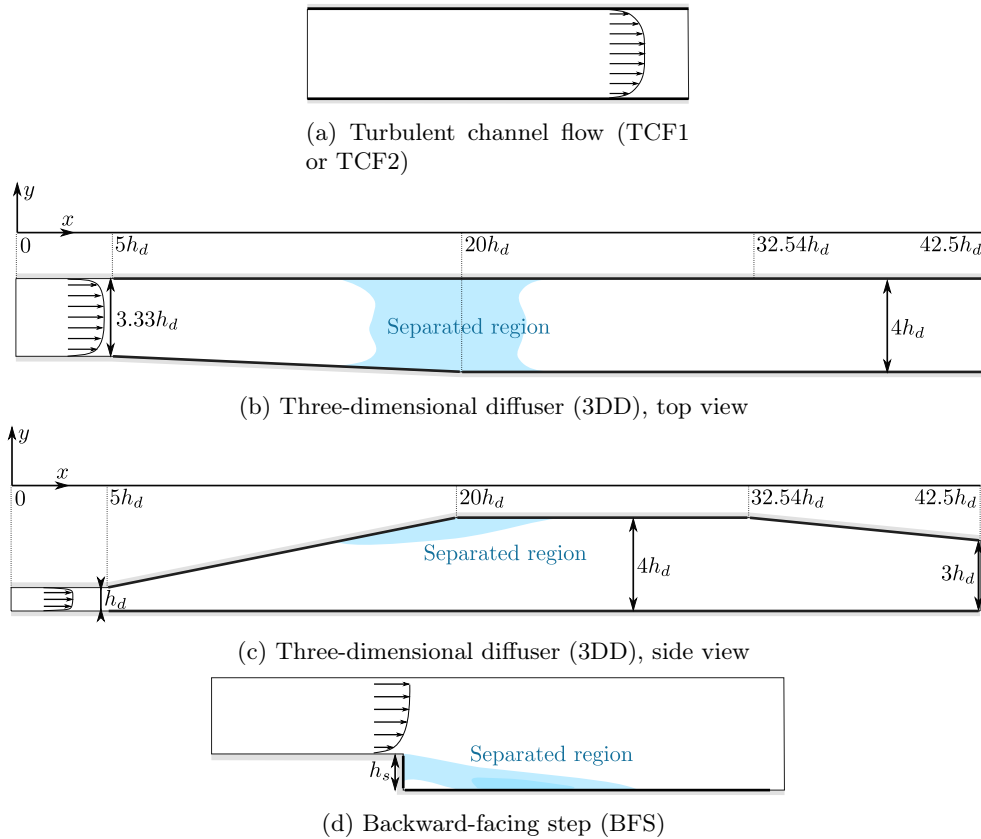


Figure 1: Geometry of the four simulations in the database. The blue areas are separated in the mean. The training data are extracted along the thick black lines along the walls.

2 Data preparation and filtering

2.1 Dataset

To learn about the wall behavior of fluids, we need a model that can discriminate the immense variety of possible flow behavior. The complexity of the task means that there is currently no dataset available to specify the problem. For the data-driven modeling to be useful in a wide array of configurations and applicable in various flow solvers, it is crucial to demonstrate that they can learn from heterogeneous databases, aggregating simulations performed by different teams and using different solvers and numerical methods. Producing such database would require a large human effort. As a first step, this study uses a database of instantaneous three-dimensional time-dependent data which aggregates four numerical simulations performed by various research groups. The training was restricted on pragmatic grounds to quasi-incompressible flows without variations of fluid properties. This simplifies the task of the model by excluding from the analysis some of the complexity that may be found in real flows, such as the effect of compressibility [34] or the complex interaction between turbulence and temperature [72, 25, 26]. Our main focus is the development of a wall model that accurately treats flow separation. The database includes both turbulent boundary layers and separated regions. The fourth simulation, which corresponds to the flow over a backward-facing step, is excluded from the training dataset for testing purposes and to allow a realistic assessment of the reliability and generalization performance of the model.

- First, the database includes the direct numerical simulation (DNS) of a fully-developed channel flow at a friction Reynolds number $Re_\tau = 180$, performed at Imperial College London [1] and hereafter denoted TCF1. The geometry is periodic in the streamwise and spanwise

direction. The domain size is $8\pi h_c \times 2h_c \times 4\pi h_c$, where h_c is the half-height of the channel. The mesh includes $62 \times 19 \times 60$ hexahedral elements with fifth-order polynomial (effectively $372 \times 114 \times 360$ solution points). The numerical schemes are based on the flux reconstruction method of Huynh [41] in the high-order code PyFR [97]. Although this numerical solver is compressible, the compressibility effects are negligible since the Mach number based on centerline velocity $Ma = 0.1$ of the simulation is low. The height of the first point off the wall is $\Delta y^+ = 0.25$ in wall units, that is using the classical scaling (+) based on the wall shear stress τ , the density ρ and the kinematic viscosity ν , namely $\Delta y^+ = \Delta y \sqrt{\tau/\rho}/\nu$.

- Second, the database includes the direct numerical simulation of a fully-developed channel flow at a friction Reynolds number $Re_\tau = 950$, performed at the Polytechnic University of Madrid [61, 23, 60] and hereafter denoted TCF2. In that case, the domain size is $2\pi h_c \times 2h_c \times \pi h_c$ and the mesh includes $512 \times 385 \times 512$ grid points. The simulation has been performed using a hybrid Fourier-Chebyshev spectral method. The height of the first point off the wall is $\Delta y^+ = 0.03$.
- Third, the database includes the high-fidelity wall-resolved simulation of the flow separating in a three-dimensional diffuser corresponding to the geometry “Diffuser 1” of Cherry *et al.* [16], performed at Barcelona Supercomputing Center [29] and hereafter denoted 3DD. The geometry is composed of three main parts: the development of a duct flow from $x/h_d = -65$ to $x/h_d = 5$, where h_d is the height of the inflow duct; a lateral and vertical expansion up to $x/h_d = 20$, from a width $L_z = 3.33h_d$ to $L_z = 4h_d$ and from a height of $L_y = h_d$ to $L_y = 4h_d$ respectively; and, finally, a vertical contraction from $L_y = h_d$ to $L_y = 4h_d$ from $x/h_d = 32.54$ to $x/h_d = 42.5$ (figures 1b and 1c). The Reynolds number based on the inlet channel height is $Re_h = U_{\text{inlet}}h_d/\nu = 10000$. The mesh includes $3139 \times 199 \times 399$ grid points. The simulation uses the low-dissipation finite element scheme of Lehmkuhl *et al.* [52]. The height of the first point off the wall in wall units is 0.5 on the top, bottom and side walls of the inlet channel.
- Fourth, the database includes the high-fidelity wall-resolved simulation of the flow over a backward-facing step [78, 79], performed at CERFACS and hereafter denoted BFS. The Reynolds number based on the inflow bulk velocity and step height h_s is $Re_h = 5100$. The inlet is located $10 h_s$ before the step and the outlet $10 h_s$ after the step. The geometry is periodic in the spanwise direction and the spanwise length of the domain is $2 h_s$. The simulation has been performed using a cell-vertex finite-element method [83] with second-order accurate convection and diffusion schemes [49].

The data for the machine learning process is extracted along the top and bottom walls of the turbulent channels (figure 1a), the top, bottom and side walls of the expanding and contracting regions of the three-dimensional diffuser (figures 1b and 1c) and the front and bottom walls of the backward-facing step (figure 1d). The channel flow simulations (TCF1 and TCF2) provides reference data for canonical fully developed wall turbulence. The three-dimensional diffuser simulation has a more complex physics that includes flow features such as separation and sidewall effects. The backward-facing step simulation also includes a region of strong separation. It is a relevant test case since all flow phenomena present in the BFS case may be found in the training simulations (TCF1, TCF2 and 3DD). While the physical phenomena involved are similar, the BFS case is not specifically related to the 3DD case, and there is, to the best of our knowledge, nothing special about the 3DD case that would make it particularly representative of other separated flows. We can thus to some extent consider the test performance on the BFS case as representative of the performance of the model in other similar flows, bearing in mind that this is a gross approximation and that given the limited size of the training database creating a truly general wall model is not realistic.

2.2 Data preparation

The instantaneous three-dimensional flow fields are preprocessed to accommodate the machine learning process. This study adheres to the classical LES formalism of Leonard [53] in which

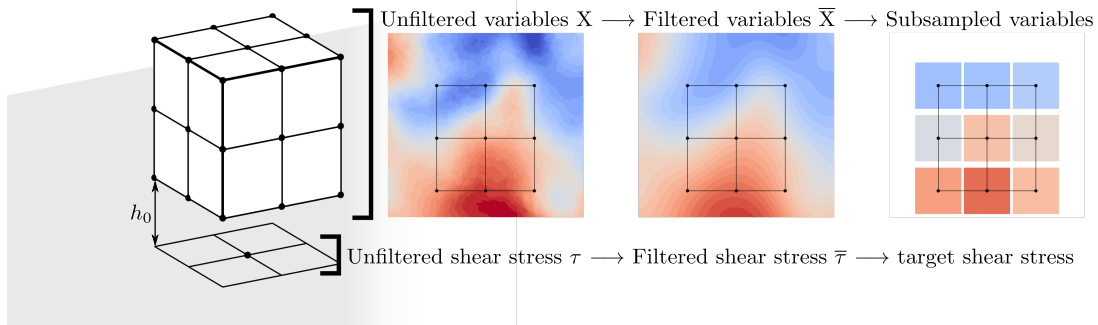


Figure 2: Preparation of the data for the machine-learning algorithm.

large-eddy simulations aim to simulate the evolution of filtered physical fields. The instantaneous numerical data are thus filtered and subsampled onto a grid with WMLES resolution. The filter, denoted $\bar{\cdot}$, is a low-pass filter intended to remove the small scales of turbulence that cannot be captured by the computational grid. The filter used to train the machine-learning wall model should be in agreement with the filter underlying the LES governing equations in *a posteriori* tests. While generally no filter is explicitly employed *a posteriori* in a large-eddy simulation, this could mean if applicable that the characteristic length $\hat{\Delta}$ associated with the filter size in the subgrid-scale model formulation should be consistent with the size of the filter used to train the machine-learning model. In practice, $\hat{\Delta}$ is almost always taken to be the mesh size and the size of the filter used to train the model should therefore corresponds to the LES mesh size.

The preparation of the data for the machine-learning algorithm is represented in figure 2. Around a selected target grid point at the wall, we first locally consider a regular WMLES grid embedded to the wall and without relationship to the fine mesh of the simulation. This WMLES grid is devised to have a cell size in the two tangential directions $\Delta_1^+ = \Delta_2^+ = 33$ in wall units in the channel flows, at the top, bottom and side walls of the inlet channel of the diffuser and for the boundary layer before the edge of the backward-facing step. In the wall-normal direction, the cell size in wall units is randomly sampled in the range $\Delta_y^+ = 25 - 75$. This range determines the range of mesh resolution that the machine-learning model can operate on, and thus indirectly the maximal Reynolds number that can be addressed *a posteriori* at moderate computational cost. The presently used range of $\Delta_y^+ = 25 - 75$ is suitable for instance in turbomachinery-flow simulations [24, 54, 70]. The two tangential axes of the WMLES grid are randomly generated, and thus not necessarily aligned with those of the corresponding highly-resolved simulation.

The fine-mesh data are filtered and interpolated onto this virtual WMLES grid. This is performed using a three-dimensional Gaussian filter centered at each simulation node with a width equal to the WMLES mesh size, the filter width being defined, following Leonard [53], as $\sqrt{6}$ times the standard deviation of the Gaussian kernel. The filtered wall shear stress is thus computed with a two-dimensional Gaussian filter along the wall. The whole process is repeated for various timesteps and for each selected target wall grid point independently. This associates for each selected wall grid point a filtered wall shear stress and a $3 \times 3 \times 3$ grid of input variables. The task of the machine-learning model will be to predict from the input variables the filtered wall shear stress.

2.3 Scaling

The selection of the input features is a critical issue in machine learning. In particular, it is useful to scale the input features in order to build with limited data a model than can generalize to flows with different length scales, different velocity scales or different fluid properties. The preparation of the input data with physically based scaling may be seen as a way to encode a *prior* to the machine-learning model, that is an initial assumption regarding the underlying data distribution that is enforced or promoted before using any training data. In that view, the training process is a mechanism that updates this prior using the observed training data, and the resulting

trained model is the “posterior”.

In order to ease comparison with literature, table 1 compares the input feature selection of the existing studies from the literature that are known to the authors. The table does not include the study of Bhaskaran *et al.* [8] or Moriya *et al.* [65] since these works did not use any input scaling. Zangeneh [103] use that same input feature selection as Lozano-Durán and Bae [58]. The scaling used in the present study differs from these scalings in the following ways:

- Yang *et al.* [102], Huang *et al.* [38] and Lozano-Durán and Bae [58] use scalings that require knowledge of the friction velocity $u_\tau = \sqrt{\tau/\rho}$, which is directly related to the target τ of the model. Although u_τ could be computed in an *a posteriori* setting, for instance using the shear stress at previous timestep, as in [58], we decided to only use input features that are known instantaneously in the present scaling.
- Yang *et al.* [102], followed by Zhou *et al.* [105], engineered the input features to ease the training process and generalization in equilibrium flows, e.g. by including the logarithm of the wall distance. Since we intend to build a model than can also operate in various flow configurations, including non-equilibrium configurations, we decided to make minimal assumptions regarding the nature of the function that should be learned by the model.
- Yang *et al.* [102] and Lozano-Durán and Bae [58] feed spatial information to the model indirectly, through the spatial derivatives at a single grid point. We decided to not explicitly feed any derivatives to the model but instead provide the input features on several grid points, leaving at the discretion of the model to internally compute the derivatives it deems fit.
- Yang *et al.* [102], Huang *et al.* [38], Lozano-Durán and Bae [58] and Zhou *et al.* [105] use scalings that involve flow-specific scales, such as a far-field velocity or a geometric length. Defining this type of scale in various different complex geometries can be a complicated matter. We decided to use purely local input features and only involve local quantities to define the scaling, which makes it trivial to apply the model in any flow configuration.

