

# Toward neural-network-based large eddy simulation: application to turbulent channel flow

Jonghwan Park<sup>1</sup> and Haechon Choi<sup>1,2,†</sup>

<sup>1</sup>Department of Mechanical Engineering, Seoul National University, Seoul 08826, Korea

<sup>2</sup>Institute of Advanced Machines and Design, Seoul National University, Seoul 08826, Korea

(Received 29 April 2020; revised 14 September 2020; accepted 19 October 2020)

A fully connected neural network (NN) is used to develop a subgrid-scale (SGS) model mapping the relation between the SGS stresses and filtered flow variables in a turbulent channel flow at  $Re_\tau = 178$ . *A priori* and *a posteriori* tests are performed to investigate its prediction performance. In *a priori* test, an NN-based SGS model with the input filtered strain rate or velocity gradient tensor at multiple points provides highest correlation coefficients between the predicted and true SGS stresses, and reasonably predicts the backscatter. However, this model provides unstable solution in *a posteriori* test, unless a special treatment such as backscatter clipping is used. On the other hand, an NN-based SGS model with the input filtered strain rate tensor at single point shows an excellent prediction capability for the mean velocity and Reynolds shear stress in *a posteriori* test, although it gives low correlation coefficients between the true and predicted SGS stresses in *a priori* test. This NN-based SGS model trained at  $Re_\tau = 178$  is applied to a turbulent channel flow at  $Re_\tau = 723$  using the same grid resolution in wall units, providing fairly good agreements of the solutions with the filtered direct numerical simulation (DNS) data. When the grid resolution in wall units is different from that of trained data, this NN-based SGS model does not perform well. This is overcome by training an NN with the datasets having two filters whose sizes are bigger and smaller than the grid size in large eddy simulation (LES). Finally, the limitations of NN-based LES to complex flow are discussed.

**Key words:** turbulence modelling

## 1. Introduction

In large eddy simulation (LES), the effect of the subgrid-scale (SGS) velocity fluctuations on the resolved flow should be modelled, and thus the aim of SGS modelling is to find the relations between the resolved flow variables and SGS stresses. A conventional approach for SGS modelling is to approximate the SGS stresses with the resolved flow variables

† Email address for correspondence: [choi@snu.ac.kr](mailto:choi@snu.ac.kr)

in an arithmetic form based on turbulence theory and hypothesis. For example, an eddy viscosity model is based on the Boussinesq hypothesis that linearly relates the SGS stress tensor  $\boldsymbol{\tau}$  with the resolved strain rate tensor  $\bar{\mathbf{S}}$ , i.e.  $\boldsymbol{\tau} - \frac{1}{3}\text{tr}(\boldsymbol{\tau})\mathbf{I} = -2\nu_t\bar{\mathbf{S}}$ , where  $\mathbf{I}$  is the identity tensor, and  $\nu_t$  is an eddy viscosity to be modelled with the resolved flow variables (see, for example, Smagorinsky 1963; Nicoud & Ducros 1999; Vreman 2004; Nicoud *et al.* 2011; Verstappen 2011; Rozema *et al.* 2015; Trias *et al.* 2015; Silvis, Remmerswaal & Verstappen 2017). Some models dynamically determine the coefficients of the eddy viscosity models (Germano *et al.* 1991; Lilly 1992; Ghosal *et al.* 1995; Piomelli & Liu 1995; Meneveau, Lund & Cabot 1996; Park *et al.* 2006; You & Moin 2007; Lee, Choi & Park 2010; Verstappen *et al.* 2010). Other types of SGS model include the similarity model (Bardina, Ferziger & Reynolds 1980; Liu, Meneveau & Katz 1994; Domaradzki & Saiki 1997), the mixed model (Bardina *et al.* 1980; Zang, Street & Koseff 1993; Liu *et al.* 1994; Vreman, Geurts & Kuerten 1994; Liu, Meneveau & Katz 1995; Salvetti & Banerjee 1995; Horiuti 1997; Akhavan *et al.* 2000), and the gradient model (Clark, Ferziger & Reynolds 1979; Liu *et al.* 1994). These models have been successfully applied to various turbulent flows, but there are still drawbacks to overcome. For example, the eddy viscosity model is purely dissipative, and thus the energy transfer from subgrid to resolved scales (i.e. backscatter) cannot be predicted. On the other hand, the scale similarity model (SSM) provides the backscatter but does not dissipate energy sufficiently, and thus simulations often diverge or produce inaccurate results. Therefore, an additional eddy-viscosity term is introduced and usually coupled with the SSM to properly dissipate the energy (Bardina *et al.* 1980; Liu *et al.* 1994; Langford & Moser 1999; Sarghini, Piomelli & Balaras 1999; Meneveau & Katz 2000; Anderson & Domaradzki 2012). The dynamic version of the eddy viscosity model can predict local backscatter with negative  $\nu_t$ , but an averaging procedure or *ad hoc* clipping on negative  $\nu_t$  is required in actual LES to avoid numerical instability (Germano *et al.* 1991; Lilly 1992; Ghosal *et al.* 1995; Meneveau *et al.* 1996; Park *et al.* 2006; Thiry & Winckelmans 2016).

An alternative approach for SGS modelling is to use high-fidelity direct numerical simulation (DNS) data. The optimal LES (Langford & Moser 1999; Völker, Moser & Venugopal 2002; Langford & Moser 2004; Zandonade, Langford & Moser 2004; Moser *et al.* 2009), based on the stochastic estimation (Adrian *et al.* 1989; Adrian 1990), is such an approach, where a prediction target, e.g. the SGS force (divergence of the SGS stress tensor), is expanded with input variables (velocity and velocity gradients). The coefficients of the input variables are found by minimizing the mean-squared error between the true and estimated values of the prediction target. Another example is to use a machine-learning algorithm such as the fully connected neural network (FCNN). The FCNN is a nonlinear function that maps the predefined input variables and prediction target, where the target can be the SGS stresses or SGS force. Like the optimal LES, the weight parameters of the FCNN are found by minimizing a given loss function such as the mean-squared error. In the case of two-dimensional decaying isotropic turbulence, Maulik *et al.* (2018) applied an FCNN-based approximate deconvolution model (Stolz & Adams 1999; Maulik & San 2017) to LES, where the filtered vorticity and streamfunction at multiple grid points were the inputs of FCNNs and the corresponding prediction targets were the deconvolved vorticity and streamfunction, respectively. This FCNN-based LES showed a better prediction of the kinetic energy spectrum than LES with the dynamic Smagorinsky model (DSM; see Germano *et al.* 1991; Lilly 1992). Maulik *et al.* (2019) used the same input together with eddy-viscosity kernels, but had the SGS force as the target. In *a posteriori* test, this FCNN model reasonably predicted the kinetic energy spectrum even though the prediction performance was not much better than those of the Smagorinsky and

Leith models (Leith 1968) with the model coefficients of  $C_s = 0.1\text{--}0.3$  in  $\nu_t = (C_s \bar{\Delta})^2 |\bar{S}|$ , where  $\bar{\Delta}$  is the grid spacing and  $|\bar{S}| = \sqrt{2\bar{S}_{ij}\bar{S}_{ij}}$ . In the case of three-dimensional forced isotropic turbulence, Vollant, Balarac & Corre (2017) used an FCNN with the target of the SGS scalar flux divergence  $\nabla \cdot (\bar{\mathbf{u}}\bar{\phi} - \bar{\mathbf{u}}\bar{\phi})$  and the input of  $\bar{\mathbf{S}}$ , where  $\bar{\mathbf{u}}$  and  $\bar{\phi}$  are the filtered velocity and passive scalar, respectively. They showed that the results from FCNN-based LES were very close to those from the filtered DNS (fDNS). Zhou *et al.* (2019) reported that using the filter size as well as the velocity gradient tensor as the input variables was beneficial to predict the SGS stresses for the flow having a filter size different from that of trained data. Xie, Wang & E (2020a) used an FCNN to predict the SGS force with the input of  $\nabla\bar{\mathbf{u}}$  at multiple grid points, and this FCNN performed better than DSM for the prediction of energy spectrum. In the case of three-dimensional decaying isotropic turbulence, Wang *et al.* (2018) adopted the velocity and its first and second derivatives for the input of FCNN to predict the SGS stresses, and showed better performance in *a posteriori* test than that of DSM. Beck, Flad & Munz (2019) used a convolutional neural network (CNN) to predict the SGS force with the input of the velocity in whole domain, and showed in *a priori* test that the CNN-based SGS model predicted the SGS force better than an FCNN-based SGS model did. In the case of compressible isotropic turbulence, Xie *et al.* (2019a) used FCNNs to predict SGS force and divergence of SGS heat flux, respectively, with the inputs of  $\nabla\tilde{\mathbf{u}}, \nabla^2\tilde{\mathbf{u}}, \nabla\tilde{T}, \nabla^2\tilde{T}, \bar{\rho}$  and  $\nabla\bar{\rho}$  at multiple grid points, where  $\rho$  is the fluid density, and  $\tilde{\mathbf{u}}$  and  $\tilde{T}$  are the mass-weighting-filtered velocity and temperature, respectively. Xie *et al.* (2019b) applied FCNNs to predict the coefficients of a mixed model with the inputs of  $|\tilde{\omega}|, \tilde{\theta}, \sqrt{\tilde{\alpha}_{ij}\tilde{\alpha}_{ij}}, \sqrt{\tilde{S}_{ij}\tilde{S}_{ij}}, |\nabla\tilde{T}|$ , where  $|\tilde{\omega}|, \tilde{\theta}, \tilde{\alpha}_{ij}$  and  $\tilde{S}_{ij}$  are the mass-weighting-filtered vorticity magnitude, velocity divergence, velocity gradient tensor and strain rate tensor, respectively. Xie *et al.* (2019c) trained FCNNs with  $\nabla\tilde{\mathbf{u}}, \nabla^2\tilde{\mathbf{u}}, \nabla\tilde{T}$  and  $\nabla^2\tilde{T}$  at multiple grid points as the inputs to predict SGS stresses and SGS heat flux, respectively. Xie *et al.* (2020b) used FCNNs to predict SGS stresses and SGS heat flux with the inputs of  $\nabla\tilde{\mathbf{u}}, \nabla\hat{\tilde{\mathbf{u}}}, \nabla\tilde{T}$  and  $\nabla\hat{\tilde{T}}$  at multiple grid points, where the filter size of  $\hat{\Delta}$  is twice that of  $\tilde{\Delta}$ . They (Xie *et al.* 2019a,b,c, 2020b) showed that the FCNN-based LES provided more accurate kinetic energy spectrum and structure function of the velocity than those based on DSM and dynamic mixed model.

Unlike for isotropic turbulence, the progress in LES with an FCNN-based SGS model has been relatively slow for turbulent channel flow. Sarghini, de Felice & Santini (2003) trained an FCNN with the input of filtered velocity gradient and  $\bar{u}'_i \bar{u}'_j$  to predict the model coefficient of the Smagorinsky model for a turbulent channel flow, where  $\bar{u}'_i$  is the instantaneous filtered velocity fluctuations. Pal (2019) trained an FCNN to predict  $\nu_t$  in the eddy viscosity model with the input of filtered velocity and strain rate tensor. In Sarghini *et al.* (2003) and Pal (2019), however, FCNNs were trained by LES data from traditional SGS models, i.e. mixed model (Bardina *et al.* 1980) and DSM, respectively, rather than by fDNS data. Wollblad & Davidson (2008) trained an FCNN with fDNS data to predict the coefficients of the truncated proper orthogonal decomposition (POD) expansion of the SGS stresses with the input of  $\bar{u}'_i$ , wall-normal gradient of  $\bar{u}'_i$ , filtered pressure ( $\bar{p}$ ) and wall-normal and spanwise gradients of  $\bar{p}$ . They showed from *a priori* test that the predicted SGS stresses were in good agreement with those from fDNS data. However, the FCNN alone was unstable in *a posteriori* test, and thus the FCNN combined with the Smagorinsky model was used to conduct LES, i.e.  $\tau_{ij} = c_b \tau_{ij}^{\text{FCNN}} + (1 - c_b) \tau_{ij}^{\text{Smag}}$ , where  $\tau_{ij}^{\text{FCNN}}$  and  $\tau_{ij}^{\text{Smag}}$  were the SGS stresses from the FCNN and Smagorinsky model (with

$C_s = 0.09$ ), respectively, and  $c_b$  was a weighting parameter needed to be tuned. Gamahara & Hattori (2017) used FCNNs to predict the SGS stresses with four input variable sets,  $\{\nabla\bar{\mathbf{u}}, y\}$ ,  $\{\nabla\bar{\mathbf{u}}\}$ ,  $\{\bar{\mathbf{S}}, \bar{\mathbf{R}}, y\}$  and  $\{\bar{\mathbf{S}}, y\}$ , where  $\bar{\mathbf{R}}$  is the filtered rotation rate tensor and  $y$  is the wall-normal distance from the wall. They showed in *a priori* test that the correlation coefficients between the true and predicted SGS stresses from  $\{\nabla\bar{\mathbf{u}}, y\}$  were highest among four input sets, and even higher than those from traditional SGS models (gradient and Smagorinsky models). However, *a posteriori* test (i.e. actual LES) with  $\{\nabla\bar{\mathbf{u}}, y\}$  did not provide any advantage over the LES with the Smagorinsky model. This kind of the inconsistency between *a priori* and *a posteriori* tests had been also observed during the development of traditional SGS models (Liu *et al.* 1994; Vreman, Geurts & Kuerten 1997; Park, Yoo & Choi 2005; Anderson & Domaradzki 2012).

