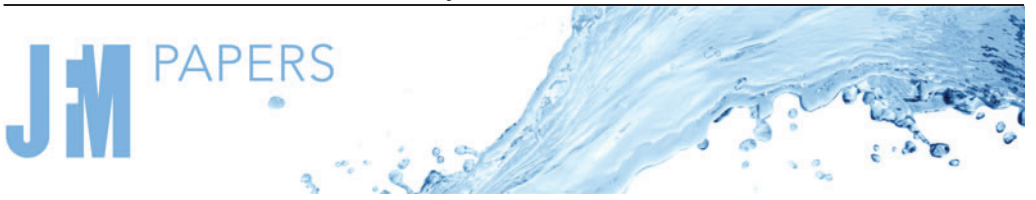
J. Fluid Mech. (2023), *vol*. 960, A4, doi:10.1017/jfm.2023.179



Artificial-neural-network-based nonlinear algebraic models for large-eddy simulation of compressible wall-bounded turbulence

Dehao Xu[1], Jianchun Wang[2,†], Changping Yu[3] and Shiyi Chen[1,2,4,†]

[1]State Key Laboratory of Turbulence and Complex Systems, College of Engineering, Peking University, Beijing 100871, PR China

[2]Department of Mechanics and Aerospace Engineering, Southern University of Science and Technology, Shenzhen 518055, PR China

[3]Laboratory of High Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, PR China

[4]Eastern Institute for Advanced Study, Ningbo 315200, PR China

(Received 2 September 2022; revised 22 February 2023; accepted 26 February 2023)

In this paper, we propose artificial-neural-network-based (ANN-based) nonlinear algebraic models for the large-eddy simulation (LES) of compressible wall-bounded turbulence. An innovative modification is applied to the invariants and the tensor bases of the nonlinear algebraic models through using the local grid widths along each direction to normalise the corresponding gradients of the flow variables. Furthermore, the dimensionless model coefficients are determined by the ANN method. The modified ANN-based nonlinear algebraic model (MANA model) has much higher correlation coefficients and much lower relative errors than the dynamic Smagorinsky model (DSM), Vreman model and wall-adapting local eddy-viscosity model in the *a priori* test. The significantly more accurate estimations of the mean subgrid-scale (SGS) fluxes of the kinetic energy and temperature variance are also obtained by the MANA models in the *a priori* test. Furthermore, in the *a posteriori* test, the MANA model can give much more accurate predictions of the flow statistics and the mean SGS fluxes of the kinetic energy and the temperature variance than other traditional eddy-viscosity models in compressible turbulent channel flows with untrained Reynolds numbers, Mach numbers and grid resolutions. The MANA model has a better performance in predicting the flow statistics in supersonic turbulent boundary layer. The MANA model can well predict both direct and inverse transfer of the kinetic energy and temperature variance, which overcomes the inherent shortcoming that the traditional eddy-viscosity models cannot predict the inverse energy transfer. Moreover, the MANA model is computationally more efficient than the DSM.

† Email addresses for correspondence: wangjc@sustech.edu.cn, chensy@sustech.edu.cn

Key words: machine learning, turbulence modelling

1. Introduction

Turbulent flows have the intrinsic multi-scale characteristics across several orders of magnitude, therefore direct numerical simulation (DNS) of turbulence at high Reynolds numbers is extremely time consuming and not tractable to solve all flow scales ranging from the dissipation range to the energy-containing range. Accordingly, the large-eddy simulation (LES) is an effective approach in which only the large-scale motions are resolved while the effect of the residual subgrid-scale (SGS) structures on the resolved large-scale motions are modelled by an SGS model (Lesieur & Metais 1996; Meneveau & Katz 2000; Sagaut 2006; Garnier, Adams & Sagaut 2009; Durbin 2018; Moser, Haering & Yalla 2021). Extensive SGS models have been proposed to reconstruct the unclosed SGS terms in previous works, and the eddy-viscosity SGS models (Smagorinsky 1963; Deardorff 1970; Nicoud & Ducros 1999; Vreman 2004) are often utilised due to the dissipative nature of the SGS stress tensor. In eddy-viscosity SGS models, the Smagorinsky model (Smagorinsky 1963; Deardorff 1970) is one of the most widely used SGS models whose model coefficient is statically calibrated by the experimental and numerical data. Germano *et al.* (1991) pioneered the famous dynamic procedure based on the Germano identity, which makes the coefficients of the dynamic Smagorinsky model (DSM) dynamically modified as the flow evolves. Subsequently, other improved and generalised dynamical versions of the Smagorinsky model were introduced by Lilly (1992), Piomelli (1993), Ghosal *et al.* (1995) and Meneveau, Lund & Cabot (1996). Furthermore, Nicoud & Ducros (1999) proposed a wall-adapting local eddy-viscosity model (WALE), which has a good performance in wall-bounded turbulence. Moreover, Vreman (2004) introduced another eddy-viscosity model with quite satisfactory results in turbulent shear flows.

The eddy-viscosity models assume that the SGS stress is linearly aligned with the filtered strain-rate tensor based on the Boussinesq hypothesis. Therefore, these eddy-viscosity models have a serious drawback that they correlate poorly with the true SGS stress (Anderson & Meneveau 1999; Da Silva & Métais 2002) and cannot predict the energy backscatter to the resolved scales (Zang, Street & Koseff 1993). Furthermore, the linear alignment assumption in the eddy-viscosity models is unphysical and cannot estimate the complicated nonlinear relation between the filtered strain-rate tensor and the SGS stress (Anderson & Meneveau 1999; Da Silva & Métais 2002; Xie *et al.* 2019*a*,*c*; Xie, Wang & Weinan 2020*a*; Xie, Yuan & Wang 2020*b*). In order to alleviate these problems, Pope (1975) derived a general nonlinear algebraic expression between the Reynolds stress and the averaged strain-rate and rotation-rate tensors with ten integrity bases based on the theory of invariants. Furthermore, Lund & Novikov (1992) reduced the ten integrity basis tensors to five, which significantly simplified the computational complexity of the nonlinear algebraic SGS model. Similarly, Wang *et al.* (2007) proposed a dynamic nonlinear tensorial diffusivity model to model the SGS scalar flux in the incompressible turbulence. Then, how to determine the model coefficients in the nonlinear algebraic models is a significant issue. In the early stages, the model coefficients were mostly determined by the experimental and numerical data. Recently, the dynamical versions of these nonlinear algebraic SGS models (Wang *et al.* 2007; Yuan *et al.* 2022) were proposed and exhibited a reasonable agreement with the filtered DNS (fDNS) data. Furthermore, Wang *et al.* (2021*b*) proposed a dynamic spatial gradient

model, which has good performance in both *a priori* and *a posteriori* tests in LES of incompressible isotropic turbulence. However, the dynamic approach requires some *ad hoc* procedures, including the averaging of model coefficients in statistically homogeneous directions and the clipping of the negative model coefficients to avoid numerical instability.

Recently, the artificial neural network (ANN) has become a popular approach to develop SGS models (Gamahara & Hattori 2017; Vollant, Balarac & Corre 2017; Beck, Flad & Munz 2019; Maulik *et al.* 2019; Xie *et al.* 2019*a*,*c*; Yang *et al.* 2019; Zhou *et al.* 2019; Xie *et al.* 2020*b*; Yuan, Xie & Wang 2020; Frezat *et al.* 2021; Park & Choi 2021; Yu *et al.* 2022). The ANN method can construct a complex nonlinear mapping between some resolved flow variables and the unresolved SGS terms. Therefore, the SGS models based on the ANN method are expected to give more accurate prediction than the traditional SGS models, and have a good performance in *a posteriori* tests of compressible and incompressible isotropic turbulence (Xie *et al.* 2019*a*,*c*, 2020*b*) and incompressible channel flow (Park & Choi 2021). It is worth noting that the ANN-based SGS models can get rid of the averaging procedure of the model coefficients in statistically homogeneous directions. Furthermore, the ANN-based SGS models in compressible and incompressible isotropic turbulence (Xie *et al.* 2019*a*,*c*, 2020*b*) and the single-point ANN-based SGS model in the incompressible channel flow (Park & Choi 2021) are free from backscatter clipping, while the multi-point ANN-based SGS model in incompressible channel flow (Park & Choi 2021) requires the backscatter clipping procedure.

The ANN method was also used to determine the model coefficients in the nonlinear algebraic models. Ling, Kurzawski & Templeton (2016) firstly used the deep neural networks to predict the model coefficients of the ten isotropic basis tensors given by Pope (1975) to obtain an improved Reynolds-averaged Navier–Stokes model. Similarly, Xie *et al.* (2020*b*) proposed ANN-based nonlinear algebraic SGS models in LES of incompressible isotropic turbulence. They constructed a neural network mapping between four dimensionless invariants and five model coefficients given by Lund & Novikov (1992), and this new model exhibited a better performance in *a priori* and *a posteriori* tests than traditional SGS models. Moreover, Vollant *et al.* (2017) proposed a new ANN-based SGS model for the LES of a passive scalar in the incompressible turbulence. Wang *et al.* (2021*a*) proposed an artificial neural network-based spatial gradient models in LES of incompressible isotropic turbulence, which used the ANN method to estimate the coefficients of the spatial gradient models.

The above studies mainly concentrated on ANN-based LES of incompressible turbulence; however, there are much fewer studies aiming to apply the ANN method to develop the well-performing SGS models for the compressible wall-bounded flows. It is worth noting that, due to the remarkable influence of compressibility, the underlying mechanisms of the compressible wall-bounded turbulence are much more complex than those in incompressible turbulence (Maeder, Adams & Kleiser 2001; Pirozzoli, Grasso & Gatski 2004; Pirozzoli, Bernardini & Grasso 2008; Duan, Beekman & Martin 2010; Pirozzoli, Bernardini & Grasso 2010; Lagha *et al.* 2011; Pirozzoli & Bernardini 2011; Xu *et al.* 2021*a*,*b*, 2022*b*,*c*; Xu, Wang & Chen 2022*a*), and the study of improved SGS models in compressible wall-bounded flows is of extraordinary importance in the aerospace industry. Recently, Yu *et al.* (2022) proposed a new eddy-viscosity SGS model constrained by the ANN-based modelling of the kinetic energy flux for LES of compressible wall-bounded turbulence. The new kinetic-energy-flux-constrained model

(KCM) showed a better performance in *a priori* and *a posteriori* tests than the traditional SGS models. However, Yu *et al.* (2022) only constructed a good SGS model for the unclosed SGS stress tensor, while advanced SGS models for the unclosed SGS heat flux are much less investigated, which are significant in the LES of compressible wall-bounded turbulence. Furthermore, the KCM developed by Yu *et al.* (2022) is essentially an eddy-viscosity model. It is assumed by KCM that the SGS stress and SGS heat flux are linearly correlated with the filtered strain-rate tensor and the gradient of the filtered temperature respectively. These linear alignment assumptions are unphysical and it can be speculated that the predicted SGS stress and SGS heat flux estimated by KCM would not have high correlations with the true SGS stress and SGS heat flux. The KCM also cannot estimate the complicated nonlinear relation between the filtered strain-rate tensor and the SGS stress as well as the gradient of the filtered temperature and the SGS heat flux.

It is worth noting that a critical issue for the ANN-based SGS model is the generalisation ability to untrained flow conditions. Park & Choi (2021) showed that an ANN-based model trained at a single grid resolution did not predict the flow statistics accurately at another grid resolution, and this conclusion was also found in Zhou *et al.* (2019). However, many studies have shown that an ANN-based SGS model trained at low Reynolds number also has a good performance at high Reynolds number (Maulik *et al.* 2019; Park & Choi 2021).

Another significant issue for the ANN-based model is the much larger computational cost, as compared with those of the conventional dynamic eddy-viscosity models. It is shown that the computational cost of the ANN-based models given in Park & Choi (2021), Wang *et al.* (2018*b*), Yuan *et al.* (2020), Xie *et al.* (2019*c*) and Xie *et al.* (2020*a*) were approximately 1.3, 1.8, 2.4, 15 and 256 times of that of DSM, respectively. Furthermore, the ANN-based model in Yu *et al.* (2022) was also time consuming due to the large number of the neurons of the hidden layer and the complex hyperbolic tangent (tanh) activation function.

Accordingly, the main objective of this study is to propose a well-performing ANN-based nonlinear algebraic model for the LES of compressible wall-bounded turbulence for the first time. The nonlinear algebraic model of the SGS stress tensor proposed by Lund & Novikov (1992) and that of the SGS passive scalar flux given by Wang *et al.* (2007) in the incompressible turbulence are modified to utilise in the compressible turbulent channel flow. The ANN method is used to estimate the dimensionless model coefficients in order to get rid of some *ad hoc* procedures such as the averaging of model coefficients in statistically homogeneous directions and the clipping of the negative model coefficients to avoid numerical instability. Furthermore, the new model is trained with compressible turbulent channel flow data at a single Reynolds number, Mach number and grid resolution, and can have better performance in predicting the flow statistics in compressible turbulent channel flows with untrained Reynolds numbers, Mach numbers and grid resolutions. Moreover, the new model can also accurately predict the flow statistics in the supersonic turbulent boundary layer. Finally, the proposed models are more computationally efficient than the DSM, and make more accurate predictions than the conventional SGS models.

The rest of the paper is organised as follows. The governing equations of the LES and the *a priori* test of the traditional SGS models are described in § 2. The structure and the *a priori* test of the ANN-based nonlinear algebraic model are shown in § 3. The *a posteriori* test of the modified ANN-based nonlinear algebraic model is presented in § 4. Finally, summary and conclusions are given in § 5.

2. The governing equations of the LES and the *a priori* test of the traditional SGS models

2.1. *The governing equations of the LES*

By applying the filtering method to Navier–Stokes equations, the following compressible dimensionless LES governing equations are derived (Garnier *et al.* 2009):

$$\frac{\partial \bar{\rho}}{\partial t} + \frac{\partial (\bar{\rho}\tilde{u}_j)}{\partial x_j} = 0, \tag{2.1}$$

$$\frac{\partial (\bar{\rho}\tilde{u}_i)}{\partial t} + \frac{\partial [\bar{\rho}\tilde{u}_i\tilde{u}_j + \bar{p}\delta_{ij}]}{\partial x_j} = \frac{1}{Re}\frac{\partial \tilde{\sigma}_{ij}}{\partial x_j} - \frac{\partial \bar{\rho}\tau_{ij}}{\partial x_j} + \frac{1}{Re}\frac{\partial}{\partial x_j}(\overline{\sigma_{ij}} - \tilde{\sigma}_{ij}), \tag{2.2}$$

$$\frac{\partial \bar{\rho}\tilde{E}}{\partial t} + \frac{\partial [(\bar{\rho}\tilde{E} + \bar{p})\tilde{u}_j]}{\partial x_j} = \frac{1}{Re}\frac{\partial (\tilde{\sigma}_{ij}\tilde{u}_i)}{\partial x_j} - \frac{1}{\alpha}\frac{\partial \tilde{q}_j}{\partial x_j} - \frac{\partial (C_p\bar{\rho}Q_j)}{\partial x_j} - \frac{\partial \mathcal{J}_j}{\partial x_j} + \frac{\partial}{\partial x_j}[(\overline{\sigma_{ij}u_i} - \tilde{\sigma}_{ij}\tilde{u}_i) + (\overline{q_j} - \tilde{q}_j)], \tag{2.3}$$

$$\bar{p} = \bar{\rho}\tilde{T}/(\gamma M^2). \tag{2.4}$$

Here, $\bar{\rho}$, $\tilde{u}_i$, $\tilde{T}$, $\bar{p}$ represent the resolved density, velocity, temperature and pressure, respectively, and $\bar{\rho}\tilde{E}$ is the Favre-averaged total energy defined later in (2.13). The filtered field for a variable f can be defined as (Martin, Piomelli & Candler 2000; Xu *et al.* 2021*b*)

$$\bar{f}(\boldsymbol{x}) = \int \mathrm{d}^3\boldsymbol{r} G_l(\boldsymbol{r}) f(\boldsymbol{x} + \boldsymbol{r}), \tag{2.5}$$

where $G_l(\boldsymbol{r}) \equiv l^{-3}G(\boldsymbol{r}/l)$ is the filter function, and $G(\boldsymbol{r})$ is a normalised window function. Here, l is the filter width associated with the wavelength of the smallest scale retained by the filtering operation. The symbol $\tilde{f} = \overline{\rho f}/\bar{\rho}$ represents the Favre filtered field.

The above compressible LES equations are non-dimensionalised by the following set of reference scales, including the reference length L_∞, free-stream density ρ_∞, velocity U_∞, temperature T_∞, pressure $p_\infty = \rho_\infty U_\infty^2$, energy per unit volume $\rho_\infty U_\infty^2$, viscosity μ_∞ and thermal conductivity κ_∞. The ratio of specific heat at constant pressure C_p to that at constant volume C_v is defined as $\gamma = C_p/C_v$ and assumed to be equal to 1.4. Moreover, $Re = \rho_\infty U_\infty L_\infty/\mu_\infty$ is the Reynolds number, $M = U_\infty/c_\infty$ is the Mach number and $\alpha = Pr\,Re(\gamma - 1)M^2$, where the Prandtl number $Pr = \mu_\infty C_p/\kappa_\infty$ is equal to 0.7.

The resolved viscous stress tensor $\tilde{\sigma}_{ij}$ and the resolved heat flux $\tilde{q}_j$ are defined as

$$\tilde{\sigma}_{ij} = 2\tilde{\mu}(\tilde{T})(\tilde{S}_{ij} - \tfrac{1}{3}\delta_{ij}\tilde{S}_{kk}) \tag{2.6}$$

and

$$\tilde{q}_j = -\kappa(\tilde{T})\frac{\partial \tilde{T}}{\partial x_j}, \tag{2.7}$$

where $\tilde{\mu} = (\tilde{T}/\tilde{T}_\infty)^{3/2}[(\tilde{T}_\infty + T_s)/(\tilde{T} + T_s)]$ is the molecular viscosity calculated by Sutherland's law with $T_s = 110.4\,\mathrm{K}$ and $\tilde{T}_\infty = 288.15\,\mathrm{K}$, and the thermal conductivity $\tilde{\kappa}$ has the same expression as $\tilde{\mu}$. Furthermore, $\tilde{S}_{ij} = (\partial\tilde{u}_i/\partial x_j + \partial\tilde{u}_j/\partial x_i)/2$ is the resolved strain-rate tensor.