To define the present scaling, the incompressibility assumption is used to encode our prior knowledge on the Mach number dependence of the flow into the model. We are specifically concerned with the dependence of the flow on the Mach number of the “real” quasi-incompressible target flow to model or simulate. Indeed, if the incompressible Navier–Stokes equations are assumed, the Mach number as seen by the flow model is strictly zero since the speed of sound is infinite. The Mach number of the “real” target flow is not involved in the incompressible Navier–Stokes equations and plays no role in the flow physics if the flow variables are properly scaled. Therefore, the incompressible Navier–Stokes equations can be described as being equivariant under a change of Mach number, as the physics of quasi-incompressible flows is invariant under a change of Mach number when appropriately scaled. The Mach number equivariance of the model prediction is enforced by scaling each input flow variable ϕ as $\hat{\phi}$ in a predefined pre-processing step of the model, while the predicted wall shear stress is unscaled in a predefined post-processing step. As discussed above, we only use the information directly available to the wall model, namely the fluid density ρ , the dynamic viscosity μ and the local height of the first off-the-wall point h_0 in the WMLES grid to define the scaling. This yields the following variable transformations: $\hat{\rho} = 1$, $\hat{\mu} = 1$, $\hat{h}_0 = 1$, $\hat{u} = (\rho h_0/\mu)u$, $\hat{\tau} = \rho(h_0/\mu)^2\tau$. Note that equivalently, the scaled velocity can be expressed as $\hat{u} = u^+ h_0^+$ and the scaled wall shear stress as $\hat{\tau} = (h_0^+)^2$, where $u^+ = u\sqrt{\rho/\tau}$ is the velocity in wall units and $h_0^+ = h_0\sqrt{\rho\tau}/\mu$ the first-point height in wall units. The scaling achieves three objectives. First, it guarantees the Mach number equivariance of the model without requiring data augmentation. Indeed, the input of the model $\hat{u} = (\rho h_0/\mu)u$ tends to a constant as the Mach number of the flow tends to zero, at constant Reynolds number and mesh resolution [71]. Second, it equalizes the scale of the wall shear stress in each simulation, facilitating the use of a loss that does not bias the regression towards one particular simulation. Third, it reduces the number of variables to consider, as with such scaling the height of the first off-the-wall point is always equal to one and can be ignored from the model inputs. The scaling may be seen a transformation, or a change of the flow scales, that modifies the Mach number of the flow but does not change its Reynolds number. Thus, the scaling does not lead to a loss of physically important information in

Table 1: Input feature selection in present work and various studies from the literature. The third column reports the number of input grid points of the model (#IGP) and, if spatial derivatives are included in the model, the effective number of grid points required to compute the model and each input spatial derivative, assuming a centered scheme. For clarity, the notations have been adapted: y is the wall-normal height, u_i is the i -th component of velocity, $S_{ij} = (\partial_i u_j + \partial_j u_i)/2$ the rate of deformation tensor, $\Omega_{ij} = (\partial_j u_i - \partial_i u_j)/2$ the rate of rotation tensor, \mathcal{R} is the aspect ratio of the first cell, L_h is the channel half-height and $L_{\mathcal{H}}$ is the hill height, ℓ_{Ω}^+ is a nondimensionalized rotation-induced length scale, $y_0 = (\nu/u_{\tau}) \exp(-\kappa B)$ is a viscous length scale, with $\kappa \approx 0.4$ the von Kármán constant and $B \approx 5$, $y^* = \nu/\sqrt{u_v^2 + u_p^2}$ is a combined length scale, with $u_v = \sqrt{\|\nu \mathbf{u}_t/y\|}$ and $u_p = [(\nu/\rho)\partial_x p]^{(1/3)}$. The velocity and length scales of [58] are $\mathcal{U}_1 = \mathcal{U}_2 = u_{\tau}$, $\mathcal{U}_3 = u_p$, $\mathcal{L}_1 = L_h$, $\mathcal{L}_2 = \mathcal{U}_2/\|\nabla u\|$ and $\mathcal{L}_3 = \nu/\mathcal{U}_3$. In [59], γ_{12} is the angle between \mathbf{u}_1 and \mathbf{u}_2 , a_1 is the magnitude of the acceleration at the first control volume, $\dot{\gamma}_1$ is the time derivative of the angle of \mathbf{u}_1 in the wall-parallel direction, k_1 and k_{m1} are the turbulent kinetic energy and mean kinetic energy, respectively at the first control volume and relative to the wall.

Study	Input features	#IGP (effective value, if different)
Yang <i>et al.</i> [102]	$\ \mathbf{u}_t^+\ /h_0^+$, $\log(h_0/y_0)/\mathbf{u}_t^+$, \mathcal{R} , $(h_0/L_h)\partial_z p^+$	1 (3, spanwise aligned)
Huang <i>et al.</i> [38]	y^+ , y^+/ℓ_{Ω}^+	1
L.-D. and Bae [58]	u_i/u_{∞} , u_i/\mathcal{U}_k , $S_{ij}\mathcal{L}_k/\mathcal{U}_k$, $\Omega_{ij}\mathcal{L}_k/\mathcal{U}_k$, $\mathcal{U}_k\mathcal{L}_k/\nu$, \dots	1 (7, stellated arrangement)
L.-D. and Bae [59]	$u_1 y_1/\nu$, $u_2 y_2/\nu$, γ_{12} , $a_1 y_1^3/\nu^2$, $\dot{\gamma}_1 y_1^2/\nu$, k_1/k_{m1}	2
Zhou <i>et al.</i> [105]	$\log(y/y^*)$, u_i/h_0 , $(h_0/L_{\mathcal{H}})\partial_k p$	1 – 6, vertically (3 – 18)
Present study	$(\rho h_0/\mu)u_i$	1 – 27, cubic arrangement

the limit of isothermal quasi-incompressible flows. In more practical terms, this implies that the model can operate for flows with different scales of length, velocity or fluid properties.

3 Network architecture

The wall shear stress is determined using a feed-forward artificial neural network as represented in figure 3. The neural network architecture is devised to make by construction the following physical assumptions: The field of velocity in the vicinity of the target wall node is a relevant feature to predict the wall shear stress; The wall shear stress should not be modified if the system undergoes a Galilean transformation; The wall shear stress prediction should be equivariant under orthogonal transformations, that is rotations and reflections. In accordance with the input-feature scaling used, the flow is assumed to be isothermal and quasi-incompressible, which implies that the model prediction should be equivariant under a change of Mach number as described in section 2.3. In addition, there are some physical assumptions that are encoded in the selection of the training database, as physical phenomena not included in the training database are implicitly assumed to not be occurring. This implies with the present training database, following section 2.1, that the model will assume a flow with negligible temperature variations and without compressibility effects such as shocks.

The neural network is a multilayer perceptron with n_h hidden layers, each composed of n_u neural units. The output of any neural i in the hidden layer ℓ is given by

$$z_{\ell}^{(i)} = \sigma \left(\sum_j w_{\ell}^{(i,j)} z_{(\ell-1)}^{(j)} + b_{\ell}^{(i)} \right), \quad (1)$$

where the weights $w_{\ell}^{(i,j)}$ and biases b_{ℓ} are the learned parameters of the network and with σ a predetermined activation function. In this work, the selected activation function is an exponential linear unit [19],

$$\sigma(x) = \begin{cases} x & \text{if } x \geq 0 \\ a(e^x - 1) & \text{otherwise} \end{cases} \quad (2)$$

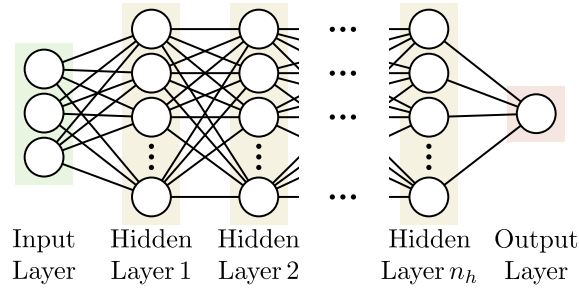


Figure 3: Graphical representation of a multilayer perceptron with three inputs and one output.

The use of a rectified linear unit or a hyperbolic tangent lead to similar results. The output layer is a linear combination of the values of last hidden layer n_h ,

$$z_o = \sum_j w_o^{(j)} z_{(n_h)}^{(j)} + b_o. \quad (3)$$

The weights of the network are initialized using a Glorot-uniform distribution [35] and trained using the Adam optimizer [46] with a base learning rate of 0.001. The loss function is defined as the mean absolute error (MAE) between the predicted wall shear stress and the reference wall shear stress obtained from the filtering of the three-dimensional instantaneous data.

The output of the model is the scaled wall shear stress $\hat{\tau}$, with τ the norm of the shear stress vector

$$\boldsymbol{\tau} = \boldsymbol{\Sigma} \cdot \mathbf{e}_n - (\mathbf{e}_n \cdot \boldsymbol{\Sigma} \cdot \mathbf{e}_n) \mathbf{e}_n, \quad (4)$$

where $\boldsymbol{\Sigma} = \mu (\nabla \mathbf{u} + (\nabla \mathbf{u})^T - (2/3)(\nabla \cdot \mathbf{u}) \mathbf{I}_d)$ is the viscous stress tensor and \mathbf{e}_n a unit wall-normal vector. The inputs of the model are the scaled components of the velocity vector $\hat{\mathbf{u}}$, given to the machine-learning model in a Cartesian coordinate system embedded to the wall. Namely, \mathbf{u} is the relative velocity of the flow with respect to the wall velocity \mathbf{u}_{wall} . The wall-normal velocity is denoted u_n and the components of the tangential velocity $\mathbf{u}_t = \mathbf{u} - u_n \mathbf{e}_n$ are denoted u_{t1} and u_{t2} . In order to assess the influence of an increasing amount of spatial information on the model performance, we consider several cases for the model input, as reported in table 2: using only the velocity at the first point off the wall (S1, S2), as in a classical algebraic wall stress model, using the velocity at three grid points at various distance to the wall (L2), using the velocity components from the nine points in a plane parallel to the wall (P2), and, finally, using the velocity at all the 27 points of the “box” above the wall (B2). It is not necessary to include the cell size in the wall-normal direction as an input feature, since it is constant and equal to $\hat{h}_0 = 1$ for each sample. The model is Galilean invariant since the inputs are constructed using relative velocities. This is not sufficient however to ensure the invariance of the wall model to orthogonal transformations, namely rotations or reflections. Therefore, the dataset is augmented by arbitrary rotations around the wall-normal direction and reflections of the axes to ensure that the predictions are statistically invariant to reflections or rotations of the coordinate system. Note that it is also possible, instead of data augmentation, to use the velocity vector at the first point off the wall to define the tangential axes of the coordinate system [102]. The dataset is decomposed into a training dataset and testing dataset using 70% of the samples associated with lower timesteps for training and the samples associated with later timesteps for testing. The number of hidden layers, the number of units per hidden layer and the mini-batch size are determined using a Bayesian optimization process [88] that evaluates 120 hyperparameter combinations for each dataset and model type investigated. The final model is the model associated with the hyperparameters that lead to the lower validation loss. For reference, the resulting number of hidden layers, number of units and total number of weights are reported for each case in table 3. The results of the hyperparameter optimization should be interpreted with care as the Bayesian process does not take into account sources of uncertainty. The number of hyperparameters should in particular be interpreted with care, since the Bayesian optimization process does not take into account the complexity of the model. The number of hyperparameters of the various models is thus not indicative of the underlying complexity of the problem.

Table 2: List of input type investigated for the neural networks.






Name	Input variables	Number of inputs
 S1	$\hat{u}_{t1}, \hat{u}_{t2}$	2×1
 S2	$\hat{u}_{t1}, \hat{u}_{t2}, \hat{u}_n$	3×1
 L2	$\hat{u}_{t1}, \hat{u}_{t2}, \hat{u}_n$	3×3
 P2	$\hat{u}_{t1}, \hat{u}_{t2}, \hat{u}_n$	3×9
 B2	$\hat{u}_{t1}, \hat{u}_{t2}, \hat{u}_n$	3×27

Table 3: Number of hidden layers n_h and number of units n_u of the multilayer perceptrons trained on the datasets TCF1 and TCF2 (TCF-), the datasets TCF1, TCF2 and 3DD (TCF3DD-) and the datasets TCF1, TCF2 and BFS (TCFBFS-), with various input types as per table 2.

Model	n_h	n_u	Number of weights
TCF-B2	10	24	7393
TCF-P2	3	26	2159
TCF-L2	5	18	1567
TCF-S2	7	12	997
TCF-S1	11	40	16561
TCF3DD-B2	11	36	16309
TCF3DD-P2	7	64	26817
TCF3DD-L2	8	56	22961
TCF3DD-S2	7	36	8173
TCF3DD-S1	3	512	527361
TCFBFS-B2	5	36	8317
TCFBFS-P2	6	28	4873
TCFBFS-L2	12	36	15049
TCFBFS-S2	3	72	10873
TCFBFS-S1	7	26	4317

4 Results

4.1 A priori tests

This section describes the training and testing of the machine-learning models *a priori*, that is based solely on the filtered numerical data. Section 4.2 presents *a posteriori* results, that is by performing simulations implementing the models.

4.1.1 Channel flows

As a first step, the approach is assessed for fully developed wall turbulence by training a model using only the channel flow data (TCF1 and TCF2). The target data is represented in figure 4a in a $u^+(y^+)$ graph, using the local filtered wall shear stress to compute the wall-unit scaling ($^+$). The scatter plots in the figure aggregates instantaneous data in the channels at various locations, various timesteps and for various WMLES mesh resolutions. For reference, the instantaneous data

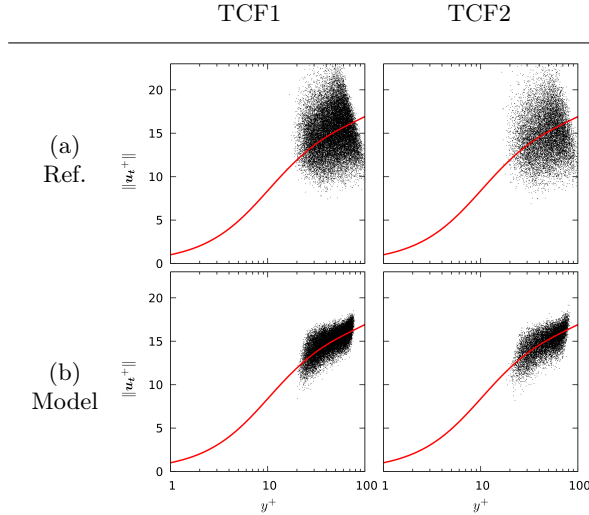


Figure 4: A priori validation: Norm of the scaled tangential velocity u_t^+ as a function of the scaled distance to the wall y^+ in the turbulent channel flows TCF1 and TCF2, using the local target wall shear stress (a) or the prediction of the TCF-B2 model to compute the wall unit scaling ($^+$). The red line is Reichardt's law, given by equation (5).

are compared to the analytical profile of Reichardt [81],

$$u^+(y^+) = \frac{1}{\kappa} \log(1 + \kappa y^+) + 7.8 \left[1 - \exp\left(\frac{-y^+}{11}\right) - \frac{y^+}{11} \exp\left(\frac{-y^+}{3}\right) \right], \quad (5)$$

with \log the natural logarithm and κ the von Kármán constant. The typical value $\kappa = 0.41$ is used, although recent works suggest that the actual value of the von Kármán constant might be $\kappa = 0.387$ [62, 32, 76]. For the two channel flow simulations, the data points follow the law of the wall in the mean but there is a large variance that corresponds to the Reynolds fluctuations of the filtered velocity. A neural-network model of type B2 (as per the classification of table 2) is trained and denoted TCF-B2 hereafter. The predictions of the neural network are reported in figure 4b. Results for models S1, S2, L2 and P2 of table 2 are not shown here for sake of clarity, but provided in appendix A. The model successfully captures the behavior of the law of the wall, but the variance around the mean is clearly underestimated. The lower variance of the tangential velocity using the TCF-B2 is a natural consequence of the loss function used to train the model. In the presence of uncertainty, the loss function favours a prediction close to the mean value because it is less likely to deviate significantly from the true value whereas extreme values, in the sense of values that deviate significantly from the mean, will incur a large penalty if the prediction is inaccurate. Note that this effect remains significant although the loss function used in the present study is the mean absolute error (MAE), which penalises extreme values less than the more common mean squared error (MSE). The scatter plot given in figure 5 shows that the linear correlation between the model output and the target wall shear stress is poor, and can result in a large misprediction for a particular data point. Indeed, the coefficient of correlation between the model output τ_m and the target wall shear stress τ_t , defined as $\rho_c = (\langle \tau_m \tau_t \rangle - \langle \tau_m \rangle \langle \tau_t \rangle) / (\sqrt{\langle \tau_t^2 \rangle - \langle \tau_t \rangle^2} \sqrt{\langle \tau_t \rangle - \langle \tau_t \rangle^2})$, is in the range 0.3 – 0.4 which indicates a low correlation. Besides, the coefficient of determination between the model output τ_m and the target wall shear stress τ_t , defined as $R_2 = 1 - \langle (\tau_m - \tau_t)^2 \rangle / \langle (\tau_t - \langle \tau_t \rangle)^2 \rangle$ with $\langle \cdot \rangle$ the arithmetic average, is below 0.2 in the two datasets. This shows that the model only captures a small portion of the variance of the fluctuating wall shear stress in the two channels. Such result does not necessarily indicate the poor training performance of the model. Indeed, there is in the context of LES wall modelling an irreducible variance as a large amount of information has been explicitly withdrawn from the model by the filtering and coarse-discretisation processes during the data-preparation step. In addition, it is more critical in the case of a fully developed channel flow to predict accurately the mean value of the wall shear stress and its large-scale variations than the small-scale fluctuations. The machine-learning model follows the mean velocity profile in the mean,

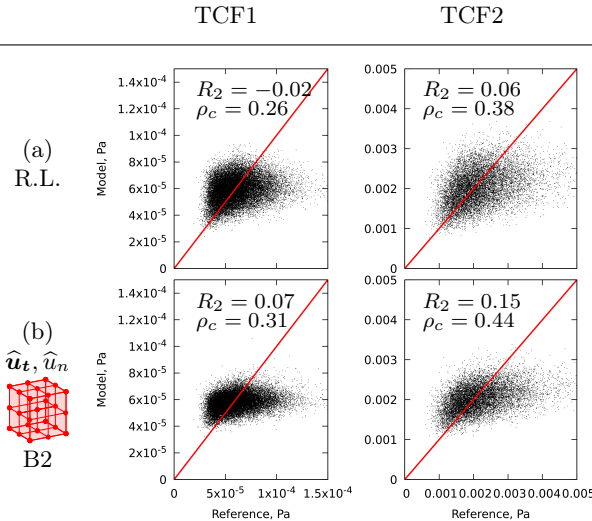


Figure 5: A priori validation: Scatter plot between the target wall shear stress and the prediction of a model based on Reichardt’s law (R.L., a) and the TCF-B2 model (b) in the turbulent channel flows TCF1 and TCF2. The red line is the identity. The coefficient of determination R_2 and the coefficient of correlation ρ_c between the model and the reference wall shear stress are also reported.

and has a residual variance in the same order of magnitude as the residual variance induced by the use of a law-of-the-wall model, which is for many purposes sufficiently accurate for wall-modeled large-eddy simulation of channel flows. This suggests that a comparable performance could be achieved in *a posteriori* tests. This will be demonstrated in section 4.2.2.