Previous studies (Wollblad & Davidson 2008; Gamahara & Hattori 2017) showed that FCNN is a promising tool for modelling SGS stresses from *a priori* test, but it is unclear why FCNN-based LESs did not perform better for a turbulent channel flow than LESs with traditional SGS models. Thus, a more systematic investigation on the SGS variables such as the SGS dissipation and transport is required to diagnose the performance of FCNN. The input variables for the FCNN should be also chosen carefully based on the characteristics of the SGS stresses. Therefore, the objective of the present study is to develop an FCNN-based SGS model for a turbulent channel flow, based on both *a priori* and *a posteriori* tests, and to find appropriate input variables for the successful LES with FCNN. We train FCNNs with different input variables such as  $\bar{\mathbf{S}}$  and  $\nabla\bar{\mathbf{u}}$ , and the target to predict is the SGS stress tensor. We also test  $\bar{\mathbf{u}}$  and  $\partial\bar{\mathbf{u}}/\partial y$  as the input for FCNN (note that these were the input variables of the optimal LES for a turbulent channel flow by Völker *et al.* 2002). The input and target data are obtained by filtering the data from DNS of a turbulent channel flow at the bulk Reynolds number of  $Re_b = 5600$  ( $Re_\tau = u_\tau\delta/\nu = 178$ ), where  $u_\tau$  is the wall-shear velocity,  $\delta$  is the channel half height and  $\nu$  is the kinematic viscosity. In *a priori* test, we examine the variations of the predicted SGS dissipation, backscatter and SGS transport with the input variables, which are known to be important variables for successful LES of a turbulent channel flow (Piomelli, Yu & Adrian 1996; Völker *et al.* 2002; Park *et al.* 2006). In *a posteriori* test, we perform LESs with FCNN-based SGS models at  $Re_\tau = 178$  and estimate their prediction performance by comparing the results with those from the fDNS data and LESs with DSM and SSM (Liu *et al.* 1994). The details about DNS and FCNN are given in § 2. The results from *a priori* and *a posteriori* tests at  $Re_\tau = 178$  are given in § 3. Applications of the FCNN trained at  $Re_\tau = 178$  to LES of a higher-Reynolds-number flow ( $Re_\tau = 723$ ) and to LES with a different grid resolution at  $Re_\tau = 178$  are also discussed in § 3, followed by conclusions in § 4.

## 2. Numerical details

### 2.1. NN-based SGS model

The governing equations for LES are the spatially filtered continuity and Navier–Stokes equations,

$$\frac{\partial\bar{u}_i}{\partial x_i} = 0, \tag{2.1}$$

$$\frac{\partial\bar{u}_i}{\partial t} + \frac{\partial\bar{u}_i\bar{u}_j}{\partial x_j} = -\frac{\partial\bar{p}}{\partial x_i} + \frac{1}{Re} \frac{\partial^2\bar{u}_i}{\partial x_j\partial x_j} - \frac{\partial\tau_{ij}}{\partial x_j}, \tag{2.2}$$

where  $x_1 (=x)$ ,  $x_2 (=y)$  and  $x_3 (=z)$  are the streamwise, wall-normal and spanwise directions, respectively,  $u_i (=u, v, w)$  are the corresponding velocity components,  $p$  is the pressure,  $t$  is time, the overbar denotes the filtering operation and  $\tau_{ij} (= \overline{u_i u_j} - \bar{u}_i \bar{u}_j)$  is the SGS stress tensor. We use a FCNN (denoted as NN hereafter) with the input of the filtered flow variables to predict  $\tau_{ij}$ . The database for training the NN is obtained by filtering the instantaneous flow fields from DNS of a turbulent channel flow at  $Re_\tau = 178$  (see § 2.2). To estimate the performance of the present NN-based SGS model, we perform two additional LESs with the DSM (Germano *et al.* 1991; Lilly 1992) and SSM (Liu *et al.* 1994). For the DSM,  $\tau_{ij} - \frac{1}{3} \tau_{kk} \delta_{ij} = -2C^2 |\bar{S}| \bar{S}_{ij}$ , where  $C^2 = -\frac{1}{2} \langle L_{ij} M_{ij} \rangle_h / \langle M_{ij} M_{ij} \rangle_h$ ,  $|\bar{S}| = \sqrt{2 \bar{S}_{ij} \bar{S}_{ij}}$ ,  $\bar{S}_{ij} = \frac{1}{2} (\partial \bar{u}_i / \partial x_j + \partial \bar{u}_j / \partial x_i)$ ,  $L_{ij} = \widetilde{\widetilde{u_i u_j}} - \widetilde{\widetilde{u}_i} \widetilde{\widetilde{u}_j}$ ,  $M_{ij} = (\bar{\Delta} / \widetilde{\Delta})^2 |\bar{S}| \widetilde{\widetilde{S}_{ij}} - \widetilde{\widetilde{S}} |\bar{S}| \bar{S}_{ij}$ ,  $\bar{\Delta}$  and  $\widetilde{\Delta} (=2\bar{\Delta})$  denote the grid and test filter sizes, respectively, and  $\langle \rangle_h$  denotes averaging in the homogeneous ( $x$  and  $z$ ) directions. For the SSM,  $\tau_{ij} = \widetilde{\widetilde{u_i u_j}} - \widetilde{\widetilde{u}_i} \widetilde{\widetilde{u}_j}$ , where  $\widetilde{\widetilde{k}}_{cut} = 0.5 \bar{k}_{cut}$  and  $k_{cut}$  is the cut-off wavenumber.

The NN adopted in the present study has two hidden layers with 128 neurons per hidden layer, and the output of the NN is the six components of  $\tau_{ij}$  (figure 1). Previous studies used one (Gamahara & Hattori 2017; Maulik & San 2017; Maulik *et al.* 2018; Zhou *et al.* 2019) or two (Sarghini *et al.* 2003; Wollblad & Davidson 2008; Vollant *et al.* 2017; Wang *et al.* 2018; Maulik *et al.* 2019; Xie *et al.* 2019a,b,c, 2020a,b) hidden layers, and Gamahara & Hattori (2017) showed that 100 neurons per hidden layer were sufficient for the accurate predictions of  $\tau_{ij}$  for a turbulent channel flow in a *a priori* test. We also tested NN with three hidden layers, but more hidden layers than two did not further improve the performance both in a *a priori* and a *a posteriori* tests (see the Appendix).

In the present NN, the output of the  $m$ th layer,  $\mathbf{h}^{(m)}$ , is as follows:

$$\left. \begin{aligned}
 h_i^{(1)} &= q_i \quad (i = 1, 2, \dots, N_q); \\
 h_j^{(2)} &= \max[0, r_j^{(2)}], r_j^{(2)} \\
 &= \gamma_j^{(2)} \left( \sum_{i=1}^{N_q} W_{ij}^{(1)(2)} h_i^{(1)} + b_j^{(2)} - \mu_j^{(2)} \right) / \left( \sigma_j^{(2)} + \beta_j^{(2)} \right) \quad (j = 1, 2, \dots, 128); \\
 h_k^{(3)} &= \max[0, r_k^{(3)}], r_k^{(3)} \\
 &= \gamma_k^{(3)} \left( \sum_{j=1}^{128} W_{jk}^{(2)(3)} h_j^{(2)} + b_k^{(3)} - \mu_k^{(3)} \right) / \left( \sigma_k^{(3)} + \beta_k^{(3)} \right) \quad (k = 1, 2, \dots, 128); \\
 h_l^{(4)} &= s_l = \sum_{k=1}^{128} W_{kl}^{(3)(4)} h_k^{(3)} + b_l^{(4)} \quad (l = 1, 2, \dots, 6),
 \end{aligned} \right\} \tag{2.3}$$

where  $q_i$  is the input,  $N_q$  is the number of input components,  $\mathbf{W}^{(m)(m+1)}$  is the weight matrix between the  $m$ th and  $(m+1)$ th layers,  $\mathbf{b}^{(m)}$  is the bias of the  $m$ th layer,  $s_l$  is the output and  $\boldsymbol{\mu}^{(m)}$ ,  $\boldsymbol{\sigma}^{(m)}$ ,  $\boldsymbol{\gamma}^{(m)}$  and  $\boldsymbol{\beta}^{(m)}$  are parameters for a batch normalization (Ioffe & Szegedy 2015). We use a rectified linear unit (ReLU; see Nair & Hinton 2010),  $\mathbf{h}^{(m)} = \max[0, \mathbf{r}^{(m)}]$ , as the activation function at the hidden layers. We also tested other typical activation functions such as sigmoid and hyperbolic tangent functions, but

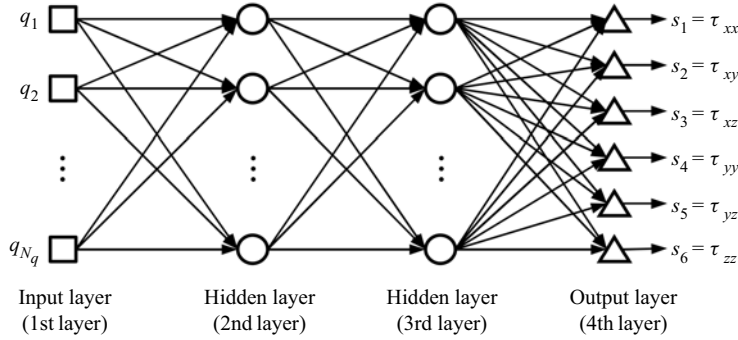


Figure 1. Schematic diagram of the present NN with two hidden layers (128 neurons per hidden layer). Here,  $\mathbf{q} (= [q_1, q_2, \dots, q_{N_q}]^T)$  is the input of NN,  $N_q$  is the number of input components (see table 1) and  $\mathbf{s} (= [s_1, s_2, \dots, s_6]^T)$  is the output of NN.

the convergence of the loss function (2.4) was faster with the ReLU than with others. Here  $\mathbf{W}^{(m)(m+1)}$ ,  $\mathbf{b}^{(m)}$ ,  $\boldsymbol{\gamma}^{(m)}$  and  $\boldsymbol{\beta}^{(m)}$  are trainable parameters that are optimized to minimize the loss function defined as

$$L = \frac{1}{2N_b} \frac{1}{6} \sum_{l=1}^6 \sum_{n=1}^{N_b} \left( s_{l,n}^{\text{fDNS}} - s_{l,n} \right)^2 + 0.005 \sum_o w_o^2, \tag{2.4}$$

where  $s_{l,n}^{\text{fDNS}}$  is the SGS stresses obtained from fDNS data,  $N_b$  is the number of minibatch data (128 in this study following Kingma & Ba 2014) and  $w_o$  denotes the components of  $\mathbf{W}^{(m)(m+1)}$ . An adaptive moment estimation (Kingma & Ba 2014), which is a variant of gradient descent method, is applied to update the trainable parameters and the gradients of the loss function with respect to those parameters are calculated through the chain rule of derivatives (Rumelhart, Hinton & Williams 1986; LeCun, Bengio & Hinton 2015). All training procedures are conducted using the Python open-source library TensorFlow.

