The Iranian Journal of Mechanical Engineering Transactions of ISME
Journal homepage: https://jmee.isme.ir/
Vol. 24, No. 2, 2023, pp. 157-172

Research Paper

DOI: https://doi.org/10.30506/jmee.2024.2015293.1330

Subgrid-scale Flux Modeling of a Passive Scalar in Turbulent Channel Flow using Artificial Neural Network

Amin Rasam *
Assistant Professor

Mehran Shirazi †
M.Sc. Student

A deep neural network (DNN) has been developed to model the subgrid-scale (SGS) flux associated with a passive scalar in incompressible turbulent channel flow. To construct the training dataset for the DNN, a direct numerical simulation (DNS) was performed for a channel flow at the friction Reynolds number $Re_\tau = 179$ *encompassing a passive scalar transport with Prandtl number* $Pr = 0.71$ *using a pseudo-spectral in-house code. The DNS data of velocity and scalar fields was filtered to obtain the SGS scalar flux vector,* q_i*, filtered scalar gradient, and filtered strain-rate tensor, which were subsequently used to train the DNN, enabling it to predict* q_i *for large-eddy simulation. A priori evaluation of the DNN's performance in predicting* q_i *revealed a close match with filtered DNS data, demonstrating correlations of up to 98%, 79% and 85% for the three components of* q_i*. Additionally, analysis of the mean SGS dissipation and its probability density function indicated promising predictions by the DNN. Notably, this study extends the applications of DNNs for predicting* q_i *to the case of turbulent channel flow.*

Keywords: Subgrid-scale scalar flux, Deep neural network, DNS, LES, Turbulent channel flow

1 Introduction

Direct numerical simulation (DNS) of turbulent flows, which resolves all temporal and spatial flow scales, demands a substantial number of grid points and is practically infeasible for numerical simulation of high Reynolds number flows [1]. Hence, turbulence modelling is often invoked for simulation of industrial turbulent flows. Two main modelling approaches are practically used, namely Reynolds-averaged Navier-Stokes (RANS) and large-eddy simulation

*Corresponding author, Assistant Professor, Faculty of Mechanical and Energy Engineering, Shahid Beheshti University, 16589-53571, Tehran, Iran, a_rasam@sbu.ac.ir, ORCID: 0000-0002-3173-7502
†MSc. Student, Faculty of Mechanical and Energy Engineering, Shahid Beheshti University,16589-53571, Tehran, Iran, mehr_shirazi@mail.sbu.ac.ir

Received: 2023/11/21, Revised: 2024/02/13, Accepted: 2024/03/06

(LES). In RANS, all turbulence is modelled and only statistics are predicted, whereas, only the small-scales need to be modelled in LES and the relevant large scales are computed. Hence, LES can also provide instantaneous turbulent flow fields and is computationally less expensive compared with the DNS. Modelling of subgrid-scale (SGS) quantities in LES is a challenging task, which has always been in the focus of research in LES [2].

Machine learning techniques have shown interesting capabilities for the extraction of data features and correlations for fluid mechanics applications [3]. They have also recently been employed for turbulence modeling applications, both in RANS [4] and LES [5]. In the context of LES, artificial neural-network (ANN) architectures have been employed for prediction of unknown SGS quantities from the known resolved ones. In LES, ANN-based architectures have been used for both direct modelling of the SGS stress and scalar flux and also computation of SGS model parameters as data-informed closure models. The work by Gamahara *et al.* [5], where an ANN was trained using filtered data obtained from DNS of turbulent channel flow to predict SGS stresses is an example of direct modelling of the SGS stresses. The influence of input parameters on the correlation of ANN-predictions with those of the filtered DNS data were examined using *a priori* tests for the plane channel flow. Dependence of ANN predictions on the input parameters was examined by Wang *et al.* [6], where filtered velocity components and their first and second derivatives were used.

Artificial neural networks have been utilized for predicting SGS scalar fluxes in turbulent flows involving passive scalar transport. Vollant *et al*. [7] used optimal estimation theory to identify the most accurate input parameters for training an ANN to predict SGS scalar flux vector in forced homogeneous isotropic turbulence (HIT).

Milani *et* al. [8] used data from filtered DNS of a jet in cross flow to train a deep neural network (DNN) incorporating embedded coordinate frame invariance. The DNN was utilized for predicting the sub-grid scale (SGS) scalar flux using a tensorial eddy-diffusivity model.

Bode *et al.* [9] used filtered DNS data from homogeneous isotropic turbulence to train an ANN for prediction of SGS scalar dissipation. Frezat *et al.* [10] used physical invariances, such as translational, rotational and Galilean invariance, to improve ANN prediction of SGS scalar flux vector in forced HIT. Akhavan-Safaei and Zayernouri [11] also used ANN for modeling of the SGS scalar flux using filtered DNS data of HIT with a uniform mean scalar gradient. They showed that their ANN outperforms traditional SGS models based on the gradient-diffusion hypothesis.

While numerous studies have tackled the prediction of SGS scalar flux vector using ANN, these applications have primarily focused on the forced HIT. In the current investigation, a DNN is trained for the prediction of the SGS scalar flux vector in turbulent channel flow with the transport of a scalar. The current test case is different from the previous studies in that it comprises an inhomogeneous direction and wall effects, which makes the prediction of SGS quantities more challenging.

The rest of the paper is organized as follows. In section (2), governing equations and definition of SGS scalar flux are described. Numerical method and simulation details of the DNS is given in section (3). Filtering of the training data set for the DNN and computation of the SGS scalar flux vector are described in section (4). Architecture of the employed ANN is introduced in section (5). Results and discussions are given in section (6), followed by concluding remarks in section (7).

2 Governing equations of a passive scalar in LES and definition of subgrid-scale flux

The LES equations are derived by filtering the governing equations of fluid motion and scalar transport. A general conserved passive scalar, θ, has the following convection diffusion governing equation [10].

$$\frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{t}} + \boldsymbol{u_j}\frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{x_j}} = \frac{\mathbf{1}}{\boldsymbol{RePr}}\frac{\partial^2 \boldsymbol{\theta}}{\partial \boldsymbol{x_j}\partial \boldsymbol{x_j}}, \tag{1}$$

where, $\boldsymbol{u_j}$ denotes the velocity vector, $\boldsymbol{Re}$ is the Reynolds number, $\boldsymbol{Pr}$ is the Prandtl number, $\boldsymbol{x_j}$ is the coordinate system and $\boldsymbol{t}$ denotes time. Summation convention over the repeated indices is implied in this equation and is used in the proceeding equations. It has to be mentioned that Eq. (1) is not related to and does not influence the Navier-Stokes equations, since $\boldsymbol{\theta}$ is a passive scalar. A passive scalar does not influence the fluid flow. Examples of a passive scalar are the concentration of a substance in the flow or temperature in a weakly heated flow [12]. Filtering the above equation leads to the following transport equation for LES

$$\begin{gathered}\frac{\partial \overline{\theta}}{\partial t} + \overline{u}_j\frac{\partial \overline{\theta}}{\partial x_j} = \frac{1}{RePr}\frac{\partial^2 \overline{\theta}}{\partial x_j \partial x_j} - \frac{\partial q_j}{\partial x_j}, \\ q_j = \overline{u_j\theta} - \overline{u_j}\,\overline{\theta}.\end{gathered} \tag{2}$$

In this context, $\overline{\cdot}$ represents a filtered quantity and $\boldsymbol{q_j}$ is the SGS scalar flux. This is an unknown vector quantity that requires modeling to close the equation for the scalar. Modelling the $\boldsymbol{q_j}$ is the subject of the current study.