4.1.2 Separated flows

To assess the ability of a neural network to predict the wall shear stress in separated regions, a neural network is trained on the three datasets TCF1, TCF2 and 3DD, the three-dimensional diffuser simulation (3DD) providing a complex physics that involves internal corner flows and three dimensional separation. The channel flow data are also included since the model should be able to handle both attached and detached flow regions. The simulation of the backward-facing step (BFS) is kept apart for testing purpose, as this allows us to assess the generalization performance of the model in a simulation distinct from the training simulations. The neural network, hereafter referred to as TCF3DD-B2 is of type B2 according to the classification of table 2. The results are presented in the figures 6b and 7b. In the figures, the analytical profile of Reichardt (5) is also given, not as a reference but as an indication of what a model with insufficient capacity would predict. On the channel flow datasets, the model performs similarly to the model TCF-B2 described in section 4.1.1, in which only data from turbulent channel flows were considered for learning. On the separated flows 3DD and BFS, the data points associated with separated regions are for the most part scattered below Reichardt’s law in the $u^+(y^+)$ graph with both the model-predicted wall shear stress (figure 6b) and the reference wall shear stress (figure 6a). However, there is based on the reference data a non-negligible amount of points falling above Reichardt’s law, a behavior not captured by the machine-learning model. Overall, the relationship between the model output and the target shear stress follows the identity line in the mean (figure 7b), whereas a model based on the law of the wall expectedly deviates from this slope (figure 7a). Coefficients of determination (R_2) of 0.59 and 0.45 are obtained in the 3DD and BFS datasets respectively, which is in both cases markedly larger than with a law-of-the-wall based model. The comparison of the R_2 coefficient across simulations should be made with care. In particular, the variance of the wall shear stress is related to mean spatial variations of the wall shear stress in the 3DD and BFS datasets, since these flows are spatially inhomogeneous, but not in the plane channels. Similarly, the coefficient of correlation ρ_c may only be interpreted in relation to a given dataset since it depends not only

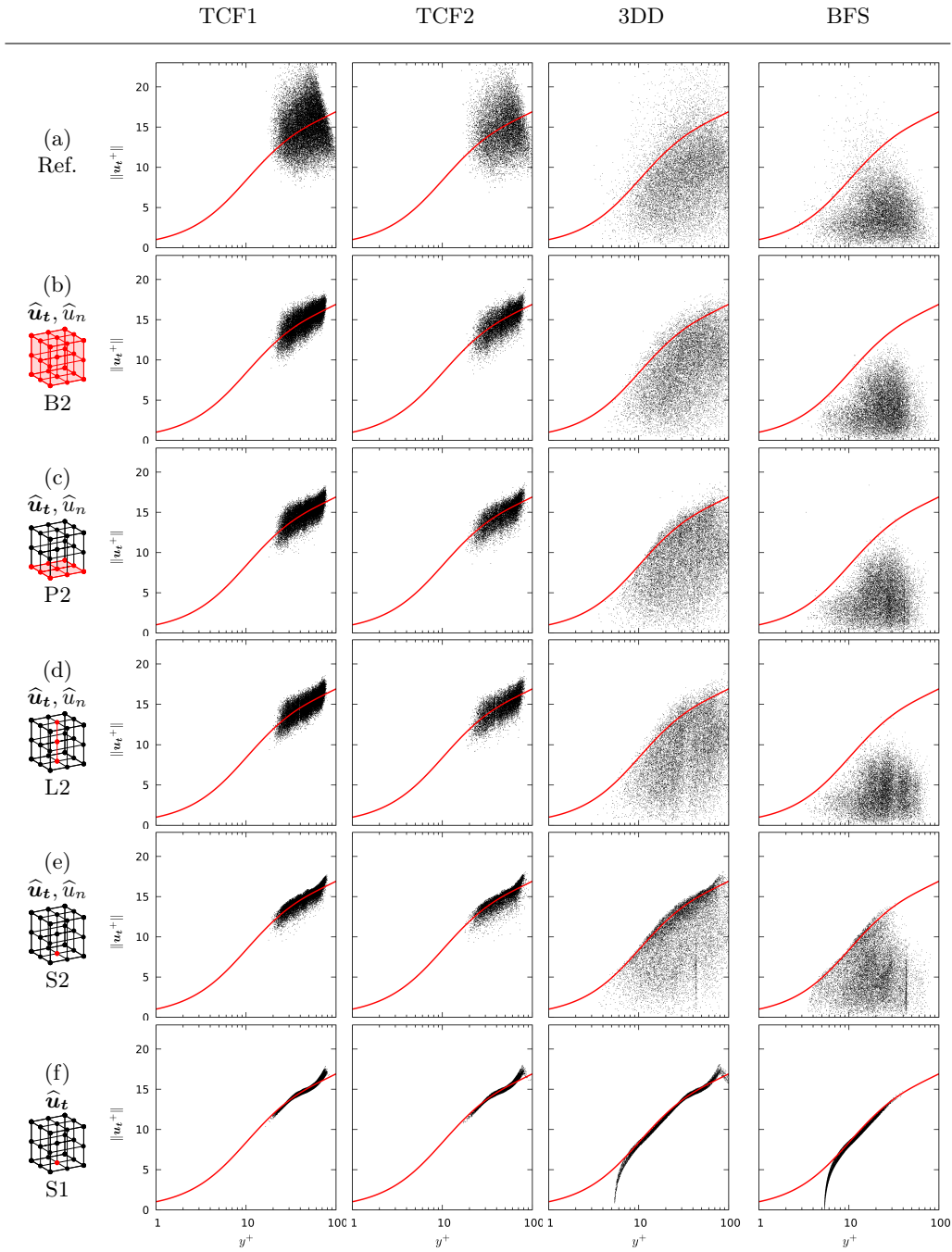


Figure 6: A priori validation: Norm of the scaled tangential velocity u_t^+ as a function of the scaled distance to the wall y^+ in the datasets TCF1, TCF2, 3DD and BFS, using the local target wall shear stress (a) or the prediction of machine-learning models trained on the datasets TCF1, TCF2 and 3DD with various input types (b)–(f) to compute the wall unit scaling ($^+$). The red line is Reichardt’s law, given by equation (5).

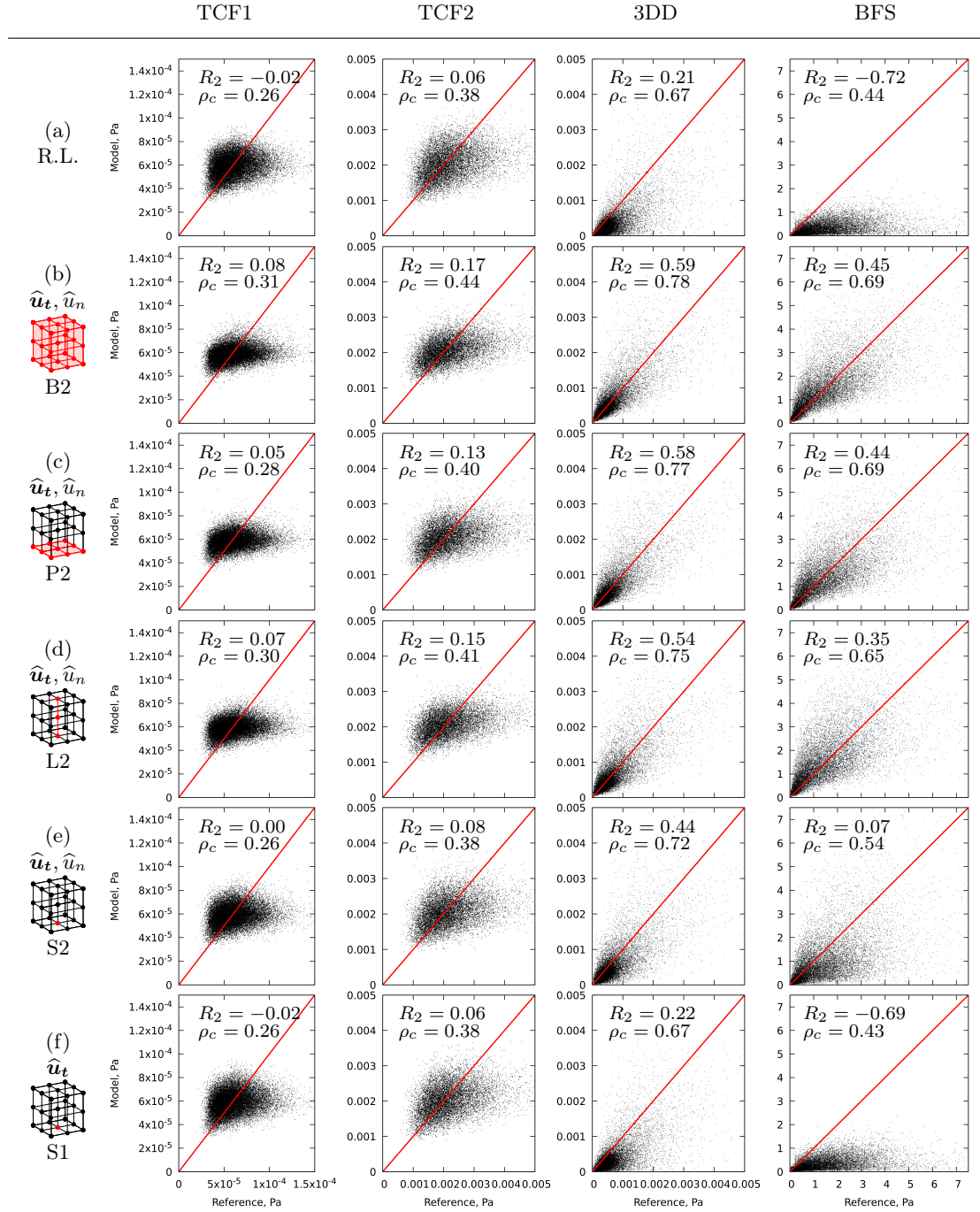


Figure 7: A priori validation: Scatter plot between the target wall shear stress and the prediction of a model based on Reichardt's law (R.L., a) and machine-learning models trained on the datasets TCF1, TCF2 and 3DD with various input types (b)–(f) in the datasets TCF1, TCF2, 3DD and BFS. The red line is the identity. The coefficient of determination R_2 and the coefficient of correlation ρ_c between the model and the reference wall shear stress is also reported.

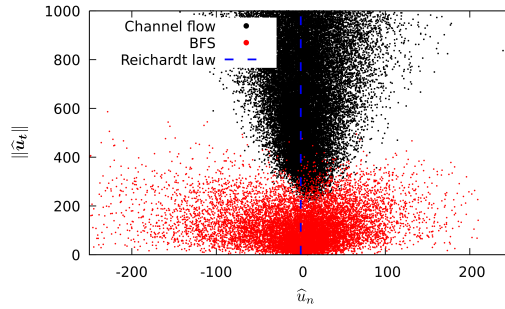


Figure 8: A priori validation: Norm of the scaled tangential velocity \hat{u}_t as a function of the scaled wall-normal velocity \hat{u}_n in the channel flows (TCF1 & TCF2) and the backward-facing step simulation (BFS). The profile implied by the use of the algebraic law of the wall of Reichardt [81] is given for reference.

on the intrinsic performance of the model but also on the underlying distribution of the data. For instance, the fact that the value of ρ_c associated with Reichardt’s law is larger in the 3DD and BFS simulations than in the TCF1 and TCF2 simulations does not imply that the model performs better in these two datasets, the opposite being in fact true. The fact that the model provides sound predictions in the BFS dataset, despite this simulation not being included in the training dataset, is encouraging for the prospect of machine-learning wall modelling as it is indicative of the fact there are similarities between the fields of two different separated flows. This is corroborated by appendix B, which shows that, similarly, a model trained on the dataset TCF1, TCF2 and BFS generalizes well to the 3DD dataset. This suggests that a relatively limited number of simulations could be sufficient to produce a model that can operate in large number of flows, provided that the physical phenomena involved are similar. This should encourage the progressive development of larger databases to specify more diversely the near-wall behavior of fluids.

4.1.3 Effect of increasing spatial information

The model TCF3DD-B2 described in section 4.1.2 makes different predictions in the two channel flows and in the two separated flows 3DD and BFS, since the model predictions are close to the law of the wall in the two channel flows but deviates from the law of the wall in the 3DD and BFS simulations. This shows that the model is able, at least to some extent, to discriminate between separated and non-separated flow regions. In order to investigate the local spatial information that is required for an accurate discrimination, we compare the predictions of machine-learning models trained with a various amount of input information, following table 2. To begin with, the input type S1 is definitely insufficient for discrimination as a model of type S1 can, assuming the model is rotationally invariant, only use the norm of tangential velocity at the first off-the-wall grid point for its prediction. In practice, the model of type S1 is very similar to the classical algebraic wall stress model (figure 6f). The model of type S2, which may in addition use the wall-normal velocity at the first point off the wall, can already discriminate to a large extent the separated regions of our datasets from the channel flow cases (figure 6e). Nevertheless, the predictions in the cases 3DD and BFS remains clearly biased towards the law of the wall (figure 6e). Adding more spatial information, either by using several grid point heights (case L2, figure 6d) or several grid points in the tangential directions (case P2, figure 6c), allows the model to more accurately discriminate the separated regions and adjust its prediction. However, the results suggest that the performance of models of type P2 and L2 is only slightly lower than that of the full model TCF3DD-B2 in our dataset. The examination of the scatter plots and determination coefficients in figure 7 suggests that the models based on a single grid points (S1 and S2) have a poor generalization performance (test score on the BFS dataset) compared to more complex models.