We choose five different input variables (corresponding to NN1–NN5), as listed in table 1. Six components of  $\bar{S}_{ij}$  and nine components of  $\bar{\alpha}_{ij} (= \partial \bar{u}_i / \partial x_j)$  at each grid point are the inputs to NN1 and NN2, respectively, and the output is six components of  $\tau_{ij}$  at the same grid location. The input  $\bar{\alpha}_{ij}$  is selected for NN2 because  $\tau_{ij}$  can be written as  $\tau_{ij} = 2\gamma \bar{\alpha}_{ik} \bar{\alpha}_{jk} + O(\gamma^2)$ , where  $\gamma(\zeta) = \int_{-\infty}^{\infty} \xi^2 G(\xi, \zeta) d\xi$  and  $G(\xi, \zeta)$  is the kernel of the filter (Bedford & Yeo 1993). On the other hand, a general class of SGS model based on the local velocity gradient (Lund & Novikov 1992; Silvis *et al.* 2019) can be expressed as  $\tau_{ij} = \sum_{k=0}^5 c^{(k)} T_{ij}^{(k)}$ , where  $c^{(k)}$  is the model coefficient,  $T_{ij}^{(0)} = \delta_{ij}$ ,  $T_{ij}^{(1)} = \bar{S}_{ij}$ ,  $T_{ij}^{(2)} = \bar{S}_{ik} \bar{S}_{kj}$ ,  $T_{ij}^{(3)} = \bar{R}_{ik} \bar{R}_{kj}$ ,  $T_{ij}^{(4)} = \bar{S}_{ik} \bar{R}_{kj} - \bar{R}_{ik} \bar{S}_{kj}$ ,  $T_{ij}^{(5)} = \bar{S}_{ik} \bar{S}_{kl} \bar{R}_{lj} - \bar{R}_{ik} \bar{S}_{kl} \bar{S}_{lj}$  and  $\bar{R}_{ij}$  is the filtered rotation rate tensor. Thus, NN1 can be regarded as an SGS model including  $T_{ij}^{(0)}$ ,  $T_{ij}^{(1)}$  and  $T_{ij}^{(2)}$ , but it directly predicts  $\tau_{ij}$  through a nonlinear process of NN rather than predicting  $c^{(k)}$ . In NN3 and NN4, a stencil of data at  $3(x) \times 3(z)$  grid points are the input, and  $\tau_{ij}$  at the center of this stencil is the output. In NN5, the filtered velocity and wall-normal velocity gradient at  $3(x) \times 3(z)$  grid points are the input variables, and the output is the same as that of NN3 and NN4. The use of a stencil of data for NN3–NN5 is motivated by the results of Xie *et al.* (2019c) that using a stencil of input variables ( $\bar{\alpha}_{ij}$  and temperature gradient) predicted  $\tau_{ij}$  better than using the same input only at one grid point. The choice of  $\bar{u}_i$  and  $\partial \bar{u}_i / \partial y$  as the input of NN5 is also motivated by the results of optimal LES by Völker *et al.* (2002), in which LES with the input of both  $\bar{u}_i$  and  $\partial \bar{u}_i / \partial y$

NN model	Input variable(s)	Input grid point(s)	Number of input components ( $N_q$ )
NN1	$\bar{S}_{ij}$	1	6
NN2	$\bar{\alpha}_{ij}$	1	9
NN3	$\bar{S}_{ij}$	$3(x) \times 3(z)$	54
NN4	$\bar{\alpha}_{ij}$	$3(x) \times 3(z)$	81
NN5	$\bar{u}_i$ and $\partial\bar{u}_i/\partial y$	$3(x) \times 3(z)$	54

Table 1. Input variables of NN models.

outperformed that with the input of  $\bar{u}_i$  alone. We also considered an NN with the input of  $\bar{u}_i$  at  $n_x(x) \times 3(y) \times n_z(z)$  grid points, where  $n_x = n_z = 3, 5, 7,$  or  $9$ . The results with these three-dimensional multiple input grid points were little different in *a priori* tests from that of NN5. As shown in § 3, the results with NN3–NN5 in *a priori* tests are better than those with NN1 and NN2 (single input grid point), but actual LES (i.e. *a posteriori* test) with NN3–NN5 are unstable. Therefore, we did not seek to adopt more input grid points. Note also that we train a single NN for all  $y$  locations using pairs of the input and output variables. The relations between these variables are different for different  $y$  locations, and thus  $y$  locations are implicitly embedded in this single NN. One may train an NN at each  $y$  location, but this procedure increases the number of NNs and the memory size. On the other hand, Gamahara & Hattori (2017) provided  $y$  locations as an additional input variable for a single NN, but found that the result of *a priori* test with  $y$  location was only slightly better than that without  $y$  location. Therefore, we do not attempt to include  $y$  location as an additional input variable in this study.

While training NN1–NN5, the input and output variables are normalized in wall units, which provides successful results because the flow variables in turbulent channel flow are well scaled in wall units (see § 3). As the performance of an NN depends on the normalization of input and output variables (see, for example, Passalis *et al.* 2019), we considered two more normalizations: one was with the centreline velocity ( $U_c$ ) and channel half height ( $\delta$ ), and the other was such that the input and output variables were scaled to have zero mean and unit variance at each  $y$  location, e.g.  $\tau_{ij}^*(x, y, z, t) = (\tau_{ij}(x, y, z, t) - \tau_{ij}^{mean}(y))/\tau_{ij}^{rms}(y)$  (no summation on  $i$  and  $j$ ), where the superscripts *mean* and *rms* denote the mean and root-mean-square (r.m.s.) values, respectively. The first normalization was not successful for the prediction of a higher-Reynolds-number flow with an NN trained at lower Reynolds number, because the near-wall flow was not properly scaled with this normalization. The second normalization requires *a priori* knowledge on  $\tau_{ij}^{mean}(y)$  and  $\tau_{ij}^{rms}(y)$  even for a higher-Reynolds-number flow to predict. Thus, we did not take the second normalization either.

Figure 2 shows the variations of the training error  $\epsilon_\tau$  with the epoch, and the correlation coefficients  $\rho_\tau$  between true and predicted SGS stresses for NN1–NN5, where  $\epsilon_\tau$  and  $\rho_\tau$  are defined as

$$\epsilon_\tau = \frac{1}{2N_{data}} \frac{1}{6} \sum_{l=1}^6 \sum_{n=1}^{N_{data}} \left( s_{l,n}^{fDNS} - s_{l,n} \right)^2, \quad (2.5)$$

$$\rho_\tau = \frac{\sum_{l=1}^6 \sum_{n=1}^{N_{data}} \left( s_{l,n}^{fDNS} s_{l,n} \right)}{\left( \sqrt{\sum_{l=1}^6 \sum_{n=1}^{N_{data}} \left( s_{l,n}^{fDNS} \right)^2} \sqrt{\sum_{l=1}^6 \sum_{n=1}^{N_{data}} \left( s_{l,n} \right)^2} \right)}. \quad (2.6)$$

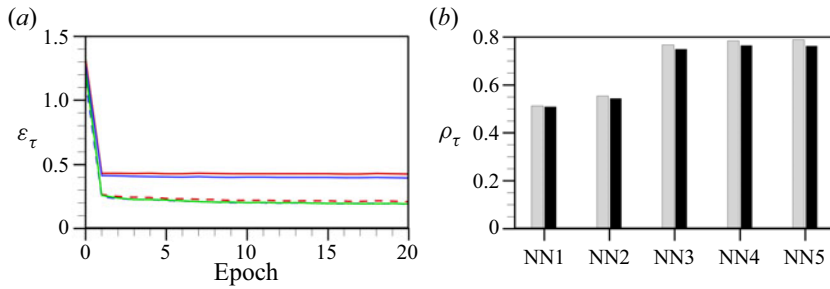


Figure 2. Training error and correlation coefficient by NN1–NN5: (a) training error versus epoch; (b) correlation coefficient. In (a), red solid line, NN1; blue solid line, NN2; red dashed line, NN3; blue dashed line, NN4; green solid line, NN5. In (b), gray and black bars are the correlation coefficients for training and test datasets, respectively, where the number of test data is the same as that of the training data ( $N_{data} = 1\,241\,600$  (§ 2.2)).

Here, one epoch denotes one sweep through the entire training dataset (Hastie, Tibshirani & Friedman 2009), and  $N_{data}$  is the number of entire training data. The training errors nearly converge at 20 epochs (figure 2a). In terms of computational time using a single graphic process unit (NVIDIA GeForce GTX 1060), about 1 min is spent for each epoch. The correlation coefficients from the training and test datasets are quite similar to each other (figure 2b), indicating that severe overfitting is not observed for NN1–NN5. The training error and correlation coefficient are smaller and larger, respectively, for NN3–NN5 than those for NN1 and NN2.

Sarghini *et al.* (2003) and Pal (2019) indicated that required computational time for their LESs with NNs was less than that with traditional SGS models. When an NN is used for obtaining the SGS stresses, its cost depends on the numbers of hidden layers and neurons therein as well as the choices of input and output variables. Actually, in the present study, the computational time required for one computational time-step advancement with NN1 is approximately 1.3 times that with a traditional SGS model such as DSM.

### 2.2. Details of DNS and input and output variables

A DNS of turbulent channel flow at  $Re_b = 5600$  ( $Re_\tau = 178$ ) is conducted to obtain the input and output of NN1–NN5 (table 1), where  $Re_b$  is the bulk Reynolds number defined by  $Re_b = U_b(2\delta)/\nu$ ,  $U_b$  is the bulk velocity and  $Re_\tau = u_\tau\delta/\nu$  is the friction Reynolds number. The Navier–Stokes and continuity equations are solved in the form of the wall-normal vorticity and the Laplacian of the wall-normal velocity, as described in Kim, Moin & Moser (1987). The dealiased Fourier and Chebyshev polynomial expansions are used in the homogeneous ( $x$  and  $z$ ) and wall-normal ( $y$ ) directions, respectively. A semi-implicit fractional step method is used for time integration, where a third-order Runge–Kutta and second-order Crank–Nicolson methods are applied to the convection and diffusion terms, respectively. A constant mass flux in a channel is maintained by adjusting the mean pressure gradient in the streamwise direction at each time step.

Table 2 lists the computational parameters of DNS, where  $N_{x_i}$  are the numbers of grid points in  $x_i$  directions,  $L_{x_i}$  are the corresponding computational domain sizes,  $\Delta x$  and  $\Delta z$  are the uniform grid spacings in  $x$  and  $z$  directions, respectively, and  $\Delta y_{min}^+$  is the smallest grid spacing at the wall in the wall-normal direction. Here  $\Delta_x$  and  $\Delta_z$  are the filter sizes in  $x$  and  $z$  directions, respectively, and they are used for obtaining fDNS data. We apply the spectral cut-off filter only in the wall-parallel ( $x$  and  $z$ ) directions as in the



$Re_b$	$Re_\tau$	$N_x, N_y, N_z$	$L_x/\delta, L_z/\delta$	$\Delta x^+, \Delta z^+, \Delta y_{min}^+$	$\Delta T^+$	$\Delta_x^+, \Delta_z^+$
5600	178	96, 97, 96	$2\pi, \pi$	11.7, 5.8, 0.1	9.4	69.9, 35.0
27600	723	192, 193, 192	$\pi, 0.5\pi$	11.8, 5.9, 0.1	11.4	71.0, 35.5

Table 2. Computational parameters of DNS. Here, the superscript + denotes the wall unit and  $\Delta T$  is the sampling time interval of the instantaneous DNS flow fields for constructing the input and output database.

previous studies (Piomelli *et al.* 1991, 1996; Völker *et al.* 2002; Park *et al.* 2006). The use of only wall-parallel filters can be justified because small scales are efficiently filtered out by wall-parallel filters and wall-normal filtering through the truncation of the Chebyshev mode violates the continuity unless the divergence-free projection is performed (Völker *et al.* 2002; Park *et al.* 2006). The Fourier coefficient of a filtered flow variable  $\hat{f}$  is defined as

$$\hat{f}(k_x, y, k_z, t) = \hat{f}(k_x, y, k_z, t) H(k_{x,cut} - |k_x|) H(k_{z,cut} - |k_z|), \quad (2.7)$$

where  $\hat{f}$  is the Fourier coefficient of an unfiltered flow variable  $f$ ,  $H$  is the Heaviside step function and  $k_{x,cut}$  and  $k_{z,cut}$  are the cut-off wavenumbers in  $x$  and  $z$  directions, respectively. The filter sizes in table 2,  $\Delta_x^+$  and  $\Delta_z^+$ , are the same as those in Park *et al.* (2006), and the corresponding cut-off wavenumbers are  $k_{x,cut} = 8(2\pi/L_x)$  and  $k_{z,cut} = 8(2\pi/L_z)$ , respectively. We use the input and output database at  $Re_\tau = 178$  to train NN1–NN5. The training data are collected at every other grid point in  $x$  and  $z$  directions to exclude highly correlated data, and at all grid points in  $y$  direction from 200 instantaneous fDNS fields. Then, the number of training data from 200 fDNS fields is  $1\,241\,600 (= 200 \times N_x^{fDNS}/2 \times N_z^{fDNS}/2 \times N_y)$ , where  $N_x^{fDNS} = L_x k_{x,cut}/\pi$  and  $N_z^{fDNS} = L_z k_{z,cut}/\pi$ . We have also tested 300 fDNS fields for training NNs, but their prediction performance for the SGS stresses is not further improved, so the number of training data used is sufficient for the present NNs. A DNS at a higher Reynolds number of  $Re_\tau = 723$  is also carried out, and its database is used to estimate the prediction capability of the present NN-based SGS model for untrained higher-Reynolds-number flow.