Another quantity of interest, is the resolved scalar intensity, $\boldsymbol{K_\theta = \overline{\theta^2}/2}$, which has the following transport equation [2].

$$\begin{gathered}\frac{\partial \boldsymbol{K_\theta}}{\partial \boldsymbol{t}} + \overline{\boldsymbol{u}}_{\boldsymbol{j}}\frac{\partial \boldsymbol{K_\theta}}{\partial \boldsymbol{x_j}} = -\frac{\mathbf{1}}{\boldsymbol{RePr}}\frac{\partial \overline{\boldsymbol{\theta}}}{\partial \boldsymbol{x_j}}\frac{\partial \overline{\boldsymbol{\theta}}}{\partial \boldsymbol{x_j}} + \frac{\mathbf{1}}{\boldsymbol{RePr}}\frac{\partial^2 \boldsymbol{K_\theta}}{\partial \boldsymbol{x_j}\partial \boldsymbol{x_j}} - \frac{\partial}{\partial \boldsymbol{x_j}}\left(\boldsymbol{q_j}\overline{\boldsymbol{\theta}}\right) + \boldsymbol{\chi}, \\ \boldsymbol{\chi} = \boldsymbol{q_j}\frac{\partial \overline{\boldsymbol{\theta}}}{\partial \boldsymbol{x_j}}.\end{gathered} \tag{3}$$

Here, $\boldsymbol{\chi}$ is called the SGS scalar dissipation and is the rate at which resolved scalar intensity is transferred from the resolved to the SGSs. It depends on the $\boldsymbol{q_j}$ prediction and a successful LES relies on its correct prediction.

3 Numerical method and simulation

The SIMSON code [13], which is a pseudo-spectral method is used to perform a DNS of incompressible turbulent channel flow. The code is based on the method used in reference [14]. Transport equations are solved for the wall-normal velocity, $\boldsymbol{v}$, and vorticity, $\boldsymbol{\omega}$. These equations, expressed in non-dimensional form, are written as

$$\frac{\partial}{\partial t}\nabla^2 v = h_v + \frac{1}{Re}\nabla^4 v, \quad \frac{\partial \omega}{\partial t} = h_\omega + \frac{1}{Re}\nabla^2 \omega, \tag{4}$$

where,

$$h_v = -\frac{\partial}{\partial y}\left(\frac{\partial H_1}{\partial x} + \frac{\partial H_3}{\partial z}\right) + \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right)H_2, \;\; h_\omega = \frac{\partial H_1}{\partial z} - \frac{\partial H_3}{\partial x}, H_i = u_j\frac{\partial u_i}{\partial x_j}, i = 1,2,3. \tag{5}$$

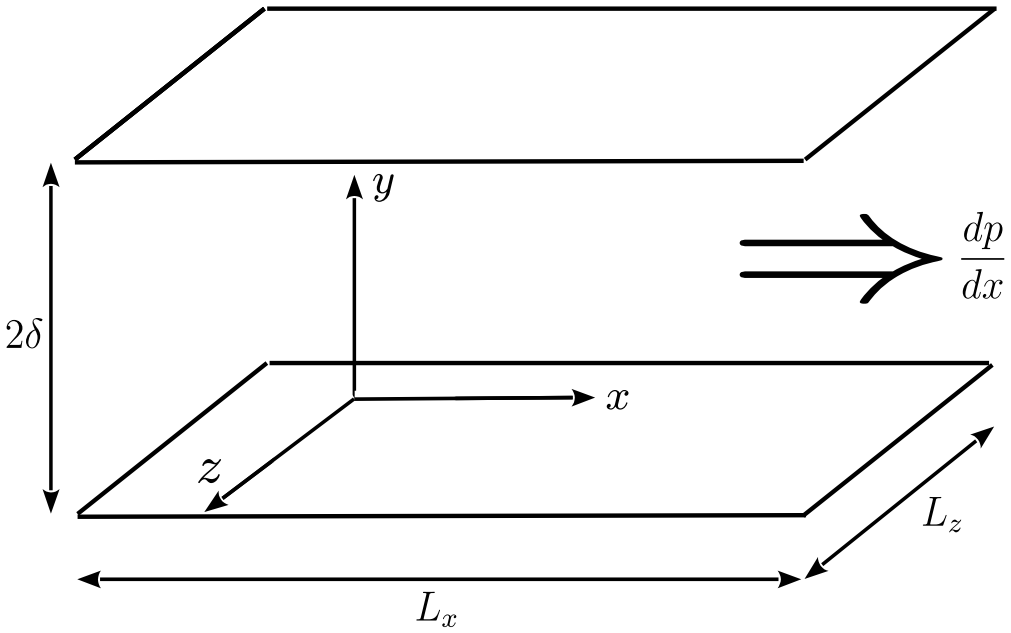


Figure 1 Geometry of the channel flow with a prescribed pressure gradient

The Reynolds number $Re_b = {u_b\delta}/{\nu}$ is based on the bulk velocity, u_b. Here, ν is the kinematic viscosity and δ is the channel half height. The channel flow schematic is illustrated in Figure (1). Fourier representation is used for the spatial discretization in the x and z directions. Chebyshev polynomials are employed for the discretization in wall-normal direction. Aliasing errors are removed using the 3/2-rule [15]. The time integration of the nonlinear terms is based on a third-order Runge-Kutta method, whereas the time discretization for the linear terms uses the Crank-Nicolson scheme. Additional information regarding the numerical approach can be seen in reference [13].

The domain size in the x, y and z directions are $L_x = 8\pi\delta$, $L_y = 2\delta$ and $L_z = 3\pi\delta$, respectively, see Figure (1). The corresponding number of grid points are $N_x = 384$, $N_y = 129$ and $N_z = 384$, leading to the following grid sizes in wall units:

$$\Delta x^+ = \frac{u_\tau \Delta \mathrm{x}}{\nu} = 11.81, \qquad \Delta y^+ = \frac{u_\tau \Delta \mathrm{y}}{\nu} = 0.05 \sim 4.4, \qquad \Delta z^+ = \frac{u_\tau \Delta \mathrm{z}}{\nu} = 4.4. \tag{6}$$

Here, Δx, $\Delta \mathrm{y}$ and Δz are the grid spacings and $u_\tau = \sqrt{\tau_{wall}/\rho}$ is the friction velocity, where τ_{wall} is the wall shear stress and ρ is the fluid density. No-slip boundary condition at the walls and periodic boundary condition in the horizontal directions are employed. Simulations are performed at $\mathrm{Re_b} = 2800$ resulting in the friction Reynolds number $\mathrm{Re_\tau} = u_\tau \delta/\nu \approx 179$. The transport equation (1) is also solved for a passive scalar, θ with $Pr = 0.71$. A constant θ boundary condition with $\theta = 0$ and 1 at the lower and upper walls, respectively, is employed.

In the proceeding sections, $\langle . \rangle$ denotes averaging of turbulence statistics in the homogeneous x and z directions. A quantity expressed in wall units, denoted by a $^+$ sign, as in Eq. (6), is non-dimensionalized with either u_τ and ν or the scalar gradient at the wall and the diffusion coefficient. Averaging is only performed over one velocity and scalar field, which was found to be adequate for the objectives of the current study. A verification of the employed numerical method, for DNS and LES of turbulent channel flows and boundary layers can be found in ref. [13, 16, 17]. A validation of the current DNS predictions of the mean velocity profile and turbulent kinetic energy is presented in the following.