Additional insights may be gained by inspecting more closely the function learned by the S2 model, as made possible by the low-dimensionality of this particular model. Figure 8 shows that the

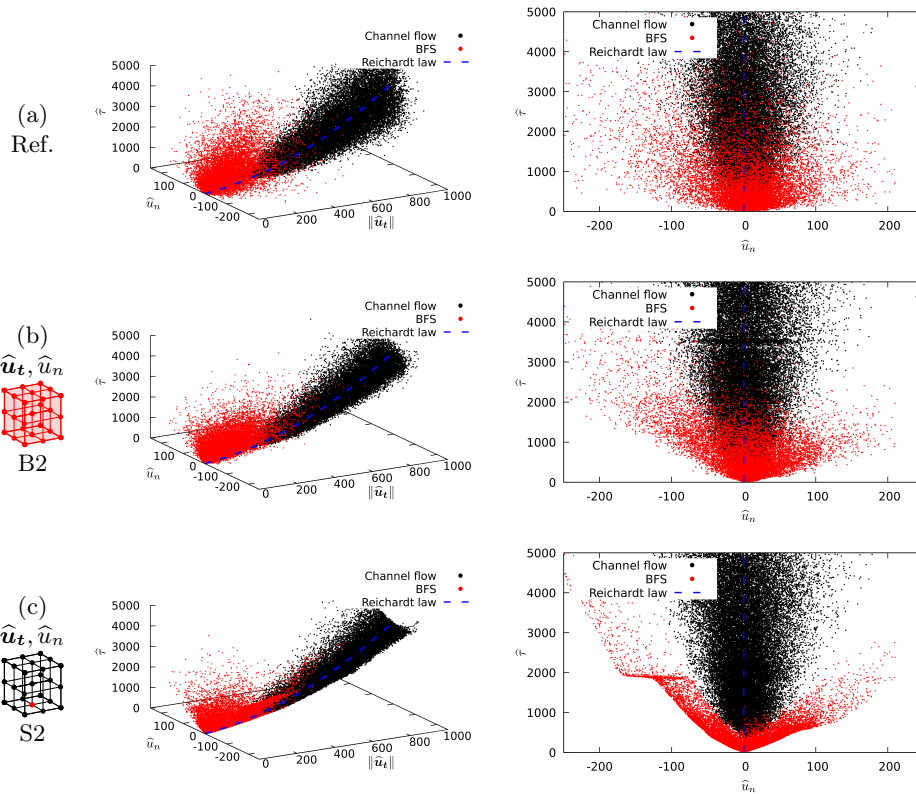


Figure 9: A priori validation: Three-dimensional representation of the relationship between the scaled wall shear stress $\hat{\tau}$, the norm of the scaled tangential velocity $\hat{\mathbf{u}}_t$ and the scaled wall-normal velocity \hat{u}_n in the channel flows (TCF1 & TCF2) and the backward-facing step simulation (BFS) using the local target wall shear stress (a) or the prediction of machine-learning models trained on the datasets TCF1, TCF2 and 3DD with input type B2 (b) or S2 (c). The profile implied by the use of the algebraic law of the wall of Reichardt [81] is given for reference.

channel flows and the separated region of the BFS case occupies different regions in the “effective” input space of the model, which since the model is statistically invariant to rotation is given by the wall-normal velocity and the norm of the tangential velocity at the first off-the-wall point. Namely, the separated regions are typically associated with a lower scaled tangential velocity \hat{u}_t and a larger scaled wall-normal velocity \hat{u}_n in absolute value. Figure 9 compares the behavior of the models in the separated and channel-flow regions. In the channel-flow regions, the S2 model is chiefly dependent on the tangential velocity (figure 9c), as the classical use of the law of the wall would suggest, whereas in the separated region the model mostly depends on the wall-normal velocity and largely ignores variations in tangential velocity. The target wall shear stress, however, is not fully described by these two input features (figure 9a). Thus, the S2 model only provides a rough approximation of the behavior of the target scaled wall shear stress and a large residual variance remains. The models using more spatial information, such as the TCF3DD-B2 model (figure 9b), lead to a more accurate approximation of the target wall shear stress behavior.

Figure 10 compares the spatially averaged wall shear stress predicted by all models investigated. Excluding the model of type S1, which reproduces the law of the wall, the machine-learning models reproduce the overall behavior of the wall-shear-stress profile in both the backward-facing step and the three-dimensional diffuser. While the model S2 provides a clear improvement compared to the use of Reichardt’s law, it consistently underpredicts the average wall shear stress, in particular on the side walls of the three-dimensional diffuser (figures 10e and 10d) and the bottom wall of the backward-facing step (figure 10a), except in the near vicinity of the step ($x/h_s = 0$). The models L2, P2 and B2 lead to similar average wall-shear-stress predictions. In the three-dimensional diffuser simulation, the average wall shear stress is satisfactory on the top and bottom walls after

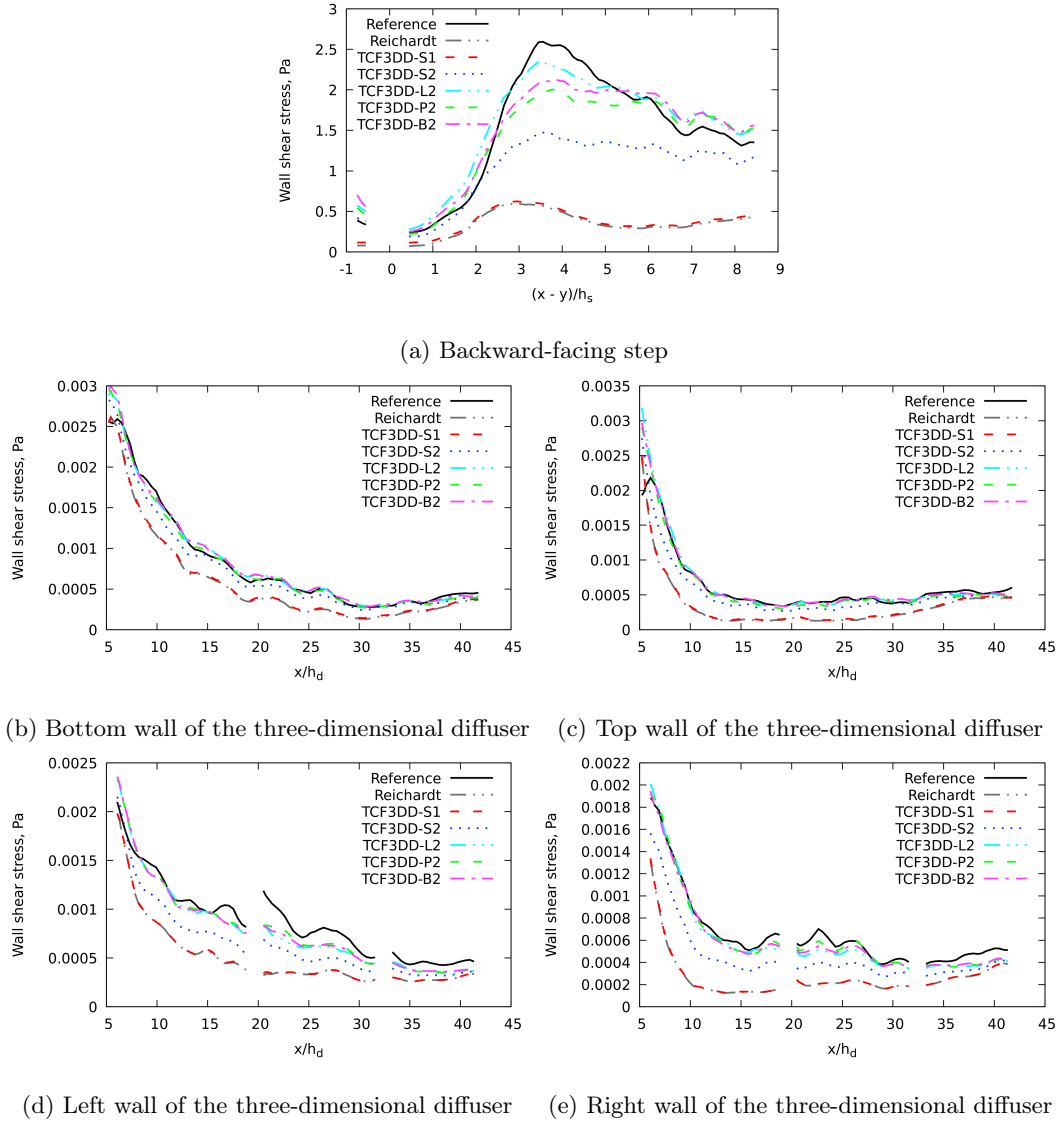


Figure 10: A priori validation: Average prediction of a model based on Reichardt's law and of machine-learning models trained on the datasets TCF1, TCF2 and 3DD with various input types. The averaging is performed in time and along the spanwise direction for the BFS case and the bottom and top walls of the 3DD case, and performed in time and along the vertical direction for the left and right walls of the 3DD case.

$x/h_d = 7$ (figures 10c and 10b). Directly after the onset of the expanding region ($x/h_d = 5$), there is a discrepancy between the model predictions and the reference filtered values which suggests that the models are not able to reproduce the physics associated with the abrupt change of geometry. The models also slightly underestimate the wall shear stress on the side walls (figures 10e and 10d), especially after the expanding region ($x/h_d = 20$). In the backward-facing step simulation (figure 10a), the average wall-shear stress is overestimated near the step and far downstream, whereas it is underestimated in the intermediate region ($x/h_s \approx 4$). In view of these *a priori* results, the simpler model formulations P2 and L2 are thus sufficient for the discrimination of the separated regions in our datasets.

4.2 A posteriori tests

The machine-learning wall models are assessed *a posteriori* by implementing the models described in section 4.1 in the compressible, unstructured and massively parallel flow solver AVBP [83]. Two configurations are studied *a posteriori*: a turbulent channel flow and the flow over a backward-facing step. The LES formalism is similar in both cases and described below.

4.2.1 Governing equations

In our numerical setup, the fluid (air) is modeled as a continuous medium which follows the compressible Navier–Stokes equations. The fluid is supposed Newtonian and obeys the ideal gas equation of state. Stokes’ hypothesis and Fourier’s law are assumed to compute the viscous stress tensor and the conductive heat flux respectively. The governing equations of the large-eddy simulations may be expressed as:

$$\partial_t \rho + \partial_j (\rho U_j) = 0, \quad (6)$$

$$\partial_t (\rho U_i) + \partial_j (\rho U_j U_i) = -\partial_i P + \partial_j \Sigma_{ij}, \quad (7)$$

$$\partial_t (\rho E) + \partial_j (\rho U_j H) = -\partial_j Q_j + \partial_j (\Sigma_{ij} U_i), \quad (8)$$

with ρ the filtered density, t the time, \mathbf{U} the Favre-filtered velocity, \mathbf{x} the Cartesian coordinate, P the filtered pressure, \mathbf{Q} the conductive heat flux, E the total energy per unit mass and $H = E + P/\rho$ the total enthalpy per unit mass. All flow variables implicitly corresponds to filtered fields in accordance with the large-eddy simulation formalism. Temperature is related to internal energy $e = E - \frac{1}{2} U_i U_i$ using the tabulated data of Stull and Prophet [92]. Σ is a tensor combining viscous stresses and subgrid-scale stresses, modeled according to an eddy-viscosity assumption,

$$\Sigma_{ij} = (\mu + \mu_{\text{sgs}}) \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} - \frac{2}{3} \frac{\partial U_k}{\partial x_k} \delta_{ij} \right), \quad (9)$$

where δ_{ij} denotes the Kronecker delta.

4.2.2 Channel flow

The relevance of the data-driven wall modeling approach is first assessed for channel flows at friction Reynolds numbers $Re_\tau = 395$, $Re_\tau = 950$ and $Re_\tau = 2000$. This includes both Reynolds numbers that are within and beyond the Reynolds numbers found in the training dataset (180 and 950). The simulations are performed by discretizing the governing equations (6)–(8) using a cell-vertex finite-element method and an explicit time stepping scheme. The second-order scheme of Lax and Wendroff [49] is used for convection, while diffusion is discretized with a second-order accurate centered scheme. The subgrid-scale viscosity is computed using the Sigma subgrid-scale model [68]. At $Re_\tau = 395$, the domain size is $5.4\pi h_c \times 2h_c \times 2.7\pi h_c$, where h_c is the half-height of the channel. At $Re_\tau = 950$ and $Re_\tau = 2000$, the domain size is $4\pi h_c \times 2h_c \times 2\pi h_c$. The WMLES meshes are regular and uniform. Two mesh refinement are defined. With the mesh refinement A, the cell size is $\Delta^+ = 33$ in wall units and the mesh contains $201 \times 23 \times 101$ grid points at $Re_\tau = 395$, $359 \times 55 \times 180$ grid points at $Re_\tau = 950$ and $755 \times 114 \times 378$ grid points at $Re_\tau = 2000$. With the

mesh refinement B, the cell size is $\Delta^+ = 33$ in wall units and the mesh contains $103 \times 12 \times 52$ grid points at $Re_\tau = 395$ and $183 \times 28 \times 92$ grid points at $Re_\tau = 950$. With both the mesh refinements A and B, the cell size is consistent with the assumption of the training procedure regarding the cell sizes, namely a first off-the-wall point below $y^+ = 75$.

The simulations are performed using the neural network TCF3DD-B2, trained on the datasets TCF1, TCF2, 3DD with the input type B2 and the neural network TCF3DD-L2, trained on the datasets TCF1, TCF2, 3DD with the input type L2. The results are compared in figure 11 to the corresponding simulations with a classical algebraic wall stress model. With the algebraic wall stress model, the logarithmic layer is shifted either upward or downward compared to the reference profile. This mismatch is well documented in the literature and is expected with this type of model [9]. At $Re_\tau = 395$ and $Re_\tau = 950$, the simulations with the machine-learning models improves upon these results. In particular, the simulation with the TCF3DD-B2 predicts the mean velocity profile accurately throughout the logarithmic layer. This improvement is possible because the machine learning model uses several grid points in the model prediction, while the mismatch is related to an unphysical correlation between the first off-the-wall point and the wall shear stress [101]. Note that using an algebraic wall stress model, the logarithmic-layer mismatch can following Kawai and Larsson [45] be prevented by using a farther imposition point for the input velocity, typically three cells within the large-eddy-simulation grid [48, 11]. The present “law-of-the-wall model” should thus be considered as a baseline of a commonly used algebraic wall model that does not attempt to correct for this mismatch. At $Re_\tau = 2000$, the velocity profiles obtained using the algebraic wall stress model and the TCF3DD-B2 model are however almost identical. The comparison of the standard deviation of streamwise velocity (figure 11) confirms that the predictions of the algebraic wall stress model and the machine-learning model are similar. Since this friction Reynolds number is not within the range of friction Reynolds number seen during training, the results characterise the ability of the machine-learning models to extrapolate in terms of Reynolds number. Owing to the universality of the mean velocity profile in a plane channel, the models are able to operate to some extent for unseen Reynolds number. However, the topology of the instantaneous velocity fields may be different, as reflected by the emergence of an outer peak in the streamwise velocity spectra at high Reynolds numbers [39, 40]. This induces the model to make a prediction close to the mean velocity profile and therefore to the “law-of-the-wall model” at $Re_\tau = 2000$.

Figure 12 shows the results, using the coarser mesh refinement B, of large-eddy simulations with the model TCF3DD-B2 at $Re_\tau = 395$ and $Re_\tau = 950$. As with the finer mesh, the machine-learning model TCF3DD-B2 is able at those two friction Reynolds numbers to prevent the logarithmic-layer mismatch and produce a mean velocity profile that is close to the reference. To give a more quantitative measure of the accuracy of the wall-modeled large-eddy simulations, the mean wall shear stress of the wall-modeled large-eddy simulations are reported in table 4. The machine-learning models provide a clear improvement compared to the algebraic shear stress model at $Re_\tau = 395$ and $Re_\tau = 950$ and matches the performance of the algebraic wall stress model at $Re_\tau = 2000$. Overall, the results show that the machine-learning strategy is capable of learning a model that generalizes to a Reynolds number not included in the training dataset, provided that it remains within the range of Reynolds number seen during training. The simulation of a channel flow at a Reynolds number greater than seen during training may also be carried out, but lead to a degraded performance because the velocity-field topology is Reynolds number dependent. Hence, machine-learning models should ideally be developed on a range of Reynolds number consistent with the target application.