### 3. Results

In § 3.1, we perform *a priori* tests for two different Reynolds numbers,  $Re_\tau = 178$  and 723, in which the SGS stresses are predicted by NN1–NN5 with the input variables from fDNS at each Reynolds number, and compared with the SGS stresses from fDNS. Note that NN1–NN5 are constructed at  $Re_\tau = 178$  and  $Re_\tau = 723$  is an untrained higher Reynolds number. The filter sizes used in *a priori* tests,  $\Delta_x^+$  and  $\Delta_z^+$ , are given in table 2. In § 3.2, *a posteriori* tests (i.e. actual LESs solving (2.1) and (2.2)) with NN1–NN5 are performed for a turbulent channel flow at  $Re_\tau = 178$  and their results are compared with those of fDNS. Furthermore, LES with NN1 (trained at  $Re_\tau = 178$ ) is carried out for a turbulent channel flow at  $Re_\tau = 723$  and its results are compared with those of fDNS. Finally, in § 3.3, we provide the results when the grid resolution in LES is different from that used in training NN1, and suggest a way to obtain good results.

#### 3.1. *A priori* test

Figure 3 shows the mean SGS shear stress  $\langle \tau_{xy} \rangle$  and dissipation  $\langle \varepsilon_{SGS} \rangle$  predicted by NN1–NN5, together with those of fDNS and from DSM and SSM, where  $\varepsilon_{SGS} = -\tau_{ij} \bar{S}_{ij}$

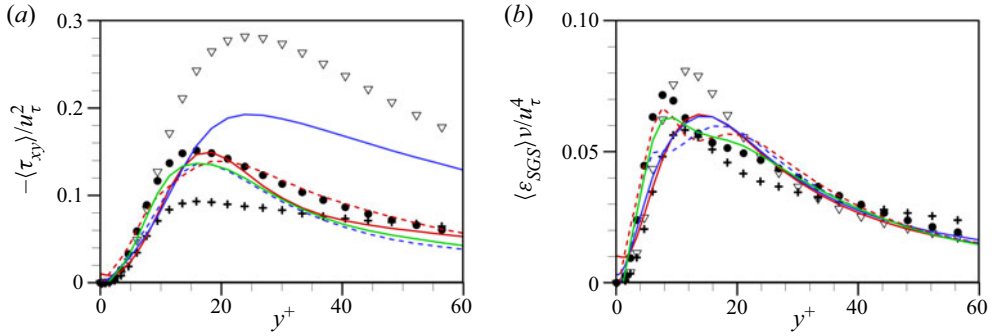


Figure 3. Mean SGS shear stress and dissipation predicted by NN1–NN5 (*a priori* test at  $Re_\tau = 178$ ): (a) mean SGS shear stress  $\langle \tau_{xy} \rangle$ ; (b) mean SGS dissipation  $\langle \varepsilon_{SGS} \rangle$ . •, fdNS; red solid line, NN1; blue solid line, NN2; red dashed line, NN3; blue dashed line, NN4; green solid line, NN5; +, DSM;  $\nabla$ , SSM.

	NN1	NN2	NN3	NN4	NN5	DSM	SSM
$\rho_{\tau_{xy}}$	0.231	0.432	0.414	0.630	0.600	0.090	-0.016
$\rho_{\varepsilon_{SGS}}$	0.358	0.472	0.507	0.624	0.576	0.165	0.081

Table 3. Correlation coefficients between the true and predicted  $\tau_{xy}$  and  $\varepsilon_{SGS}$ .

and  $\langle \rangle$  denotes the averaging in the homogeneous directions and time. Predictions of  $\langle \tau_{xy} \rangle$  by NNs (except that by NN2) are better than those by DSM and SSM, and NN5 provides an excellent prediction of  $\langle \varepsilon_{SGS} \rangle$  albeit other NN models are also good in the estimation of  $\langle \varepsilon_{SGS} \rangle$ . Table 3 lists the correlation coefficients  $\rho$  between the true and predicted  $\tau_{xy}$  and  $\varepsilon_{SGS}$ , respectively. The values of  $\tau_{xy}$  predicted by DSM and SSM have very low correlations with true  $\tau_{xy}$ , as reported by Liu *et al.* (1994) and Park *et al.* (2005). On the other hand, NN1–NN5 have much higher correlations of  $\tau_{xy}$  and  $\varepsilon_{SGS}$  than those by DSM and SSM, indicating that instantaneous  $\tau_{xy}$  and  $\varepsilon_{SGS}$  are relatively well captured by NN1–NN5. These SGS variables are even better predicted by having the input variables at multiple grid points (NN3–NN5) than at single grid point (NN1 and NN2). As we show in the following, however, high correlation coefficients of  $\tau_{xy}$  and  $\varepsilon_{SGS}$  in *a priori* test do not necessarily guarantee excellent prediction performance in actual LES.

Figure 4(a) shows the mean SGS transport  $\langle T_{SGS} \rangle$ , where  $T_{SGS} = \partial(\tau_{ij}\bar{u}_i)/\partial x_j$ . Völker *et al.* (2002) indicated that a good prediction of  $\langle T_{SGS} \rangle$  is necessary for an accurate LES, and the optimal LES provided good representation of  $\langle T_{SGS} \rangle$  in *a posteriori* test. Among NN models considered, NN5 shows the best agreement of  $\langle T_{SGS} \rangle$  with that of fdNS, but NN1 and NN2 are not good at accurately predicting  $\langle T_{SGS} \rangle$  although they are still better than SSM. Figure 4(b) shows the mean backward SGS dissipation (backscatter, i.e. energy transfer from subgrid to resolved scales),  $\langle \varepsilon_{SGS}^- \rangle = \frac{1}{2} \langle \varepsilon_{SGS} - |\varepsilon_{SGS}| \rangle$ .  $\langle \varepsilon_{SGS}^- \rangle = 0$  for DSM owing to the averaging procedure in determining the model coefficient. The mean backscatters from SSM and NN3–NN5 show reasonable agreements with that of fdNS, but NN1 and NN2 severely underpredict the backscatter. An accurate prediction of backscatter is important in wall-bounded flows, because it is related to the bursting and sweep events (Härtel *et al.* 1994; Piomelli *et al.* 1996). However, SGS models with non-negligible backscatter such as SSM do not properly dissipate energy and incur numerical instability in actual LES (Liu *et al.* 1994; Akhavan *et al.* 2000; Meneveau &

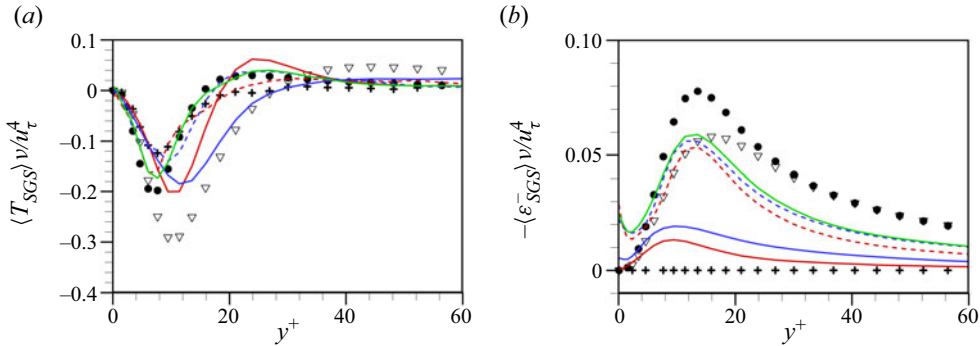


Figure 4. Mean SGS transport and backward SGS dissipation predicted by NN1–NN5 (*a priori* test at  $Re_\tau = 178$ ): (a) mean SGS transport  $\langle T_{SGS} \rangle$ ; (b) mean backward SGS dissipation (backscatter)  $\langle \varepsilon_{SGS}^- \rangle$ . •, fDNS; red solid line, NN1; blue solid line, NN2; red dashed line, NN3; blue dashed line, NN4; green solid line, NN5; +, DSM;  $\nabla$ , SSM.

Katz 2000; Anderson & Domaradzki 2012). For this reason, some NN-based SGS models suggested in the previous studies clipped the backscatter to be zero for ensuring stable LES results (Maulik *et al.* 2018, 2019; Zhou *et al.* 2019). Therefore, the accuracy and stability in the solution from LES with NN3–NN5 may not be guaranteed, even if these models properly predict the backscatter and produce high correlation coefficients between the true and predicted SGS stresses.

Figure 5 shows the statistics from *a priori* test for  $Re_\tau = 723$  with NN-based SGS models trained at  $Re_\tau = 178$ . The statistics predicted by NN1–NN5 for  $Re_\tau = 723$  show very similar behaviours to those for  $Re_\tau = 178$ , except for an underprediction of  $\langle \tau_{xy} \rangle$  by NN1 (similar to that by DSM) which does not degrade its prediction capability in a *posteriori* test (see § 3.2).

### 3.2. *A posteriori* test

In this section, *a posteriori* tests (i.e. actual LESs) with NN-based SGS models are conducted for a turbulent channel flow with a constant mass flow rate ( $Re_b = 5600$  or 27 600). Numerical methods for solving the filtered Navier–Stokes and continuity equations are the same as those of DNS described in § 2.2. Table 4 shows the computational parameters of LES. The grid resolution for the cases of LES178 is the same as that of Park *et al.* (2006). The cases of LES178 have nearly the same grid resolutions in wall units (because of slightly different values of  $Re_\tau$ ) in  $x$  and  $z$  directions as those of fDNS used in training NNs, and the cases of LES178c and LES178f use larger and smaller grid sizes in  $x$  and  $z$  directions than those of trained data, respectively. In the case of LES723 ( $Re_\tau = 723$ ), the grid sizes in wall units in  $x$  and  $z$  directions are nearly the same as those of trained data.

In the present LESs with NN3–NN5 and SSM, we clip the SGS stresses to be zero wherever backscatter occurs, i.e.  $\tau_{ij} = 0$  when  $\varepsilon_{SGS} < 0$ , as done in the previous studies (Maulik *et al.* 2018, 2019; Zhou *et al.* 2019). Otherwise, the solution diverges. While removing the backscatter, we rescale the SGS stresses to maintain the net amount of SGS dissipation in the computational domain  $V$  as follows:

$$\tau_{ij}^* = \frac{1}{2} [1 + \text{sign}(\varepsilon_{SGS})] \tau_{ij} \cdot \frac{\int_V \varepsilon_{SGS} dV}{\frac{1}{2} \int_V (\varepsilon_{SGS} + |\varepsilon_{SGS}|) dV}. \quad (3.1)$$

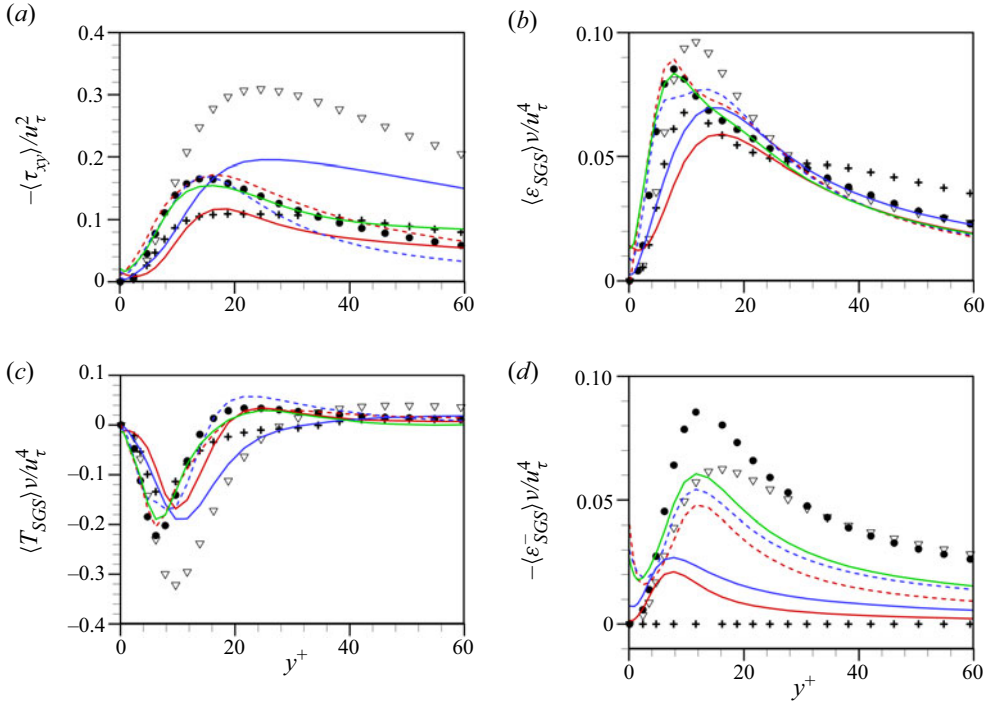


Figure 5. Statistics from *a priori* test at  $Re_\tau = 723$ : (a) mean SGS shear stress  $\langle \tau_{xy} \rangle$ ; (b) mean SGS dissipation  $\langle \epsilon_{SGS} \rangle$ ; (c) mean SGS transport  $\langle T_{SGS} \rangle$ ; (d) mean backscatter  $\langle \epsilon_{SGS}^- \rangle$ .  $\bullet$ , fdNS; red solid line, NN1; blue solid line, NN2; red dashed line, NN3; blue dashed line, NN4; green solid line, NN5; +, DSM;  $\nabla$ , SSM. Here, NN1–NN5 are trained with fdNS at  $Re_\tau = 178$ .