As was done in many previous studies, including Ragab, Sheen & Sreedhar (1992), Piomelli (1999), Dubois, Domaradzki & Honein (2002), Kawai & Larsson (2013) and Yu *et al.* (2022), the unclosed terms $(1/Re)(\partial/\partial x_j)(\overline{\sigma_{ij}} - \tilde{\sigma}_{ij})$ in (2.2) and $(\partial/\partial x_j)[(\overline{\sigma_{ij}u_i} - \tilde{\sigma}_{ij}\tilde{u}_i) + (\overline{q_j} - \tilde{q}_j)]$ in (2.3) can be neglected, which is justified by the inviscid criterion introduced in Aluie (2013) and Zhao & Aluie (2018). Therefore, (2.2) and (2.3) can be simplified to

$$\frac{\partial\left(\bar{\rho}\tilde{u}_i\right)}{\partial t} + \frac{\partial[\bar{\rho}\tilde{u}_i\tilde{u}_j + \bar{p}\delta_{ij}]}{\partial x_j} = \frac{1}{Re}\frac{\partial\tilde{\sigma}_{ij}}{\partial x_j} - \frac{\partial\bar{\rho}\tau_{ij}}{\partial x_j}, \tag{2.8}$$

$$\frac{\partial\bar{\rho}\tilde{E}}{\partial t} + \frac{\partial[(\bar{\rho}\tilde{E} + \bar{p})\tilde{u}_j]}{\partial x_j} = \frac{1}{Re}\frac{\partial(\tilde{\sigma}_{ij}\tilde{u}_i)}{\partial x_j} - \frac{1}{\alpha}\frac{\partial\tilde{q}_j}{\partial x_j} - \frac{\partial(C_p\bar{\rho}Q_j)}{\partial x_j} - \frac{\partial\mathcal{J}_j}{\partial x_j}. \tag{2.9}$$

It is found that there are three unclosed terms in the above compressible LES equations, including the SGS stress tensor

$$\tau_{ij} = \widetilde{u_iu_j} - \tilde{u}_i\tilde{u}_j, \tag{2.10}$$

the SGS heat flux

$$Q_j = \widetilde{u_jT} - \tilde{u}_j\tilde{T}, \tag{2.11}$$

and the SGS turbulent diffusion

$$\mathcal{J}_j = \frac{1}{2}\bar{\rho}\left(\widetilde{u_ju_iu_i} - \tilde{u}_j\widetilde{u_iu_i}\right). \tag{2.12}$$

Moreover, $\bar{\rho}\tilde{E}$ is the Favre-averaged total energy, and can be defined as

$$\bar{\rho}\tilde{E} = \frac{\bar{p}}{\gamma - 1} + \frac{1}{2}\bar{\rho}\tilde{u}_i\tilde{u}_i + \frac{1}{2}\bar{\rho}\tau_{ii}. \tag{2.13}$$

Therefore, the resolved temperature $\tilde{T}$ can be calculated by

$$\tilde{T} = (\gamma - 1)\,\gamma M^2\left[\tilde{E} - \frac{1}{2}\tilde{u}_i\tilde{u}_i - \frac{\tau_{ii}}{2}\right]. \tag{2.14}$$

Furthermore, it is suggested that the SGS turbulent diffusion can be approximated by $\mathcal{J}_j \approx \bar{\rho}\tau_{ij}\tilde{u}_i$ (Martin *et al.* 2000; Jiang *et al.* 2013; Yu *et al.* 2022). Therefore, only the SGS stress tensor τ_{ij} and the SGS heat flux Q_j need to be modelled based on the resolved variables.

It is noted that the subscripts i and j of τ_{ij} and Q_j satisfy $i, j = 1, 2, 3$, where the numbers 1, 2, 3 represent the streamwise (x), wall-normal (y) and spanwise (z) directions, respectively.

2.2. *The a priori test of the traditional SGS models*

It is worth noting that the eddy-viscosity models are the most famous SGS models for the SGS stress tensor τ_{ij} and the SGS heat flux Q_j. The DSM (Moin *et al.* 1991; Lilly 1992), the Vreman model (Vreman 2004; Sayadi & Moin 2012) and the WALE model (Nicoud & Ducros 1999; Garnier *et al.* 2009) are the widely used SGS models, which are used to compared with the new proposed models in the present study.

The eddy-viscosity models of the SGS stress tensor τ_{ij} and the SGS heat flux Q_j can be written as follows:

$$\tau_{ij} - \tfrac{1}{3}\delta_{ij}\tau_{kk} = -2\mu_{SGS}(\tilde{S}_{ij} - \tfrac{1}{3}\delta_{ij}\tilde{S}_{kk}), \tag{2.15}$$

$$\tau_{kk} = 2\mu_{I,SGS}\left|\tilde{S}\right|, \tag{2.16}$$

$$Q_j = -\frac{\mu_{SGS}}{Pr_t}\frac{\partial \tilde{T}}{\partial x_j}. \tag{2.17}$$

In the DSM, the SGS eddy-viscosity coefficients μ_{SGS} and $\mu_{I,SGS}$ are respectively defined as (Moin *et al.* 1991; Lilly 1992)

$$\mu_{SGS} = C_{sm}^2\Delta^2\left|\tilde{S}\right| \tag{2.18}$$

and

$$\mu_{I,SGS} = C_I\Delta^2\left|\tilde{S}\right|, \tag{2.19}$$

with

$$\left|\tilde{S}\right| = \sqrt{2\tilde{S}_{ij}\tilde{S}_{ij}}. \tag{2.20}$$

Here, $\Delta = (\Delta_x\Delta_y\Delta_z)^{1/3}$ is a characteristic length scale of local grid widths, and Δ_x, Δ_y and Δ_z are the local grid widths along the streamwise, wall-normal and spanwise directions, respectively. The model coefficients C_{sm}^2, C_I and μ_{SGS}/Pr_t in the DSM can be solved dynamically based on the Germano identity (Germano *et al.* 1991; Moin *et al.* 1991; Lilly 1992; Xie *et al.* 2019*b*).

In the Vreman model, the SGS eddy-viscosity coefficients μ_{SGS} and $\mu_{I,SGS}$ are respectively defined as (Vreman 2004; Sayadi & Moin 2012)

$$\mu_{SGS} = C_v D_v \tag{2.21}$$

and

$$\mu_{I,SGS} = C_I D_v, \tag{2.22}$$

with

$$D_v = \sqrt{\frac{\Pi_\beta}{\tilde{\alpha}_{ij}\tilde{\alpha}_{ij}}}, \tag{2.23}$$

$$\Pi_\beta = \beta_{11}\beta_{22} - \beta_{12}^2 + \beta_{11}\beta_{33} - \beta_{13}^2 + \beta_{22}\beta_{33} - \beta_{23}^2, \tag{2.24}$$

$$\tilde{\alpha}_{ij} = \frac{\partial \tilde{u}_j}{\partial x_i}, \tag{2.25}$$

$$\beta_{ij} = \sum_{m=1}^{3}\Delta_m^2\tilde{\alpha}_{mi}\tilde{\alpha}_{mj}, \tag{2.26}$$

where Δ_m is the local grid width along the mth direction. The model coefficient C_v is empirically defined as $C_v = 2.5C_{sm}^2$, where $C_{sm} = 0.1$ (Sayadi & Moin 2012; Yu *et al.* 2022).

In the WALE model, the SGS eddy-viscosity coefficients μ_{SGS} and $\mu_{I,SGS}$ are respectively defined as (Nicoud & Ducros 1999; Garnier *et al.* 2009)

$$\mu_{SGS} = C_w \Delta^2 D_w \tag{2.27}$$

and

$$\mu_{I,SGS} = C_I \Delta^2 D_w, \tag{2.28}$$

with

$$D_w = \frac{\left(\tilde{\mathcal{S}}_{ij}^d \tilde{\mathcal{S}}_{ij}^d\right)^{3/2}}{\left(\tilde{S}_{ij}\tilde{S}_{ij}\right)^{5/2} + \left(\tilde{\mathcal{S}}_{ij}^d \tilde{\mathcal{S}}_{ij}^d\right)^{5/4}}, \tag{2.29}$$

$$\tilde{\mathcal{S}}_{ij}^d = \frac{1}{2}\left(\tilde{\alpha}_{ij}^2 + \tilde{\alpha}_{ji}^2\right) - \frac{1}{3}\tilde{\alpha}_{kk}^2 \delta_{ij}, \tag{2.30}$$

$$\tilde{\alpha}_{ij}^2 = \tilde{\alpha}_{ik}\tilde{\alpha}_{kj}. \tag{2.31}$$

The model coefficient C_w is empirically defined as $C_w = 10.6C_{sm}^2$, where $C_{sm} = 0.1$ (Garnier *et al.* 2009; Yu *et al.* 2022).

Furthermore, it is worth noting that, consistent with the assumption in Sayadi & Moin (2012) and Yu *et al.* (2022), we also assume $C_I = 0$ in the Vreman and WALE models, and assume the SGS Prandtl number $Pr_t = 0.9$ empirically in the Vreman and WALE models (Sayadi & Moin 2012; Yu *et al.* 2022).

In order to examine the performance of the three traditional SGS models listed above, the *a priori* test is performed for DNS data of a temporally compressible isothermal-wall turbulent channel flow (Coleman, Kim & Moser 1995). The governing parameters of the DNS data are listed as follows: the Reynolds number is $Re = 3000$ and the Mach number is $M = 1.5$. Therefore, we denote this DNS data as 'R1M15' for brevity. The computational domain of R1M15 is a box with a size of $4\pi \times 2 \times \frac{4}{3}\pi$. The grid resolutions are $384 \times 193 \times 128$, and $\Delta X^+ \times \Delta Y_w^+ \times \Delta Z^+ = 7.13 \times 0.30 \times 7.13$, where $\Delta X^+ = \Delta X/\delta_v$, $\Delta Y_w^+ = \Delta Y_w/\delta_v$ and $\Delta Z^+ = \Delta Z/\delta_v$ are the normalised spacing of the streamwise direction, the first point off the wall and the spanwise direction in the DNS data of R1M15, respectively. Here, $\delta_v = \mu_w/(\rho_w u_\tau)$ is the viscous length scale, where μ_w and ρ_w are the viscosity and density on the wall, respectively. Moreover, $u_\tau = \sqrt{\tau_w/\rho_w}$ is the friction velocity and $\tau_w = (\mu \partial \langle u \rangle_{xz}/\partial y)_w$ is the wall shear stress, where $\langle \cdot \rangle_{xz}$ represents the average in the streamwise and spanwise directions. In the *a priori* test, the DNS data in R1M15 are filtered in the streamwise and spanwise directions with a top-hat filter (Martin *et al.* 2000; Xu *et al.* 2021*b*), which can be calculated in one dimension by (Martin *et al.* 2000; Xu *et al.* 2021*b*)

$$\bar{f}_i = \frac{1}{4n}\left(f_{i-n} + 2\sum_{j=i-n+1}^{i+n-1} f_j + f_{i+n}\right), \tag{2.32}$$

where the filter width is $l = 2n\Delta$. In the following *a priori* test of the three traditional SGS models, the filtered streamwise and spanwise widths are $\Delta x = n_x \Delta X$ and $\Delta z = n_z \Delta Z$, respectively, where the filtered streamwise and spanwise sizes are $n_x = 8$ and $n_z = 4$, respectively. It is noted that the values of n_x and n_z need to be even numbers.

In *a priori* test, the correlation coefficient C and the relative error E_r are calculated to evaluate the difference between the true value (H^{real}) of the unclosed SGS terms obtained

from the fDNS data and the predicted value (H^{model}) calculated by the SGS models, and they can be expressed respectively as

$$C(H) = \frac{\left\langle \left(H^{real} - \langle H^{real}\rangle_{xz}\right)\left(H^{model} - \langle H^{model}\rangle_{xz}\right)\right\rangle_{xz}}{\left\langle \left(H^{real} - \langle H^{real}\rangle_{xz}\right)^2\right\rangle_{xz}^{1/2}\left\langle \left(H^{model} - \langle H^{model}\rangle_{xz}\right)^2\right\rangle_{xz}^{1/2}}, \tag{2.33}$$

$$E_r(H) = \frac{\left\langle \left(H^{real} - H^{model}\right)^2\right\rangle_{xz}^{1/2}}{\left\langle \left(H^{real}\right)^2\right\rangle_{xz}^{1/2}}. \tag{2.34}$$

The SGS models with high correlation coefficients and low relative errors suggest a well-performing modelling.

The correlation coefficients C and the relative errors E_r of the DSM, Vreman model and WALE model along wall-normal direction in *a priori* tests with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15 are shown in figures 1, 2 and 3, respectively. The lighter sections of each line in figures 1, 2 and 3(*a*,*c*) represent the 95 % confidence intervals for the correlation coefficients calculated by using the Fisher transformation (Bonett & Wright 2000). The eddy-viscosity models assume that the SGS stress is linearly aligned with the filtered strain-rate tensor based on the Boussinesq hypothesis; similarly, the SGS heat flux is also assumed to be linearly proportional to the gradient of the resolved temperature. However, the linear alignment assumptions are unphysical and contradictory with the complicated nonlinear relations between the SGS stress and the filtered strain-rate tensor as well as between the SGS heat flux and the gradient of the filtered temperature. Therefore, it is shown that the correlation coefficients C of the three traditional eddy-viscosity models are almost lower than 0.6, indicating that the traditional eddy-viscosity models correlate poorly with the true unclosed SGS terms (Anderson & Meneveau 1999; Da Silva & Métais 2002; Xie *et al.* 2019*a*,*c*, 2020*a*,*b*). These low correlations of the three traditional eddy-viscosity models can be ascribed to the unphysical linear alignment assumptions in the eddy-viscosity models. To be specific, the correlation coefficients in DSM are relatively small in the near-wall region, and slightly increase as y^+ increases. However, in the Vreman model and WALE model, the correlation coefficients in the near-wall region are similar to those in the far-wall region. These observations can be attributed to the reasons that the DSM cannot capture the correct behaviour in the vicinity of the wall, resulting in an extremely high damping of fluctuations near the wall, while the Vreman model and WALE model are constructed to guarantee that the dissipation is relatively small in the near-wall region. Therefore, the correlation coefficients of Vreman model and WALE model are slightly larger than those of DSM near the wall. Furthermore, it is found that the correlation coefficients of the diagonal components of the SGS stress (i.e. τ_{11}, τ_{22} and τ_{33}) are slightly larger in DSM than those in Vreman model and WALE model. This can be ascribed to the reason that the model coefficients C_I of the SGS eddy viscosity $\mu_{I,SGS}$ are zero in the Vreman model and WALE model, while C_I is dynamically solved based on the Germano identity (Moin *et al.* 1991; Lilly 1992; Xie *et al.* 2019*b*) in DSM. Moreover, the relative errors E_r of the three traditional eddy-viscosity models are almost equal to 1.0, suggesting that the traditional eddy-viscosity models have very large errors when predicting the true unclosed SGS terms. The above observations are consistent with many previous studies (Anderson & Meneveau 1999; Da Silva & Métais 2002; Xie *et al.* 2019*a*, 2020*a*,*b*; Yuan *et al.* 2020).

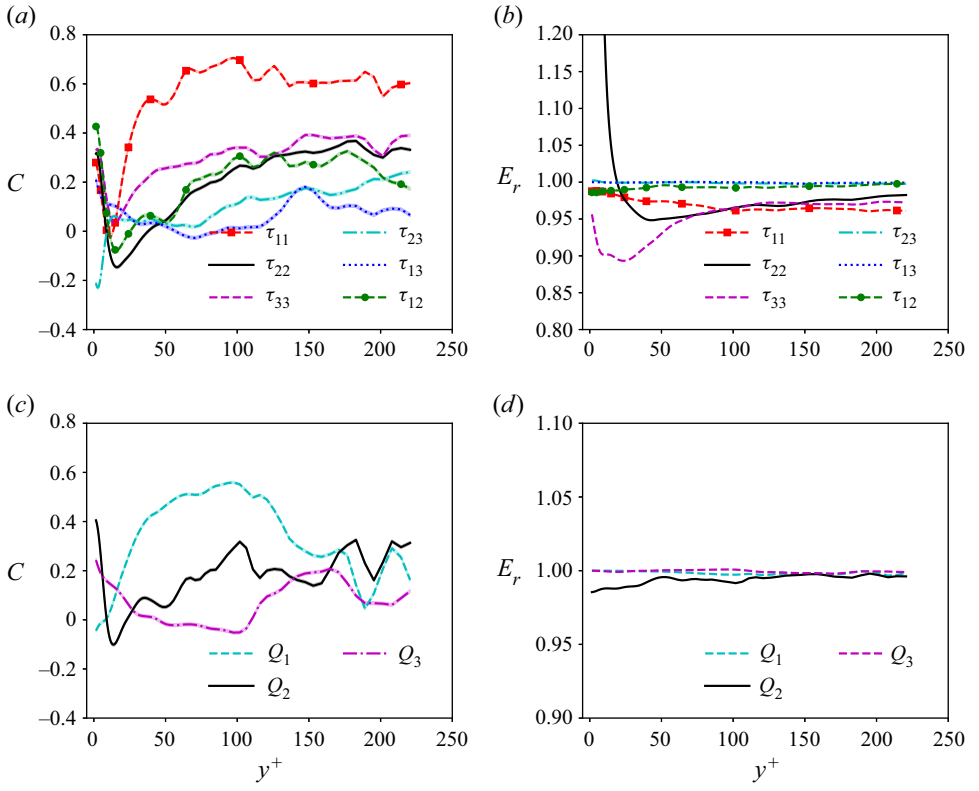


Figure 1. (*a*,*c*) The correlation coefficients C and (*b*,*d*) the relative errors E_r of DSM along the wall-normal direction in *a priori* test with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15. The lighter sections of each line in (*a*,*c*) represent the 95 % confidence intervals for the correlation coefficients.