4.2.3 Backward-facing step

In this section, the relevance of the machine-learning wall model is assessed a posteriori for the simulation of a backward-facing step. The results are compared to a reference high-fidelity simulation that reproduces the backward-facing step configuration of Le *et al.* [50]. Namely, the Reynolds number based on the inflow bulk velocity and step height h_s is $Re_h = 5100$ and the expansion ratio $ER = (L_y + h_s)/L_y$, where L_y is the inlet channel height, is equal to 1.2. At the inlet, located $20 h_s$ before the step, the mean velocity profile of Spalart [89] imposed with the characteristic boundary condition of Poinso and Lele [77]. In addition, isotropic fluctuations are

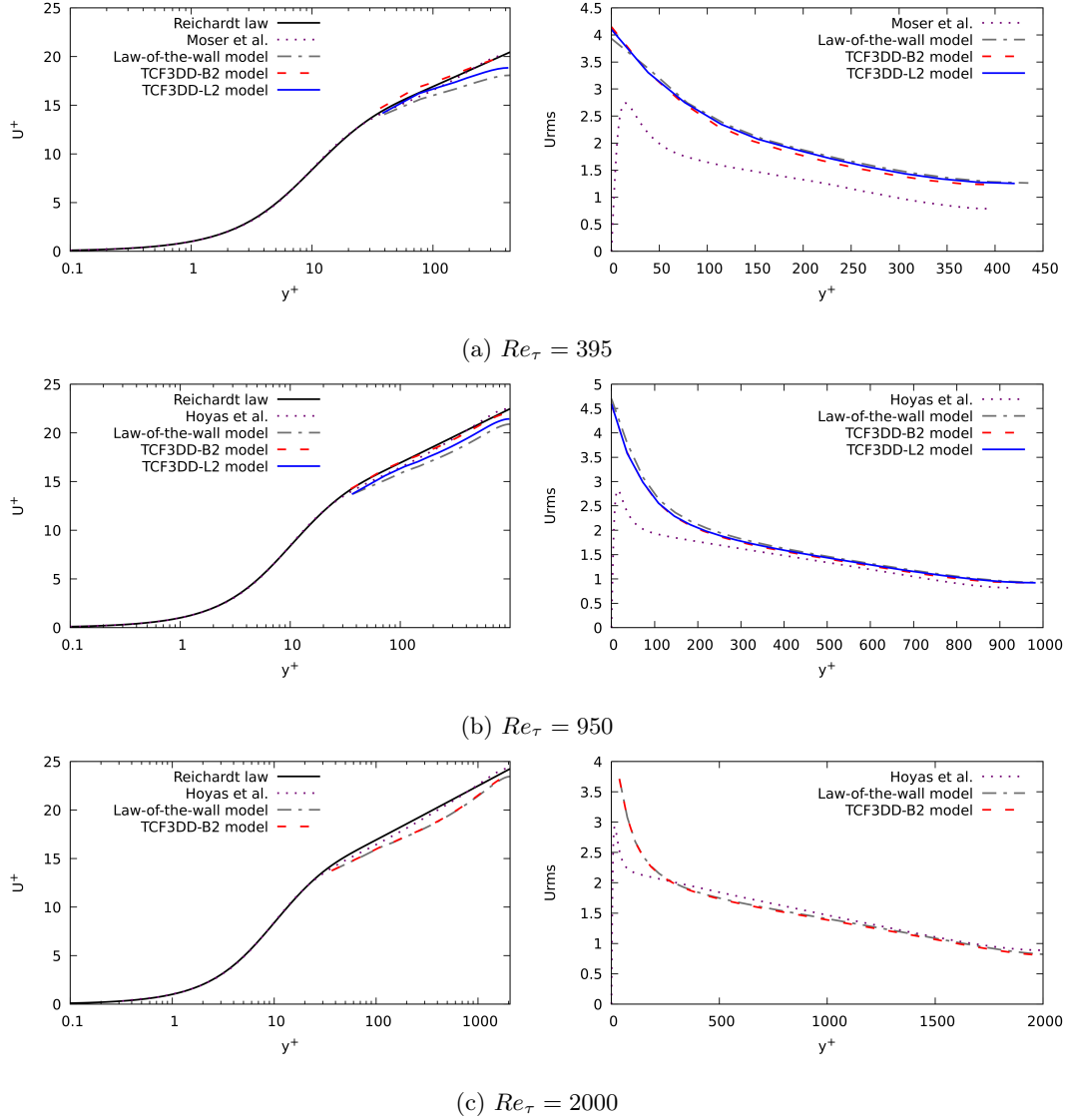


Figure 11: A posteriori validation: Mean streamwise velocity (left) and standard deviation of streamwise velocity (right) in large-eddy simulations of a channel flow at the friction Reynolds numbers $Re_\tau = 395$, $Re_\tau = 950$ and $Re_\tau = 2000$ with the mesh refinement A using an algebraic wall stress model, the TCF3DD-B2 machine-learning model and the TCF3DD-L2 machine-learning model. The unfiltered direct numerical simulation profiles of Moser *et al.* [66] and Hoyas and Jiménez [37] are given for comparison. Note that the unfiltered direct numerical simulation profiles of Moser *et al.* [66] or Hoyas and Jiménez [37] do not provide a target reference for the standard deviation of streamwise velocity in figure 11b, since a large part of the turbulence kinetic energy is contained within subgrid scales in the WMLES.

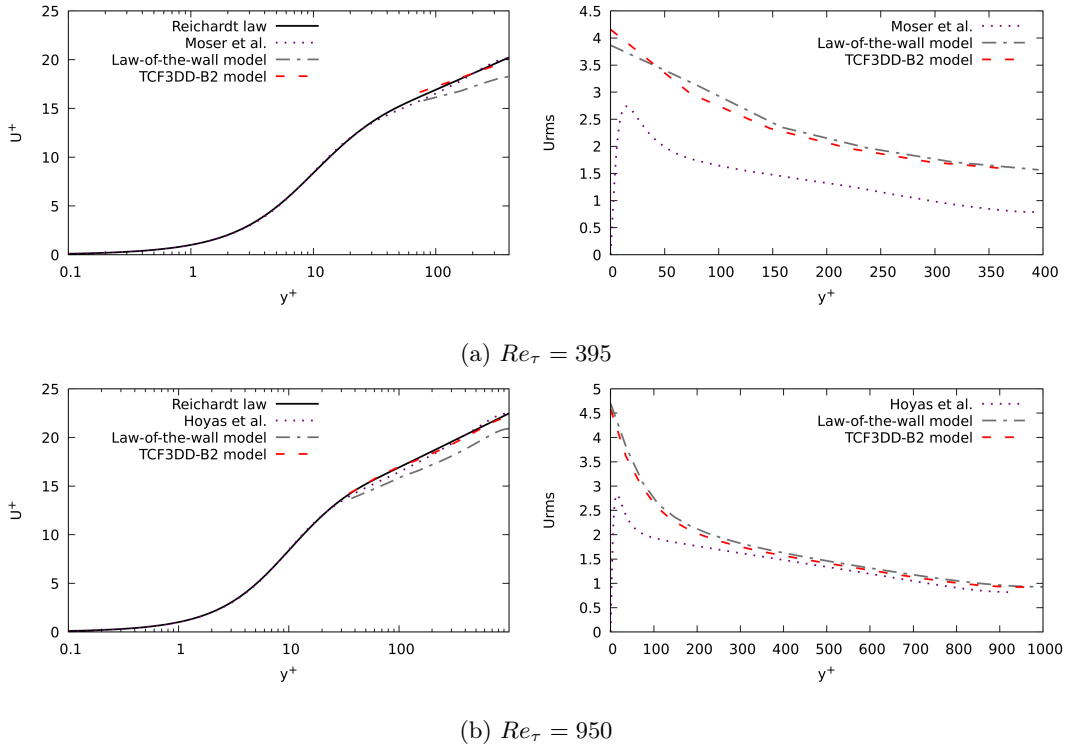


Figure 12: A posteriori validation: Mean streamwise velocity (left) and standard deviation of streamwise velocity (right) in large-eddy simulations of a channel flow at the friction Reynolds numbers $Re_\tau = 395$, $Re_\tau = 950$ and $Re_\tau = 2000$ with the mesh refinement B using an algebraic wall stress model and the TCF3DD-B2 machine-learning model. The unfiltered direct numerical simulation profiles of Moser *et al.* [66] and Hoyas and Jiménez [37] are given for comparison.

injected using the approach of Kraichnan [47], following the turbulence kinetic energy profile of Spalart [89] in the mean. A constant pressure of 1 bar is imposed at the outlet, located $20 h_s$ after the step. A symmetric boundary condition is applied at the upper boundary of the computational domain, which imposes a zero shear stress and a zero wall-normal velocity. The geometry is periodic in the spanwise direction and the spanwise length of the domain is $4 h_s$. The mesh contains 114.3 million grid points in total and the height of the first point off the wall is $\Delta y^+ = 0.4$ for the boundary layer before the edge of the step. In the wall-modeled large-eddy simulation, the inlet is located $10 h_s$ before the step and the mean streamwise velocity and turbulence kinetic energy profile of the wall-resolved simulation are imposed. The WMLES mesh has been refined near the interface between the inlet boundary layer and the separated region to ensure a proper prediction of the shear layer, while maintaining a coarse mesh resolution near the walls. A close-up view of the mesh around the refined region is given in figure 13. The mesh contains 0.9 million grid points in total. The cell size on the boundary layer before the edge of the step is $\Delta x^+ = \Delta z^+ = 33$ in the streamwise and spanwise directions, and $\Delta y^+ = 35$ in the wall-normal direction. The numerical method is identical to that of section 4.2.2 and the same subgrid-scale model is used.

The simulation is performed using a classical algebraic wall stress model, the TCF3DD-B2 neural network model and the TCF3DD-L2 neural network model. The mean streamwise velocity of each case is compared in figure 14. With both the machine-learning model and the algebraic wall stress model, the maximum backflow region is located farther from the step than in the reference wall-resolved numerical simulation (figure 14). The machine-learning model slightly improves the shape of the mean velocity profile at the onset of the mean backflow region ($x/h_s = 0 - 2$) compared to the algebraic wall stress model. This is confirmed by the streamlines of mean velocity, given in figure 15. There is no secondary recirculation region using the algebraic law-of-the-wall model, but there is small secondary recirculation region with the TCF3DD-B2 model and

Table 4: Mean nondimensionalized wall shear stress $\tau/(\rho u_c^2) = (u_\tau/u_c)^2$, where u_c is the centerline velocity, in large-eddy simulations of a channel flow at the friction Reynolds numbers $Re_\tau = 395$, $Re_\tau = 950$ and $Re_\tau = 2000$ with an algebraic wall stress model and a machine-learning wall model.

	$Re_\tau = 395$	$Re_\tau = 950$	$Re_\tau = 2000$
Reference: { Moser <i>et al.</i> [66]	0.00247	—	—
{ Hoyas and Jiménez [37]	—	0.00199	0.00169
Law-of-the-wall model, mesh refinement A	0.00306 (+24%)	0.00229 (+15%)	0.00182 (+8%)
TCF3DD-L2 model, mesh refinement A	0.00282 (+14%)	0.00218 (+9%)	N/D
TCF3DD-B2 model, mesh refinement A	0.00253 (+2%)	0.00207 (+4%)	0.00180 (+7%)
Law-of-the-wall model, mesh refinement B	0.00299 (+21%)	0.00249 (+25%)	N/D
TCF3DD-B2 model, mesh refinement B	0.00260 (+5%)	0.00222 (+11%)	N/D

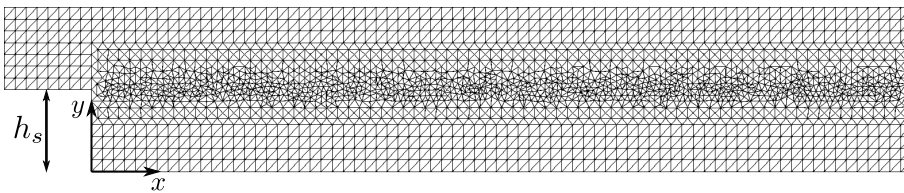


Figure 13: Cross-section of the mesh around the refined region in the x - y plane.

the TCF3DD-L2 model. Both the amplitude and size of this secondary recirculation region are greatly underestimated compared to the reference simulation however. In the boundary layer upstream of the step, the TCF3DD-L2 machine-learning model provide a prediction that is close to that of our algebraic law-of-the-wall model, whereas the TCF3DD-B2 machine-learning model provide a prediction that is farther from the reference (figure 16). This difference does not affect significantly the analysis downstream of the step. Indeed, the effect of the shear stress model on the mean velocity profiles downstream of the step is overall small, as the mean velocity predominantly results from the shear layer induced by the step. Figure 17 shows that there is no clear difference in the mean velocity profiles between the TCF3DD-L2 model, the TCF3DD-B2 model and the algebraic wall stress model. Figure 18 compares the predictions of the machine-learning model in a $u^+(y^+)$ graph for a particular timestep of the WMLES simulations of the channel flow and the backward-facing step. The model predictions are close to the law of the wall in the channel flow simulation, while in the backward-facing step simulation the model predictions are scattered below Reichardt’s law in a manner that is qualitatively in accordance with the *a priori* results of figure 6. This demonstrates the ability of the model to leverage the local spatial information to discriminate attached wall turbulence and separated regions in an *a posteriori* setting.

The profile of mean wall shear stress along the wall is compared in figure 19. The mean wall shear stress is computed by averaging the instantaneous shear stress norm in time and in the spanwise direction. The magnitude of the shear stress downstream of the step up to $x/h_s = 2$ is well predicted in the two wall-modeled large-eddy simulations. Farther from the step, the wall shear stress is significantly underestimated with the algebraic wall stress model compared to the reference simulation, while the simulations with the machine-learning models leads to a larger wall shear stress. Although the discrepancy is greatly reduced with both the TCF3DD-B2 and TCF3DD-L2 models, the peak wall shear stress remains underestimated and the wall shear stress increase at $x/h_s = 2$ is less sharp than in the reference simulation. The profile of the mean streamwise wall shear stress (figure 20) confirms the underestimation for the peak of the mean streamwise wall shear stress. The present wall-modelled large-eddy simulations are compared to the numerical results of Chen *et al.* [15] and Shi *et al.* [85] in figure 20, which use alternative wall-modelling approaches suitable for non-equilibrium conditions. The results are not directly comparable since the mesh and the numerical method varies between the different studies. In particular, Shi *et al.* [85] used a mesh that is more than twice a refined as ours. The simulations of Chen *et al.* [15] compare favorably to ours for the prediction of the peak of the mean streamwise wall shear stress,

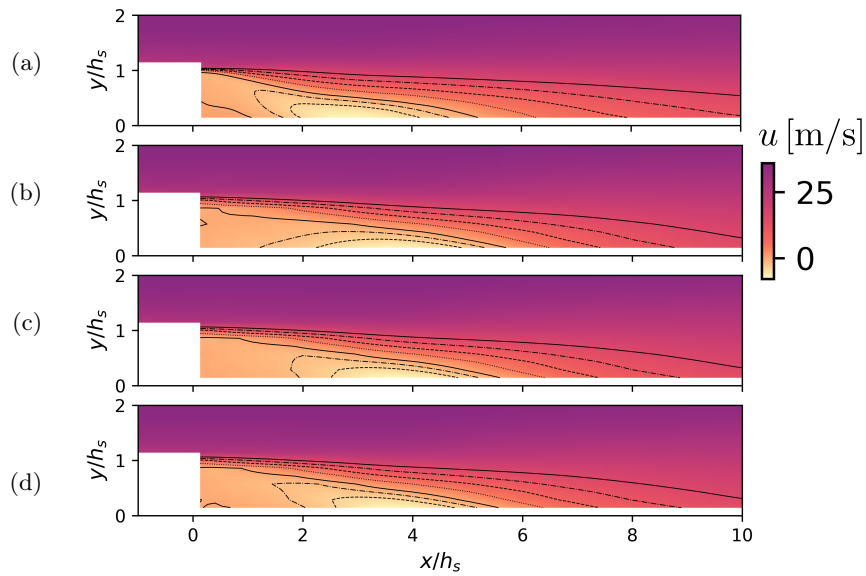


Figure 14: A posteriori validation: Mean streamwise velocity in the flow over a backward-facing step as predicted by the reference wall-resolved simulation (a), a large-eddy simulations with an algebraic wall stress model (b), the machine-learning model TCF3DD-B2 (c) and the machine-learning model TCF3DD-L2 (d). The contour lines denote the level sets -4 , -2 , 0 , 4 , 8 , 12 and 16 m/s. Values within the first WMLES cell are not provided since the wall-modeled large-eddy simulations do not provide a physical velocity at the wall.

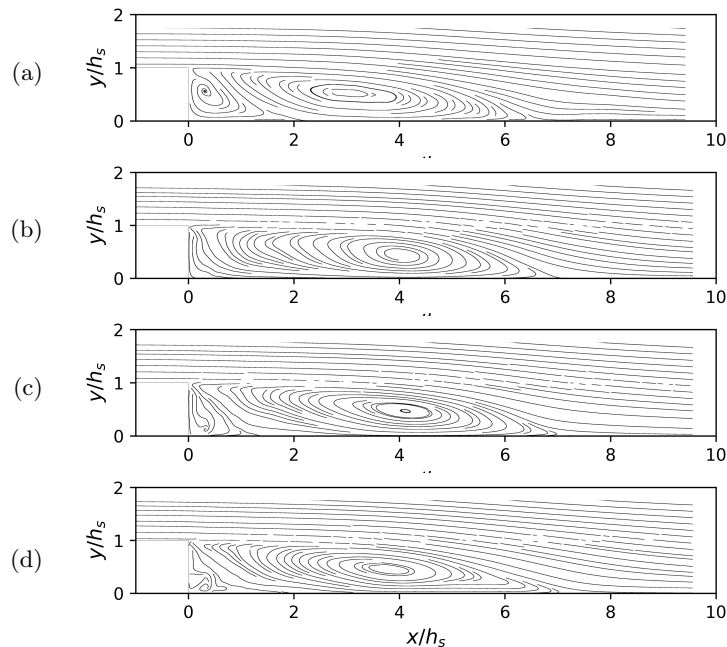


Figure 15: A posteriori validation: Streamlines of mean velocity in the flow over a backward-facing step as predicted by the reference wall-resolved simulation (a), a large-eddy simulations with an algebraic wall stress model (b), the machine-learning model TCF3DD-B2 (c) and the machine-learning model TCF3DD-L2 (d).