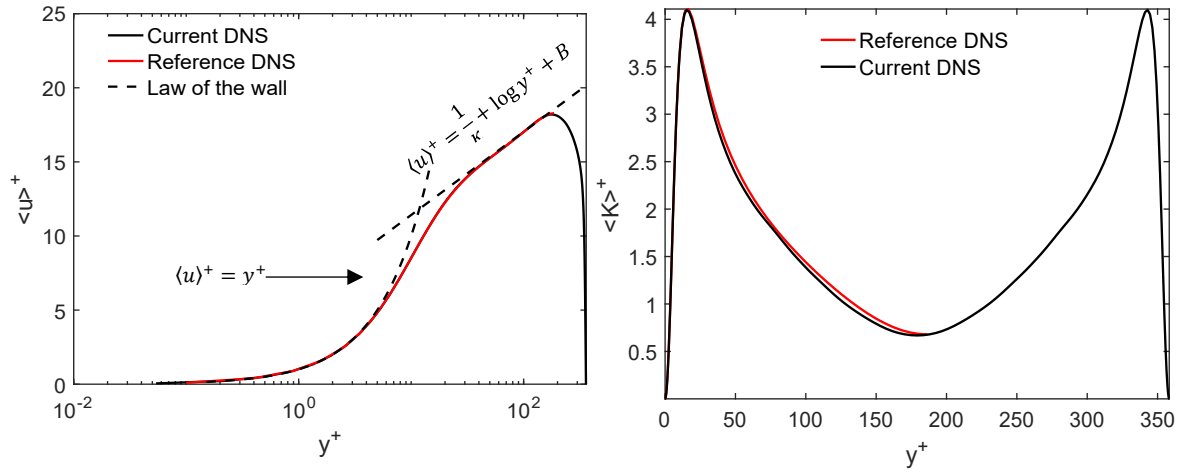


Figure 2 Mean streamwise velocity, $\langle u\rangle^+$, **(a)** and TKE, $\langle k\rangle^+$, **(b)** for the reference DNS (red solid line) and current DNS (black solid line), expressed in wall units. For clarity, the reference DNS [18] is only plotted for half of the channel height. The law of the wall for the logarithmic region is denoted by $\langle u\rangle^+ = \frac{1}{\kappa} log\, y^+ + B$, where $\kappa = 0.41$ and $B = 5.8$ and for the viscous sublayer is denoted by $\langle u\rangle^+ = y^+$.

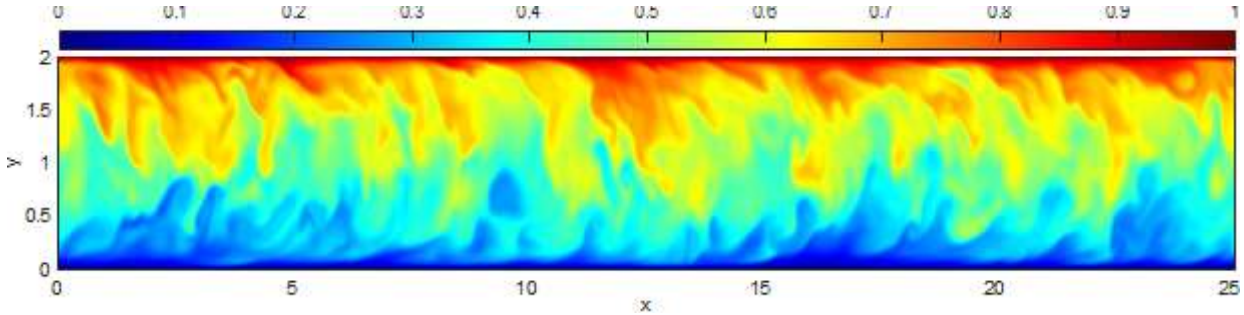
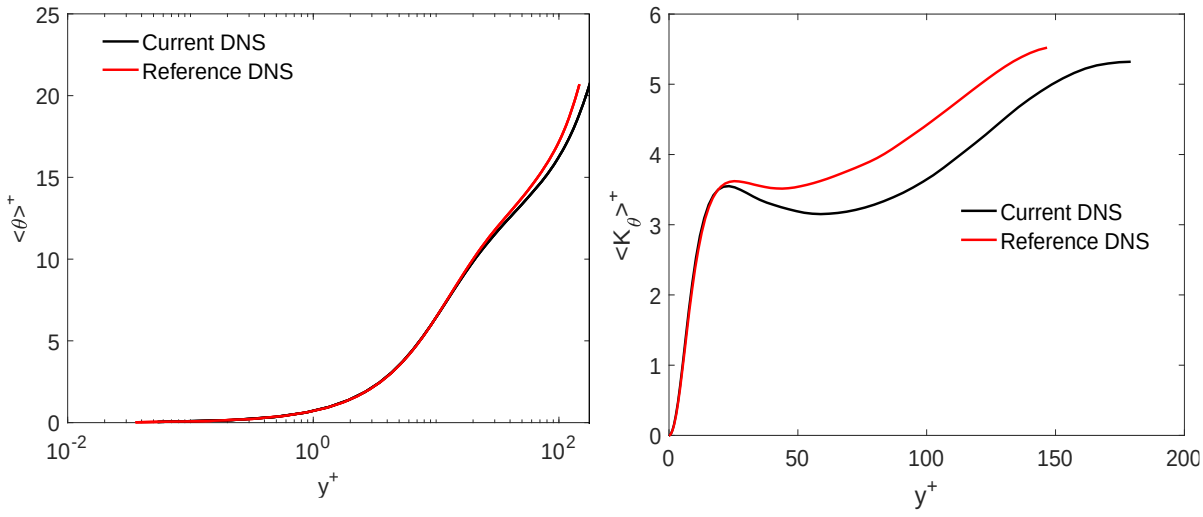


Figure 3 A snapshot of the instantaneous scalar field, θ, for the current DNS in an x-y plane.

A comparison of the mean velocity, $\langle u\rangle^+$, in wall units, between the current DNS prediction and the reference DNS prediction of Moser *et al.* [18], at the same Re_τ, and the law of the wall is given in Figure (2a). It is observed that the current DNS predictions of $\langle u\rangle^+$ are in excellent agreement with the reference DNS, across the whole channel height. The prediction also follows the linear velocity profile in the viscous sublayer, $y^+ < 5$, and the logarithmic law of the wall for $y^+ > 30$. The turbulent kinetic energy (TKE) in wall units, $\langle K\rangle^+$, is also presented in Figure (2b) is compared with the DNS prediction of Moser *et al.* [18], where also excellent agreement between the two DNSs are observed.

A snapshot of the instantaneous scalar, θ, in an $x - y$ plane is given in Figure (3), to provide an overview of the scalar transfer phenomena and its dynamics in the current DNS. Figure shows inclined instantaneous flow structures highlighted by the iso-contours of the scalar.

A comparison of the mean scalar in wall units, $\langle\theta\rangle^+$, and the mean scalar intensity in wall units, $\langle K_\theta\rangle^+$, is given in Figure (4a) and (4b) , between the current DNS and those of reference [19]. Since the reference DNS is at a different Reynolds number, i.e., $Re_\tau \approx 150$, the predictions are slightly different for $\langle\theta\rangle^+$ close to the channel center, see Figure (4a). Predictions of $\langle K_\theta\rangle^+$ are similar for $y^+ < 25$ for both DNSs, although each has a different Reynolds number. Predictions become significantly different for $y^+ > 25$, see Figure (4b). However, a similar trend between the current and reference predictions is observed in the two figures. It has to be pointed out that suitable DNS data for the current test case was not found for the scalar in the open literature. Hence, the current comparison for the scalar predictions with literature at a lower Reynolds number is only qualitative.