3. The structure and *a priori* test of the ANN-based nonlinear algebraic model

3.1. *The original ANN-based nonlinear algebraic model*

In the SGS modelling, the constitutive relation of the unclosed SGS terms can be considered as the function of the local filtered quantities, including the filtered strain-rate tensor $\tilde{S}_{ij}$, the filtered rotation-rate tensor $\tilde{\Omega}_{ij}$ and the filtered temperature gradients $\partial\tilde{T}/\partial x_j$, namely (Pope 1975; Lund & Novikov 1992; Wang *et al.* 2007; Vollant *et al.* 2017)

$$\tau_{ij} = f(\tilde{S}_{ij}, \tilde{\Omega}_{ij}), \tag{3.1}$$

$$Q_j = f\left(\tilde{S}_{ij}, \tilde{\Omega}_{ij}, \frac{\partial\tilde{T}}{\partial x_j}\right), \tag{3.2}$$

where the filtered rotation-rate tensor $\tilde{\Omega}_{ij}$ is given by

$$\tilde{\Omega}_{ij} = \frac{1}{2}\left(\frac{\partial\tilde{u}_i}{\partial x_j} - \frac{\partial\tilde{u}_j}{\partial x_i}\right). \tag{3.3}$$

The SGS stress tensor τ_{ij} can be divided into anisotropic and isotropic parts, respectively: $\tau_{ij} = \tau_{ij}^A + \tau_{ij}^I$, where the anisotropic part is $\tau_{ij}^A = \tau_{ij} - \frac{1}{3}\delta_{ij}\tau_{kk}$, and the isotropic part is $\tau_{ij}^I = \frac{1}{3}\delta_{ij}\tau_{kk}$.

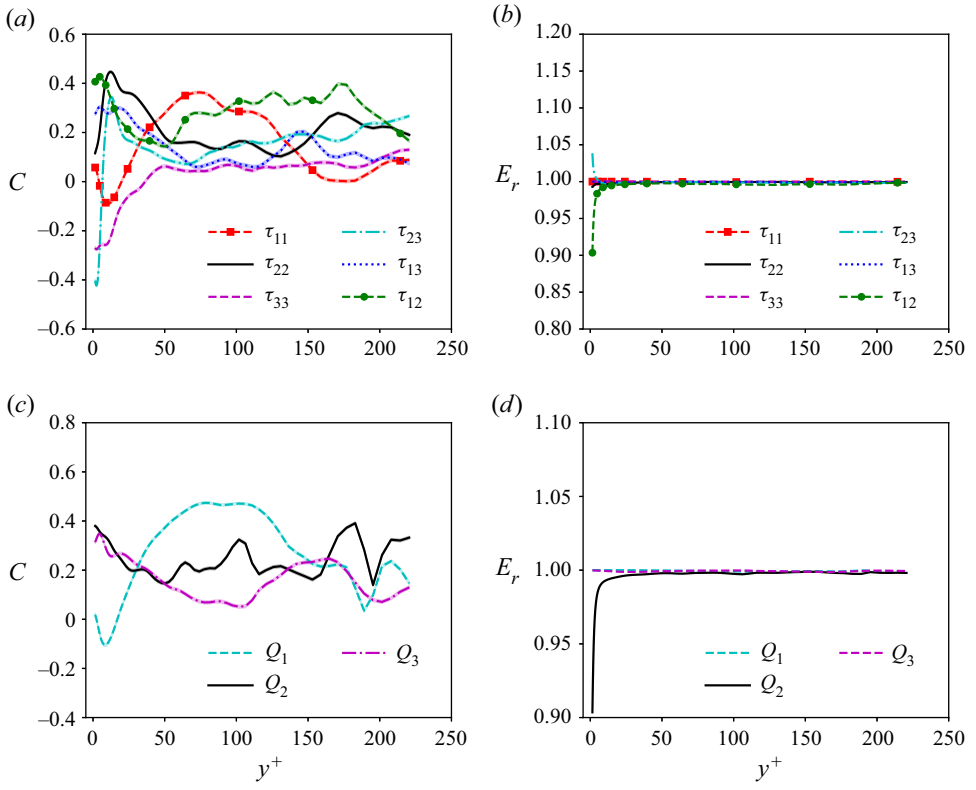


Figure 2. (*a*,*c*) The correlation coefficients C and (*b*,*d*) the relative errors E_r of the Vreman model along the wall-normal direction in *a priori* test with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15. The shaded areas under each line in (*a*,*c*) represent the 95 % confidence intervals for the correlation coefficients.

In order to get rid of the unphysical linear alignment assumption in the traditional eddy-viscosity models, the nonlinear algebraic models of the SGS stress (Lund & Novikov 1992) and SGS heat flux (Vollant *et al.* 2017) were proposed to construct the nonlinear relation between the unclosed SGS terms and the local filtered quantities.

Similar to the previous analysis of Lund & Novikov (1992), the anisotropic SGS stress can be approximated as (Lund & Novikov 1992; Xie *et al.* 2020*b*)

$$\tau_{ij}^A = \Delta^2 \left(C_1^A \left\| \tilde{\boldsymbol{S}} \right\| \boldsymbol{T}_1^A + C_2^A \boldsymbol{T}_2^A + C_3^A \boldsymbol{T}_3^A + C_4^A \boldsymbol{T}_4^A + C_5^A \frac{1}{\left\| \tilde{\boldsymbol{S}} \right\|} \boldsymbol{T}_5^A \right), \tag{3.4}$$

where

$$\boldsymbol{T}_1^A = \tilde{\boldsymbol{S}} - \tfrac{1}{3}\boldsymbol{I} \cdot \mathrm{Tr}(\tilde{\boldsymbol{S}}), \tag{3.5}$$

$$\boldsymbol{T}_2^A = \tilde{\boldsymbol{S}}\tilde{\boldsymbol{\Omega}} - \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{S}}, \tag{3.6}$$

$$\boldsymbol{T}_3^A = \tilde{\boldsymbol{S}}^2 - \tfrac{1}{3}\boldsymbol{I} \cdot \mathrm{Tr}(\tilde{\boldsymbol{S}}^2), \tag{3.7}$$

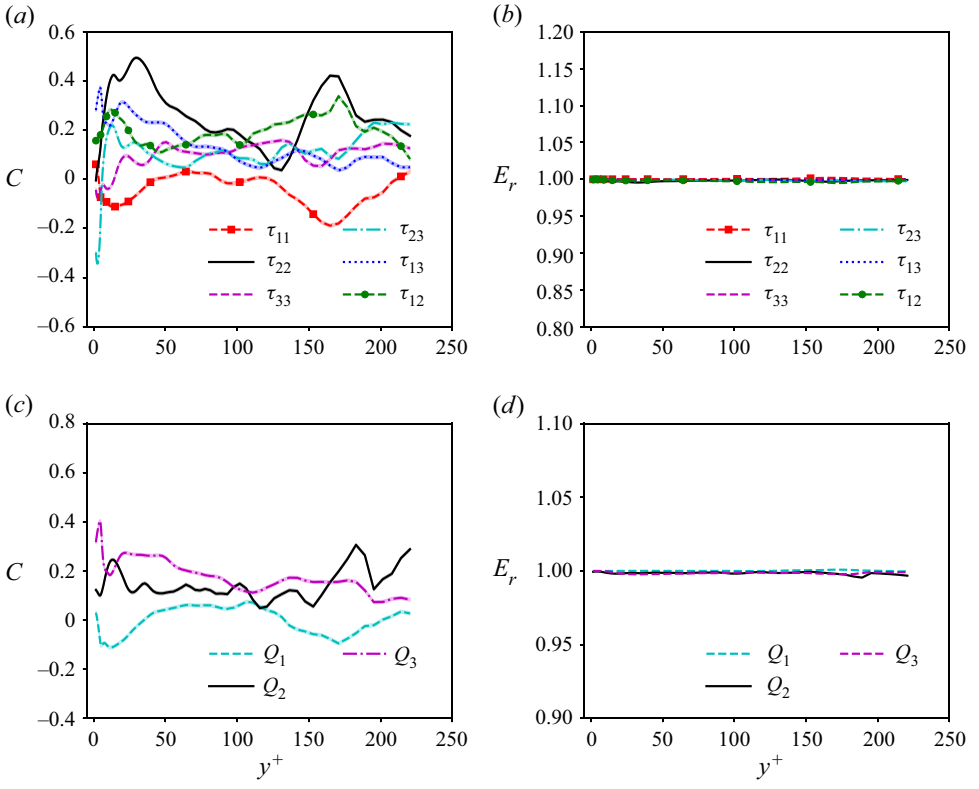


Figure 3. (*a*,*c*) The correlation coefficients C and (*b*,*d*) the relative errors E_r of the WALE model along the wall-normal direction in *a priori* test with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15. The shaded areas under each line in (*a*,*c*) represent the 95 % confidence intervals for the correlation coefficients.

$$\boldsymbol{T}_4^A = \tilde{\boldsymbol{\Omega}}^2 - \tfrac{1}{3}\boldsymbol{I}\cdot\mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2), \tag{3.8}$$

$$\boldsymbol{T}_5^A = \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{S}}^2 - \tilde{\boldsymbol{S}}^2\tilde{\boldsymbol{\Omega}}. \tag{3.9}$$

It is noted that, for brevity and simplicity of the tensorial polynomials, the matrix multiplications for the tensor contractions are expressed as

$$\tilde{\boldsymbol{S}}^2 = \tilde{S}_{ik}\tilde{S}_{kj}, \quad \tilde{\boldsymbol{S}}\tilde{\boldsymbol{\Omega}} = \tilde{S}_{ik}\tilde{\Omega}_{kj}, \quad \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{S}}^2 = \tilde{\Omega}_{ik}\tilde{S}_{kl}\tilde{S}_{lj}, \ \mathrm{Tr}(\tilde{\boldsymbol{S}}^2) = \tilde{S}_{ik}\tilde{S}_{ki}. \tag{3.10a–c}$$

Furthermore, the isotropic SGS stress can be approximated as

$$\tau_{kk} = \Delta^2\left(C_0^I\left\|\tilde{\boldsymbol{S}}\right\|^2 + C_1^I\left\|\tilde{\boldsymbol{S}}\right\|\boldsymbol{T}_1^I + C_2^I\boldsymbol{T}_2^I + C_3^I\boldsymbol{T}_3^I\right), \tag{3.11}$$

where

$$\boldsymbol{T}_1^I = \boldsymbol{I}\cdot\mathrm{Tr}(\tilde{\boldsymbol{S}}), \tag{3.12}$$

$$\boldsymbol{T}_2^I = \boldsymbol{I}\cdot\mathrm{Tr}(\tilde{\boldsymbol{S}}^2), \tag{3.13}$$

$$\boldsymbol{T}_3^I = \boldsymbol{I}\cdot\mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2). \tag{3.14}$$

Here, $\Delta = (\Delta_x \Delta_y \Delta_z)^{1/3}$ is the characteristic length scale of the grid width and $\|\tilde{\boldsymbol{S}}\| = \sqrt{\mathrm{Tr}(\tilde{\boldsymbol{S}}^2)}$.

It has been shown that these dimensionless model coefficients C_i^A and C_i^I in (3.4) and (3.11) are functions of the following five invariants (Pope 1975; Lund & Novikov 1992; Xie *et al.* 2020*b*):

$$\lambda_1 = \mathrm{Tr}(\tilde{\boldsymbol{S}}^2), \quad \lambda_2 = \mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2), \quad \lambda_3 = \mathrm{Tr}(\tilde{\boldsymbol{S}}^3), \quad \lambda_4 = \mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2\tilde{\boldsymbol{S}}), \quad \lambda_5 = \mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2\tilde{\boldsymbol{S}}^2), \tag{3.15a–e}$$

where λ_1 is the strain-rate magnitude, λ_2 is the enstrophy magnitude, λ_3 is associated with the self-amplification of the strain rate, λ_4 is the vortex stretching and λ_5 is the magnitude of the vortex stretching vector (Chamecki, Meneveau & Parlange 2007). Furthermore, we can reduce the above five invariants to four dimensionless ones using one or several of them appropriately in denominators (Lund & Novikov 1992; Chamecki *et al.* 2007; Xie *et al.* 2020*b*), and the four dimensionless invariants can be expressed as

$$d_1 = \frac{\mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2)}{\mathrm{Tr}(\tilde{\boldsymbol{S}}^2)}, \quad d_2 = \frac{\mathrm{Tr}(\tilde{\boldsymbol{S}}^3)}{\mathrm{Tr}(\tilde{\boldsymbol{S}}^2)^{3/2}}, \quad d_3 = \frac{\mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2\tilde{\boldsymbol{S}})}{\mathrm{Tr}(\tilde{\boldsymbol{S}}^2)^{1/2}\mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2)}, \quad d_4 = \frac{\mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2\tilde{\boldsymbol{S}}^2)}{\mathrm{Tr}(\tilde{\boldsymbol{S}}^2)\mathrm{Tr}(\tilde{\boldsymbol{\Omega}}^2)}. \tag{3.16a–d}$$

On the other hand, similar to the previous analysis of Vollant *et al.* (2017), the SGS heat flux can be approximated as

$$Q_j = \Delta^2 \left(C_1^Q \left|\tilde{\boldsymbol{S}}\right| \frac{\partial \tilde{T}}{\partial x_j} + C_2^Q \tilde{S}_{jk} \frac{\partial \tilde{T}}{\partial x_k} + C_3^Q \tilde{\Omega}_{jk} \frac{\partial \tilde{T}}{\partial x_k} + C_4^Q \frac{\tilde{S}_{jk}\tilde{S}_{kl}}{\left|\tilde{\boldsymbol{S}}\right|} \frac{\partial \tilde{T}}{\partial x_l} \right), \tag{3.17}$$

where $|\tilde{\boldsymbol{S}}| = \sqrt{2\tilde{S}_{ij}\tilde{S}_{ij}}$. The model coefficients C_i^Q in (3.17) depend on the principal invariants of tensor $\tilde{S}_{ij}$ and vector $\tilde{v}_i = \Delta^2 |\tilde{\boldsymbol{S}}| \frac{\partial \tilde{T}}{\partial x_i}$, which are expressed as (Noll 1967; Vollant *et al.* 2017)

$$q_1 = \tilde{S}_{ii}, \quad q_2 = \tilde{S}_{ij}\tilde{S}_{ji}, \quad q_3 = \tilde{S}_{ik}\tilde{S}_{kl}\tilde{S}_{li}, \quad q_4 = \tilde{v}_i\tilde{v}_i, \quad q_5 = \tilde{v}_i\tilde{S}_{ik}\tilde{v}_k, \quad q_6 = \tilde{v}_i\tilde{S}_{ik}\tilde{S}_{kj}\tilde{v}_j. \tag{3.18a–f}$$

Similar to the DSM, the dimensionless model coefficients C_i^A, C_i^I and C_i^Q in (3.4), (3.11) and (3.17) respectively can also be solved dynamically based on the Germano identity (Germano *et al.* 1991; Lilly 1992; Yuan *et al.* 2022). Therefore, this original dynamic nonlinear algebraic model is named the 'ODNA model' for brevity. The correlation coefficients C and the relative errors E_r of the ODNA model along wall-normal direction in *a priori* test with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15 are shown in figure 4. It is found that, in the far-wall region ($y^+ > 50$), the correlation coefficients of the ODNA model are all larger than 0.4, and are significantly larger than those of the aforementioned three eddy-viscosity SGS models (DSM, Vreman and WALE). This observation indicates that the nonlinear algebraic model has a better performance than the traditional eddy-viscosity SGS models. However, the correlation coefficients of the ODNA model are pretty small near the wall, mainly due to the neglect of the anisotropy of the grid widths in wall-bounded flows. Moreover, the relative errors of the ODNA model are approximately equal to 0.98 in most components, which is only slightly smaller than the traditional eddy-viscosity SGS models.

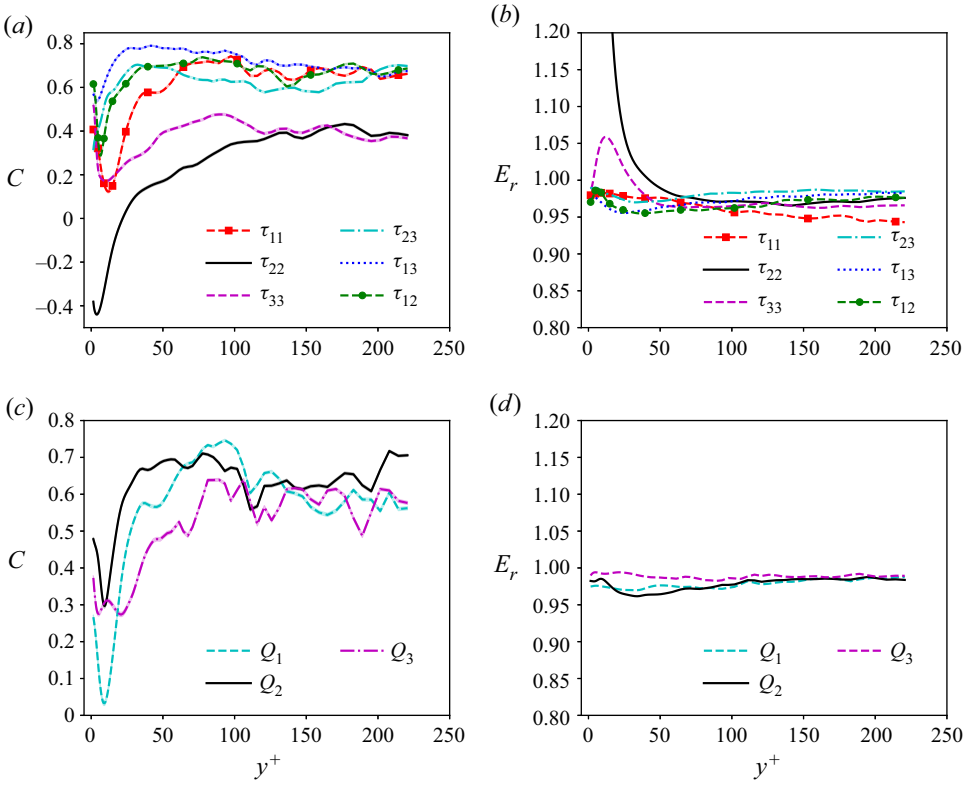


Figure 4. (*a*,*c*) The correlation coefficients C and (*b*,*d*) the relative errors E_r of the ODNA model along wall-normal direction in *a priori* test with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15. The lighter sections of each line in (*a*,*c*) represent the 95 % confidence intervals for the correlation coefficients.