Case	$Re_b$	$(N_x, N_z, N_y)$	$(\Delta x^+, \Delta z^+)$	SGS model	$Re_\tau$ (w/o clipping)	$Re_\tau$ (with clipping)
LES178	5600	(16, 16, 49)	(69.9, 35.0)	NN1	181	175
				NN2	177	170
				NN3	diverged	183
				NN4	diverged	177
				NN5	diverged	178
				DSM	174	—
				SSM	diverged	176
LES178c	5600	(12, 12, 49)	(93.2, 46.6)	no	195	—
				NN1	175	175
				DSM	171	—
LES178f	5600	(24, 24, 49)	(46.6, 23.3)	no	190	—
				NN1	192	176
				DSM	177	—
LES723	27 600	(32, 32, 97)	(71.0, 35.5)	no	193	—
				NN1	729	—
				DSM	707	—
—	—	—	—	no	763	—

Table 4. Computational parameters of LES. Here, the computations are performed at constant mass flow rates (i.e.  $Re_b = 5600$  and  $27\,600$ ) and  $Re_\tau$  given in this table are the results of LESs. For  $Re_b = 5600$  and  $27\,600$ , the domain sizes in  $x$  and  $z$  directions are  $2\pi\delta \times \pi\delta$  and  $\pi\delta \times 0.5\pi\delta$ , respectively.  $\Delta y_{min}^+ = 0.4$  for all simulations, and  $\Delta y_{min}^+$ ,  $\Delta x^+$  and  $\Delta z^+$  in this table are computed with  $u_\tau$  from DNS (table 2).

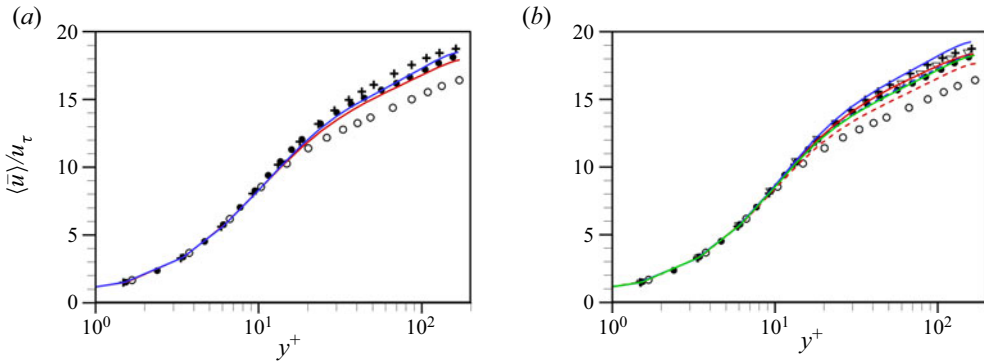


Figure 6. Mean velocity profiles from LES178 (*a posteriori* test): (a) without clipping the backscatter; (b) with clipping the backscatter. •, fDNS; red solid line, NN1; blue solid line, NN2; red dashed line, NN3; blue dashed line, NN4; green solid line, NN5; +, DSM; ∇, SSM; o, no SGS model. LESs with NN3–NN5 and SSM without clipping diverged.

This backscatter clipping and rescaling on  $\tau_{ij}$  is similar to that of Akhavan *et al.* (2000) in their development of dynamic two-component model. For the cases of LESs with NN1 and NN2, we obtain stable solutions without any special treatment such as the clipping, wall damping or averaging over homogeneous directions and, thus, we perform LESs with and without clipping, respectively. In LES with DSM, an averaging procedure is included to determine the model coefficient, as mostly done in previous studies. As the present simulations are conducted for a constant mass flow rate in a channel, the wall-shear velocity or  $Re_\tau$  changes depending on the choice of SGS models. Those  $Re_\tau$  are listed in table 4. For  $Re_b = 5600$ ,  $Re_\tau$  from LES178 are well predicted by NN1 and NN2 even without clipping (less than 2% error) and by NN3–NN5 with clipping (less than 3% error). On the other hand,  $Re_\tau$  from no SGS model has about 10% error.

Figure 6 shows the mean velocity profiles from LES178 for various SGS models without and with clipping the backscatter, respectively. Without clipping, LESs with NN1 and NN2 show excellent predictions of the mean velocity, but those with NN3–NN5 and SSM diverge. On the other hand, with clipping, LESs with all the SGS models considered provide very good predictions of the mean velocity, which clearly indicates that backscatter incurs numerical instability in LES. Therefore, in the following, we present the results of LESs with clipping for NN3–NN5, and without clipping for NN1 and NN2, respectively.

Figure 7 shows the statistics of various turbulence quantities from LES178 with NN1–NN5, together with those of fDNS and from LESs with DSM and SSM. All NNs considered show good predictions of the r.m.s. velocity fluctuations (figure 7a). While LES without the SGS model (i.e. coarse DNS) fortuitously well predicts  $\bar{u}_{rms}$  owing to overpredicted friction velocity, LES with DSM overpredicts it (Park *et al.* 2006). As DSM determines the model coefficient  $C^2(y)$  to be uniform in the homogeneous directions ignoring the locality of  $C^2(x, y, z)$ , its prediction performance of local SGS dissipation is degraded and may result in the overprediction of  $\bar{u}_{rms}$ . For the predictions of the Reynolds shear stress and SGS shear stress, NN1 performs the best among all the SGS models considered (figures 7b and 7c). On the other hand, NN2 underpredicts the Reynolds shear stress and significantly overpredicts the SGS shear stress. This result is consistent with that of *a priori* test (figure 3a). The overprediction of  $\langle \tau_{xy} \rangle$  results in the underprediction of  $-\langle \bar{u}'\bar{v}' \rangle$  from the total shear stress equation,  $d\langle \bar{u}^+ \rangle / dy^+ - \langle \bar{u}'\bar{v}' \rangle / u_\tau^2 - \langle \tau_{xy} \rangle / u_\tau^2 = 1 - y/\delta$ . NN3–NN5 slightly overpredict the Reynolds shear stress

but underpredict the SGS shear stress. These NN models (NN3–NN5) are forced not to produce the backscatter owing to the clipping as described before. NN1 and NN2 provide backscatter but underpredict it (figure 7d). Note that DSM and SSM also require an averaging over the homogeneous directions and clipping the backscatter, respectively, for stable solution, and thus  $\varepsilon_{SGS}^- = 0$ . Therefore, NN1 is the most-promising SGS model for LES of turbulent channel flow among the NN models considered, even though NN3–NN5 show better prediction performance in *a priori* test. NN1 also shows the best prediction of the mean SGS transport  $\langle T_{SGS} \rangle$  (figure 7e), confirming that a good prediction of  $\langle T_{SGS} \rangle$  is necessary for a successful LES (Völker *et al.* 2002). On the other hand, LESs with all SGS models underpredict the mean SGS dissipation  $\langle \varepsilon_{SGS} \rangle$  (figure 7f), unlike the results of *a priori* test (figure 3b), indicating that an excellent prediction of  $\langle \varepsilon_{SGS} \rangle$  is not a necessary condition for the accurate prediction of the turbulence statistics in LES of turbulent channel flow, as also reported by Park *et al.* (2006).

Figure 8 shows the instantaneous vortical structures identified by the iso-surfaces of  $\lambda_2 = -0.005u_\tau^4/\nu^2$  (Jeong & Hussain 1995). As compared with the flow field from DNS, the arches of the hairpin-like vortices disappear in the fDNS flow field, caused by the larger filter size in the  $x$  direction ( $\Delta_x^+ \approx 70$ ) than the diameter of the arch ( $d^+ \approx 20$ ) (Park *et al.* 2006). The instantaneous flow fields from LESs with DSM and NN1 are similar to that of fDNS, whereas more vortical structures are observed from no SGS model due to insufficient dissipation. As NN1 produces the best results among the NN models considered, we provide the results from NN1 hereafter. Figure 9 shows one-dimensional energy spectra of the velocity fluctuations at  $y^+ = 30$  from LES with NN1, together with those of fDNS and from LES with DSM. Overall agreements of the velocity spectra from NN1 with those of fDNS are very good, such as those from DSM.

Now, we apply NN1 to a turbulent channel flow at a higher Reynolds number of  $Re_b = 27\,600$  ( $Re_\tau = 723$  from DNS). LES is conducted at nearly the same resolution in wall units as that of trained data at  $Re_\tau = 178$  (see table 4). The predictions of  $Re_\tau$  from NN1 and DSM are excellent, showing about 0.8% and 2.2% errors, respectively, whereas the error from no SGS model is about 6%. Figures 10 and 11 show the turbulence statistics and energy spectra from LES723 with NN1, respectively, together with those of fDNS and from LESs with DSM and no SGS model. As shown, NN1 accurately predicts the turbulence statistics and energy spectra even at higher Reynolds number, even though the training is performed at a lower Reynolds number of 178. This result indicates that an NN-based SGS model trained at a lower-Reynolds-number flow maintains their prediction performance for a higher-Reynolds-number flow, once the grid resolution in wall units is kept to be nearly the same (Gamahara & Hattori 2017).

### 3.3. LES with a grid resolution different from that of trained data

We test the performance of NN1 when the grid resolution in LES is different from that of trained data. We consider two different grid resolutions (LES178c and LES178f) as listed in table 4. LESs with NN1 are conducted without and with clipping the backscatter, respectively, to examine how the clipping affects the turbulence statistics for the cases with different resolutions. With LES178c,  $Re_\tau$  is well predicted with and without clipping, whereas  $Re_\tau$  is overpredicted with LES178f by about 8% without clipping but becomes closer to that of DNS with clipping. Predictions of  $Re_\tau$  by DSM are not very good with coarser grids but become very good with denser grids, whereas no SGS model overpredicts  $Re_\tau$ .

Figure 12 shows the changes in the turbulence statistics from NN1 due to different grid resolutions (LES178c and LES178f), together with the statistics from fDNS and

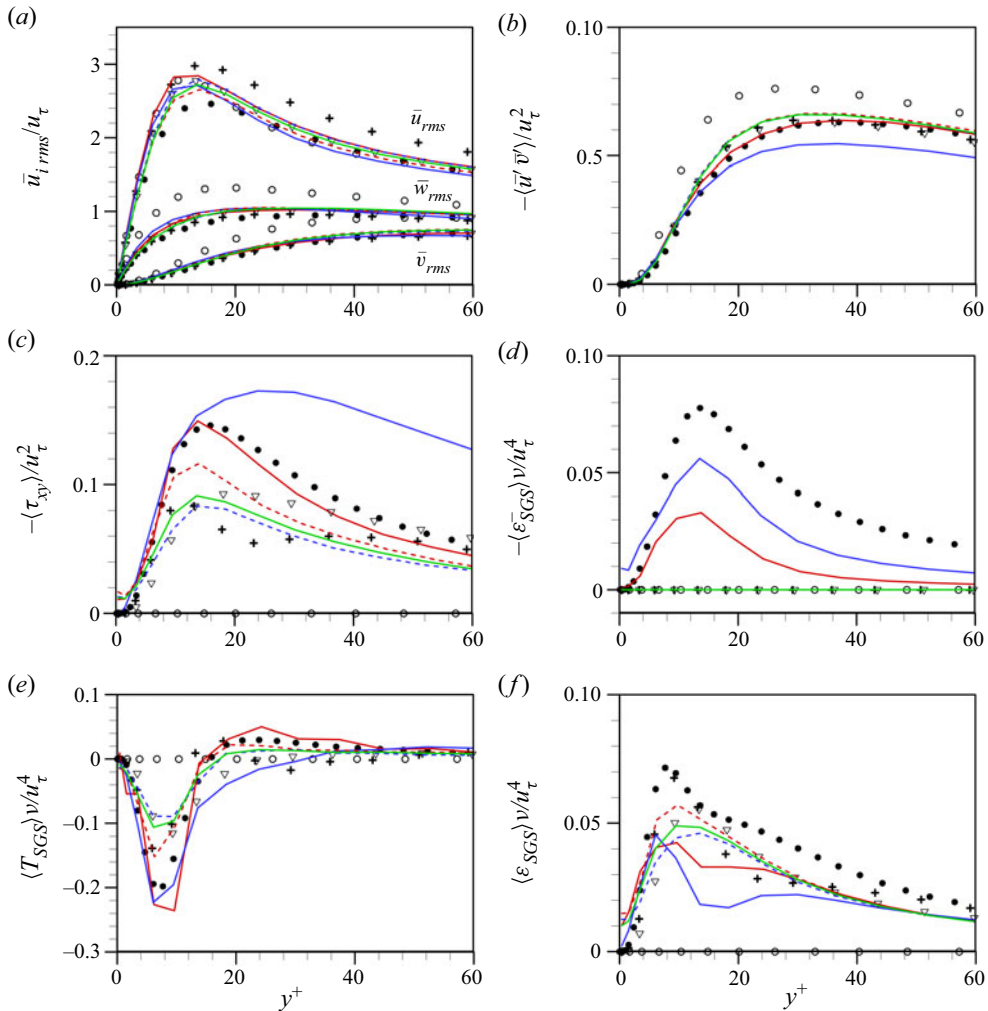


Figure 7. Turbulence statistics from LES178 (*a posteriori* test): (a) rms velocity fluctuations; (b) Reynolds shear stress; (c) mean SGS shear stress; (d) mean backscatter; (e) mean SGS transport; (f) mean SGS dissipation. •, fDNS; red solid line, NN1; blue solid line, NN2; red dashed line, NN3; blue dashed line, NN4; green solid line, NN5; +, DSM; ▽, SSM; ○, no SGS model. Note that the results of NN3–NN5 and SSM are obtained with clipping the backscatter.