(a) **(b)**

Figure 4 Comparison of the mean scalar in wall units, $\langle\theta\rangle^+$, (a) and mean scalar intensity in wall units, $\langle K_\theta\rangle^+$, (b) between the current DNS at $Re_\tau \approx 179$ and reference DNS [2] at $Re_\tau \approx 150$.

4 Filtering of the training dataset and computation of subgrid-scale scalar flux

Filtering of the dataset is carried out using a top-hat (box) filter. The filtering operation employs the following convolution integral [2].

$$\overline{\boldsymbol{f}}(\boldsymbol{x}) = \int_{\boldsymbol{D}} \boldsymbol{f}(\boldsymbol{p})\boldsymbol{G}_{\overline{\boldsymbol{\Delta}}}(\boldsymbol{x}-\boldsymbol{p})\boldsymbol{dp}, \tag{7}$$

where, $\boldsymbol{f}$ is the function to be filtered, the kernel of the filter is dented by $\boldsymbol{G}_{\overline{\boldsymbol{\Delta}}}$, $\overline{\boldsymbol{\Delta}}$ is the size of the filter and $\boldsymbol{D}$ is the domain of integration. The employed filter has the following discrete formulation

$$\bar{\boldsymbol{f}}_{\boldsymbol{i}} = \frac{\boldsymbol{1}}{\boldsymbol{2n}}\left(\boldsymbol{f}_{\boldsymbol{i-n/2}} + \boldsymbol{f}_{\boldsymbol{i+n/2}}\right) + \frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i-n/2+1}}^{\boldsymbol{i+n/2-1}} \boldsymbol{f}_{\boldsymbol{i}}, \tag{8}$$

where, $n = \frac{\overline{\Delta}}{\Delta}$ is the filter to grid size ratio and f_i is a function of interest at a discrete position. The filters are applied in the horizontal plane (x-z), as is common in LES of turbulent channel flow with non-uniform grids in the wall-normal direction. Filtering in the wall-normal direction in physical space using the employed top hat filter should be avoided, since it violates the continuity equation, due to the non-uniform grid spacing in that direction [20-23].

The filter is applied to a single flow and scalar field of the DNS database, which comprise $Nx \times Ny \times Nz$ grid points, equal to $384 \times 129 \times 384 = 19{,}021{,}829$. The volume of the data is found to be sufficient for the purpose of the training of the DNN. It has to be pointed out that a sharp spectral filter has also been employed for preparation of the training data. But it led to lower correlations between the DNN predictions and the filtered DNS data, compared with the top hat filter. Hence, the top hat filter was chosen for this study. Furthermore, time filtering has not been performed in obtaining the resolved and SGS quantities in the current study. Finally, it has to be pointed out that the de-aliasing, performed in the DNS code does not affect the current filtering. In fact, the effect that de-aliasing has on LES in a pseudo-spectral

method is very limited, even at coarse resolutions, see e.g. [17]. Since only *a priori* analysis is performed in the current study, further investigation of this matter is not performed here.

To compute the SGS scalar flux, q_i, the filtered velocity $\overline{u}_i$, filtered scalar, $\overline{\theta}$, and the filtered product $\overline{u_i\theta}$ are first computed using the filtering operation in Eq. (8). Then, q_i is computed using Eq. (2). The velocity-gradient tensor $\partial u_i/\partial x_j$ and scalar-gradient vector $\partial\theta/\partial x_j$ are computed by differentiating the velocity and scalar fields in the Fourier space. The Fourier transformation is performed using the fast Fourier transform algorithm. Then, the resolved (filtered) velocity gradient, $\partial\overline{u}_i/\partial x_j$ and resolved scalar-gradient vector, $\partial\overline{\theta}/\partial x_j$ are computed using Eq. (8).

5 Architecture of the artificial neural network

Figure (5) illustrates the schematic of the employed artificial neural network (ANN). The ANN is constructed as a deep neural network (DNN), using the Keras API and the Tensor Flow [24]. It consists of input and output layers and two hidden ones, each with 128 neurons. Two activation functions, namely the Re Fiectified Linear Unit (ReLU) and Sigmoid [25] are tested for the hidden layers, whereas the output layer uses a linear activation function. The backpropagation algorithm together with Adam optimization algorithm [26] is used for the DNN. The selection of the number of hidden layers and neurons aligns with the recommendations outlined in the referenced literature [20]. Weights and biases of each layer is iteratively updated and corrected during the learning process. The loss function is formulated using mean absolute error (MAE), as expressed below:

$$MAE_{train} = \left(\frac{1}{m}\right)\sum_{i=1}^{m}\left|\left(Y_i - \widehat{Y}_i\right)\right|, \tag{9}$$

where m is the number of samples, Y_i is the DNN output value and $\widehat{Y}_i$ represents the exact value of the output.

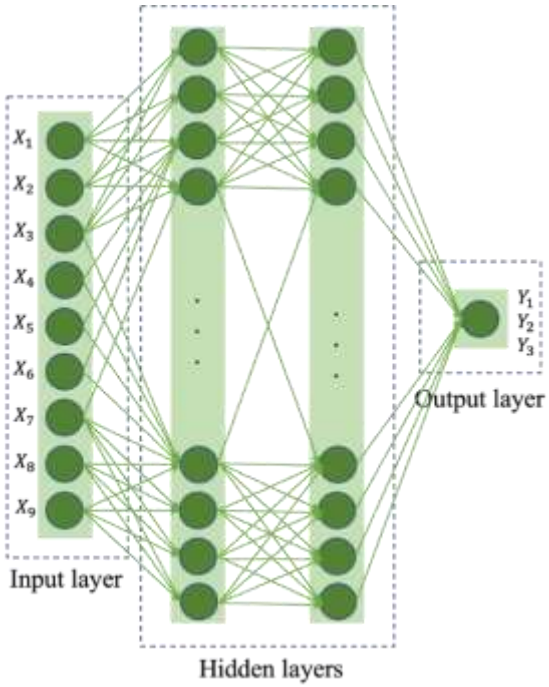


Figure 5 Schematic of the employed DNN

To train the DNN, DNS data is filtered, as explained in the previous section. Six different cases are considered, each provide different weights and biases. These cases are given in Table (1). The filtered DNS dataset is further divided to training (80%) and testing datasets (20%). The training dataset consists of 80% and the testing dataset includes 20% of the filtered DNS dataset, which are randomly chosen. The training and testing datasets undergo normalization to achieve zero mean and unit variance, a process known to enhance the convergence rate of the training [10]. The Typical loss curves for the training and testing data is provided in Figure (6).