It is shown in figure 4 that, when the dimensionless model coefficients C_i^A, C_i^I and C_i^Q are solved dynamically based on the Germano identity (Germano *et al.* 1991; Lilly 1992; Yuan *et al.* 2022), the ODNA model cannot achieve a very good performance in *a priori* tests. Therefore, the dimensionless model coefficients C_i^A, C_i^I and C_i^Q in (3.4), (3.11) and (3.17) respectively can also be determined by the ANN method (Xie *et al.* 2020*b*). The schematic diagram of the ANN structure is shown in figure 5. It is shown that the ANN is composed of multiple layers and each layer has many neurons. The neurons in the lth layer receive the inputs X_j^{l-1} from the $(l-1)$th layer and then transmit them to the outputs X_i^l activated by the nonlinear activation function. The transfer function from the $(l-1)$th layer to the lth layer can be expressed as

$$X_i^{(l)} = \sigma\left[b_i^{(l)} + \sum_j W_{ij}^{(l)} X_j^{(l-1)}\right], \tag{3.19}$$

where $\sigma[\cdot]$ is the nonlinear activation function, $W_{ij}^{(l)}$ and $b_i^{(l)}$ are the weights and biases in the lth layer, respectively. It is noted that the ANN method is only used to generate the nonlinear relation between the input features and the model coefficients C_i^A, C_i^I and C_i^Q. Then the generated model coefficients C_i^A, C_i^I and C_i^Q are multiplied with the

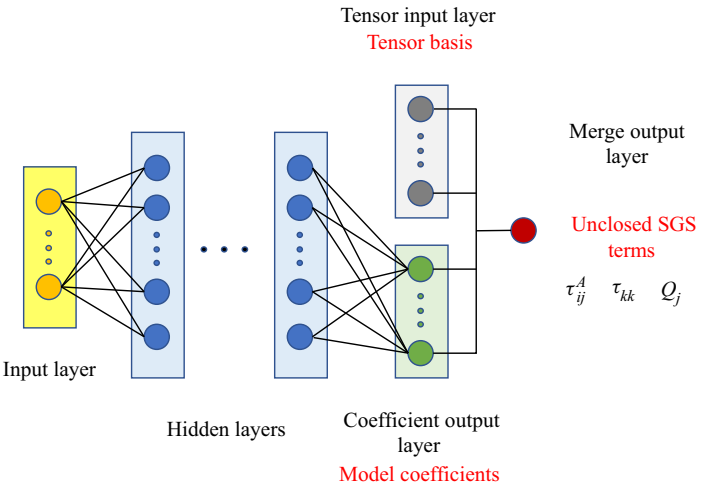


Figure 5. The schematic diagram of the ANN structure.

corresponding tensor bases to get predicted values of the unclosed SGS terms τ_{ij}^A, τ_{kk} and Q_j, respectively. Then the discrepancy between the predicted values of the unclosed SGS terms and the true values of the unclosed SGS terms are the loss function which is minimised during the training process. This original ANN-based nonlinear algebraic model is named the 'OANA model'.

There are four layers of neurons ($N_i : N_h : N_h : N_o$) between the input features and the coefficient output layer with a leaky-rectified linear unit (leaky-ReLU) activation function (Ling *et al.* 2016; Xie *et al.* 2020*b*) $\sigma(x) = \max[-0.1x, x]$, where N_i, N_h and N_o are the numbers of neurons of the input layer, hidden layers and coefficient output layer, respectively. It is noted that the numbers of neurons of the hidden layers N_h can be arbitrarily chosen, while the numbers of the neurons of the coefficient output layer N_o is a fixed number which is equal to the number of model coefficients. It is worth noting that the leaky-ReLU activation is known to perform better than the ReLU activation (Xu *et al.* 2015), due to the fact that the leaky-ReLU activation is capable of resolving the gradient vanishing problem of the ReLU activation. The input features, the variable of the coefficient output layer, the variable of the merge output layer and the values of N_i and N_o of the OANA model for the unclosed SGS terms τ_{ij}^A, τ_{kk} and Q_j are respectively listed in table 1. It is noted that, except for the invariants shown in (3.16*a–d*) and (3.18*a–f*), another feature $Re_l = \bar{\rho}(\tilde{u}^2 + \tilde{v}^2 + \tilde{w}^2)^{1/2}h/\tilde{\mu}$ is also included to represent the local Reynolds number, and this feature Re_l can introduce one of the key quantities of wall-bounded flows, the wall-normal distance from the wall h, into the input features (Yu *et al.* 2022). The coefficient output layer has N_o neurons and represents the model coefficients C_i^A, C_i^I and C_i^Q for the unclosed SGS terms τ_{ij}^A, τ_{kk} and Q_j, respectively. The tensor input layer also has N_o neurons for the N_o tensor bases shown in (3.4), (3.11) and (3.17), respectively, where $N_o = 5$ for the τ_{ij}^A term, $N_o = 4$ for the τ_{kk} term, $N_o = 4$ for the Q_j term. Each model coefficient C_i^A, C_i^I and C_i^Q in the coefficient output layer is multiplied with the corresponding tensor bases in the tensor input layer to generate the predicted values of the unclosed SGS term τ_{ij}^A, τ_{kk} and Q_j as (3.4), (3.11) and (3.17) expressed, respectively. The differences between the predicted values and the true values of the unclosed SGS terms

Input features	Coefficient output layer	Merge output layer	N_i	N_o
$\{d_1, d_2, d_3, d_4, Re_l\}$	C_i^A $(i = 1, \ldots, 5)$	τ_{ij}^A	5	5
$\{d_1, d_2, d_3, d_4, Re_l\}$	C_i^I $(i = 0, \ldots, 3)$	τ_{kk}	5	4
$\{q_1, q_2, q_3, q_4, q_5, q_6, Re_l\}$	C_i^Q $(i = 1, \ldots, 4)$	Q_j	7	4

Table 1. The input features, the variable of the coefficient output layer, the variable of the merge output layer and the values of N_i and N_o of the OANA model for the unclosed SGS terms τ_{ij}^A, τ_{kk} and Q_j, respectively. Here, the input features d_1–d_4 are the invariants shown in (3.16*a–d*), and the input features q_1–q_6 are the invariants shown in (3.18*a–f*). Furthermore, $Re_l = \bar{\rho}(\tilde{u}^2 + \tilde{v}^2 + \tilde{w}^2)^{1/2}h/\tilde{\mu}$ represents the local Reynolds number, where h is the wall-normal distance from the wall.

are minimised by the optimisation algorithm during the training process. The unclosed SGS terms τ_{ij}^A, τ_{kk} and Q_j have six, one and three components, respectively, therefore we should train ten ANNs in total.

The OANA model is trained with fDNS data of the turbulent channel flow case R1M15, with the filtered streamwise and spanwise sizes $n_x = 8$ and $n_z = 4$, respectively. The input variables are rescaled by the mean and standard deviation of the input filtered variables, which can be given by

$$m^\star = \frac{m - \langle m \rangle_{xz}}{\sqrt{\left\langle \left(m - \langle m \rangle_{xz}\right)^2 \right\rangle_{xz}}}, \tag{3.20}$$

where $m^\star$ represents the normalisation of the input variable m. Due to the strong anisotropy in the wall-normal direction of the wall-bounded flows, the input variables are normalised by the mean and standard deviation in the streamwise–spanwise wall-parallel plane at each wall-normal location. Furthermore, the true unclosed SGS terms τ_{ij}^A, τ_{kk} and Q_j and their corresponding tensor bases are normalised by the r.m.s. (root mean square) values of the gradient models for τ_{ij}^A, τ_{kk} and Q_j at each wall-normal location, which can be expressed as

$$\tau_{ij}^{A\star} = \frac{\tau_{ij}^A}{\sqrt{\left\langle \left(\tau_{ij}^{A\,GM}\right)^2 \right\rangle_{xz}}}, \quad \tau_{kk}^\star = \frac{\tau_{kk}}{\sqrt{\left\langle \left(\tau_{kk}{}^{GM}\right)^2 \right\rangle_{xz}}}, \quad Q_j^\star = \frac{Q_j}{\sqrt{\left\langle \left(Q_j{}^{GM}\right)^2 \right\rangle_{xz}}}; \tag{3.21a–c}$$

$$\boldsymbol{T}^\star_{\tau_{ij}^A} = \frac{\boldsymbol{T}_{\tau_{ij}^A}}{\sqrt{\left\langle \left(\tau_{ij}^{A\,GM}\right)^2 \right\rangle_{xz}}}, \quad \boldsymbol{T}^\star_{\tau_{kk}} = \frac{\boldsymbol{T}_{\tau_{kk}}}{\sqrt{\left\langle \left(\tau_{kk}{}^{GM}\right)^2 \right\rangle_{xz}}}, \quad \boldsymbol{T}^\star_{Q_j} = \frac{\boldsymbol{T}_{Q_j}}{\sqrt{\left\langle \left(Q_j{}^{GM}\right)^2 \right\rangle_{xz}}}, \tag{3.22a–c}$$

where $\boldsymbol{T}^\star_{\tau_{ij}^A}$, $\boldsymbol{T}^\star_{\tau_{kk}}$ and $\boldsymbol{T}^\star_{Q_j}$ are the tensor bases of the unclosed terms τ_{ij}^A, τ_{kk} and Q_j in (3.4), (3.11) and (3.17), respectively. Moreover, the gradient models $\tau_{ij}^{A\,GM}$, $\tau_{kk}{}^{GM}$ and $Q_j{}^{GM}$ for

τ_{ij}^{A}, τ_{kk} and Q_j respectively can be expressed as (Clark 1979)

$$\tau_{ij}^{A\,GM} = \frac{1}{12}\Delta_l^2 \frac{\partial \tilde{u}_i}{\partial x_l}\frac{\partial \tilde{u}_j}{\partial x_l} - \frac{1}{3}\delta_{ij}\frac{1}{12}\Delta_l^2 \frac{\partial \tilde{u}_k}{\partial x_l}\frac{\partial \tilde{u}_k}{\partial x_l}, \tag{3.23}$$

$$\tau_{kk}{}^{GM} = \frac{1}{12}\Delta_l^2 \frac{\partial \tilde{u}_k}{\partial x_l}\frac{\partial \tilde{u}_k}{\partial x_l}, \tag{3.24}$$

$$Q_j{}^{GM} = \frac{1}{12}\Delta_l^2 \frac{\partial \tilde{u}_j}{\partial x_l}\frac{\partial \tilde{T}}{\partial x_l}. \tag{3.25}$$

During the training process, the weights $W_{ij}^{(l)}$ and biases $b_i^{(l)}$ are optimised to minimise the mean-squared error loss function defined as

$$L = \frac{1}{N_{batch}}\sum_{n=1}^{N_{batch}} \left(f_{true} - f_{predict}\right)^2, \tag{3.26}$$

where f represents τ_{ij}^{A}, τ_{kk} and Q_j in their corresponding ANN training process. Furthermore, f_{true} is the true unclosed SGS term (i.e. τ_{ij}^{A}, τ_{kk} and Q_j) obtained from the fDNS data, and $f_{predict}$ is the predicted unclosed SGS term by the ANN. The size of the mini-batch N_{batch} is set to be 1000.

A total of almost 1.8×10^7 samples of the fDNS are used for the ANN training procedure. The dataset is divided into the training, validating and testing sets to suppress parameter overfitting of the ANN: 60 % of the samples are used as the training set, 15 % of the samples are used as the validating set and 25 % of the samples are used as the testing set. Furthermore, the samples at each wall-normal location are randomly extracted to be the training, validating and testing sets. The weights of the ANN are initialised by the Glorot-uniform algorithm (Glorot & Bengio 2010), and the biases are initialised to be zero. Finally, the ANN is trained by the Adam algorithm (Kingma & Ba 2014) to update $W_{ij}^{(l)}$ and $b_i^{(l)}$ for 2000 epochs with an initial learning rate of 0.001. The learning rate will further decrease by 90 % once the validated loss stops decreasing in each 10 epochs. The influences of the size of the mini-batch N_{batch}, different optimisers and activation functions on the performance of the OANA model in *a priori* tests are checked in the Appendix.

We train two sets of the OANA model with different numbers of neurons of the hidden layers N_h, and name them 'OANA-1' for $N_h = 10$ and 'OANA-2' for $N_h = 20$. The learning curves of the training and validation losses of the OANA-1 and OANA-2 models for τ_{ij}^{A}, τ_{kk} and Q_j are shown in figure 6. It is shown that the training and validation losses of most components decrease drastically in the initial several epochs, and then converge quickly and gradually reach stationarity. Furthermore, the learning curves of the training and validation losses are close to each other, indicating that the ANN models are well trained and remove overfitting phenomena. However, the training and validation losses of the wall-normal components (i.e. τ_{23}^{A}, τ_{12}^{A}, Q_2) oscillate drastically, although they decrease quickly in the initial several epochs. This indicates that the wall-normal components τ_{23}^{A}, τ_{12}^{A} and Q_2 are relatively hard to train. Moreover, the weights and biases of the OANA model used in the following *a priori* test are obtained from the epoch where the validation losses reach the minimum values. It is also found that the training and validation losses are all pretty large, suggesting the relatively large errors of the OANA model.

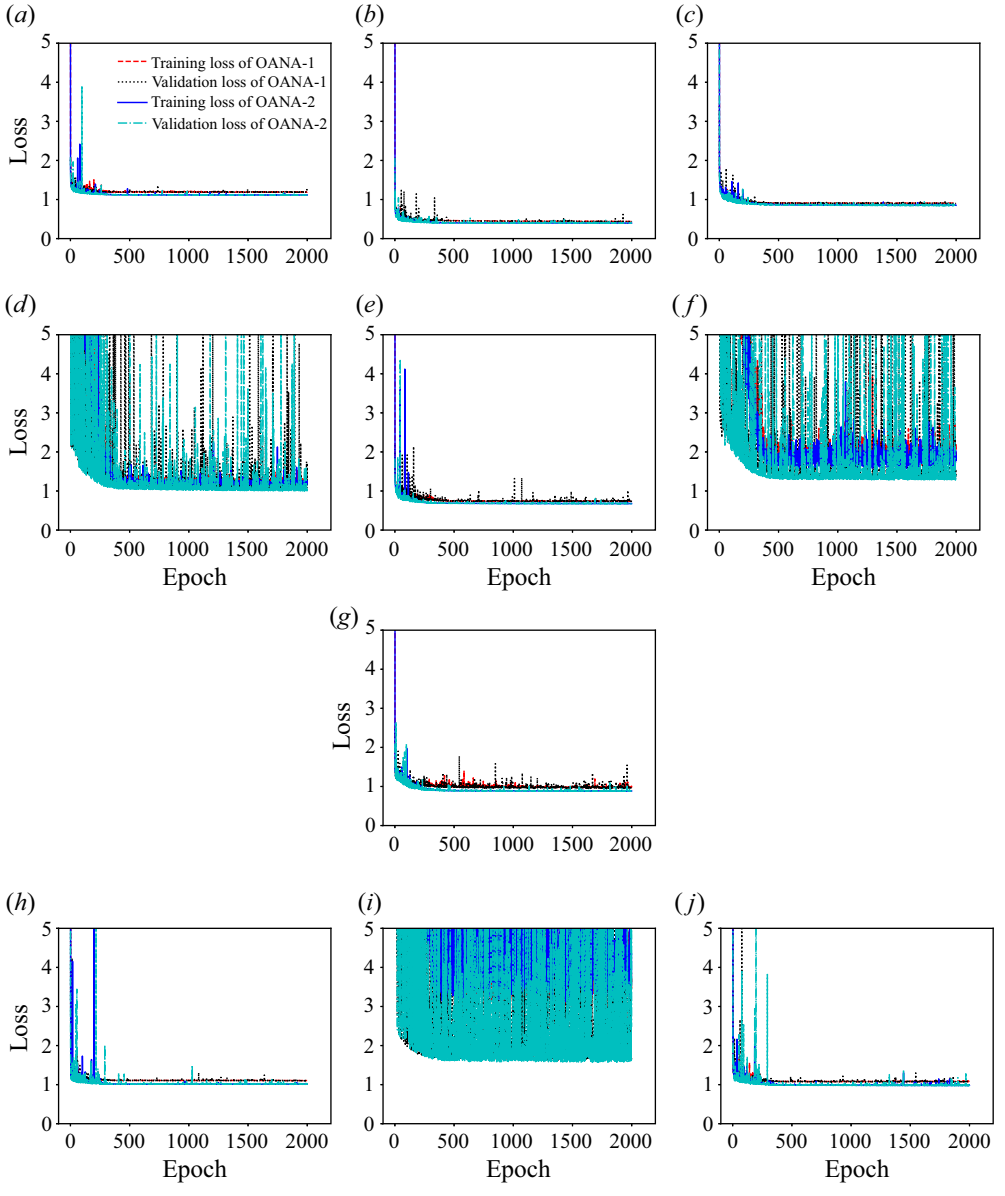


Figure 6. The learning curves of the training and validation loss of the OANA-1 and OANA-2 models for (*a*) τ_{11}^A, (*b*) τ_{22}^A, (*c*) τ_{33}^A, (*d*) τ_{23}^A, (*e*) τ_{13}^A, (*f*) τ_{12}^A, (*g*) τ_{kk}, (*h*) Q_1, (*i*) Q_2 and (*j*) Q_3.

In order to evaluate the performance of the OANA model in an *a priori* test, the correlation coefficients C and the relative errors E_r of the OANA-2 model in the testing set along wall-normal direction are shown in figure 7. It is noted that the correlation coefficients C and the relative errors E_r of the OANA-2 model in the training set and validating set are similar to those in the testing set, therefore we only exhibit the results of the testing set to evaluate the performance of the OANA-2 model in the untrained data. It is shown that the correlation coefficients C of most components are larger than 0.6, except for those of the components τ_{11}^A and τ_{33}^A, which are approximately larger than 0.4.