LES with DSM. When the grid resolution is coarser (LES178c) than that of trained data, NN1 predicts the mean velocity quite well, but significantly overpredicts the r.m.s. velocity fluctuations and Reynolds shear stress, which is similar to the results from DSM. The backscatter clipping does not improve the results. When the grid resolution is finer (LES178f) than that of trained data, NN1 without clipping significantly underpredicts the mean velocity owing to the increased wall-shear velocity ( $Re_\tau$ ), but reasonably predicts the r.m.s. velocity fluctuations and Reynolds shear stress. As NN1 is trained with  $\bar{S}_{ij}$  and  $\tau_{ij}$  at a given grid resolution, it provides a (trained) amount of energy transfer between the larger and smaller scales than the grid size. Although LES178f is performed at a finer grid resolution, NN1 still provides an amount of energy transfer trained at a coarser grid resolution. This may cause the increase in the amount of energy transfer and accordingly

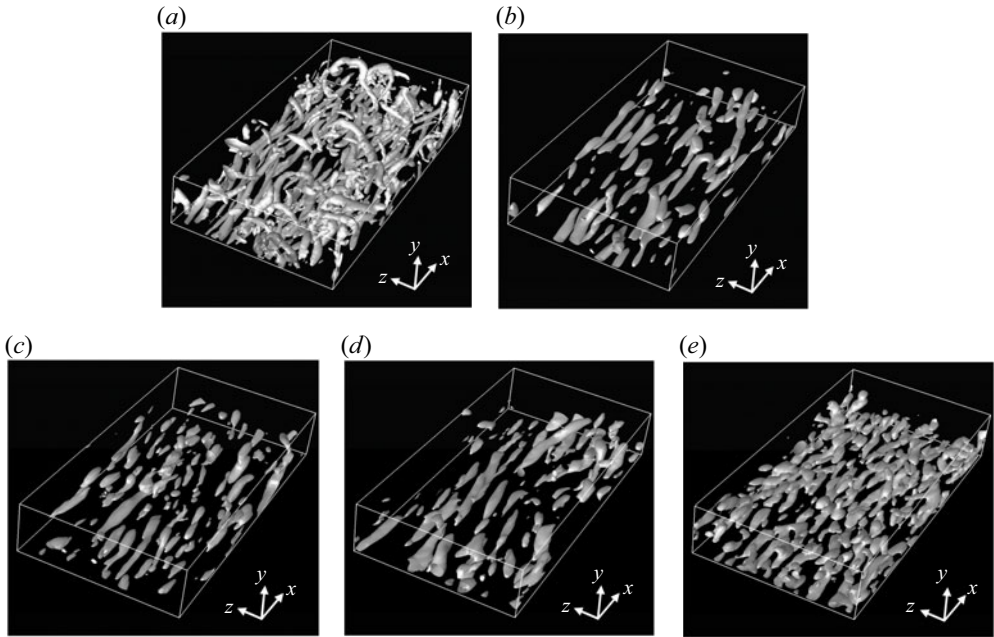


Figure 8. Instantaneous vortical structures from LES178 (*a posteriori* test): (a) DNS; (b) fDNS; (c) NN1; (d) DSM; (e) no SGS model. For the visual clarity, the vortical structures from fDNS and LES are plotted at the same grid resolutions in  $x$  and  $z$  directions as those of DNS by padding high wavenumber components of the velocity with zeros.

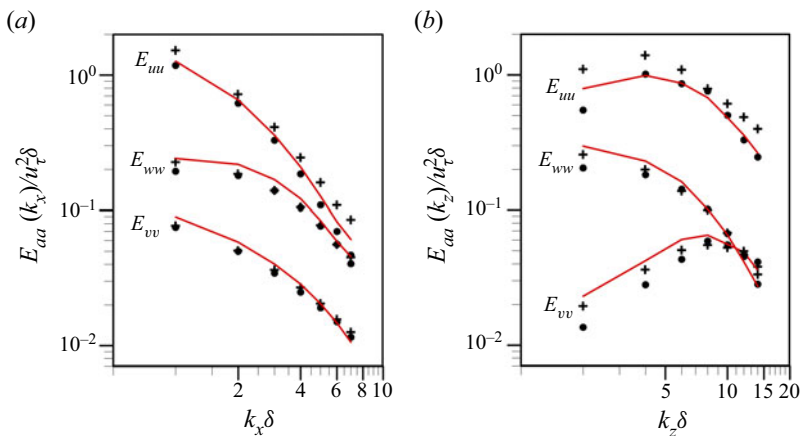


Figure 9. One-dimensional energy spectra of the velocity fluctuations at  $y^+ = 30$  from LES178 (*a posteriori* test): (a) streamwise wavenumber; (b) spanwise wavenumber.  $\bullet$ , fDNS; red solid line, NN1;  $+$ , DSM.

in the wall-shear velocity. On the other hand, when the grid resolution is coarser than that of trained data, the trained amount of energy transfer given to the grid scale is smaller than the real one. For this reason, with clipping the backscatter, changes in the turbulence statistics including the mean velocity are notable for LES178f but not for LES178c.

From this result, it is clear that the NN-based LES requires a special treatment when the grid resolution is different from that of trained data. Thus, we consider an NN1 trained by



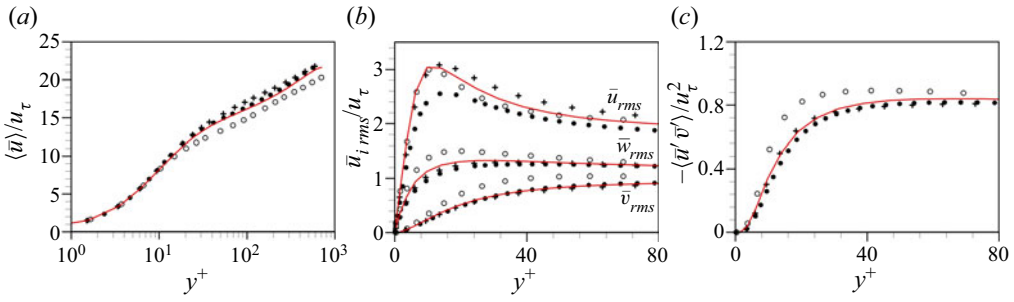


Figure 10. Turbulence statistics from LES723 (*a posteriori* test): (a) mean velocity; (b) r.m.s. velocity fluctuations; (c) Reynolds shear stress. ●, fDNS; red solid line, NN1; +, DSM; ○, no SGS model.

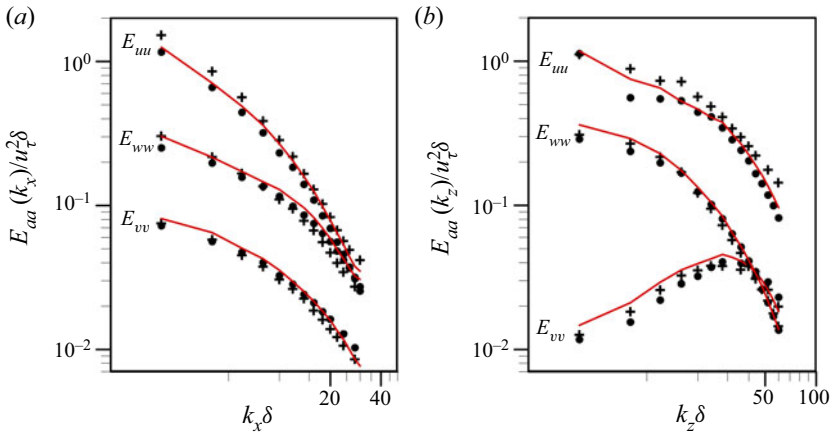


Figure 11. One-dimensional energy spectra of the velocity fluctuations at  $y^+ = 30$  from LES723 (*a posteriori* test): (a) streamwise wavenumber; (b) spanwise wavenumber. ●, fDNS; red solid line, NN1; +, DSM.

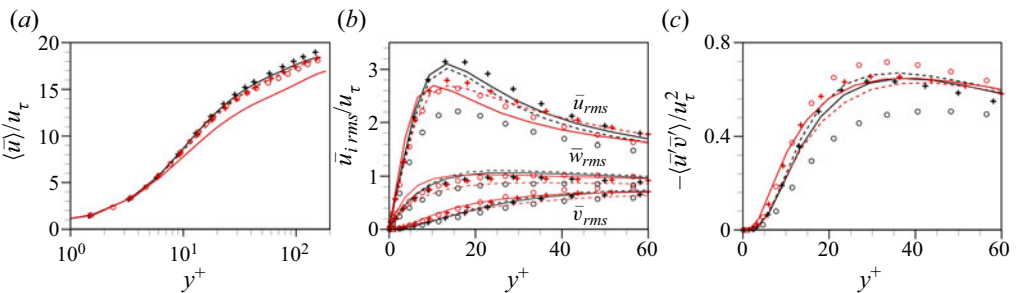


Figure 12. Changes in the turbulence statistics due to different grid resolutions (LES178c and LES178f) (*a posteriori* test): (a) mean velocity; (b) r.m.s. velocity fluctuations; (c) Reynolds shear stress. ○, fDNS; solid line, NN1 without clipping the backscatter; dashed line, NN1 with clipping the backscatter; +, DSM; black lines and symbols are from LES178c; and red lines and symbols are from LES178f.

two fDNS datasets from two different filter sizes with the input and output variables of  $\bar{S}_{ij}$  and  $\tau_{ij}$ . Here, we do not include the filter size as an additional input variable. Table 5 lists the details of various NN1s considered in the present study. NN16 was already tested by LES178c and LES178f. NN12 and NN24 are trained using fDNS datasets having the same

NN1 model	fDNS(s) for trained data	LES(s)
NN <sub>16</sub>	fDNS <sub>16</sub>	LES178c and LES178f
NN <sub>12</sub>	fDNS <sub>12</sub>	LES178c
NN <sub>8,16</sub>	fDNS <sub>8</sub> and fDNS <sub>16</sub>	LES178c
NN <sub>24</sub>	fDNS <sub>24</sub>	LES178f
NN <sub>16,32</sub>	fDNS <sub>16</sub> and fDNS <sub>32</sub>	LES178f

Table 5. NN1 trained with different fDNS datasets. Here, fDNS<sub>N</sub> denotes the fDNS data with the number of grid points  $N$  ( $=N_x = N_z$ ). Note that the numbers of grid points ( $N_x \times N_z$ ) for LES178c and LES178f are  $12 \times 12$  and  $24 \times 24$ , respectively, as listed in table 4.

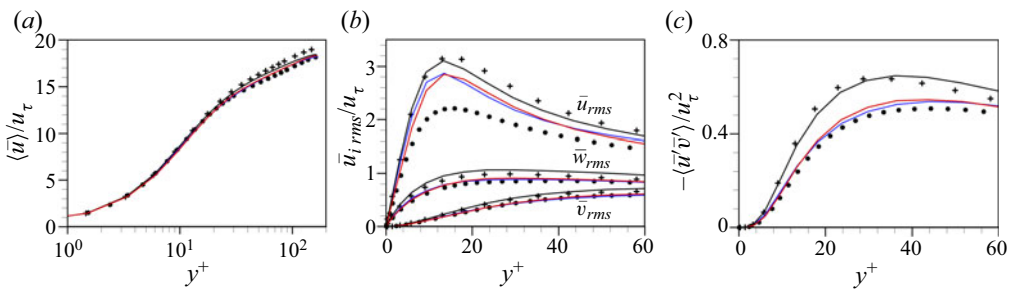


Figure 13. Turbulence statistics from LES178c (*a posteriori* test): (a) mean velocity; (b) r.m.s. velocity fluctuations; (c) Reynolds shear stress. ●, fDNS; black solid line, NN<sub>16</sub>; blue solid line, NN<sub>12</sub>; red solid line, NN<sub>8,16</sub>; +, DSM.

grid resolutions as those of LES178c and LES178f, respectively. On the other hand, NN<sub>8,16</sub> is trained by both fDNS datasets with two filters corresponding to larger and smaller sizes ( $N = 8$  and  $16$ ) than the grid resolution in LES178c ( $N = 12$ ), and NN<sub>16,32</sub> is trained by two fDNS datasets with  $N = 16$  and  $32$  ( $N = 24$  for LES178f), respectively.