The DNN has 9 inputs, which include the filtered strain-rate tensor, $\overline{S}_{ij}$, and scalar gradient vector $\partial\overline{\theta}/\partial x_j$ as

$$\mathrm{X}_{1-9} = \left\{\overline{S}_{11}, \overline{S}_{12}, \overline{S}_{13}, \overline{S}_{22}, \overline{S}_{23}, \overline{S}_{33}, \frac{\partial\overline{\theta}}{\partial x_1}, \frac{\partial\overline{\theta}}{\partial x_2}, \frac{\partial\overline{\theta}}{\partial x_3}\right\} \tag{10}$$

where, $\overline{S}_{ij}$ is a symmetric tensor defined as

$$\overline{\mathrm{S}}_{ij} = \frac{1}{2}\left(\frac{\partial\overline{u}_i}{\partial x_j} + \frac{\partial\overline{u}_j}{\partial x_i}\right). \tag{11}$$

Here, velocity gradients are computed in Fourier space. This choice of input parameters is based on the findings of Vollant *et al.* [7] for DNN predictions of SGS scalar flux for the case of HIT. To assess the accuracy of the DNN-predicted q_i, denoted by q_i^{DNN}, correlation coefficients CC_i between q_i^{DNN} and the SGS scalar flux vector obtained from the filtered DNS data, denoted by q_i^{DNS}, are computed using the following formulation

$$CC_i = \frac{\langle\left(q_i^{DNS} - \langle q_i^{DNS}\rangle\right)\left(q_i^{DNN} - \langle q_i^{DNN}\rangle\right)\rangle}{\left[\langle\left(q_i^{DNS} - \langle q_i^{DNS}\rangle\right)^2\rangle\right]^{0.5}\left[\langle\left(q_i^{DNN} - \langle q_i^{DNN}\rangle\right)^2\rangle\right]^{0.5}}, \quad i = 1,2,3 \tag{12}$$

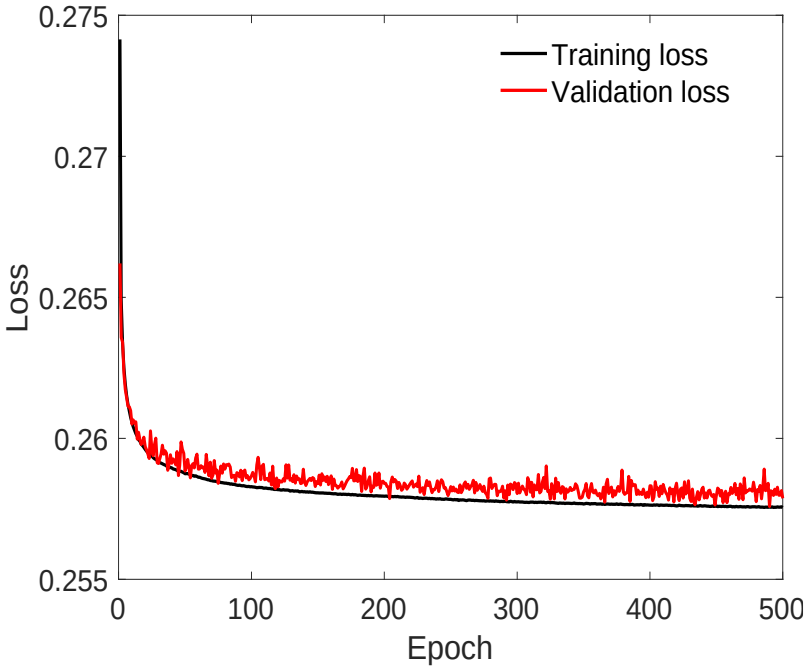


Figure 6 Training and testing loss curves for the current DNN of case D4, see Table (1).

Table 1 Effect of the DNN configuration, filter size and type on the correlation coefficient between q_i^{DNN} and q_i^{DNS}, i.e., CC_1, CC_2 and CC_3.

case	Number of hidden layers	CC_1	CC_2	CC_3	Activation functions	Number of neurons	Filter size	$\left(\frac{K_{SGS}}{K}\right)$ %
D1	1	0.9200	0.6337	0.6710	ReLu	128	8Δ	26.08
D2	2	0.9256	0.6668	0.7004				
D3	3	0.9273	0.6772	0.7082				
D4	2	0.9231	0.6552	0.6905	Sigmoid		4Δ	12.11
D5	2	0.9683	0.7463	0.8031	ReLu		4Δ	12.11
D6	2	0.9871	0.7937	0.8509	ReLu		2Δ	5.23

Obviously, a higher correlation coefficient indicates a q_i^{DNN} that more closely follows the q_i^{DNS} predicted by the filtered DNS data. The predicted CC_i are provided in Table (1). It can be observed that CC_1 is always higher than 0.92, whereas CC_2 and CC_3 are always larger than 0.69. This observation indicates that prediction of q_1, which is an order of magnitude larger than q_2, is easier for the DNN. In the current simulations, the mean SGS scalar flux in the spanwise direction, $\langle q_3 \rangle$, is zero. This is due to the absence of the mean scalar transport and scalar gradient in the z direction. Hence, $\langle q_3 \rangle$ is also not presented and discussed hereafter. However, the CC_3 can be computed due to the non-zero q_3 fluctuations and are presented in Table (1).

To quantify the effect of the filter size on the SGS content, the ratio of the SGS TKE to the total TKE, i.e., K_{SGS}/K, is given in percentage in Table (1). It can be observed that K_{SGS}/K % lies between 5.23 and 26.08 percent for $2\Delta \leq n \leq 8\Delta$.

The filter size ratio, n has a significant impact on the performance of the DNN. Comparing cases D2, D5, and D6 with $n = 8, 4$ and 2, respectively, which have the same DNN configuration but the filter size varies, one can find out that the highest correlation coefficients are obtained for the lowest n, showing that a closer relation between the inputs and outputs of the DNN exist at lower filter size ratios. However, since $n = 4$ is closer to practical LES resolutions [17], it is chosen for the final training of the DNN and the results presented in the next section.

Exploring the impact of the number of hidden layers on the predicted SGS scalar flux in cases D1, D2, and D3 reveals that increasing the number of hidden layers, from one in case D1 to two in case D2, and subsequently to three in case D3, results in an augmentation of the correlation coefficient CC_i. However, a fair prediction of q_i can be obtained with two hidden layers, which is also in correspondence with the findings of ref. [20] for the prediction of SGS stresses in a similar test case using DNN.

A comparison between the ReLU and Sigmoid activation functions in cases D4 and D5 shows that the ReLU function for the hidden layers gives higher correlation values. Hence, it is chosen for the forthcoming DNN predictions.

With these considerations, the DNN in case D5 with two hidden layers, ReLU activation parameter and filter size ratio $\boldsymbol{n = 4}$ is used for the prediction of SGS scalar fluxes in the next section. A flow chart of the whole process of the SGS scalar flux computations is also given in Figure (7).