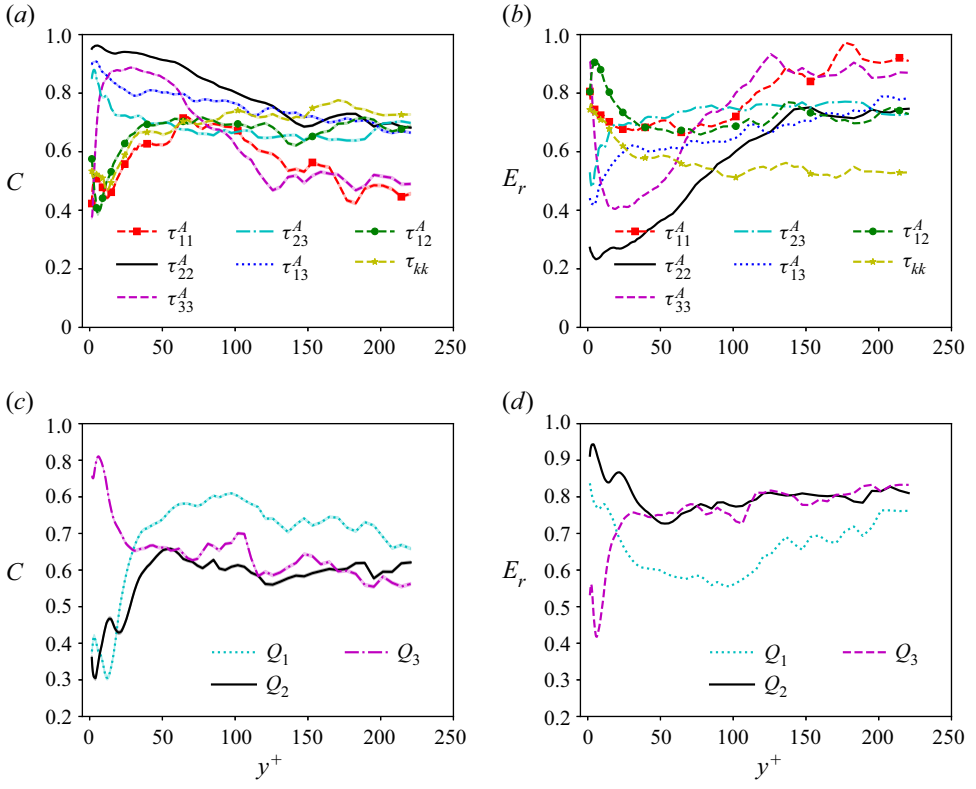


Figure 7. (*a*,*c*) The correlation coefficients C and (*b*,*d*) the relative errors E_r of the OANA-2 model in the testing set along wall-normal direction in *a priori* test with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15. The lighter sections of each line in (*a*,*c*) represent the 95 % confidence intervals for the correlation coefficients.

Even though the correlation coefficients C of OANA-2 model are not very high, they are still significantly larger than those of the aforementioned three eddy-viscosity SGS models (DSM, Vreman and WALE), which are mostly smaller than 0.4. Furthermore, the relative errors E_r are approximately equal to 0.8 in most components, which is relatively high but still much smaller than those of the aforementioned three eddy-viscosity SGS models (DSM, Vreman and WALE). On the other hand, it is also found that the OANA-2 model has significantly larger correlation coefficients than the ODNA model near the wall ($y^+ < 50$), while the correlation coefficients of OANA-2 model are similar to those of ODNA model far from the wall ($y^+ > 50$), which indicates that the ANN method can significantly enhance the correlation coefficients near the wall, while it makes a relatively small enhancement in the far-wall region compared with the dynamic method when predicting the model coefficients. Moreover, the relative errors of OANA-2 model are much smaller than those of the ODNA model, which suggests that the ANN method can significantly decrease the relative errors compared with the dynamics method. It is noted that increasing the numbers of hidden layers and neurons only gives a slight improvement of the model performance. Therefore, it is concluded that, although the OANA model has larger correlation coefficients, smaller relative errors and better performance than those of the aforementioned three eddy-viscosity SGS models (DSM, Vreman and WALE) and the ODNA model, it is still not a well-performing model.

3.2. *The modified ANN-based nonlinear algebraic model*

The main reason for the poor performance of the OANA model and ODNA model is that they ignore the anisotropy of the grid widths in the wall-bounded flows, especially in the wall-normal direction. In the near-wall region, the wall-normal grid width Δ_y is much smaller than the streamwise grid width Δ_x and the spanwise grid width Δ_z, which further leads to the observation that the wall-normal components τ_{23}^A, τ_{12}^A and Q_2 have strongly oscillating training and validation losses in the OANA model. Therefore, we propose a modified version of the nonlinear algebraic model. The core idea of the modified nonlinear algebraic model can be explained as follows: the local grid widths along the streamwise, wall-normal and spanwise directions Δ_x, Δ_y and Δ_z are used to normalise the corresponding gradients of the flow variables (including u_i and T), which can be expressed as

$$\frac{\partial u_i}{\partial x^*} \doteq \frac{\partial u_i}{\partial\,(x/\Delta_x)}, \quad \frac{\partial u_i}{\partial y^*} \doteq \frac{\partial u_i}{\partial(\,y/\Delta_y)}, \quad \frac{\partial u_i}{\partial z^*} \doteq \frac{\partial u_i}{\partial\,(z/\Delta_z)}; \tag{3.27a–c}$$

and

$$\frac{\partial T}{\partial x^*} \doteq \frac{\partial T}{\partial\,(x/\Delta_x)}, \quad \frac{\partial T}{\partial y^*} \doteq \frac{\partial T}{\partial(\,y/\Delta_y)}, \quad \frac{\partial T}{\partial z^*} \doteq \frac{\partial T}{\partial\,(z/\Delta_z)}; \tag{3.28a–c}$$

where x^*, y^* and z^* are the normalised variables, which are normalised by the corresponding local grid widths Δ_x, Δ_y and Δ_z, respectively. To be specific, the original gradients of the flow variables (including u_i and T) are first calculated by the sixth-order central-difference scheme, and then multiplied by the corresponding local grid widths to generate the modified gradients in (3.27a–c) and (3.28a–c). Therefore, the modified filtered strain-rate tensor $\widetilde{S^*}_{ij}$ and modified filtered rotation-rate tensor $\widetilde{\Omega^*}_{ij}$ become

$$\widetilde{S^*}_{ij} = \frac{1}{2}\left(\frac{\partial \tilde{u}_i}{\partial x_j^*} + \frac{\partial \tilde{u}_j}{\partial x_i^*}\right), \tag{3.29}$$

$$\widetilde{\Omega^*}_{ij} = \frac{1}{2}\left(\frac{\partial \tilde{u}_i}{\partial x_j^*} - \frac{\partial \tilde{u}_j}{\partial x_i^*}\right). \tag{3.30}$$

Then, the original nonlinear algebraic model can be modified as follows. The anisotropic SGS stress can be approximated as

$$\tau_{ij}^A = \Delta^{*2}\left(C_1^{*A}\left\|\widetilde{\boldsymbol{S}^*}\right\|\boldsymbol{T}_1^{*A} + C_2^{*A}\boldsymbol{T}_2^{*A} + C_3^{*A}\boldsymbol{T}_3^{*A} + C_4^{*A}\boldsymbol{T}_4^{*A} + C_5^{*A}\frac{1}{\left\|\widetilde{\boldsymbol{S}^*}\right\|}\boldsymbol{T}_5^{*A}\right), \tag{3.31}$$

where

$$\boldsymbol{T}_1^{*A} = \widetilde{\boldsymbol{S}^*} - \frac{1}{3}\boldsymbol{I}\cdot\mathrm{Tr}\left(\widetilde{\boldsymbol{S}^*}\right), \tag{3.32}$$

$$\boldsymbol{T}_2^{*A} = \widetilde{\boldsymbol{S}^*}\widetilde{\boldsymbol{\Omega}^*} - \widetilde{\boldsymbol{\Omega}^*}\widetilde{\boldsymbol{S}^*}, \tag{3.33}$$

$$\boldsymbol{T}_3^{*A} = \widetilde{\boldsymbol{S}^*}^2 - \frac{1}{3}\boldsymbol{I}\cdot\mathrm{Tr}\left(\widetilde{\boldsymbol{S}^*}^2\right), \tag{3.34}$$

$$\boldsymbol{T}_4^{*A} = \widetilde{\boldsymbol{\Omega}^*}^2 - \frac{1}{3}\boldsymbol{I}\cdot\mathrm{Tr}\left(\widetilde{\boldsymbol{\Omega}^*}^2\right), \tag{3.35}$$

$$\boldsymbol{T}_5^{*A} = \widetilde{\boldsymbol{\Omega}^*}\widetilde{\boldsymbol{S}^*}^2 - \widetilde{\boldsymbol{S}^*}^2\widetilde{\boldsymbol{\Omega}^*}. \tag{3.36}$$

The isotropic SGS stress can be approximated as

$$\tau_{kk} = \Delta^{*2}\left(C_0^{*I}\left\|\widetilde{\boldsymbol{S}}^*\right\|^2 + C_1^{*I}\left\|\widetilde{\boldsymbol{S}}^*\right\|\boldsymbol{T}_1^{*I} + C_2^{*I}\boldsymbol{T}_2^{*I} + C_3^{*I}\boldsymbol{T}_3^{*I}\right), \tag{3.37}$$

where

$$\boldsymbol{T}_1^{*I} = \boldsymbol{I}\cdot\mathrm{Tr}\left(\widetilde{\boldsymbol{S}}^*\right), \tag{3.38}$$

$$\boldsymbol{T}_2^{*I} = \boldsymbol{I}\cdot\mathrm{Tr}\left(\widetilde{\boldsymbol{S}}^{*2}\right), \tag{3.39}$$

$$\boldsymbol{T}_3^{*I} = \boldsymbol{I}\cdot\mathrm{Tr}\left(\widetilde{\boldsymbol{\Omega}}^{*2}\right). \tag{3.40}$$

Here, $\Delta^* = 1$ is the normalised characteristic length scale and $\|\widetilde{\boldsymbol{S}}^*\| = \sqrt{\mathrm{Tr}(\widetilde{\boldsymbol{S}}^{*2})}$.

Moreover, the modified model coefficients C_i^{*A} and C_i^{*I} in (3.31) and (3.37) are functions of the modified four dimensionless invariants, which can be expressed as

$$\left.\begin{aligned} d_1^* &= \frac{\mathrm{Tr}\left(\widetilde{\boldsymbol{\Omega}}^{*2}\right)}{\mathrm{Tr}\left(\widetilde{\boldsymbol{S}}^{*2}\right)}, \quad d_2^* = \frac{\mathrm{Tr}\left(\widetilde{\boldsymbol{S}}^{*3}\right)}{\mathrm{Tr}\left(\widetilde{\boldsymbol{S}}^{*2}\right)^{3/2}}, \\ d_3^* &= \frac{\mathrm{Tr}\left(\widetilde{\boldsymbol{\Omega}}^{*2}\widetilde{\boldsymbol{S}}^*\right)}{\mathrm{Tr}\left(\widetilde{\boldsymbol{S}}^{*2}\right)^{1/2}\mathrm{Tr}\left(\widetilde{\boldsymbol{\Omega}}^{*2}\right)}, \quad d_4^* = \frac{\mathrm{Tr}\left(\widetilde{\boldsymbol{\Omega}}^{*2}\widetilde{\boldsymbol{S}}^{*2}\right)}{\mathrm{Tr}\left(\widetilde{\boldsymbol{S}}^{*2}\right)\mathrm{Tr}\left(\widetilde{\boldsymbol{\Omega}}^{*2}\right)}. \end{aligned}\right\} \tag{3.41}$$

Similarly, the SGS heat flux can be approximated as

$$Q_j = \Delta^{*2}\left(C_1^{*Q}\left|\widetilde{\boldsymbol{S}}^*\right|\frac{\partial\tilde{T}}{\partial x_j^*} + C_2^{*Q}\widetilde{S}^*_{jk}\frac{\partial\tilde{T}}{\partial x_k^*} + C_3^{*Q}\widetilde{\Omega}^*_{jk}\frac{\partial\tilde{T}}{\partial x_k^*} + C_4^{*Q}\frac{\widetilde{S}^*_{jk}\widetilde{S}^*_{kl}}{\left|\widetilde{\boldsymbol{S}}^*\right|}\frac{\partial\tilde{T}}{\partial x_l^*}\right), \tag{3.42}$$

where $|\widetilde{\boldsymbol{S}}^*| = \sqrt{2\widetilde{S}^*_{ij}\widetilde{S}^*_{ij}}$. The model coefficients C_i^{*Q} in (3.42) depend on the principal invariants of tensor $\widetilde{S}^*_{ij}$ and vector $\tilde{v}^*_i = \Delta^{*2}|\widetilde{\boldsymbol{S}}^*|(\partial\tilde{T}/\partial x_i^*)$, which are written as

$$\left.\begin{aligned} q_1^* &= \widetilde{S}^*_{ii}, \quad q_2^* = \widetilde{S}^*_{ij}\widetilde{S}^*_{ji}, \quad q_3^* = \widetilde{S}^*_{ik}\widetilde{S}^*_{kl}\widetilde{S}^*_{li}, \\ q_4^* &= \tilde{v}^*_i\tilde{v}^*_i, \quad q_5^* = \tilde{v}^*_i\widetilde{S}^*_{ik}\tilde{v}^*_k, \quad q_6^* = \tilde{v}^*_i\widetilde{S}^*_{ik}\widetilde{S}^*_{kj}\tilde{v}^*_j. \end{aligned}\right\} \tag{3.43}$$

Similar to the ODNA model, the modified dimensionless model coefficients C_i^{*A}, C_i^{*I} and C_i^{*Q} in (3.31), (3.37) and (3.42), respectively, can also be solved dynamically based on the Germano identity (Germano *et al.* 1991; Lilly 1992; Yuan *et al.* 2022). Therefore, this modified dynamic nonlinear algebraic model is named the 'MDNA model'. The correlation coefficients C and the relative errors E_r of the MDNA model along the wall-normal direction in *a priori* tests with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15 are shown in figure 8. It is found that, compared with the correlation coefficients of the ODNA model in figure 4(*a*,*c*), the correlation coefficients of the MDNA model have significantly larger values than those of the ODNA model. The enhancement is large in the near-wall region ($y^+ < 50$) where the anisotropy of the grid widths is significantly strong. As y^+ further increases, the enhancement of the MDNA model becomes weaker, mainly due to

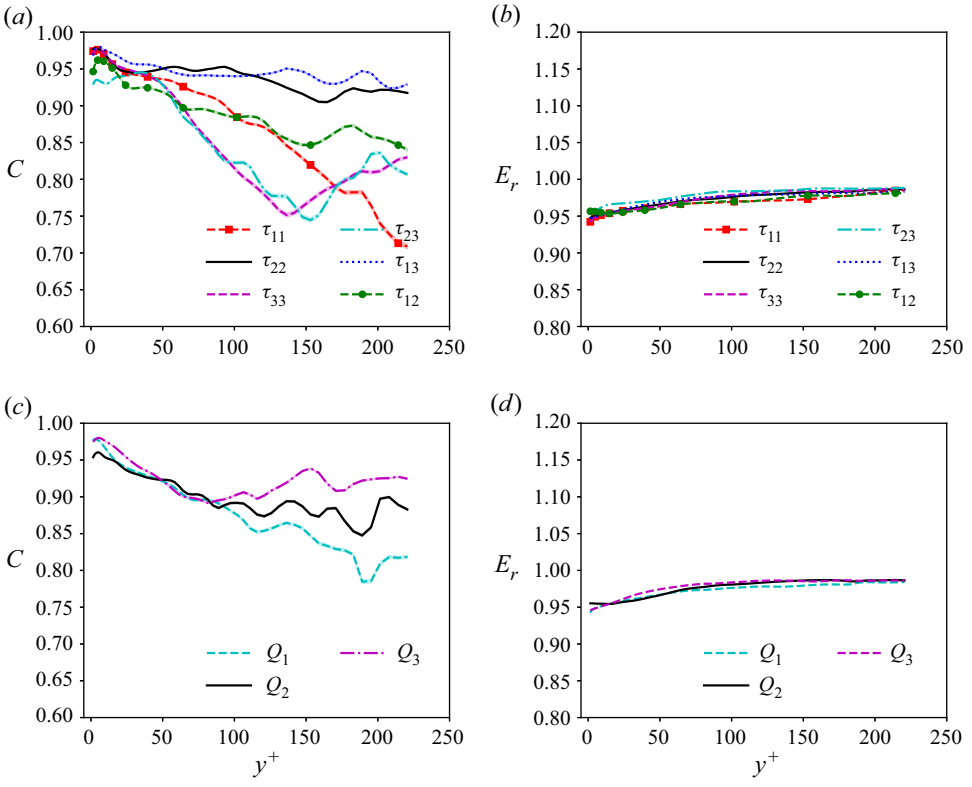


Figure 8. (*a*,*c*) The correlation coefficients C and (*b*,*d*) the relative errors E_r of MDNA model along the wall-normal direction in *a priori* tests with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15. The lighter sections of each line in (*a*,*c*) represent the 95 % confidence intervals for the correlation coefficients.

the fact that the anisotropy of the grid widths also becomes weaker. These observations indicate that the above modification of using the local grid widths along each direction to normalise the corresponding gradients of the flow variables can significantly enhance the correlation coefficients of the dynamic nonlinear algebraic model, especially in the near-wall region. Moreover, the relative errors of the MDNA model are close to 0.95, which are slightly smaller than those of the ODNA model. This observation suggests that the above modification can slightly decrease the relative errors of the dynamic nonlinear algebraic model.