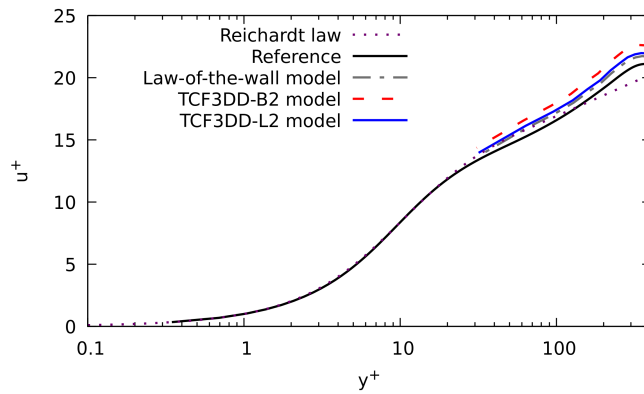


Figure 16: A posteriori validation: Mean streamwise velocity upstream of the step ($x/h_s = -1$) in the backward-facing step simulation in the reference simulation and the wall-modelled large-eddy simulations.

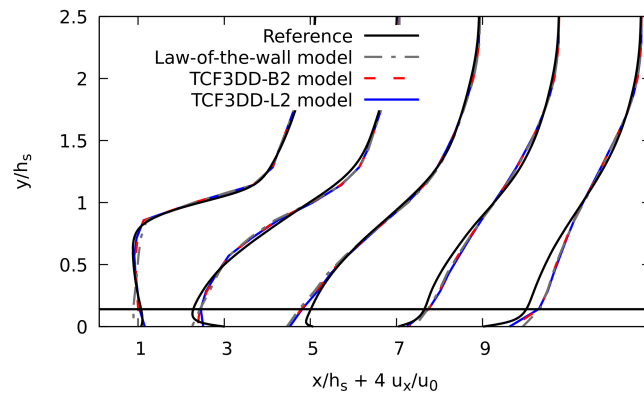


Figure 17: A posteriori validation: Profile of mean streamwise velocity in the flow over a backward-facing step at the location $x/h_s = 1, 3, 5, 7, 9$, scaled by the freestream velocity u_0 , as predicted by large-eddy simulations with an algebraic wall stress model, the machine-learning model TCF3DD-B2 and the machine-learning model TCF3DD-L2. The horizontal solid line gives the height of the first point off the wall.

but predicts a reattachment point further downstream than in the reference simulation. Conversely, the streamwise location of the reattachment point is underestimated in both of our large-eddy simulations. The difference in behaviour can be explained by the modelling paradigm. Indeed, in the present study the machine-learning procedure only models the norm of the shear stress vector while its orientation is, as in the algebraic wall stress model, given by the tangential velocity at the first point above the wall. This limitation may be observed in figure 21, which compares the mean streamwise velocity at the height of the first WMLES cell and the wall shear stress in the reference simulation and the wall-modelled large-eddy simulation with the TCF3DD-L2 model. There is offset in the order of h_s between the locations where the wall shear stress and the tangential velocity at the first point above the wall reach zero in the reference simulation, which may not be reproduced in the wall-modelled large-eddy simulation with the present approach. The development of machine-learning wall models that consider the prediction of the wall shear stress vector is necessary to remedy this inherent limitation of the present paradigm. To give a more quantitative measure of the accuracy of the wall-modeled large-eddy simulations, the mean squared error between the shear stress of the wall-modeled large-eddy simulations and the reference wall-resolved simulation is given in table 5. The machine-learning models provides a clear improvement compared to the algebraic shear stress model relying on RANS assumptions and intended for a

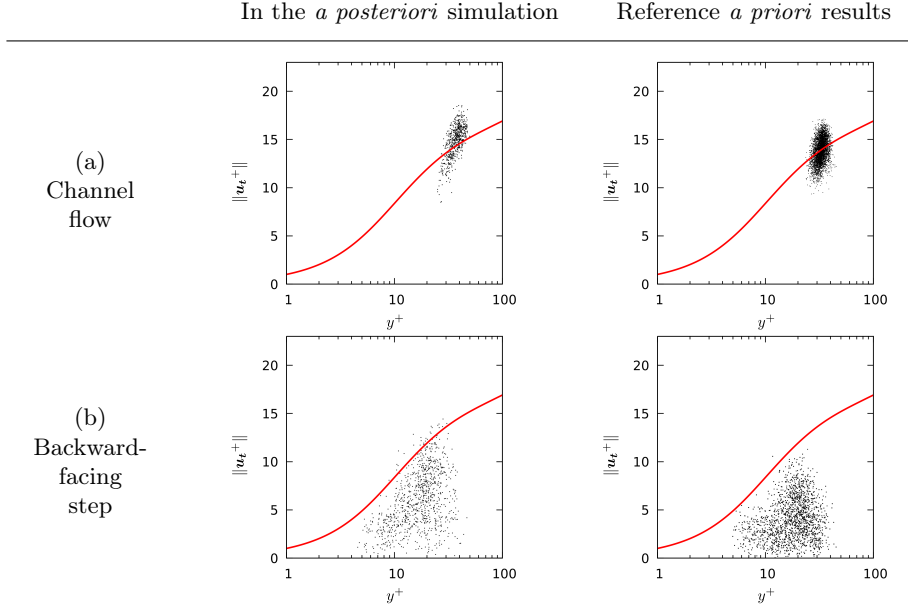


Figure 18: *A posteriori* validation: Norm of the scaled tangential velocity \mathbf{u}_t^+ as a function of the scaled distance to the wall y^+ according to the prediction of the TCF3DD-B2 machine-learning model for a particular timestep of the *a posteriori* simulations of the channel flow at $Re_\tau = 395$ (top left) and of the backward-facing step (bottom left). For reference, the corresponding prediction based on the reference *a priori* database are also given for the channel flows at $Re_\tau = 180$ and 950 (top right) and the backward-facing step (bottom right). The red line is Reichardt's law, given by equation (5).

Table 5: Integral measure of the disagreement between the reference wall shear stress in the flow over a backward-facing step and the wall shear stress predicted by large-eddy simulations with an algebraic wall stress model, the TCF3DD-B2 machine-learning wall model and the TCF3DD-L2 machine-learning wall model.

Simulation	$\frac{\int_0^{14h_s} (\tau - \tau^{\text{ref}})^2 dx}{\int_0^{14h_s} (\tau^{\text{ref}})^2 dx}$	$\frac{\int_0^{14h_s} (\tau_x - \tau_x^{\text{ref}})^2 dx}{\int_0^{14h_s} (\tau_x^{\text{ref}})^2 dx}$
WMLES with law-of-the-wall model	0.223772	0.26975
WMLES with TCF3DD-B2 model	0.0378455	0.209642
WMLES with TCF3DD-L2 model	0.0200867	0.214579

fully turbulent flat plate without pressure gradient for the prediction of the wall shear stress in the backward-facing step case in a WMLES framework. With regard to the comparison between the TCF3DD-B2 model and the TCF3DD-L2 model, there is no clear superiority of one over the other overall. Indeed, the TCF3DD-L2 model provides more accurate predictions than the TCF3DD-B2 model in the backward-facing step case, whereas the TCF3DD-B2 model provides more accurate predictions than the TCF3DD-L2 model in the channel flows.

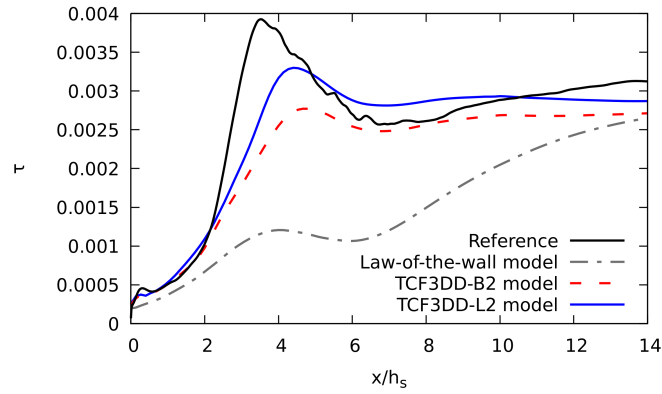


Figure 19: A posteriori validation: Mean wall shear stress profile in the flow over a backward-facing step as predicted by large-eddy simulations with an algebraic wall stress model, the TCF3DD-B2 machine-learning model and the TCF3DD-L2 machine-learning model.

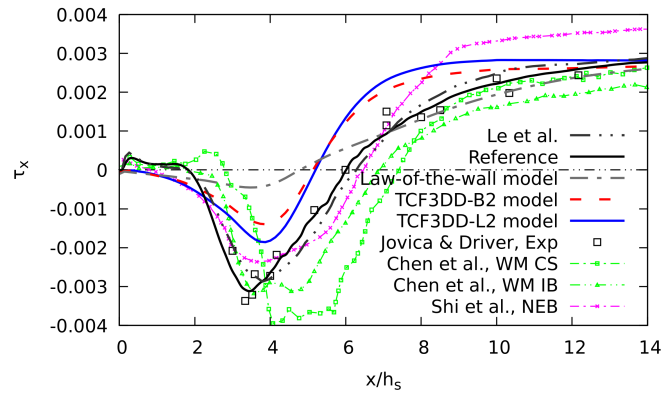


Figure 20: A posteriori validation: Mean streamwise component of the wall shear stress vector in the flow over a backward-facing step as predicted by large-eddy simulations with an algebraic wall stress model, the TCF3DD-B2 machine-learning model and the TCF3DD-L2 machine-learning model. The experimental results of Jovic and Driver [44] and the direct numerical simulation results of Le *et al.* [50] are also given for reference, as well as the wall-modelled large-eddy simulations of Chen *et al.* [15] and Shi *et al.* [85].

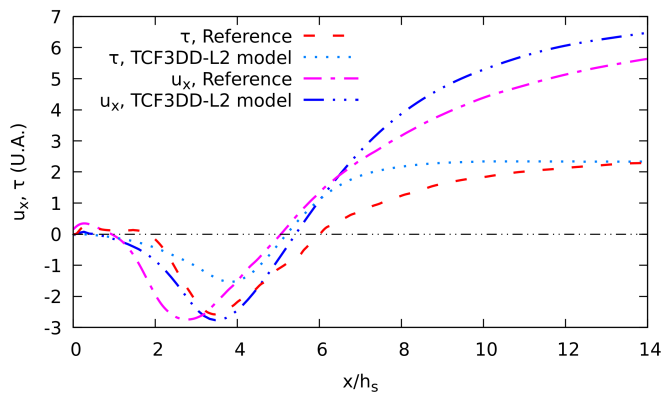


Figure 21: A posteriori validation: Mean streamwise velocity at the height of the first WMLES cell and mean streamwise wall shear stress in the reference simulation and the wall-modelled large-eddy simulation with the TCF3DD-L2 model.

5 Conclusion

The wall modeling approach developed in this paper uses a database of instantaneous three-dimensional flow snapshots to learn the relationship between the wall shear stress and the velocity components at one or several grid points above the wall. The procedure is data-driven and learns the velocity profile purely from the training data with minimal feature engineering, only encoding the incompressibility assumption and accordingly the Mach number equivariance of the flow at the quasi-incompressible limit. The relevance of the modeling approach has been demonstrated for turbulent separated flows by training feed-forward artificial neural networks with the filtered and subsampled data from four high-fidelity numerical simulations. The resulting model leverages the three velocity components at several spatial locations to discriminate between regions of fully developed wall turbulence and detached flow regions. The model generalizes well to a flow configuration not included in the training dataset. The improvements compared to an algebraic wall stress model based on statistical equilibrium assumption are demonstrated *a posteriori* for the wall-modeled large-eddy simulation of a channel flow and of the flow over a backward-facing step. These results suggest that data-driven wall models can be relevant and shed light on the importance of the availability of public direct numerical simulation datasets.

Acknowledgments

This work was granted access to the computer resources of IDRIS under the allocation A0102A06074 made by GENCI. This project has received funding from the European Union’s Horizon 2020 research and innovation program (grant agreement n° 814837). We are grateful to O. Lehmkuhl and A. Miró for providing us the data for the three-dimensional diffuser simulation and L. Agostini and P. Vincent for providing us the data for the channel flow at $Re_\tau = 180$.

A Additional *a priori* results for channel flows

Section 4.1 performed an *a priori* validation of the machine-learning procedure for a turbulent channel flow using a model of type B2 (table 2). This section presents the corresponding results for models of type S1, S2, L2 and P2 (figures 22 and 23).

B Additional *a priori* results for separated flows

Section 4.1 demonstrated the generalization capability of the machine-learning procedure by evaluating the performance in the backward-facing step dataset (BFS) of a model trained on the datasets TCF1, TCF2 and 3DD. This section performs the opposite analysis, that is evaluating the generalization performance by training a model on the datasets TCF1, TCF2 and BFS and testing the model on the three-dimensional diffuser dataset (3DD). The results are similar overall (figures 24 and 25).

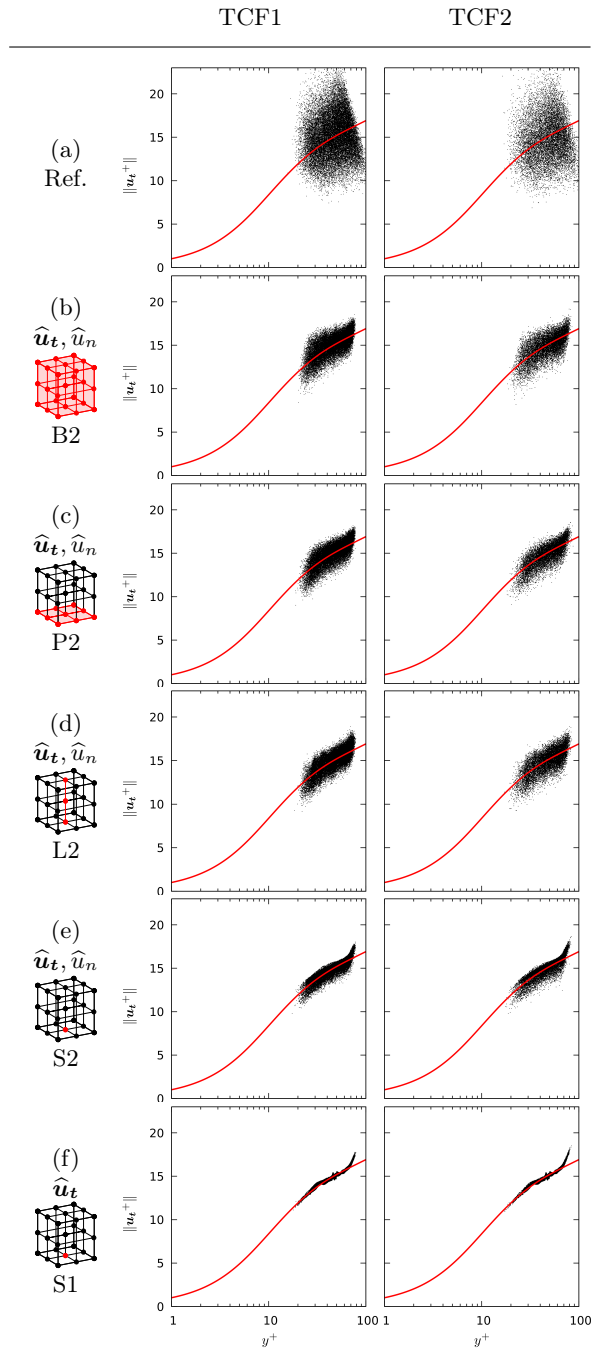


Figure 22: A priori validation: Norm of the scaled tangential velocity \mathbf{u}_t^+ as a function of the scaled distance to the wall y^+ in the turbulent channel flows TCF1 and TCF2, using the local target wall shear stress (a) or the prediction of machine-learning models trained on the turbulent channel flows TCF1 and TCF2 with various input types (b)–(f) to compute the wall unit scaling ($^+$). The red line is Reichardt’s law, given by equation (5).