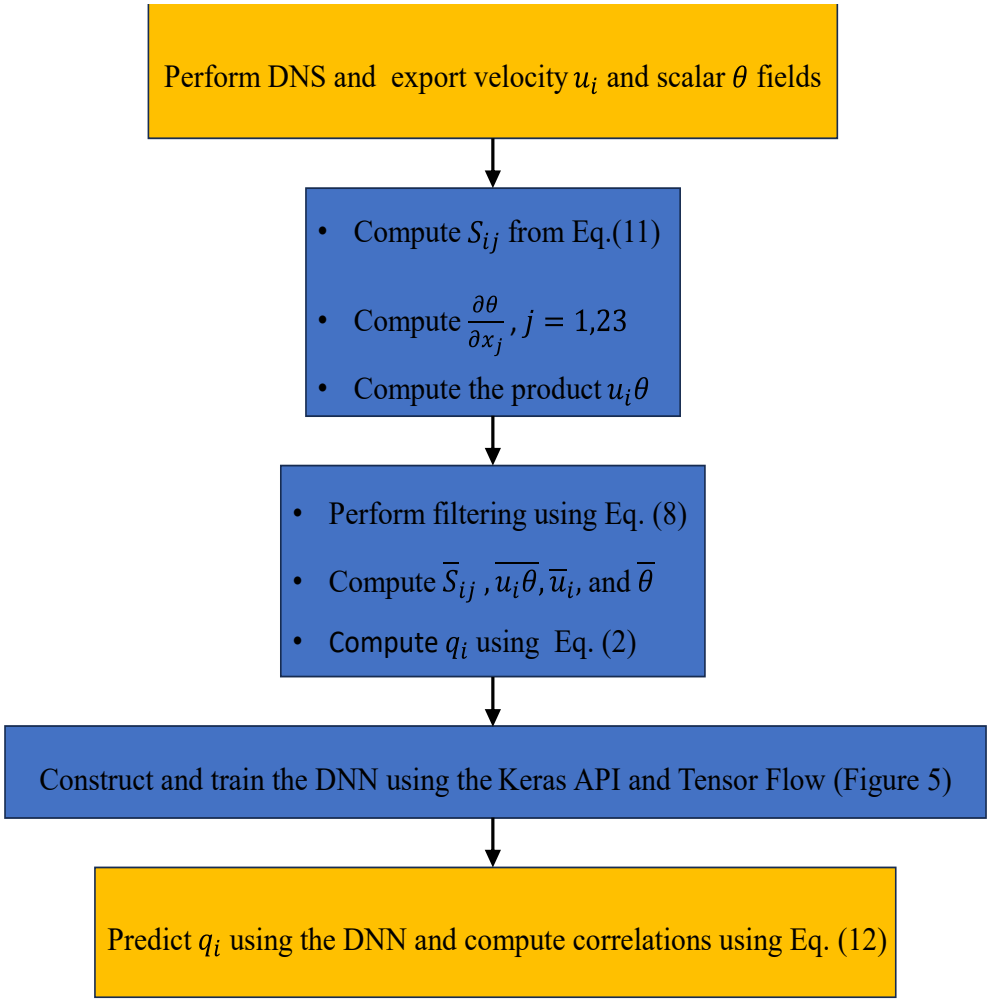


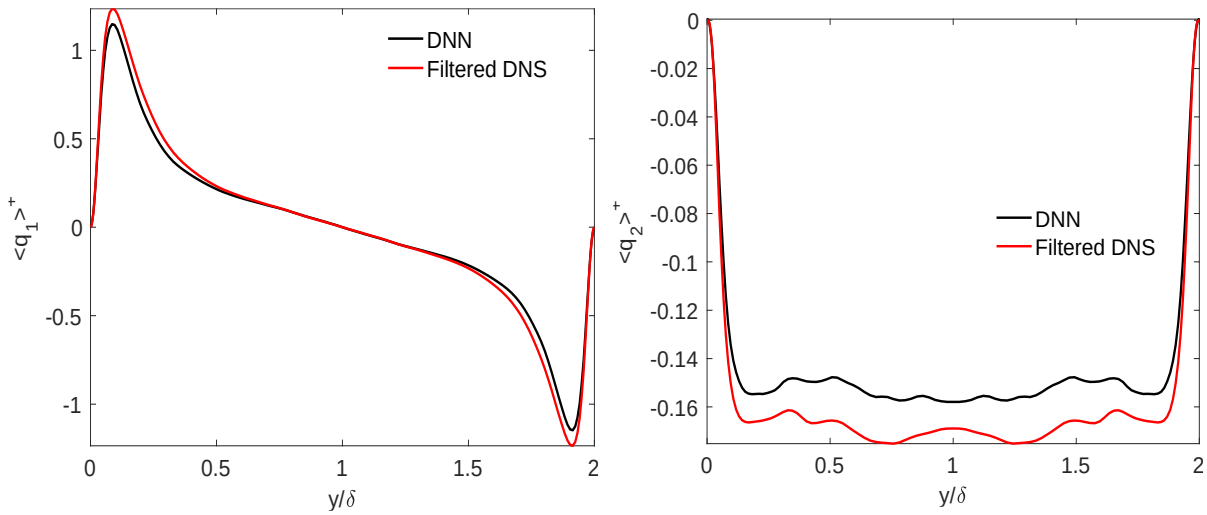
Figure 7 Flow chart showing the steps taken for the current computations

6 Results and discussions

Figure (8b) presents the mean SGS scalar fluxes in the streamwise and wall-normal directions, denoted as $\langle q_1 \rangle^+$ and $\langle q_2 \rangle^+$, respectively, predicted by the DNN and obtained from the filtered DNS data. Notably, $\langle q_1 \rangle^+$ exhibits a peak at the vicinity of the wall, demonstrating the turbulence transport of the scalar in this area. In contrast to $\langle q_2 \rangle^+$, $\langle q_1 \rangle^+$ cannot be modeled solely using the scalar gradient vector. Hence, SGS models relying on the well-known gradient-diffusion hypothesis fail to predict this component [16]. As an example of such models, recall the eddy-diffusivity model with the following formulation

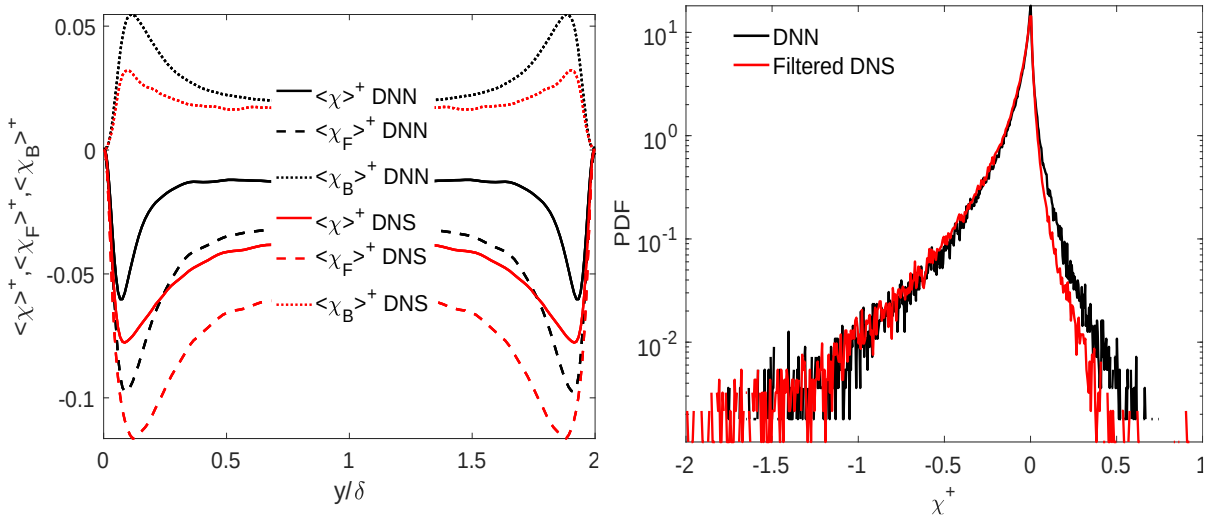
$$q_i = -\frac{\nu_t}{\mathrm{Pr_t}} \frac{\partial \overline{\theta}}{\partial x_i} \tag{13}$$

where, ν_t and Pr_t are the SGS eddy viscosity and SGS turbulent Prandtl number, respectively, which are model parameters. Since the mean scalar gradient in the streamwise direction, i.e., $\langle \partial\overline{\theta}/\partial x_1 \rangle = 0$, is zero in the current turbulent channel flow, such models predict $\langle q\_1 \rangle = 0$, whereas it has a significant value at the vicinity of the wall, as was already described. In contrast to such models, the DNN performs well in prediction of $\langle q_1 \rangle^+$, with a close agreement with the filtered DNS data with a small under-prediction of the near-wall peak value, see Figure (8a) and (8b). The DNN also gives a reasonable prediction of $\langle q_2 \rangle^+$ in Figure (8b).