It is shown in figure 8 that the MDNA model has a better performance than the eddy-viscosity models (DSM, Vreman and WALE) and the ODNA model, while has a similar performance to the OANA model. It is found in figures 4 and 7 that the ANN-based model has a better performance than the dynamic model. Therefore, it can be speculated that the ANN-based version of the modified nonlinear algebraic model will have a better performance. Similar to the OANA model, the modified dimensionless model coefficients C_i^{*A}, C_i^{*I} and C_i^{*Q} in (3.31), (3.37) and (3.42), respectively, can be determined by the ANN method, therefore the modified ANN-based nonlinear algebraic model is named the 'MANA model'. The ANN schematic, parameter setting of the ANN and the training process of the MANA model are all the same as those in OANA model shown in § 3.1, except for that the input features are changed to the corresponding invariants, which are shown in table 2. It is noted that the MANA model considers the anisotropy of the grid

Input features	Coefficient output layer	Merge output layer	N_i	N_o
$\{d_1^*, d_2^*, d_3^*, d_4^*, Re_l\}$	C_i^A $(i = 1, \ldots, 5)$	τ_{ij}^A	5	5
$\{d_1^*, d_2^*, d_3^*, d_4^*, Re_l\}$	C_i^I $(i = 0, \ldots, 3)$	τ_{kk}	5	4
$\{q_1^*, q_2^*, q_3^*, q_4^*, q_5^*, q_6^*, Re_l\}$	C_i^Q $(i = 1, \ldots, 4)$	Q_j	7	4

Table 2. The input features, the variable of the coefficient output layer, the variable of the merge output layer and the values of N_i and N_o of the MANA model for the unclosed SGS terms τ_{ij}^A, τ_{kk} and Q_j, respectively. Here, the input features d_1^*–d_4^* are the invariants shown in (3.41), and the input features q_1^*–q_6^* are the invariants shown in (3.43).

widths in wall-bounded flows, and it is expected that the MANA model will have a better performance in the *a priori* and *a posteriori* tests. The influences of the size of the mini-batch N_{batch}, different optimisers and activation functions on the performance of the MANA model in *a priori* test are checked in the Appendix.

We train three sets of MANA model with different numbers of neurons of the hidden layers N_h, and name them 'MANA-1' for $N_h = 10$, 'MANA-2' for $N_h = 20$ and 'MANA-3' for $N_h = 40$. The learning curves of the training and validation losses of the MANA-1, MANA-2 and MANA-3 models for τ_{ij}^A, τ_{kk} and Q_j are shown in figure 9. It is found that the training and validation losses are close to each other, and converge well, which indicates that the MANA models are well trained and avoid overfitting phenomena. Furthermore, the strong oscillation phenomena observed in the training and validation losses of the wall-normal components τ_{23}^A, τ_{12}^A and Q_2 of OANA model (figure 6) disappear in the validation losses of the MANA model, which indicates that the above modification in the MANA model can alleviate the effect of the anisotropy of the grid widths in wall-bounded flows. The weights and biases of the MANA model used in the following *a priori* test are acquired from the epoch where the validation losses reach the minimum values. It is also found that the training and validation losses slightly decrease as the number of neurons of the hidden layers increases, indicating a slightly better performance of the MANA model with a larger number of hidden layer neurons.

In order to explicitly evaluate the performance of the MANA model in *a priori* tests, the correlation coefficients C and the relative errors E_r of the MANA-1 model in the testing set along the wall-normal direction are shown in figure 10. It is found that the correlation coefficients of τ_{ij}^A, τ_{kk} and Q_j in the MANA-1 model are all larger than 0.91, which are much higher than the ODNA model, OANA model, MDNA model and the three traditional eddy-viscosity models. Furthermore, the relative errors in the MANA-1 model are all smaller than 0.4, which are much smaller than the ODNA model, OANA model, MDNA model and the three traditional eddy-viscosity models. These observations indicate that the MANA model has a much better performance than the ODNA model, OANA model, MDNA model and the three traditional eddy-viscosity models in *a priori* tests. Furthermore, in order to examine the influence of the numbers of hidden layer neurons, the correlation coefficients C and the relative errors E_r of τ_{11}^A and Q_1 are chosen as representatives to compare the performances of the MANA-1, MANA-2 and MANA-3 models in figure 11. It is found the numbers of hidden layer neurons have little influence on the performance of the MANA model in *a priori* tests. Therefore, considering the computational cost of the MANA model in *a posteriori* tests, we use the MANA-1 model with $N_h = 10$ as the final MANA model in *a posteriori* tests. The weights and biases of

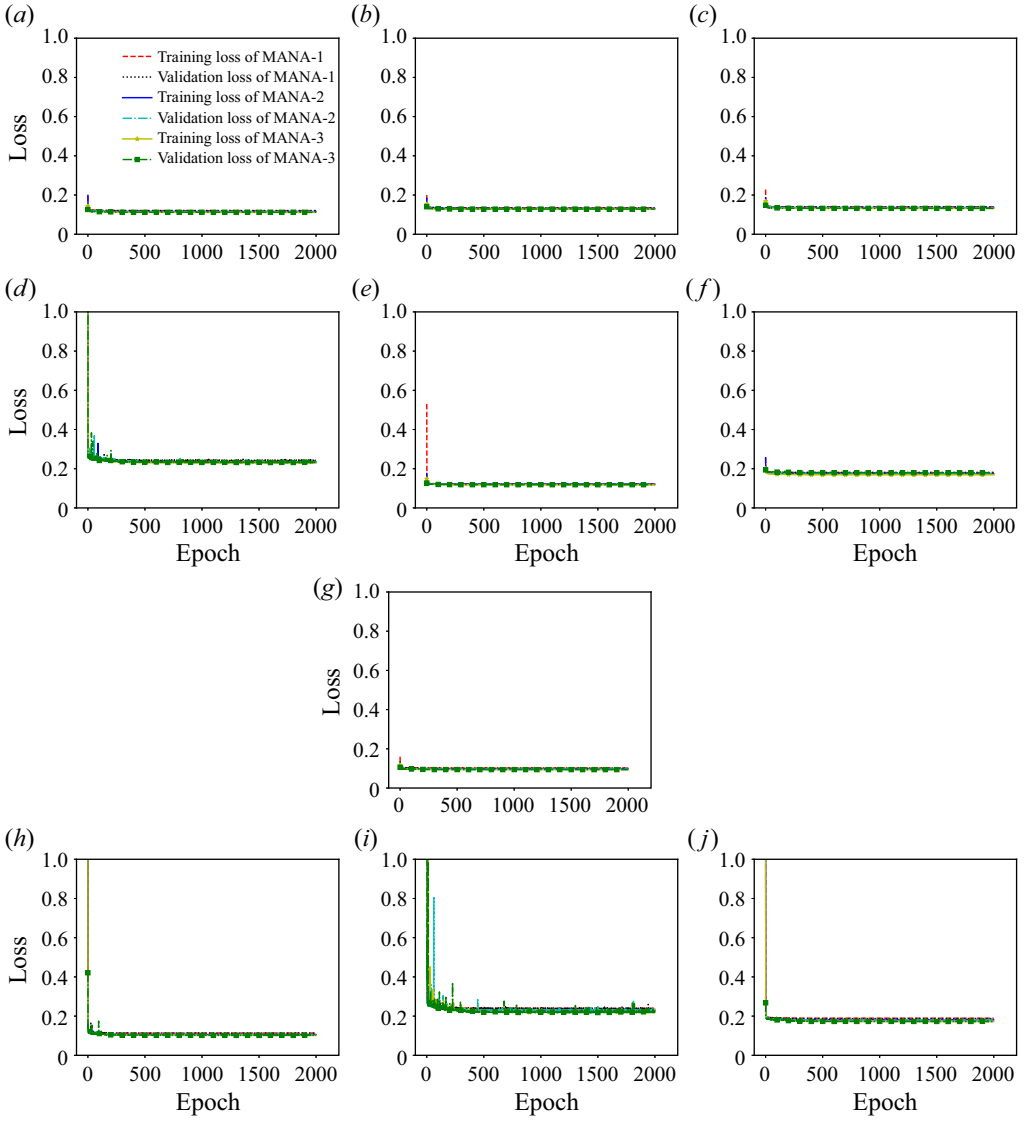


Figure 9. The learning curves of the training and validation losses of the MANA-1, MANA-2 and MANA-3 models for (*a*) τ_{11}^A, (*b*) τ_{22}^A, (*c*) τ_{33}^A, (*d*) τ_{23}^A, (*e*) τ_{13}^A, (*f*) τ_{12}^A, (*g*) τ_{kk}, (*h*) Q_1, (*i*) Q_2 and (*j*) Q_3.

the MANA model used in the *a posteriori* tests are given in the supplementary material available at https://doi.org/10.1017/jfm.2023.179.

The SGS flux of the kinetic energy can be defined as $\Pi_\tau = -\tau_{ij}\partial\tilde{u}_i/\partial x_j$, which represents the energy transfer rate of the kinetic energy from the large scales to the small scales (Eyink 2005; Aluie & Eyink 2009; Eyink & Aluie 2009; Wang *et al.* 2018*a*; Xu *et al.* 2021*b*); similarly, the SGS flux of the temperature variance can be written as $\Pi_Q = -Q_j\partial\tilde{T}/\partial x_j$, which indicates the energy transfer rate of the temperature variance from the large scales to the small scales (Jiménez *et al.* 2001; Vollant *et al.* 2017). Furthermore, the positive values of the SGS fluxes of the kinetic energy and the temperature variance suggest the direct energy transfer from the large scales to the small scales, while the

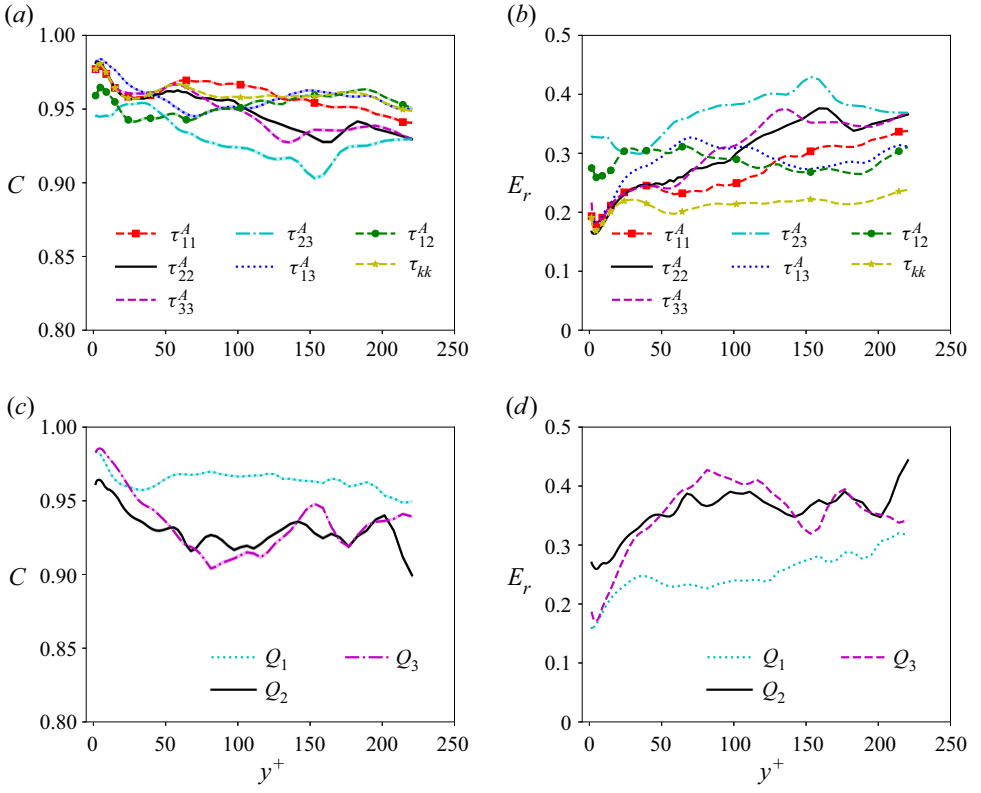


Figure 10. (*a*,*c*) The correlation coefficients C and (*b*,*d*) the relative errors E_r of the MANA-1 model in the testing set along the wall-normal direction in *a priori* test with filtered sizes $n_x = 8$ and $n_z = 4$ in R1M15. The shaded areas under each line in (*a*,*c*) represent the 95 % confidence intervals for the correlation coefficients.

negative values represent the inverse energy transfer from the small scales to the large scales (Eyink 2005; Aluie & Eyink 2009; Eyink & Aluie 2009; Wang *et al.* 2018*a*; Xu *et al.* 2021*b*). The normalised SGS fluxes of the kinetic energy and the temperature variance can be expressed as $\Pi^+_\tau = (-\tau_{ij}\partial\tilde{u}_i/\partial x_j)/(u^3_\tau/\delta_\nu)$ and $\Pi^+_Q = (-Q_j\partial\tilde{T}/\partial x_j)/(u_\tau T^2_w/\delta_\nu)$, respectively. The streamwise–spanwise averages of the normalised SGS fluxes of the kinetic energy and the temperature variance $\langle\Pi^+_\tau\rangle_{xz}$ and $\langle\Pi^+_Q\rangle_{xz}$ along the wall-normal direction for the fDNS, Vreman, DSM, WALE, MDNA and MANA models in the *a priori* tests are shown in figure 12. It is shown that the traditional eddy-viscosity models (Vreman, DSM and WALE) and the dynamic nonlinear algebraic model (MDNA) significantly underestimate the mean normalised SGS fluxes of the kinetic energy and the temperature variance $\langle\Pi^+_\tau\rangle_{xz}$ and $\langle\Pi^+_Q\rangle_{xz}$, while the $\langle\Pi^+_\tau\rangle_{xz}$ and $\langle\Pi^+_Q\rangle_{xz}$ profiles of the MANA model almost collapse to those of the fDNS result, which indicates that the MANA model has a much better performance in predicting the mean normalised SGS fluxes of the kinetic energy and the temperature variance compared with the traditional eddy-viscosity models and the dynamic nonlinear algebraic model (MDNA) in *a priori* tests.

3.3. *Discussion about the performance of different models*

It is shown that the traditional eddy-viscosity models (DSM, Vreman and WALE) have poor performances in *a priori* tests, which can be ascribed to the unphysical assumption

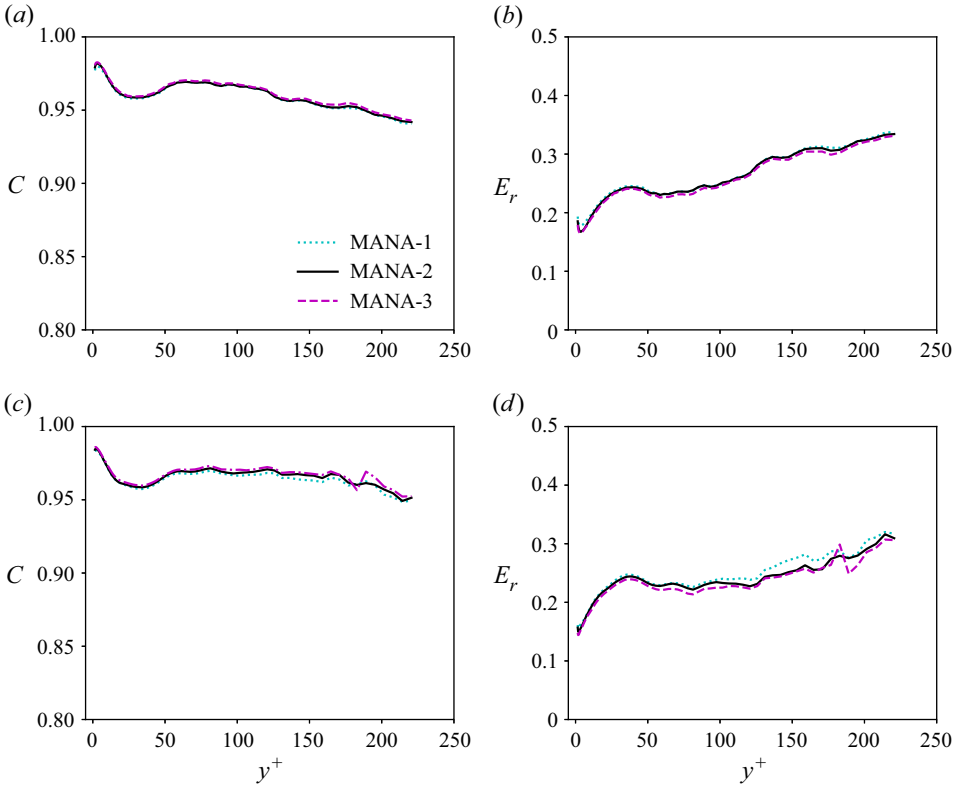


Figure 11. (*a*,*b*) The correlation coefficients C and the relative errors E_r of τ_{11}^A in the MANA-1, MANA-2 and MANA-3 models in the testing set along the wall-normal direction. (*c*,*d*) The correlation coefficients C and the relative errors E_r of Q_1 in the MANA-1, MANA-2 and MANA-3 models in the testing set along the wall-normal direction.

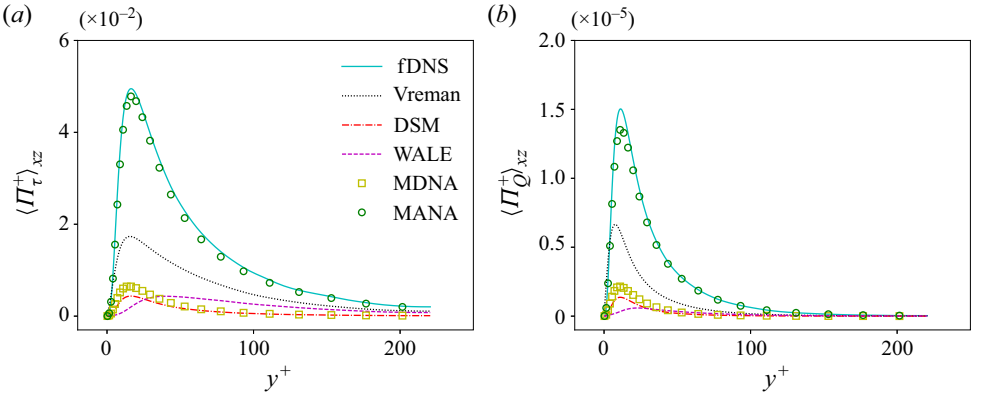


Figure 12. The streamwise–spanwise averages of the normalised SGS fluxes of the kinetic energy and the temperature variance (*a*) $\langle \Pi_\tau^+ \rangle_{xz}$ and (*b*) $\langle \Pi_Q^+ \rangle_{xz}$ along the wall-normal direction for the fDNS, Vreman, DSM, WALE, MDNA and MANA models in the *a priori* test.

that the SGS stress and SGS heat flux are linearly aligned with the filtered strain-rate tensor and the gradient of the filtered temperature, respectively (Anderson & Meneveau 1999; Da Silva & Métais 2002; Xie *et al.* 2019*a*,*c*, 2020*a*,*b*). These issues can be addressed by the nonlinear algebraic models. It is found that the ODNA model has larger correlations and smaller relative errors than the traditional eddy-viscosity models (DSM, Vreman and WALE), indicating that the nonlinear algebraic model has a better performance than the eddy-viscosity models (DSM, Vreman and WALE). However, the ODNA model ignores the anisotropy of the grid widths in wall-bounded flows, especially in the wall-normal direction. Therefore, an innovative modification is applied on the invariants and the tensor bases of the nonlinear algebraic models through using the local grid widths along each direction to normalise the corresponding gradients of the flow variables. After applying this modification, the correlation coefficients of the MDNA model are significantly enhanced compared with those of the ODNA model, especially in the near-wall region. The relative errors of the MDNA model are also smaller than those of the ODNA model. When the model coefficients are estimated by the ANN method rather than the dynamic method, the model performance in *a priori* tests is significantly enhanced. Therefore, it is concluded that the nonlinear algebraic expressions of the model, the innovative modification of the expressions of the invariants and tensor bases and the model coefficients estimated by the ANN method all make significant contributions to the much better performance of the MANA model in *a priori* tests.