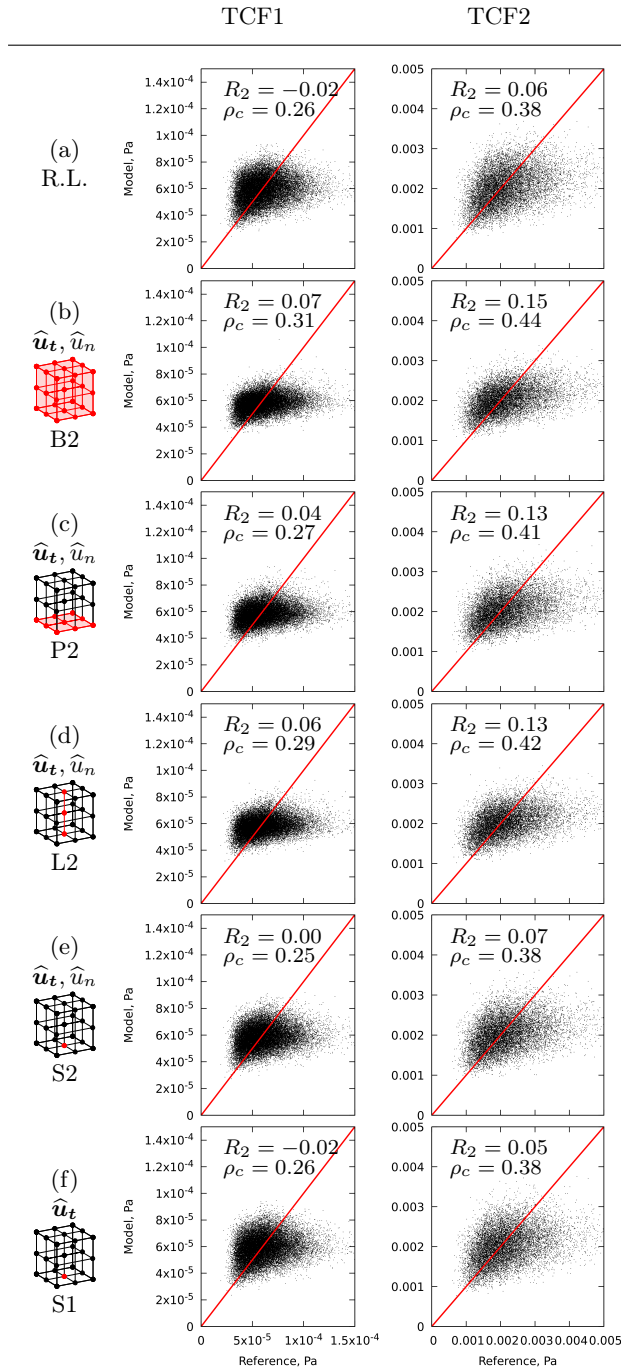


Figure 23: A priori validation: Scatter plot between the target wall shear stress and the prediction of a model based on Reichardt's law (R.L., a) and machine-learning models trained on the turbulent channel flows TCF1 and TCF2 with various input types (b)–(f). The red line is the identity. The coefficient of determination R_2 and the coefficient of correlation ρ_c between the model and the reference wall shear stress are also reported.

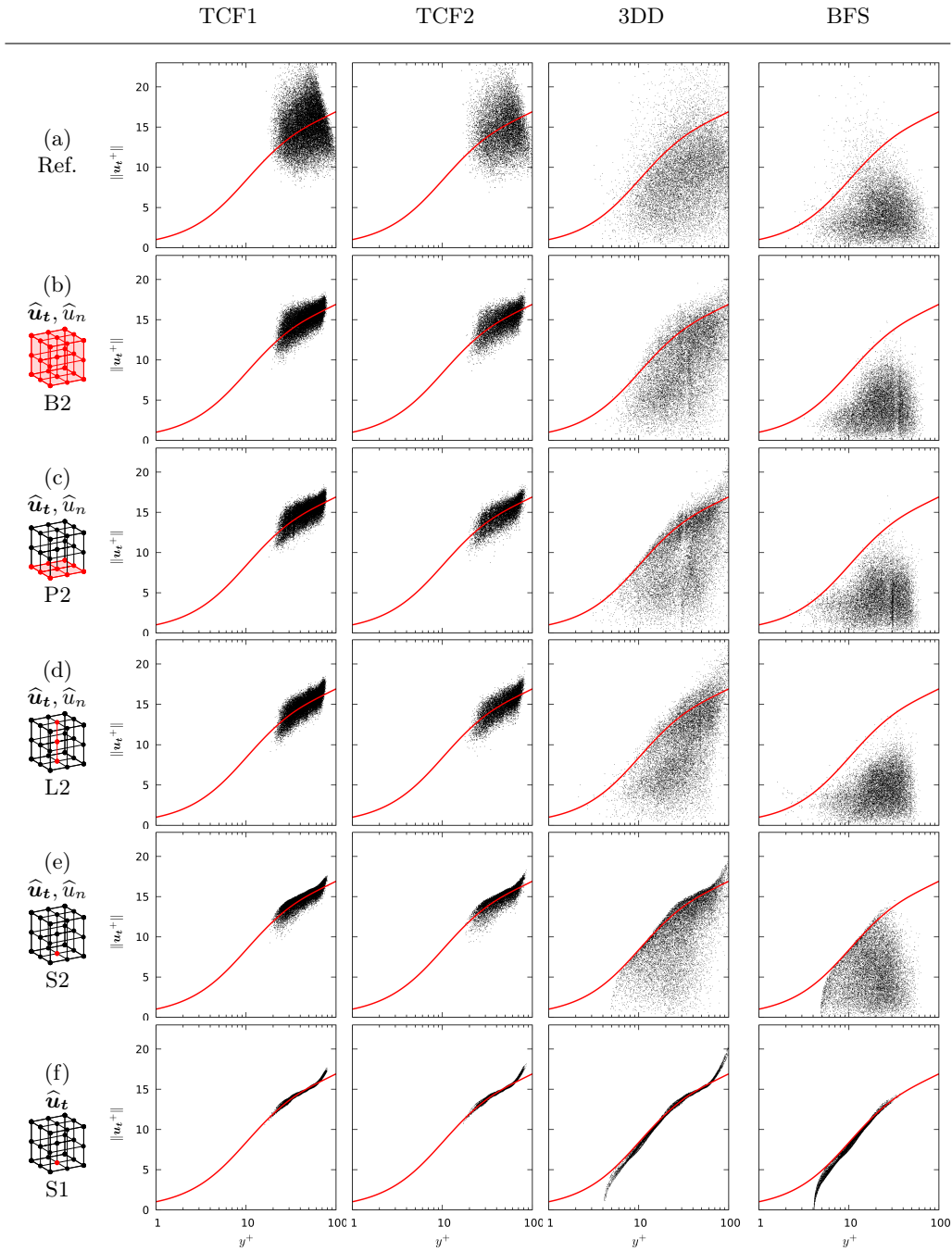


Figure 24: A priori validation: Norm of the scaled tangential velocity \mathbf{u}_t^+ as a function of the scaled distance to the wall y^+ in the datasets TCF1, TCF2, 3DD and BFS, using the local target wall shear stress (a) or the prediction of machine-learning models trained on the datasets TCF1, TCF2 and BFS with various input types (b)–(f) to compute the wall unit scaling ($^+$). The red line is Reichardt’s law, given by equation (5).

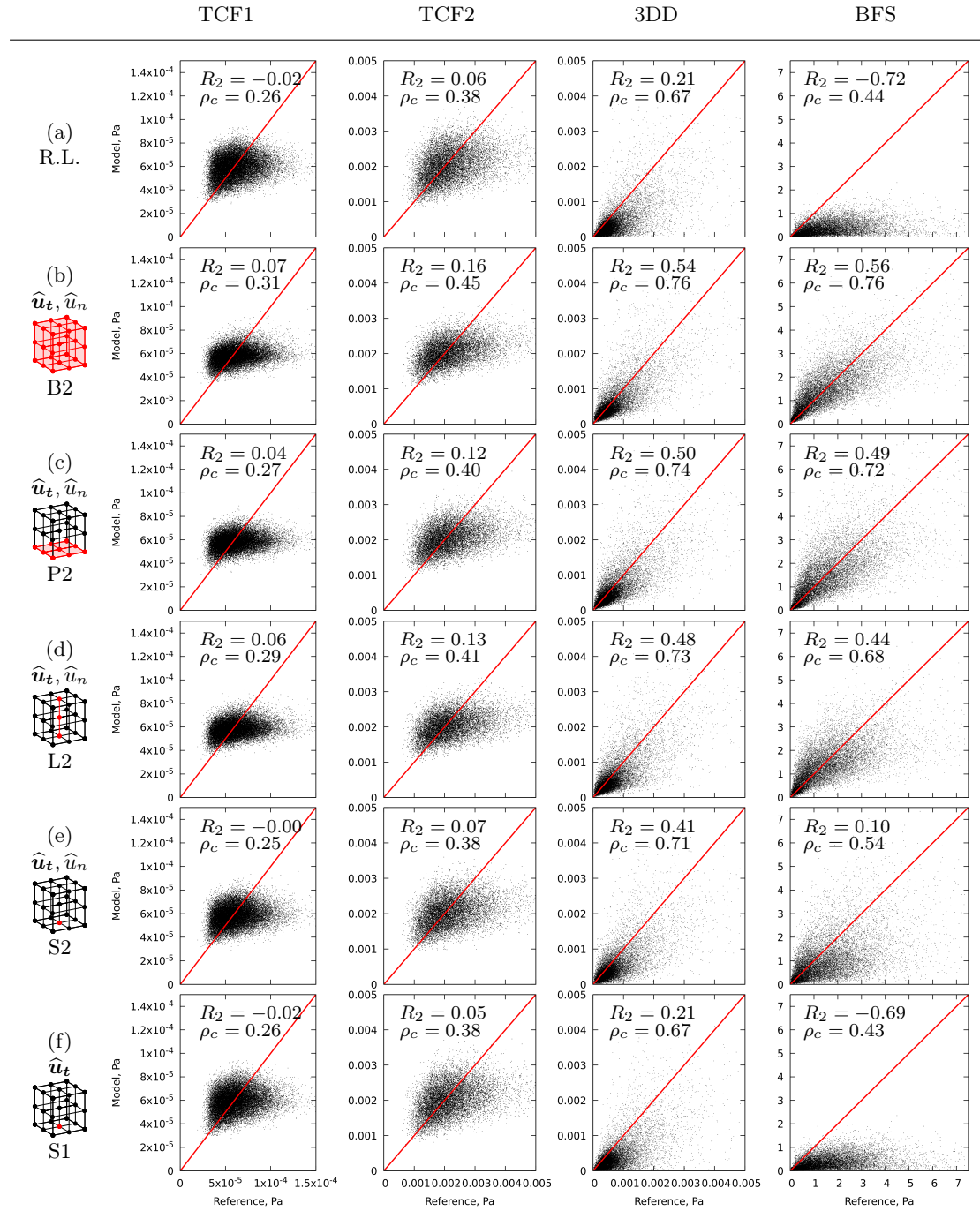


Figure 25: A priori validation: Scatter plot between the target wall shear stress and the prediction of a model based on Reichardt's law (R.L., a) and machine-learning models trained on the datasets TCF1, TCF2 and BFS with various input types (b)–(f) in the datasets TCF1, TCF2, 3DD and BFS. The red line is the identity. The coefficient of determination R_2 and the coefficient of correlation ρ_c between the model and the reference wall shear stress is also reported.

References

- [1] L. Agostini and P. Vincent. Channel flow data, GitHub repository, https://github.com/LionelAgo/channel_postprocessing, 2020.
- [2] H. J. Bae, A. Lozano-Durán, S. T. Bose, and P. Moin. Dynamic slip wall model for large-eddy simulation. *Journal of Fluid Mechanics*, 859:400, 2019.
- [3] J. S. Baggett. Some modeling requirements for wall models in large eddy simulation. *Annual Research Briefs*, 1997:265–275, 1997.
- [4] E. Balaras, C. Benocci, and U. Piomelli. Two-layer approximate boundary conditions for large-eddy simulations. *AIAA journal*, 34(6):1111–1119, 1996.
- [5] R. A. Baurle, C.-J. Tam, J. R. Edwards, and H. A. Hassan. Hybrid simulation approach for cavity flows: blending, algorithm, and boundary treatment issues. *AIAA journal*, 41(8):1463–1480, 2003.
- [6] I. Bermejo-Moreno, J. Larsson, L. Campo, J. Bodart, D. Helmer, F. Ham, and J. K. Eaton. Wall-modeled large-eddy simulations of shock/turbulent-boundary layer interaction in a duct. *Bulletin of the American Physical Society*, 57, 2012.
- [7] I. Bermejo-Moreno, L. Campo, J. Larsson, J. Bodart, D. Helmer, and J. K. Eaton. Confinement effects in shock wave/turbulent boundary layer interactions through wall-modelled large-eddy simulations. *Journal of Fluid Mechanics*, 758:5, 2014.
- [8] R. Bhaskaran, R. Kannan, B. Barr, and S. Priebe. Science-guided machine learning for wall-modeled large eddy simulation. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 1809–1816. IEEE, 2021.
- [9] S. Blanchard, N. Odier, L. Gicquel, B. Cuenot, and F. Nicoud. Stochastic forcing for sub-grid scale models in wall-modeled large-eddy simulation. *Physics of Fluids*, 33(9):095123, 2021.
- [10] S. T. Bose and P. Moin. A dynamic slip boundary condition for wall-modeled large-eddy simulation. *Physics of Fluids*, 26(1):015104, 2014.
- [11] S. T. Bose and G. I. Park. Wall-modeled large-eddy simulation for complex turbulent flows. *Annual Review of Fluid Mechanics*, 50:535–561, 2018.
- [12] S. L. Brunton, B. R. Noack, and P. Koumoutsakos. Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics*, 52:477–508, 2020.
- [13] W. Cabot and M. P. Moin. Approximate wall boundary conditions in the large-eddy simulation of high Reynolds number flow. *Flow, Turbulence and Combustion*, 63(1-4):269–291, 2000.
- [14] M. Catchirayer, J.-F. Boussuge, P. Sagaut, M. Montagnac, D. Papadogiannis, and X. Garnaud. Extended integral wall-model for large-eddy simulations of compressible wall-bounded turbulent flows. *Physics of Fluids*, 30(6):065106, 2018.
- [15] Z. L. Chen, S. Hickel, A. Devesa, J. Berland, and N. A. Adams. Wall modeling for implicit large-eddy simulation and immersed-interface methods. *Theoretical and Computational Fluid Dynamics*, 28:1–21, 2014.
- [16] E. M. Cherry, C. J. Elkins, and J. K. Eaton. Geometric sensitivity of three-dimensional separated flows. *International Journal of Heat and Fluid Flow*, 29(3):803–811, 2008.
- [17] H. Choi and P. Moin. Grid-point requirements for large eddy simulation: Chapman’s estimates revisited. *Physics of Fluids*, 24(1):011702, 2012.
- [18] D. Chung and D. I. Pullin. Large-eddy simulation and wall modelling of turbulent channel flow. *Journal of Fluid Mechanics*, 631:281–309, 2009.
- [19] D.-A. Clevert, T. Unterthiner, and S. Hochreiter. Fast and accurate deep network learning by exponential linear units (ELUs). In *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico*, 2016.
- [20] L. Davidson and S.-H. Peng. Hybrid LES-RANS modelling: a one-equation SGS model combined with $\text{ak-}\omega$ model for predicting recirculating flows. *International Journal for Numerical Methods in Fluids*, 43(9):1003–1018, 2003.
- [21] J. W. Deardorff. A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. *Journal of Fluid Mechanics*, 41(2):453–480, 1970.
- [22] S. Deck. Recent improvements in the zonal detached eddy simulation (zdes) formulation. *Theoretical and Computational Fluid Dynamics*, 26(6):523–550, 2012.
- [23] J. C. Del Álamo and J. Jiménez. Spectra of the very large anisotropic scales in turbulent channels. *Physics of Fluids*, 15(6):L41–L44, 2003.
- [24] J. Dombard, F. Duchaine, L. Gicquel, N. Odier, K. Leroy, N. Buffaz, S. Le-Guyader, J. Démolis, S. Richard, and T. Grosnickel. Evaluation of the capacity of rans/urans/les in predicting the performance of a high-pressure turbine: Effect of load and off design condition. In *Turbo Expo: Power for Land, Sea, and Air*, volume 84089, page V02CT35A041. American Society of Mechanical Engineers, 2020.