(a) **(b)**

Figure 8 Comparison of the mean streamwise $\langle q_1 \rangle^+$(a) and wall-normal $\langle q_2 \rangle^+$ (b) SGS scalar fluxes in wall units, between the DNN predictions and filtered DNS data.



(a) **(b)**

Figure 9 A comparison of the mean SGS dissipation of scalar in wall units, $\langle \chi \rangle^+$(a) and its probability distribution function (PDF) at $y^+ \approx 12$ (b), between the DNN predictions and filtered DNS data.

There is a close correspondence for $\frac{y}{\delta} < 0.2$, between the DNN prediction and filtered DNS data, whereas the DNN underpredicts $\langle q_2 \rangle^+$ by approximately 13% in the middle of the channel. A quantity of major importance in SGS modeling of the passive scalar, is the SGS dissipation, χ, of the scalar intensity, K_θ, explained in section (2), where the governing equation for $K_\theta$was given in Eq. (3). The mean SGS scalar dissipation in wall units, $\langle \chi \rangle^+$ is given in Figure (9a) and (9b) which compares the DNN prediction with the data obtained from the filtered DNS scalar filed. Although $\langle \chi \rangle$ is always negative in wall-bounded flows, hence, $\langle K_\theta \rangle$ is transferred to the SGS scales from the resolved ones by the action of χ, it can instantaneously become positive. Obviously, K_θ is transferred from the SGS scales to the resolved ones, when χ is positive, which is a well-known phenomenon.

Therefore, K_θ can also be transferred instantaneously from the SGSs to the resolved scales. A negative χ is termed forward scatter of K_θ, denoted by χ_F and a positive χ is called backscatter of K_θ, denoted by χ_B. To find out the performance of the DNN in the prediction of SGS scalar dissipation, a comparison between the DNN predictions of χ, χ_F and χ_B with filtered DNS

values is given in Figure (9) for the mean values. One should note that the following relation between the SGS dissipation and its forward- and backscatter holds.

$$\chi = \chi_F + \chi_B \tag{14}$$

A peak in $\langle\chi\rangle^+$ exists close to the wall, where K_θ also has a peak (Figure not shown). One can observe that the DNN over-predicts both χ_F and χ_B and under-predicts χ, compared with the filtered DNS data, for $\frac{y}{\delta} < 0.5$. However, its prediction of χ_F reasonably follows χ of the filtered DNS. It has to be mentioned that backscatter, which is the transfer of K_θ from the SGS to the resolved scales, may lead to instability in the numerical simulations. Hence, it would be appropriate to limit the backscatter, i.e., χ_B , when the DNN is used to perform an LES. By doing that, reasonable χ predictions will be obtained, which is necessary for a successful LES. It has to be pointed out that most SGS models do not take into account the SGS backscatter. Recently, the importance of including the SGS backscatter of energy has been discussed for the SGS stress models for LES of HIT, which shows that inclusion of backscatter in SGS models may not be necessary [27].

The probability density function (PDF) of χ is also presented in Figure (9b) at $y^+ \approx 12$ to comprehend the discussion. This $y^+$is chosen at the vicinity of the peak of K_θ in the buffer layer, where the peak of turbulence activity lies. Hence, it is an important location for the analysis of SGS models. The PDF of χ is skewed towards negative values, as is expected [28], showing the tendency of negative SGS dissipation predictions and that the mean SGS dissipation is indeed negative. Figure shows that the PDF of the DNN-predicted χ is reasonably close to that of the filtered DNS. The positive tail is over-predicted by the DNN, which is in agreement with the over-prediction of χ_B by the DNN, presented earlier.

It has to be pointed out that the noisy character of the tails of the PDF are due to the number of data points used to compute the PDF in Figure (9b). The PDF tails show the probability of more rare events, compared with the central part of the PDF, hence, less samples are available for plotting the PDF tails. Similar plots of the PDF of χ can be found in the literature, see e.g. [11].

7 Concluding remarks

A DNN was employed for the prediction of the SGS scalar flux vector for LES of turbulent channel flow. The test case is an important canonical test case for SGS model analysis. A DNS was performed at $Re_\tau \approx 179$ with a scalar transport with $Pr = 0.71$ using a highly accurate, very low dissipation pseudo-spectral method. The DNS data base was further filtered to obtain the data needed for training the DNN to predict the SGS scalar flux vector, q_i. The DNN was trained using over 19 million data points. The trained DNN was able to predict q_i with correlations of its individual components with the filtered DNS data reaching up to 92%. The highest correlation coefficient was observed for the streamwise SGS scalar flux, $\langle q_1 \rangle$. The DNN was also found to give excellent prediction of $\langle q_1 \rangle$, a component that cannot be predicted by eddy-diffusivity SGS models. Reasonable prediction of the wall-normal SGS scalar flux, $\langle q_2 \rangle$, compared with the filtered DNS data, was also observed, especially at the vicinity of the walls, where the peak of turbulence activity is located.

A detailed analysis of the SGS dissipation of scalar intensity and its mean forward and backscatter was also provided and the results were compared with those of the filtered DNS data. This is a quantity of major interest in assessment of SGS model performance. The analysis was carried out both for the mean values and probability density function (PDF). It was observed that the DNN underpredicted the mean SGS dissipation, but its forward scatter was in reasonable agreement with the mean SGS dissipation obtained from filtered DNS data. Hence,

backscatter needs to be restricted, if an LES with the current DNN model is performed, as is the case for most other SGS models. The PDF of the SGS dissipation was discussed at a certain wall distance, in the buffer layer, which exhibited the expected skewed character and confirmed the findings observed for the mean SGS dissipation.

References

[1] S. A. Orszag, "Analytical Theories of Turbulence," *Journal of Fluid Mechanics,* Vol. 41, No. 2, pp. 363-386, 1970, doi: https://doi.org/10.1017/S0022112070000642.

[2] P. Sagaut, *Large Eddy Simulation for Incompressible Flows: An Introduction*. Springer Science & Business Media, 2005, https://doi.org/10.1007/b137536.

[3] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine Learning for Fluid Mechanics," *Annual review of fluid mechanics,* Vol. 52, pp. 477-508, 2020, doi: https://doi.org/10.1146/annurev-fluid-010719-060214.

[4] K. Duraisamy, G. Iaccarino, and H. Xiao, "Turbulence Modeling in the Age of Data," *Annual review of fluid mechanics,* Vol. 51, pp. 357-377, 2019, doi: https://doi.org/10.1146/annurev-fluid-010518-040547.

[5] M. Gamahara and Y. Hattori, "Searching for Turbulence Models by Artificial Neural Network," *Physical Review Fluids,* Vol. 2, No. 5, p. 054604, 2017, doi: https://doi.org/10.48550/arXiv.1607.01042.

[6] Z. Wang, K. Luo, D. Li, J. Tan, and J. Fan, "Investigations of Data-Driven Closure for Subgrid-Scale Stress in Large-Eddy Simulation," *Physics of Fluids,* Vol. 30, No. 12, 2018, doi: https://doi.org/10.1063/1.5054835.

[7] A. Vollant, G. Balarac, and C. Corre, "Subgrid-Scale Scalar Flux Modelling Based on Optimal Estimation Theory and Machine-Learning Procedures," *Journal of Turbulence,* Vol. 18, No. 9, pp. 854-878, 2017, doi: https://doi.org/10.1080/14685248.2017.1334907.

[8] P. M. Milani, J. Ling, and J. K. Eaton, "Turbulent Scalar Flux in Inclined Jets in Crossflow: Counter Gradient Transport and Deep Learning Modelling," *Journal of Fluid Mechanics,* Vol. 906, p. A27, 2021, doi: https://doi.org/10.1017/jfm.2020.820.