4. The *a posteriori* test of the MANA model

4.1. *Application in the compressible turbulent channel flow with untrained grid resolution*

Initially, the performance of the proposed MANA model is tested in the compressible turbulent channel flow case R1M15 with an untrained grid resolution. The filtered Navier–Stokes equations (2.1), (2.8), (2.9), (2.4) are solved by a high-order finite-difference solver in Cartesian coordinates: a sixth-order central-difference scheme is applied for the discretisation of both the convective and viscous terms, and the LES equations are temporally integrated by the third-order Runge–Kutta scheme (Shu & Osher 1988). Similar to the grid setting in Jiang *et al.* (2013), the grid resolution of $64 \times 65 \times 64$ and streamwise and spanwise filtered sizes $n_x = 6$ and $n_z = 2$ are used for LES of compressible turbulent channel flow using the MANA model, implicit large-eddy simulation (ILES) method, DSM, Vreman model and WALE model. It is noted that, when the grid numbers in statistically homogeneous directions are small, it is found that the MDNA model is sometimes unstable in *a posteriori* tests. Moreover, the MDNA model is much computationally expensive, almost 2.54 times the cost of DSM when simulating the same 1000 steps. Therefore, the performance of the MDNA model is not exhibited in *a posteriori* tests. The streamwise and spanwise filtered sizes in the *a posteriori* tests are different from those in the training process, where $n_x = 8$ and $n_z = 4$ are used, respectively. The details of the grid settings of the DNS and LES are listed in table 3. Furthermore, the computational cost of the MANA model is compared with the traditional eddy-viscosity models and ILES. The computational times for simulating the same 1000 steps are compared and listed in table 4, where the ratios of computational time in all five LES models are shown and the computational time is normalised by that of DSM. It is found that the dynamic model (DSM) is much more time consuming than the constant coefficient models (Vreman, WALE). However, the proposed MANA model

	$N_x \times N_y \times N_z$	Δx^+	Δy_w^+	Δz^+
DNS	$384 \times 193 \times 128$	7.13	0.30	7.13
ILES	$64 \times 65 \times 64$	43.4	0.90	14.5
Vreman	$64 \times 65 \times 64$	41.7	0.86	13.9
DSM	$64 \times 65 \times 64$	40.9	0.85	13.6
WALE	$64 \times 65 \times 64$	42.4	0.88	14.1
MANA	$64 \times 65 \times 64$	43.2	0.90	14.4

Table 3. The details of the grid settings of the DNS, ILES, Vreman model, DSM, WALE model and MANA model in the compressible channel flow case R1M15 with $Re = 3000$ and $M = 1.5$. Here, N_x, N_y and N_z are the grid resolutions in the streamwise, wall-normal and spanwise directions, respectively.

Model	ILES	Vreman	DSM	WALE	MANA
Ratio of computational time	0.26	0.28	1.0	0.29	0.79

Table 4. Ratios of computational time for simulating the same 1000 steps for all five LES models. The computational time is normalised by that of DSM. The simulations are testing on 80 2.40 GHz Intel Xeon Gold 6148 CPUs. The computational time of DSM is approximately 15.4 s.

is computationally much cheaper than DSM, which is also much cheaper than the other ANN-based models proposed previously in Park & Choi (2021), Wang *et al.* (2018*b*), Yuan *et al.* (2020), Xie *et al.* (2019*c*) and Xie *et al.* (2020*a*).

The van Driest transformed velocity, U_{vd}^+, is defined as

$$U_{vd}^+ = \int_0^{U^+} \left(\langle\bar{\rho}\rangle_{xz} / \langle\bar{\rho}_w\rangle_{xz}\right)^{1/2} \mathrm{d}U^+, \tag{4.1}$$

where $U^+ = \langle\tilde{u}\rangle_{xz}/u_\tau$. The van Driest transformed velocity U_{vd}^+ and the normalised mean temperature profile $\langle\tilde{T}\rangle_{xz}/T_w$ along the wall-normal direction for fDNS and LES with different SGS models in R1M15 are plotted in figure 13. It is noted that the fDNS results in the R1M15 case are calculated by applying the top-hat filter with the streamwise and spanwise filtered sizes $n_x = 6$ and $n_z = 2$ to the DNS flow fields. It is shown in figure 13(*a*) that, in the viscous sublayer ($y^+ < 5$), the van Driest transformed velocity U_{vd}^+ increases linearly with y^+. Furthermore, the results of all SGS models collapse to the fDNS result in the viscous sublayer and the lower part of the buffer layer where $y^+ < 10$. In the log-law region, the U_{vd}^+ profile of the MANA model is fully consistent with that of the fDNS result. However, the U_{vd}^+ profiles of the traditional eddy-viscosity models (Vreman, DSM and WALE) are significantly larger than that of the fDNS result, and the U_{vd}^+ profile of ILES is slightly lower than that of the fDNS result. The above observations indicate that the MANA model provides the proper SGS dissipation, while the traditional eddy-viscosity models give excessive SGS dissipation, and the ILES lacks SGS dissipation, consistent with previous studies on LES of isotropic turbulence and turbulent channel flow (Xie *et al.* 2019*a*,*c*; Yu *et al.* 2022).

It is shown in figure 13(*b*) that all LES results collapse to the mean temperature profile of the fDNS result in most regions of the channel, except for some regions in the centre of the channel. It is found that the mean temperature profiles of the traditional eddy-viscosity models in the centre of the channel are slightly larger than that of the fDNS result, while

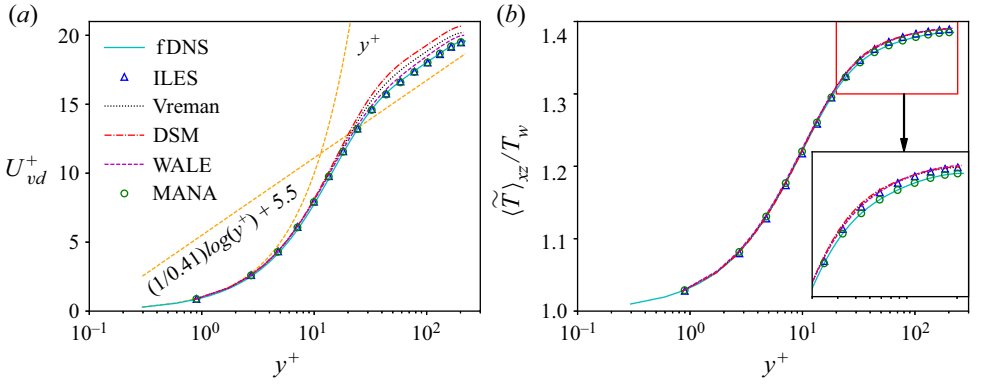


Figure 13. (*a*) The van Driest transformed velocity U^+_{vd} and (*b*) the normalised mean temperature profile $\langle\tilde{T}\rangle_{xz}/T_w$ along the wall-normal direction for fDNS and LES with different SGS models in R1M15. It is noted that the fDNS results in the R1M15 case are calculated by applying the top-hat filter with the streamwise and spanwise filtered sizes $n_x = 6$ and $n_z = 2$ to the DNS flow fields.

the mean temperature profile of the MANA model is fully consistent with that of the fDNS result.

The normalised r.m.s. values of the resolved fluctuating velocities are defined as $\tilde{u}^+_{i,rms} = \langle(\tilde{u}_i - \langle\tilde{u}_i\rangle_{xz})^2\rangle^{1/2}_{xz}/u_\tau$, which represent the resolved turbulence intensities. Furthermore, $\tilde{R}^+_{uv} = \langle(\tilde{u} - \langle\tilde{u}\rangle_{xz})(\tilde{v} - \langle\tilde{v}\rangle_{xz})\rangle_{xz}/u^2_\tau$ represents the normalised resolved Reynolds shear stress. The normalised r.m.s. values of the resolved fluctuating velocities $\tilde{u}^+_{i,rms}$ and the normalised resolved Reynolds shear stress $\tilde{R}^+_{uv}$ along the wall-normal direction for fDNS and LES with different SGS models in R1M15 are shown in figure 14. In figure 14(*a*), it is found that the ILES and the traditional eddy-viscosity models including Vreman, DSM and WALE, exhibit larger turbulence intensities than that of the fDNS result, while the $\tilde{u}^+_{rms}$ profile of the proposed MANA model fully collapses to that of the fDNS result. Similarly, the $\tilde{v}^+_{rms}$, $\tilde{w}^+_{rms}$ and $\tilde{R}^+_{uv}$ profiles of ILES, Vreman, DSM and WALE in figure 14(*b–d*) exhibit significant deviations from those of the fDNS result, where the peak values and the wall-normal locations of the peak values are much larger in ILES and the traditional eddy-viscosity models. However, the $\tilde{v}^+_{rms}$, $\tilde{w}^+_{rms}$ and $\tilde{R}^+_{uv}$ profiles of the MANA model are fully consistent with those of the fDNS result. To sum up, The MANA model can predict the resolved turbulent intensities and resolved Reynolds shear stress accurately and exhibits a much better performance than the ILES and the traditional eddy-viscosity models. In particular, the peak values and the locations of the peak values of the resolved turbulent intensities and resolved Reynolds shear stress are well predicted in the MANA model, while larger peak values and peak value locations further away from the wall for the resolved turbulent intensities are observed in the ILES and the traditional eddy-viscosity models.

The visualisations of the instantaneous vortical structures based on the Q-criterion (Hunt *et al.* 1988) for fDNS and LES with different SGS models in R1M15 are shown in figure 15. The structures are coloured by the wall-normal distance h, and the isosurfaces of the instantaneous vortical structures represent $Q = 0.4$. It is shown in figure 15(*a*) that the vortical structures exhibit abundant long tube-like structures in the fDNS result, and the vortical structures near the wall are much thinner and longer than those far from the wall, mainly due to the strong mean shear near the wall. Furthermore, the vortical

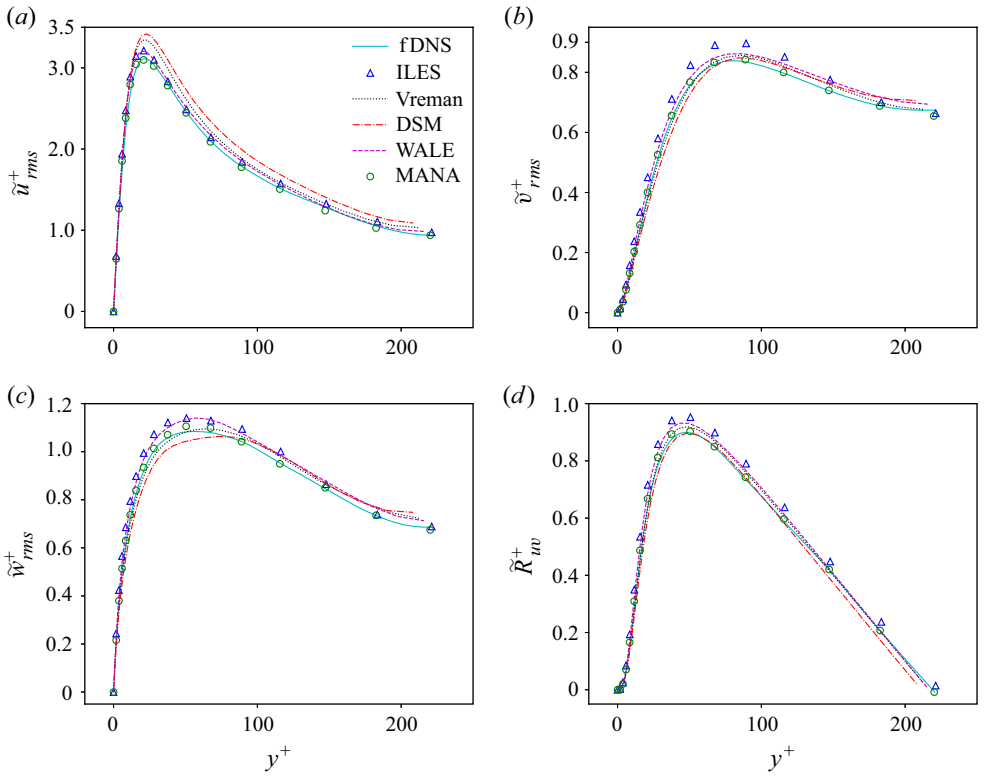


Figure 14. The normalised r.m.s. values of the resolved fluctuating velocities and the normalised resolved Reynolds shear stress along the wall-normal direction for fDNS and LES with different SGS models in R1M15. (*a*) The streamwise resolved fluctuating velocities $\tilde{u}^+_{rms}$, (*b*) the wall-normal resolved fluctuating velocities $\tilde{v}^+_{rms}$, (*c*) the spanwise resolved fluctuating velocities $\tilde{w}^+_{rms}$ and (*d*) the resolved Reynolds shear stress $\tilde{R}^+_{uv}$.

structures predicted by all SGS models are much sparser than those in the fDNS result, which arises from the sparser grid resolutions in the LES with different SGS models. Among the LES models, the MANA model exhibits more small-scale structures, while other LES models show fatter and longer tube-like structures, which indicates that the MANA model predicts more small-scale turbulent fluctuations. Accordingly, the above observations indicate that the MANA model is better than the other LES models in accurately predicting the small-scale turbulent structures.

The normalised r.m.s. value of the resolved fluctuating temperature is defined as $\tilde{T}^+_{rms} = \langle(\tilde{T} - \langle\tilde{T}\rangle_{xz})^2\rangle^{1/2}_{xz}/T_w$, and the normalised resolved streamwise Reynolds heat flux is defined as $\tilde{R}^+_{uT} = \langle(\tilde{u} - \langle\tilde{u}\rangle_{xz})(\tilde{T} - \langle\tilde{T}\rangle_{xz})\rangle_{xz}/(u_\tau T_w)$. The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along the wall-normal direction from fDNS and LES with different SGS models in R1M15 are shown in figure 16. It is shown that $\tilde{T}^+_{rms}$ and $\tilde{R}^+_{uT}$ in the ILES and the traditional eddy-viscosity models are quite close to the fDNS results, except for much larger peak values of $\tilde{T}^+_{rms}$ and $\tilde{R}^+_{uT}$. However, the $\tilde{T}^+_{rms}$ and $\tilde{R}^+_{uT}$ profiles in the MANA model perfectly collapse to those in the fDNS fields, indicating the accurate predictions of the MANA model for the thermodynamic statistics.

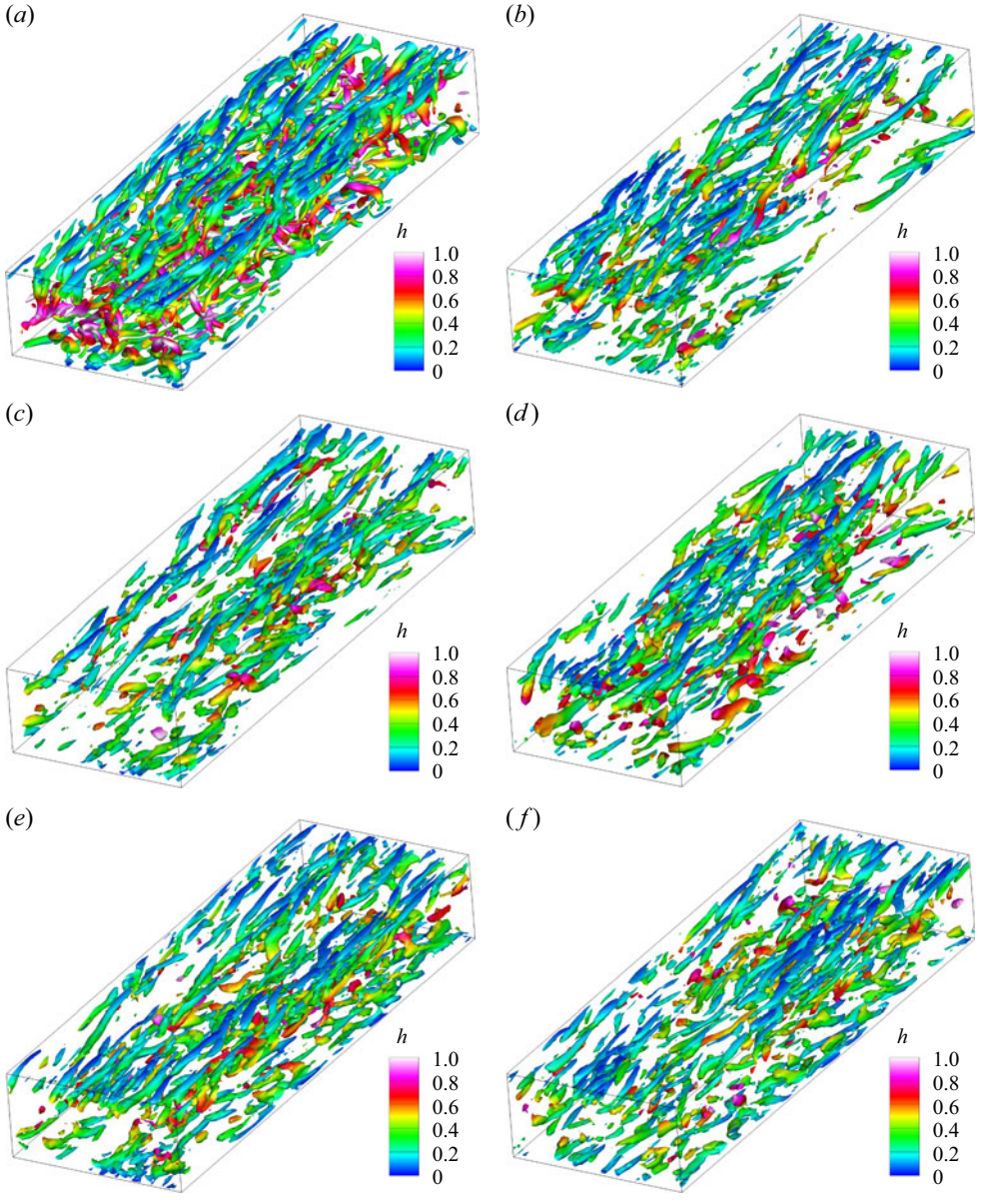


Figure 15. Visualisation of the instantaneous vortical structures based on the Q-criterion (Hunt, Wray & Moin 1988) in the (*a*) fDNS, (*b*) ILES, (*c*) Vreman, (*d*) DSM, (*e*) WALE and (*f*) MANA models. The structures are coloured by the wall-normal distance h, and the isosurfaces of the instantaneous vortical structures represent $Q = 0.4$. It is noted that the fDNS result is generated by applying streamwise and spanwise filtering with filter sizes $n_x = 6$ and $n_z = 2$ to the DNS flow fields; however, the grid resolution remains as $384 \times 193 \times 128$.