- [25] D. Dupuy, A. Toutant, and F. Bataille. Turbulence kinetic energy exchanges in flows with highly variable fluid properties. *Journal of Fluid Mechanics*, 834:5–54, 2018.
- [26] D. Dupuy, A. Toutant, and F. Bataille. Effect of the Reynolds number on turbulence kinetic energy exchanges in flows with highly variable fluid properties. *Physics of Fluids*, 31(1):015104, 2019.
- [27] D. Dupuy, L. Gicquel, N. Odier, F. Duchaine, and T. Arts. Analysis of the effect of intermittency in a high-pressure turbine blade. *Physics of Fluids*, 32(9):095101, 2020.
- [28] K. Duraisamy, G. Iaccarino, and H. Xiao. Turbulence modeling in the age of data. *Annual Review of Fluid Mechanics*, 51:357–377, 2019.
- [29] Ercoftac. Direct numerical simulation datasets, Ercoftac wiki, https://www.kbwiki.ercoftac.org/w/index.php/DNS_Index, 2022.
- [30] L. Fu, M. Karp, S. T. Bose, P. Moin, and J. Urzay. Equilibrium wall-modeled LES of shock-induced aerodynamic heating in hypersonic boundary layers. *Center for Turbulence Research Annual Research Briefs*, pages 171–181, 2018.
- [31] L. Fu, M. Karp, S. T. Bose, P. Moin, and J. Urzay. Shock-induced heating and transition to turbulence in a hypersonic boundary layer. *Journal of Fluid Mechanics*, 909, 2021.
- [32] N. Furuichi, Y. Terao, Y. Wada, and Y. Tsuji. Further experiments for mean velocity profile of pipe flow at high reynolds number. *Physics of Fluids*, 30(5):055101, 2018.
- [33] M. Gamahara and Y. Hattori. Searching for turbulence models by artificial neural network. *Physical Review Fluids*, 2(5):054604, 2017.
- [34] T. B. Gatski and J. P. Bonnet. *Compressibility, turbulence and high speed flow*. Academic Press, 2013.
- [35] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings, 2010.
- [36] Y. Han, Y. He, and J. Le. Modification to improved delayed detached-eddy simulation regarding the log-layer mismatch. *AIAA Journal*, 58(2):712–721, 2020.
- [37] S. Hoyas and J. Jiménez. Reynolds number effects on the Reynolds-stress budgets in turbulent channels. *Physics of Fluids*, 20(10):101511, 2008.
- [38] X. L. D. Huang, X. Yang, and R. Kunz. Wall-modeled large-eddy simulations of spanwise rotating turbulent channels — Comparing a physics-based approach and a data-based approach. *Physics of Fluids*, 31(12):125105, 2019.
- [39] N. Hutchins and I. Marusic. Evidence of very long meandering features in the logarithmic region of turbulent boundary layers. *Journal of Fluid Mechanics*, 579:1–28, 2007.
- [40] N. Hutchins and I. Marusic. Large-scale influences in near-wall turbulence. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1852):647–664, 2007.
- [41] H. T. Huynh. A flux reconstruction approach to high-order schemes including discontinuous Galerkin methods. In *18th AIAA Computational Fluid Dynamics Conference*, page 4079, 2007.
- [42] M. Inoue and D. I. Pullin. Large-eddy simulation of the zero-pressure-gradient turbulent boundary layer up to $Re_\theta = O(10^{12})$. *Journal of Fluid Mechanics*, 686:507–533, 2011.
- [43] J. Jiménez and C. Vasco. Approximate lateral boundary conditions for turbulent simulations. In *Proceedings of the Summer Program*, pages 399–412, 1998.
- [44] S. Jovic and D. M. Driver. Backward-facing step measurements at low reynolds number, $reh=5000$. *NASA STI/Recon Technical Report N*, 94:33290, 1994.
- [45] S. Kawai and J. Larsson. Wall-modeling in large eddy simulation: Length scales, grid resolution, and accuracy. *Physics of Fluids*, 24(1):015105, 2012.
- [46] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, USA*, 2015.
- [47] R. H. Kraichnan. Diffusion by a random velocity field. *Physics of Fluids*, 13(1):22–31, 1970.
- [48] J. Larsson, S. Kawai, J. Bodart, and I. Bermejo-Moreno. Large eddy simulation with modeled wall-stress: recent progress and future directions. *Mechanical Engineering Reviews*, 3(1):15–00418, 2016.
- [49] P. Lax and B. Wendroff. Systems of conservation laws. *Communications on Pure and Applied Mathematics*, 13(2):217–237, 1960.
- [50] H. Le, P. Moin, and J. Kim. Direct numerical simulation of turbulent flow over a backward-facing step. *Journal of Fluid Mechanics*, 330:349–374, 1997.
- [51] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [52] O. Lehmkuhl, G. Houzeaux, H. Owen, G. Chrysokentis, and I. Rodríguez. A low-dissipation finite element scheme for scale resolving simulations of turbulent flows. *Journal of Computational Physics*, 390:51–65, 2019.

- [53] A. Leonard. Energy cascade in large eddy simulations of turbulent fluid flows. *Advances in Geophysics*, 18A:237–248, 1974.
- [54] T. Leonard, M. Sanjose, S. Moreau, and F. Duchaine. Large eddy simulation of a scale-model turbofan for fan noise source diagnostic. In *22nd AIAA/CEAS Aeroacoustics Conference*, page 3000, 2016.
- [55] Q. Li, X. Chen, G. Wang, and Y. Liu. A dynamic version of the improved delayed detached-eddy simulation based on the differential reynolds-stress model. *Physics of Fluids*, 34(11):115112, 2022.
- [56] J. Ling, R. Jones, and J. Templeton. Machine learning strategies for systems with invariance properties. *Journal of Computational Physics*, 318:22–35, 2016.
- [57] J. Ling, A. Kurzawski, and J. Templeton. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, 807:155–166, 2016.
- [58] A. Lozano-Durán and H. J. Bae. Self-critical machine-learning wall-modeled LES for external aerodynamics. *Center for Turbulence Research Annual Research Briefs*, pages 197–210, 2020.
- [59] A. Lozano-Durán and H. J. Bae. Building-block-flow wall model for large-eddy simulation. *arXiv preprint arXiv:2211.07879*, 2022.
- [60] A. Lozano-Durán and J. Jiménez. Effect of the computational domain on direct simulations of turbulent channels up to $Re_\tau = 4200$. *Physics of Fluids*, 26(1):011702, 2014.
- [61] A. Lozano-Durán and J. Jiménez. Turbulent channel flow at $Re_\tau = 950$, DNS database of wall-bounded turbulent flows, http://hal.dmt.upm.es/raw_database/Channels/Re950/2pipi/, 2015.
- [62] I. Marusic, J. P. Monty, M. Hultmark, and A. J. Smits. On the logarithmic region in wall turbulence. *Journal of Fluid Mechanics*, 716, 2013.
- [63] R. Maulik and O. San. A neural network approach for the blind deconvolution of turbulent flows. *Journal of Fluid Mechanics*, 831:151–181, 2017.
- [64] R. Maulik, O. San, A. Rasheed, and P. Vedula. Data-driven deconvolution for large eddy simulations of kraichnan turbulence. *Physics of Fluids*, 30(12):125109, 2018.
- [65] N. Moriya, K. Fukami, Y. Nabae, M. Morimoto, T. Nakamura, and K. Fukagata. Inserting machine-learned virtual wall velocity for large-eddy simulation of turbulent channel flows. *arXiv preprint arXiv:2106.09271*, 2021.
- [66] R. D. Moser, J. Kim, and N. N. Mansour. Direct numerical simulation of turbulent channel flow up to $Re_\tau = 590$. *Physics of Fluids*, 11(4):943–945, 1999.
- [67] F. Nicoud, G. Winckelmans, D. Carati, J. Baggett, and W. Cabot. Boundary conditions for LES away from the wall. In *Proceedings of the Summer Program*, page 413. Citeseer, 1998.
- [68] F. Nicoud, H. Baya Toda, O. Cabrit, S. Bose, and J. Lee. Using singular values to build a subgrid-scale model for large eddy simulations. *Physics of Fluids*, 23(8):085106, 2011.
- [69] G. Novati, H. L. de Laroussilhe, and P. Koumoutsakos. Automating turbulence modelling by multi-agent reinforcement learning. *Nature Machine Intelligence*, pages 1–10, 2021.
- [70] N. Odier, A. Thacker, M. Harnieh, G. Staffelbach, L. Gicquel, F. Duchaine, N. G. Rosa, and J.-D. Müller. A mesh adaptation strategy for complex wall-modeled turbomachinery LES. *Computers & Fluids*, 214:104766, 2021.
- [71] S. Paolucci. On the filtering of sound from the Navier–Stokes equations. Technical Report SAND82-8257, Sandia National Lab., Livermore, CA, United States, 1982.
- [72] A. Patel, J. Peeters, B. Boersma, and R. Pecnik. Semi-local scaling and turbulence modulation in variable property turbulent channel flows. *Physics of Fluids*, 27(9):095101, 2015.
- [73] U. Piomelli. Wall-layer models for large-eddy simulations. *Progress in Aerospace Sciences*, 44(6):437–446, 2008.
- [74] U. Piomelli and E. Balaras. Wall-layer models for large-eddy simulations. *Annual Review of Fluid Mechanics*, 34(1):349–374, 2002.
- [75] U. Piomelli, J. Ferziger, P. Moin, and J. Kim. New approximate boundary conditions for large eddy simulations of wall-bounded flows. *Physics of Fluids A: Fluid Dynamics*, 1(6):1061–1068, 1989.
- [76] S. Pirozzoli, J. Romero, M. Fatica, R. Verzicco, and P. Orlandi. One-point statistics for turbulent pipe flow up to $Re_\tau \approx 6000$. *Journal of fluid mechanics*, 926, 2021.
- [77] T. J. Poinso and S. K. Lele. Boundary conditions for direct simulations of compressible viscous flows. *Journal of Computational Physics*, 101(1):104–129, 1992.
- [78] P. Pouech, F. Duchaine, and T. Poinso. Ignition of a premixed methane-air flow over a turbulent backward-facing step by direct numerical simulation. In *17th International Conference on Numerical Combustion*, Aachen, Germany, 2019.
- [79] P. Pouech, F. Duchaine, and T. Poinso. Premixed flame ignition in high-speed flows over a backward facing step. *Combustion and Flame*, 229:111398, 2021.
- [80] S. Radhakrishnan, U. Piomelli, A. Keating, and A. S. Lopes. Reynolds-averaged and large-eddy simulations of turbulent non-equilibrium flows. *Journal of Turbulence*, (7):N63, 2006.

- [81] H. Reichardt. Vollständige darstellung der turbulenten geschwindigkeitsverteilung in glatten leitungen. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 31(7):208–219, 1951.
- [82] F. Sarghini, G. De Felice, and S. Santini. Neural networks based subgrid scale modeling in large eddy simulations. *Computers & Fluids*, 32(1):97–108, 2003.
- [83] T. Schönfeld and M. Rudgyard. Steady and unsteady flow simulations using the hybrid flow solver AVBP. *AIAA journal*, 37(11):1378–1385, 1999.
- [84] U. Schumann. Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli. *Journal of Computational Physics*, 18(4):376–404, 1975.
- [85] B. Shi, Z. Xu, and S. Wang. A non-equilibrium slip wall model for large-eddy simulation with an immersed boundary method. *AIP Advances*, 12(9):095014, 2022.
- [86] M. L. Shur, P. R. Spalart, M. K. Strelets, and A. K. Travin. A hybrid RANS-LES approach with delayed-DES and wall-modelled LES capabilities. *International Journal of Heat and Fluid Flow*, 29(6):1638–1649, 2008.
- [87] A. J. Smits, B. J. McKeon, and I. Marusic. High-Reynolds number wall turbulence. *Annual Review of Fluid Mechanics*, 43, 2011.
- [88] J. Snoek, H. Larochelle, and R. P. Adams. Practical bayesian optimization of machine learning algorithms. In *Proceedings of the 25th International Conference on Neural Information Processing Systems-Volume 2*, pages 2951–2959, 2012.
- [89] P. R. Spalart. Direct simulation of a turbulent boundary layer up to $\text{Re}_\theta = 1410$. *Journal of Fluid Mechanics*, 187:61–98, 1988.
- [90] P. R. Spalart. Detached-eddy simulation. *Annual Review of Fluid Mechanics*, 41:181–202, 2009.
- [91] P. R. Spalart, S. Deck, M. L. Shur, K. D. Squires, M. Kh. Strelets, and A. Travin. A new version of detached-eddy simulation, resistant to ambiguous grid densities. *Theoretical and computational fluid dynamics*, 20:181–195, 2006.
- [92] D. R. Stull and H. Prophet. JANAF thermochemical tables, 2nd edition. Technical Report NSRDS-NBS 37, US National Bureau of Standards, 1971.
- [93] L. Temmerman, M. Hadžić, M. A. Leschziner, and K. Hanjalić. A hybrid two-layer URANS-LES approach for large eddy simulation at high reynolds numbers. *International Journal of Heat and Fluid Flow*, 26(2):173–190, 2005.
- [94] J. Vaquero, N. Renard, and S. Deck. Outer layer turbulence dynamics in a high-reynolds-number boundary layer up to recovering from mild separation. *Journal of Fluid Mechanics*, 942, 2022.
- [95] A. Volland, G. Balarac, and C. Corre. Subgrid-scale scalar flux modelling based on optimal estimation theory and machine-learning procedures. *Journal of Turbulence*, 18(9):854–878, 2017.
- [96] J.-X. Wang, J.-L. Wu, and H. Xiao. Physics-informed machine learning approach for reconstructing reynolds stress modeling discrepancies based on DNS data. *Physical Review Fluids*, 2(3):034603, 2017.
- [97] F. D. Witherden, A. M. Farrington, and P. E. Vincent. PyFR: An open source framework for solving advection–diffusion type problems on streaming architectures using the flux reconstruction approach. *Computer Physics Communications*, 185(11):3028–3040, 2014.
- [98] J.-L. Wu, H. Xiao, and E. Paterson. Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework. *Physical Review Fluids*, 3(7):074602, 2018.
- [99] C. Xie, J. Wang, H. Li, M. Wan, and S. Chen. Artificial neural network mixed model for large eddy simulation of compressible isotropic turbulence. *Physics of Fluids*, 31(8):085112, 2019.
- [100] X. I. A. Yang, J. Sadique, R. Mittal, and C. Meneveau. Integral wall model for large eddy simulations of wall-bounded turbulent flows. *Physics of Fluids*, 27(2):025112, 2015.
- [101] X. I. A. Yang, G. I. Park, and P. Moin. Log-layer mismatch and modeling of the fluctuating wall stress in wall-modeled large-eddy simulations. *Physical Review Fluids*, 2(10):104601, 2017.
- [102] X. I. A. Yang, S. Zafar, J.-X. Wang, and H. Xiao. Predictive large-eddy-simulation wall modeling via physics-informed neural networks. *Physical Review Fluids*, 4(3):034602, 2019.
- [103] R. Zangeneh. Data-driven model for improving wall-modeled large-eddy simulation of supersonic turbulent flows with separation. *Physics of Fluids*, 33(12):126103, 2021.
- [104] Z. Zhou, G. He, S. Wang, and G. Jin. Subgrid-scale model for large-eddy simulation of isotropic turbulent flows using an artificial neural network. *Computers & Fluids*, 195:104319, 2019.
- [105] Z. Zhou, G. He, and X. Yang. Wall model based on neural networks for LES of turbulent flows over periodic hills. *Physical Review Fluids*, 6(5):054610, 2021.