[9] M. Bode, M. Gauding, K. Kleinheinz, and H. Pitsch, "Deep Learning at Scale for Subgrid Modeling in Turbulent Flows: Regression and Reconstruction," in *International Conference on High Performance Computing*, 2019: Springer, pp. 541-560, doi: https://doi.org/10.1007/978-3-030-34356-9_50.

[10] H. Frezat, G. Balarac, J. Le Sommer, R. Fablet, and R. Lguensat, "Physical Invariance in Neural Networks for Subgrid-Scale Scalar Flux Modeling," *Physical Review Fluids,* Vol. 6, No. 2, p. 024607, 2021, doi: https://doi.org/10.1103/PhysRevFluids.6.024607.

[11] A. Akhavan-Safaei and M. Zayernouri, "Deep Learning of Subgrid-Scale Dynamics in Scalar Turbulence and Generalization to Other Transport Regimes," *Journal of Machine Learning for Modeling and Computing,* Vol. 5, No. 1, 2024, doi: https://doi.org/10.1615/JMachLearnModelComput.2023048824.

[12] Z. Warhaft, "Passive Scalars in Turbulent Flows," *Annual Review of Fluid Mechanics,* Vol. 32, No. 1, pp. 203-240, 2000, doi: https://doi.org/10.1146/annurev.fluid.32.1.203.

[13] M. Chevalier, P. Schlatter, A. Lundbladh, and D. S. Henningson, *Simson: A Pseudo-Spectral Solver for Incompressible Boundary Layer Flows*. 2007. [Online]. Available: https://www.kth.se/social/files/5b2654fb56be5b702282cec3/chevalier_schlatter_lundbladh_henningson_2007.pdfhttps://www.kth.se/social/files/5b2654fb56be5b702282cec3/chevalier_schlatter_lundbladh_henningson_2007.pdf.

[14] J. Kim, P. Moin, and R. Moser, "Turbulence Statistics in Fully Developed Channel Flow at Low Reynolds Number," *Journal of fluid mechanics,* Vol. 177, pp. 133-166, 1987, doi: https://doi.org/10.1017/S0022112087000892.

[15] M. Y. Hussaini and T. A. Zang, "Spectral Methods in Fluid Dynamics," *Annual review of fluid mechanics,* Vol. 19, No. 1, pp. 339-367, 1987.

[16] A. Rasam, G. Brethouwer, and A. V. Johansson, "An Explicit Algebraic Model for the Subgrid-Scale Passive Scalar Flux," *Journal of Fluid Mechanics,* Vol. 721, pp. 541-577, 2013, doi: https://doi.org/10.1017/jfm.2013.81.

[17] A. Rasam, G. Brethouwer, P. Schlatter, Q. Li, and A. V. Johansson, "Effects of Modelling, Resolution and Anisotropy of Subgrid-Scales on Large Eddy Simulations of Channel Flow," *Journal of turbulence,* No. 12, p. N10, 2011, doi: https://doi.org/10.1080/14685248.2010.541920.

[18] R. D. Moser, J. Kim, and N. N. Mansour, "Direct Numerical Simulation of Turbulent Channel Flow up to Re T= 590," *Physics of fluids,* Vol. 11, No. 4, pp. 943-945, 1999, doi: https://doi.org/10.1063/1.869966.

[19] N. Kasagi and O. Iida, "Progress in Direct Numerical Simulation of Turbulent Heat Transfer," *Proceedings of the 5th ASME/JSME Joint Thermal Engineering Conference,* pp. 15-19, 1999, doi: https://doi.org/10.1016/S0142-727X(97)10017-0.

[20] J. Park and H. Choi, "Toward Neural-Network-Based Large Eddy Simulation: Application to Turbulent Channel Flow," *Journal of Fluid Mechanics,* Vol. 914, p. A16, 2021, doi: https://doi.org/10.1017/jfm.2020.931.

[21] P. C. Di Leoni, T. A. Zaki, G. Karniadakis, and C. Meneveau, "Two-Point Stress–Strain-Rate Correlation Structure and Non-Local Eddy Viscosity in Turbulent Flows," *Journal of Fluid Mechanics,* vol. 914, p. A6, 2021, doi: https://doi.org/10.1017/jfm.2020.977.

[22] N. Park, S. Lee, J. Lee, and H. Choi, "A Dynamic Subgrid-Scale Eddy Viscosity Model with a Global Model Coefficient," *Physics of Fluids,* Vol. 18, No. 12, 2006, doi: https://doi.org/10.1063/1.2401626.

[23] S. Völker, R. D. Moser, and P. Venugopal, "Optimal Large Eddy Simulation of Turbulent Channel Flow Based on Direct Numerical Simulation Statistical Data," *Physics of Fluids,* Vol. 14, No. 10, pp. 3675-3691, 2002, doi: https://doi.org/10.1063/1.1503803.

[24] M. Abadi *et al.*, "Tensorflow}: A System for {Large-Scale} Machine Learning," in *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 2016, pp. 265-283, doi: https://doi.org/10.48550/arXiv.1605.08695.

[25] S. Haykin, *Neural Networks and Learning Machines, 3/E*. Pearson Education India, 2009. [Online]. Available: https://dai.fmph.uniba.sk/courses/NN/haykin.neural-networks.3ed.2009.pdfhttps://dai.fmph.uniba.sk/courses/NN/haykin.neural-networks.3ed.2009.pdf.

[26] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv preprint arXiv:1412.6980,* 2014, doi: https://doi.org/10.48550/arXiv.1412.6980.

[27] A. Vela-Martín, "Subgrid-Scale Models of Isotropic Turbulence Need Not Produce Energy Backscatter," *Journal of Fluid Mechanics,* Vol. 937, p. A14, 2022, doi: https://doi.org/10.1017/jfm.2022.123.

[28] A. Rasam, G. Brethouwer, and A. Johansson, "A Stochastic Extension of the Explicit Algebraic Subgrid-Scale Models," *Physics of fluids,* Vol. 26, No. 5, 2014, doi: https://doi.org/10.1063/1.4879436.

Nomenclature

English Symbol

CC_i	Correlation coefficient
K	Resolved turbulent kinetic energy
K_θ	Resolved scalar intensity
n	Filter size ratio
N	Number of grid points
Pr	Prandtl number
q_i	Subgrid-scale scalar flux vector
Re_b	Bulk Reynolds number
Re_τ	Friction Reynolds number
t	Time
S	Strain-rate tensor
u_b	Bulk velocity
u_i	Instantaneous velocity vector
u_τ	Friction velocity
v	Wall-normal velocity
x_i	Coordinate system
x, y, z	Coordinate distance
X_i	Input
Y_i	Output

Subscripts

i	Coordinate direction
x,y,z	Coordinate direction

Superscripts

$^+$	Wall units

Greek symbols

θ	Passive scalar
χ	Subgrid-scale scalar dissipation
χ_F	Forward scatter
χ_B	Backscatter
ω	Wall-normal vorticity
ν	Kinematic viscosity
τ	Shear stress
ρ	Density
δ	Channel half height
Δ	Grid size
$\overline{\Delta}$	Filter size

Abbreviations

ANN	Artificial neural network
DNN	Deep neural network
DNS	Direct numerical simulation
MAE	Mean absolute error
LES	Large-eddy simulation
PDF	Probability density function
RANS	Reynolds-averaged Navier-Stokes
ReLU	Rectified linear unit
SGS	Subgrid scale
HIT	Homogeneous isotropic turbulence
TKE	Turbulent kinetic energy