Now, we conduct LES178c with NN<sub>12</sub> and NN<sub>8,16</sub>, and LES178f with NN<sub>24</sub> and NN<sub>16,32</sub>, respectively, and compare the results with those of fDNS and from LESs with NN<sub>16</sub> and DSM. All LESs with NNs are conducted without clipping the backscatter. Figure 13 shows the results of LES178c. As shown, LES with NN<sub>8,16</sub> provides much more accurate predictions of the r.m.s. velocity fluctuations and Reynolds shear stress than those from NN<sub>16</sub> and DSM, showing the performance almost similar to that from NN<sub>12</sub>. In the case of LES178f (figure 14), NN<sub>16,32</sub> shows better prediction performance for the mean velocity, r.m.s. velocity fluctuations and Reynolds shear stress than those from NN<sub>16</sub>, and has similar predictions to those from NN<sub>24</sub>. We have also tested NN<sub>8,16</sub> for LES178f ( $N = 24$ ) and NN<sub>16,32</sub> for LES178c ( $N = 12$ ), respectively. In these cases, LESs with NN<sub>8,16</sub> and NN<sub>16,32</sub> do not show better performance than that with NN<sub>16</sub>. Therefore, when the resolution in LES is not similar to that of trained data, it is suggested that the datasets having two different resolutions, coarser and finer than that of LES, should be constructed and used to train an NN for successful LES.

#### 4. Conclusions

We have applied a fully connected NN to the development of a SGS model of predicting the SGS stresses for a turbulent channel flow, and conducted *a priori* and *a posteriori* tests to estimate its prediction performance. Five different NNs with different input variables have

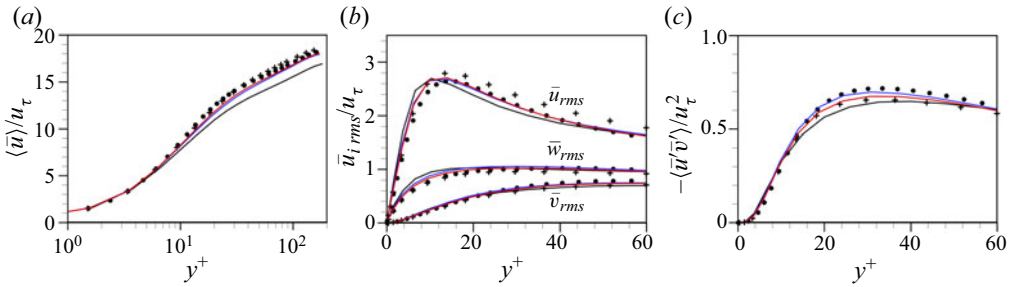


Figure 14. Turbulence statistics from LES178f (*a posteriori* test): (a) mean velocity; (b) r.m.s. velocity fluctuations; (c) Reynolds shear stress. ●, fDNS; black solid line, NN<sub>16</sub>; blue solid line, NN<sub>24</sub>; red solid line, NN<sub>16,32</sub>; +, DSM.

been trained with fDNS data at  $Re_\tau = 178$  using a spectral cut-off filter, where the input variables considered were the strain-rate tensor at single and multiple grid points (NN1 and NN3, respectively), velocity gradient tensor at single and multiple points (NN2 and NN4, respectively) and the velocity and wall-normal velocity gradient vectors at multiple points (NN5), respectively.

In *a priori* tests, the NN-based SGS models with the input variables at multiple grid points (NN3, NN4 and NN5) had higher correlations between the true and predicted SGS stresses, and better predicted backscatter than those with the input variables at single grid point (NN1 and NN2). However, actual LESs (*i.e. a posteriori* tests) with NN3–NN5 were unstable unless a special treatment such as the backscatter clipping was taken. On the other hand, NN1 and NN2 showed excellent prediction performance without any *ad hoc* clipping or wall damping function, although the correlations between the true and predicted SGS stresses were relatively low. Among NN models considered, NN1 (input of the strain-rate tensor at single grid point) performed best, and thus we applied NN1 (trained at  $Re_\tau = 178$ ) to LES at a higher Reynolds number of  $Re_\tau = 723$  with the same grid resolution in wall units, providing successful results. Finally, we applied NN1 to LESs at  $Re_\tau = 178$  with coarser and finer grid resolutions, respectively. Although the results were generally good as compared with those from LES with the DSM, they clearly showed a limitation in accurately predicting the turbulence statistics when LES was conducted with a resolution different from that used for training NN. To overcome this limitation, NN1 was trained by fDNS datasets with two filter sizes (larger and smaller than the grid size in LES), providing a successful result. Therefore, once multiple filtered datasets with various filter sizes are constructed and used to train an NN, one may expect a successful NN-based LES for turbulent channel flow, even if the grid resolution at hand is different from those used to construct the NN.

Now, let us discuss current limitations of NN-based LES and future research directions. Some limitations were also reported in Wollblad & Davidson (2008), Gamahara & Hattori (2017) and Zhou *et al.* (2019). First, the performance of NN-based SGS model depends on the input variables. In the present study, we considered the filtered strain-rate tensor  $\bar{S}_{ij}$  and filtered velocity gradient tensor  $\bar{\alpha}_{ij}$  as input variables, and showed that  $\bar{S}_{ij}$  performs better than  $\bar{\alpha}_{ij}$ . As the SGS stress tensor  $\tau_{ij}$  is a symmetric tensor, one may also consider other combinations of  $\bar{S}_{ij}$  and  $\bar{R}_{ij}$  (filtered rotation rate tensor) as input variables, as described in § 2.1. Thus, a further study in this direction is needed. Second, the results of *a priori* and *a posteriori* tests on NN-based SGS models are inconsistent with each other. Traditional physics-based SGS models have also the

same inconsistency. That is, some traditional SGS models having a poor performance in *a priori* test perform very well in *a posteriori* test. However, this poor performance in *a priori* test does not mean the failure of such models, but indicates the fundamental limitation of *a priori* test itself (Park *et al.* 2005). The present NN-based SGS model is constructed using a database containing static (i.e. instantaneous) flow information, thus lacking dynamic (i.e. temporal) information of filtered flow variables which is important in actual LES (i.e. *a posteriori* test). Therefore, the present model is not free from the inconsistency observed in traditional SGS models, and a database containing more static information does not necessarily provide better output. In this regard, a different approach of constructing NN-based SGS models may be searched for. In traditional physics-based SGS modelling, Meneveau *et al.* (1996) proposed to accumulate the flow information over flow pathlines and constructed a Lagrangian dynamic SGS model. Thus, a Lagrangian approach or reinforcement learning with a target statistics may be a way to overcome this inconsistency. To the best of the authors' knowledge, there has been no attempt to construct such an NN-based SGS model. This approach may provide an improved performance in NN-based LES. Third, an NN-based SGS model should be trained by databases containing different flow characteristics such as shear-driven, rotation-driven and separated flow characteristics. The present SGS model was trained by a database of turbulent channel flow, and thus may not be applicable to other types of flows. Thus, more databases should be generated and used for training an NN. Here, we do not mean that almost all the flow databases should be trained for successful LES, but we suggest that some representative flow databases such as rotating channel flow, flow over a backward-facing step, flow over a circular cylinder and jet may be sufficient to build a successful NN for flow inside/over a complex geometry. However, how to combine different flow databases in an NN-based SGS model is still a difficult problem. The present NN-based SGS model was trained by the input and output variables normalized by wall units, but it may not be applicable to complex flow (e.g. a circular cylinder) because this flow cannot be scaled in wall units. To overcome this limitation, one should develop a universal non-dimensionalization of input and output variables for different flow types. This is an important task for the use of NN-based SGS model to flow inside/over a complex geometry.

**Funding.** This work is supported by the National Research Foundation through the Ministry of Science and ICT (grant numbers 2019R1A2C2086237 and 2017M2A8A4018482). The computing resources are provided by the KISTI Super Computing Center (grant number KSC-2019-CRE-0114).

**Declaration of interests.** The authors report no conflict of interest.

**Author ORCIDs.**

 Haecheon Choi <https://orcid.org/0000-0003-0696-847X>.

## Appendix. *A priori* and *a posteriori* tests by NNs with different numbers of hidden layers

Figure 15 shows the effects of the number of hidden layers of NNs (NN1, NN3 and NN5) on the mean SGS and Reynolds shear stresses from *a priori* and *a posteriori* tests, respectively. The Reynolds shear stress from NN3 and NN5 are obtained from LES with clipping the backscatter. For all NNs considered, one hidden layer is not sufficient for accurately predicting the mean SGS shear stress, and at least two hidden layers are required. In actual LES, one hidden layer seems to be sufficient for NN1 and NN3, and two hidden layers are required for NN5. Therefore, two hidden layers are taken for the present study for all NNs considered.

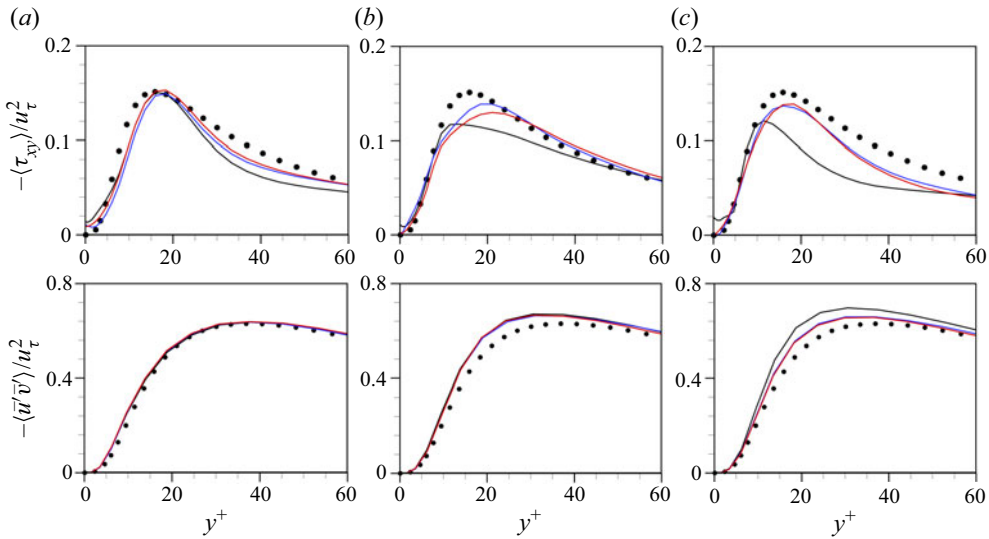


Figure 15. Effects of the number of hidden layers ( $N_{hl}$ ) on the mean SGS shear stress (*a priori* test) and Reynolds shear stress (*a posteriori* test; LES178): (a) NN1; (b) NN3; (c) NN5. ●, DNS; black solid line,  $N_{hl} = 1$ ; blue solid line,  $N_{hl} = 2$ ; red solid line,  $N_{hl} = 3$ .

#### REFERENCES

- ADRIAN, R.J. 1990 Stochastic estimation of sub-grid scale motions. *Appl. Mech. Rev.* **43**, 214–218.
- ADRIAN, R.J., JONES, B.G., CHUNG, M.K., HASSAN, Y., NITHIANANDAN, C.K. & TUNG, A.T.-C. 1989 Approximation of turbulent conditional averages by stochastic estimation. *Phys. Fluids* **1**, 992–998.
- AKHAVAN, R., ANSARI, A., KANG, S. & MANGIAVACCHI, N. 2000 Subgrid-scale interactions in a numerically simulated planar turbulent jet and implications for modelling. *J. Fluid Mech.* **408**, 83–120.
- ANDERSON, B.W. & DOMARADZKI, J.A. 2012 A subgrid-scale model for large-eddy simulation based on the physics of interscale energy transfer in turbulence. *Phys. Fluids* **24**, 065104.
- BARDINA, J., FERZIGER, J.H. & REYNOLDS, W.C. 1980 Improved subgrid scale models for large-eddy simulation. *AIAA Paper* 80-1357. American Institute of Aeronautics and Astronautics.
- BECK, A., FLAD, D. & MUNZ, C. 2019 Deep neural networks for data-driven LES closure models. *J. Comput. Phys.* **398**, 108910.
- BEDFORD, K.W. & YEO, W.K. 1993 Conjunctive filtering procedures in surface water flow and transport. In *Large Eddy Simulation of Complex Engineering and Geophysical Flows* (ed. B. Galperin & S.A. Orszag), pp. 513–539. Cambridge University Press.
- CLARK, R.A., FERZIGER, J.H. & REYNOLDS, W.C. 1979 Evaluation of subgrid-scale models using an accurately simulated turbulent flow. *J. Fluid Mech.* **91**, 1–16.
- DOMARADZKI, J.A. & SAIKI, E.M. 1997 A subgrid-scale model based on the estimation of unresolved scales of turbulence. *Phys. Fluids* **9**, 2148–2164.
- GAMAHARA, M. & HATTORI, Y. 2017 Searching for turbulence models by artificial neural network. *Phys. Rev. Fluids* **2**, 054604.
- GERMANO, M., PIOMELLI, U., MOIN, P. & CABOT, W.H. 1991 A dynamic subgrid-scale eddy viscosity model. *Phys. Fluids* **3**, 1760–1765.
- GHOSAL, S., LUND, T.S., MOIN, P. & AKSELVOLL, K. 1995 A dynamic localization model for large-eddy simulation of turbulent flows. *J. Fluid Mech.* **286**, 229–255.
- HÄRTEL, C., KLEISER, L., UNGER, F. & FRIEDRICH, R. 1994 Subgrid-scale energy transfer in the near-wall region of turbulent flows. *Phys. Fluids* **6**, 3130–3143.
- HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J.H. 2009 *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edn. Springer.
- HORIUTI, K. 1997 A new dynamic two-parameter mixed model for large-eddy simulation. *Phys. Fluids* **9**, 3443–3464.
- IOFFE, S. & SZEGEDY, C. 2015 Batch normalization: accelerating deep network training by reducing internal covariate shift. [arXiv:1502.03167](https://arxiv.org/abs/1502.03167).

- JEONG, J. & HUSSAIN, F. 1995 On the identification of a vortex. *J. Fluid Mech.* **285**, 69–94.
- KIM, J., MOIN, P. & MOSER, R. 1987 Turbulence statistics in fully developed channel flow at low Reynolds number. *J. Fluid Mech.* **177**, 133–166.
- KINGMA, D.P. & BA, J. 2014 Adam: a method for stochastic optimization. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- LANGFORD, J.A. & MOSER, R.D. 1999 Optimal LES formulations for isotropic turbulence. *J. Fluid Mech.* **398**, 321–346.
- LANGFORD, J.A. & MOSER, R.D. 2004 Optimal large-eddy simulation results for isotropic turbulence. *J. Fluid Mech.* **521**, 273–294.
- LECUN, Y., BENGIO, Y. & HINTON, G.E. 2015 Deep learning. *Nature* **521**, 436–444.
- LEE, J., CHOI, H. & PARK, N. 2010 Dynamic global model for large eddy simulation of transient flow. *Phys. Fluids* **22**, 075106.
- LEITH, C.E. 1968 Diffusion approximation for two-dimensional turbulence. *Phys. Fluids* **11**, 671–672.
- LILLY, D.K. 1992 A proposed modification of the Germano subgrid-scale closure method. *Phys. Fluids* **4**, 633–635.
- LIU, S., MENEVEAU, C. & KATZ, J. 1994 On the properties of similarity subgrid-scale models as deduced from measurements in a turbulent jet. *J. Fluid Mech.* **275**, 83–119.
- LIU, S., MENEVEAU, C. & KATZ, J. 1995 Experimental study of similarity subgrid-scale models of turbulence in the far-field of a jet. *Appl. Sci. Res.* **54**, 177–190.
- LUND, T.S. & NOVIKOV, E.A. 1992 Parameterization of subgrid-scale stress by the velocity gradient tensor. In *Annual Research Briefs, Center for Turbulence Research, Stanford University*, pp. 27–43.
- MAULIK, R. & SAN, O. 2017 A neural network approach for the blind deconvolution of turbulent flows. *J. Fluid Mech.* **831**, 151–181.
- MAULIK, R., SAN, O., RASHEED, A. & VEDULA, P. 2018 Data-driven deconvolution for large eddy simulations of Kraichnan turbulence. *Phys. Fluids* **30**, 125109.
- MAULIK, R., SAN, O., RASHEED, A. & VEDULA, P. 2019 Subgrid modelling for two-dimensional turbulence using neural networks. *J. Fluid Mech.* **858**, 122–144.
- MENEVEAU, C. & KATZ, J. 2000 Scale-invariance and turbulence models for large-eddy simulation. *Annu. Rev. Fluid Mech.* **32**, 1–32.
- MENEVEAU, C., LUND, T.S. & CABOT, W.H. 1996 A Lagrangian dynamic subgrid-scale model of turbulence. *J. Fluid Mech.* **319**, 353–385.
- MOSER, R.D., MALAYA, N.P., CHANG, H., ZANDONADE, P.S., VEDULA, P., BHATTACHARYA, A. & HASELBACHER, A. 2009 Theoretically based optimal large-eddy simulation. *Phys. Fluids* **21**, 105104.
- NAIR, V. & HINTON, G.E. 2010 Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning* (ed. J. Fürnkranz & T. Joachims), pp. 807–814. Omnipress.
- NICOUD, F. & DUCROS, F. 1999 Subgrid-scale stress modelling based on the square of the velocity gradient tensor. *Flow Turbul. Combust.* **62**, 183–200.
- NICOUD, F., TODA, H.B., CABRIT, O., BOSE, S. & LEE, J. 2011 Using singular values to build a subgrid-scale model for large eddy simulations. *Phys. Fluids* **23**, 085106.
- PAL, A. 2019 Deep learning parameterization of subgrid scales in wall-bounded turbulent flows. [arXiv:1905.12765](https://arxiv.org/abs/1905.12765).
- PARK, N., LEE, S., LEE, J. & CHOI, H. 2006 A dynamic subgrid-scale eddy viscosity model with a global model coefficient. *Phys. Fluids* **18**, 125109.
- PARK, N., YOO, J.Y. & CHOI, H. 2005 Toward improved consistency of *a priori* tests with *a posteriori* tests in large eddy simulation. *Phys. Fluids* **17**, 015103.
- PASSALIS, N., TEFA, A., KANNIAINEN, J., GABBOUJ, M. & IOSIFIDIS, A. 2019 Deep adaptive input normalization for time series forecasting. [arXiv:1902.07892](https://arxiv.org/abs/1902.07892).
- PIOMELLI, U., CABOT, W.H., MOIN, P. & LEE, S. 1991 Subgrid-scale backscatter in turbulent and transitional flows. *Phys. Fluids* **3**, 1766–1771.
- PIOMELLI, U. & LIU, J. 1995 Large-eddy simulation of rotating channel flows using a localized dynamic model. *Phys. Fluids* **7**, 839–848.
- PIOMELLI, U., YU, Y. & ADRIAN, R.J. 1996 Subgrid-scale energy transfer and near-wall turbulence structure. *Phys. Fluids* **8**, 215–224.
- ROZEMA, W., BAE, H.J., MOIN, P. & VERSTAPPEN, R. 2015 Minimum-dissipation models for large-eddy simulation. *Phys. Fluids* **27**, 085107.
- RUMELHART, D.E., HINTON, G.E. & WILLIAMS, R.J. 1986 Learning representations by back-propagating errors. *Nature* **323**, 533–536.
- SALVETTI, M.V. & BANERJEE, S. 1995 *A priori* tests of a new dynamic subgrid-scale model for finite-difference large-eddy simulations. *Phys. Fluids* **7**, 2831–2847.

- SARGHINI, F., DE FELICE, G. & SANTINI, S. 2003 Neural networks based subgrid scale modeling in large eddy simulations. *Comput. Fluids* **32**, 97–108.
- SARGHINI, F., PIOMELLI, U. & BALARAS, E. 1999 Scale-similar models for large-eddy simulations. *Phys. Fluids* **11**, 1596–1607.
- SILVIS, M.H., BAE, H.J., TRIAS, F.X., ABKAR, M. & VERSTAPPEN, R. 2019 A nonlinear subgrid-scale model for large-eddy simulations of rotating turbulent flows. [arXiv:1904.12748](https://arxiv.org/abs/1904.12748).
- SILVIS, M.H., REMMERSWAAL, R.A. & VERSTAPPEN, R. 2017 Physical consistency of subgrid-scale models for large-eddy simulation of incompressible turbulent flows. *Phys. Fluids* **29**, 015105.
- SMAGORINSKY, J. 1963 General circulation experiments with the primitive equations. *Mon. Weath. Rev.* **91**, 99–164.
- STOLZ, S. & ADAMS, N.A. 1999 An approximate deconvolution procedure for large-eddy simulation. *Phys. Fluids* **11**, 1699–1701.
- THIRY, O. & WINCKELMANS, G. 2016 A mixed multiscale model better accounting for the cross term of the subgrid-scale stress and for backscatter. *Phys. Fluids* **28**, 025111.
- TRIAS, F.X., FOLCH, D., GOROBETS, A. & OLIVA, A. 2015 Building proper invariants for eddy-viscosity subgrid-scale models. *Phys. Fluids* **27**, 065103.
- VERSTAPPEN, R. 2011 When does eddy viscosity damp subfilter scales sufficiently? *J. Sci. Comput.* **49**, 94–110.
- VERSTAPPEN, R.W.C.P., BOSE, S.T., LEE, J., CHOI, H. & MOIN, P. 2010 A dynamic eddy-viscosity model based on the invariants of the rate-of-strain. In *Proceedings of the Summer Program 2010 (Center for Turbulence Research, Stanford University)*, pp. 183–192.
- VÖLKER, S., MOSER, R.D. & VENUGOPAL, P. 2002 Optimal large eddy simulation of turbulent channel flow based on direct numerical simulation statistical data. *Phys. Fluids* **14**, 3675–3691.
- VOLLANT, A., BALARAC, G. & CORRE, C. 2017 Subgrid-scale scalar flux modelling based on optimal estimation theory and machine-learning procedure. *J. Turbul.* **18**, 854–878.
- VREMAN, A.W. 2004 An eddy-viscosity subgrid-scale model for turbulent shear flow: algebraic theory and applications. *Phys. Fluids* **16**, 3670–3681.
- VREMAN, B., GEURTS, B. & KUERTEN, H. 1994 On the formulation of the dynamic mixed subgrid-scale model. *Phys. Fluids* **6**, 4057–4059.
- VREMAN, B., GEURTS, B. & KUERTEN, H. 1997 Large-eddy simulation of the turbulent mixing layer. *J. Fluid Mech.* **339**, 357–390.
- WANG, Z., LUO, K., LI, D., TAN, J. & FAN, J. 2018 Investigations of data-driven closure for subgrid-scale stress in large-eddy simulation. *Phys. Fluids* **30**, 125101.
- WOLLBLAD, C. & DAVIDSON, L. 2008 POD based reconstruction of subgrid stresses for wall bounded flows using neural networks. *Flow Turbul. Combust.* **81**, 77–96.
- XIE, C., LI, K., MA, C. & WANG, J. 2019a Modeling subgrid-scale force and divergence of heat flux of compressible isotropic turbulence by artificial neural network. *Phys. Rev. Fluids* **4**, 104605.
- XIE, C., WANG, J. & E, W. 2020a Modeling subgrid-scale forces by spatial artificial neural networks in large eddy simulation of turbulence. *Phys. Rev. Fluids* **5**, 054606.
- XIE, C., WANG, J., LI, H., WAN, M. & CHEN, S. 2019b Artificial neural network mixed model for large eddy simulation of compressible isotropic turbulence. *Phys. Fluids* **31**, 085112.
- XIE, C., WANG, J., LI, H., WAN, M. & CHEN, S. 2020b Spatially multi-scale artificial neural network model for large eddy simulation of compressible isotropic turbulence. *AIP Adv.* **10**, 015044.
- XIE, C., WANG, J., LI, K. & MA, C. 2019c Artificial neural network approach to large-eddy simulation of compressible isotropic turbulence. *Phys. Rev. E* **99**, 053113.
- YOU, D. & MOIN, P. 2007 A dynamic global-coefficient subgrid-scale eddy-viscosity model for large-eddy simulation in complex geometries. *Phys. Fluids* **19**, 065110.
- ZANDONADE, P.S., LANGFORD, J.A. & MOSER, R.D. 2004 Finite-volume optimal large-eddy simulation of isotropic turbulence. *Phys. Fluids* **16**, 2255–2271.
- ZANG, Y., STREET, R.L. & KOSEFF, J.R. 1993 A dynamic mixed subgrid-scale model and its application to turbulent recirculating flows. *Phys. Fluids* **5**, 3186–3196.
- ZHOU, Z., HE, G., WANG, S. & JIN, G. 2019 Subgrid-scale model for large-eddy simulation of isotropic turbulent flows using an artificial neural network. *Comput. Fluids* **195**, 104319.