The streamwise–spanwise averages of the normalised SGS fluxes of the kinetic energy and the temperature variance $\langle \Pi_\tau^+ \rangle_{xz}$ and $\langle \Pi_Q^+ \rangle_{xz}$ along the wall-normal direction for the fDNS, ILES, Vreman, DSM, WALE and MANA models in *a posteriori* tests in R1M15 are depicted in figure 17. The ILES is the LES with no SGS model, therefore the normalised SGS fluxes of the kinetic energy and the temperature variance $\langle \Pi_\tau^+ \rangle_{xz}$

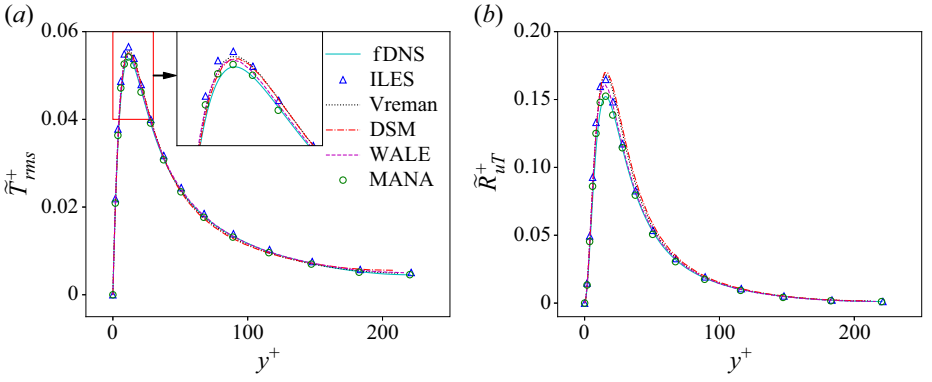


Figure 16. (*a*) The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and (*b*) the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along the wall-normal direction from fDNS and LES with different SGS models in R1M15.

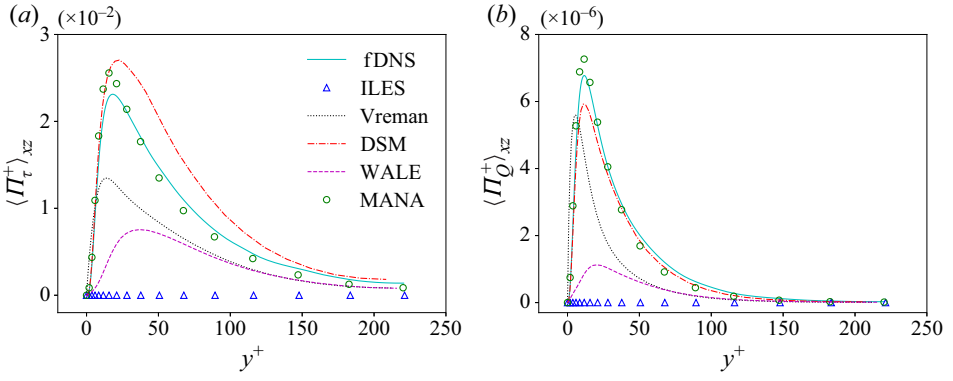


Figure 17. The streamwise–spanwise averages of the normalised SGS fluxes of the kinetic energy and the temperature variance (*a*) $\langle \Pi^+_\tau \rangle_{xz}$ and (*b*) $\langle \Pi^+_Q \rangle_{xz}$ along the wall-normal direction for fDNS, ILES, Vreman, DSM, WALE and MANA models in *a posteriori* tests in R1M15.

and $\langle \Pi^+_Q \rangle_{xz}$ are zero along the wall-normal direction. The Vreman and WALE models significantly underestimate the mean SGS fluxes of the kinetic energy and the temperature variance, while the DSM strongly overestimates the mean kinetic energy flux and underestimates the mean flux of the temperature variance. On the contrary, the MANA model accurately predicts the kinetic energy and the temperature variance transfer rate along the wall-normal direction, except for slightly larger values in the buffer layer. Furthermore, the instantaneous wall-parallel contour of the normalised SGS fluxes of the kinetic energy and the temperature variance Π^+_τ and Π^+_Q for fDNS and LES with different SGS models at $y^+ = 15$ in *a posteriori* tests in R1M15 are shown in figures 18 and 19, respectively. It is found in figures 18 and 19 that most regions exhibit positive values of Π^+_τ and Π^+_Q, indicating that the compressible turbulent channel flow is dominated by the direct transfer of the kinetic energy and the temperature variance from large scales to small scales. However, the negative values of Π^+_τ and Π^+_Q also appear in some regions, suggesting that the inverse transfer of the kinetic energy and the temperature variance from small scales to large scales also exist in the compressible turbulent channel flow.

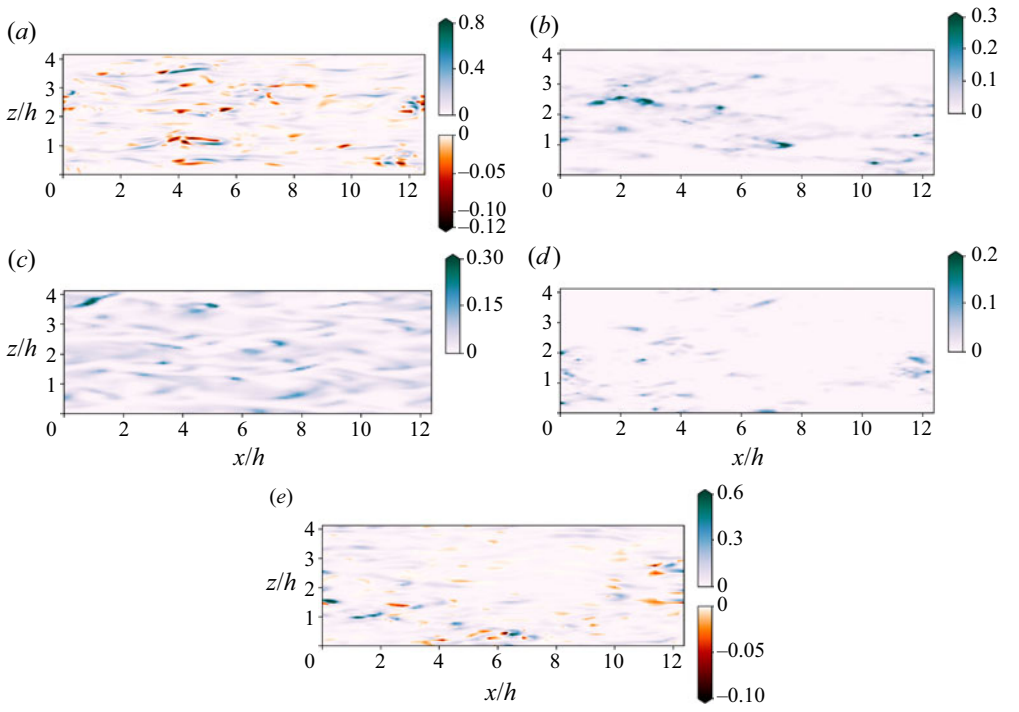


Figure 18. The instantaneous wall-parallel contour of the normalised flux of the kinetic energy Π_τ^+ for (*a*) fDNS, (*b*) Vreman, (*c*) DSM, (*d*) WALE and (*e*) MANA models at $y^+ = 15$ in *a posteriori* tests in R1M15.

The above observations are consistent with those in compressible isotropic turbulence and hypersonic turbulent boundary layers (Eyink 2005; Aluie & Eyink 2009; Eyink & Aluie 2009; Wang *et al.* 2018*a*; Xu *et al.* 2021*b*). It is also found that only the positive values of Π_τ^+ and Π_Q^+ appear in the contours of the Vreman, DSM and WALE models, which indicates that the eddy-viscosity models cannot predict the inverse transfer of the kinetic energy and the temperature variance (Xie *et al.* 2019*a*,*c*). On the contrary, both positive and negative values of Π_τ^+ and Π_Q^+ appear in the contour of the MANA model, suggesting that the MANA model can well predict both the direct and inverse transfer of the kinetic energy and the temperature variance.

The probability density functions (p.d.f.s) of the normalised SGS fluxes of the kinetic energy and the temperature variance at $y^+ = 15$ for fDNS, Vreman, DSM, WALE and MANA models in *a posteriori* tests in R1M15 are plotted in figure 20. It is found that the values of Π_τ^+ and Π_Q^+ for Vreman, DSM and WALE models are all larger than zero, and the right tails of the p.d.f.s are significantly shorter than those of the fDNS fields, indicating that the Vreman, DSM and WALE models remarkably underestimate the normalised SGS fluxes of the kinetic energy and the temperature variance Π_τ^+ and Π_Q^+. However, the p.d.f.s of Π_τ^+ and Π_Q^+ for the MANA model collapse well to those of the fDNS result, indicating that the MANA model can accurately estimate the normalised SGS fluxes of the kinetic energy and the temperature variance.

In a nutshell, according to the observations in this subsection, it is shown that the newly proposed MANA model can accurately predict the second-order statistics of both velocity and thermodynamic variables and the SGS fluxes of the kinetic energy and the

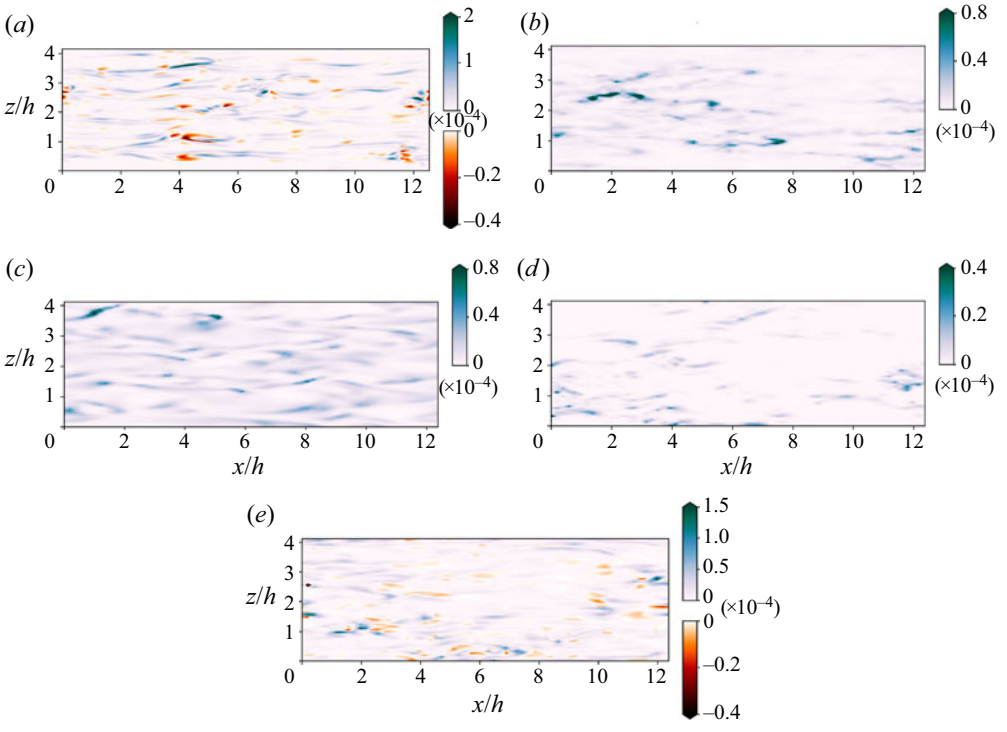


Figure 19. The instantaneous wall-parallel contour of the normalised flux of the temperature variance Π_Q^+ for (*a*) fDNS, (*b*) Vreman, (*c*) DSM, (*d*) WALE and (*e*) MANA models at $y^+ = 15$ in *a posteriori* tests in R1M15. It is noted that the colour bar value should be multiplied by $\times 10^{-4}$.

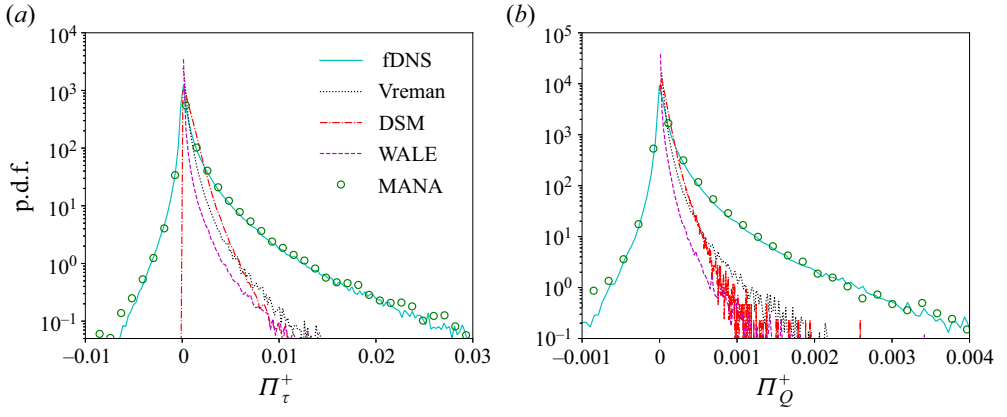


Figure 20. The probability density functions (p.d.f.s) of the normalised SGS fluxes of the kinetic energy and the temperature variance (*a*) Π_τ^+ and (*b*) Π_Q^+ at $y^+ = 15$ for the fDNS, Vreman, DSM, WALE and MANA models in *a posteriori* tests in R1M15.

temperature variance, and has remarkably better performances than the ILES and the traditional eddy-viscosity models. Furthermore, the MANA model is also much more computationally efficient than the DSM. More importantly, the proposed MANA model can accurately predict the second-order flow statistics for grid resolutions totally different

	$N_x \times N_y \times N_z$	Δx^+	Δy_w^+	Δz^+
DNS	$400 \times 210 \times 320$	14.5	0.66	6.05
ILES	$100 \times 101 \times 80$	55.5	1.14	23.1
Vreman	$100 \times 101 \times 80$	55.8	1.15	23.2
DSM	$100 \times 101 \times 80$	55.5	1.14	23.1
WALE	$100 \times 101 \times 80$	55.2	1.14	23.0
MANA	$100 \times 101 \times 80$	58.1	1.20	24.2

Table 5. The details of the grid settings of the DNS, ILES, Vreman model, DSM, WALE model and MANA model in the compressible channel flow case R2M30 with $Re = 4880$ and $M = 3.0$.

from the grid resolution used in the training process, and this merit is a great advantage over many other ANN-based SGS models proposed previously. The good performance of the MANA model in R1M15 with the untrained grid resolution is mainly owing to the elaborately designed structures of the input and tensor bases of the MANA model. The local grid widths in three directions are used to normalise the corresponding gradients of the flow variables, and the information of the grid resolutions is inherently embedded in the input and tensor bases of the MANA model, which is expected to have a good performance for untrained grid resolutions.

4.2. *Application in the compressible turbulent channel flow with untrained Reynolds numbers and Mach numbers*

The MANA model is also tested in compressible turbulent channel flows with untrained Reynolds numbers and Mach numbers. Two sets of compressible channel flow with different Reynolds numbers and the Mach numbers are tested as representatives: one is with the Reynolds number $Re = 4880$ and the Mach number $M = 3.0$ named 'R2M30', and another is with the Reynolds number $Re = 7000$ and the Mach number $M = 1.0$ named 'R3M10'.

The details of the grid settings of the DNS and LES with different SGS models in case R2M30 are listed in table 5. It is noted that the LES grid resolutions have the streamwise and spanwise filtered sizes $n_x = 4$ and $n_z = 4$ compared with the DNS grid resolution. Accordingly, the fDNS result is obtained by utilising the top-hat filter on the DNS data with the streamwise and spanwise filtered sizes $n_x = 4$ and $n_z = 4$, which are also different from the filtered sizes in the training process.

The normalised r.m.s. values of the resolved fluctuating velocities $\tilde{u}_{i,rms}^+$ and the normalised resolved Reynolds shear stress $\tilde{R}_{uv}^+$ along the wall-normal direction for the fDNS and different LES models in R2M30 are shown in figure 21. The ILES and the traditional eddy-viscosity models exhibit remarkably larger peak values of $\tilde{u}_{rms}^+$, $\tilde{v}_{rms}^+$ and $\tilde{R}_{uv}^+$ than those of the fDNS result, while the MANA model gives accurate predictions of $\tilde{u}_{rms}^+$, $\tilde{v}_{rms}^+$ and $\tilde{R}_{uv}^+$, except for slightly larger peak values of $\tilde{v}_{rms}^+$ and $\tilde{R}_{uv}^+$ in the buffer layer. For the $\tilde{w}_{rms}^+$ profiles, the ILES and the traditional eddy-viscosity models significantly underestimate the peak values, while the MANA model gives a much more accurate prediction, except for the slight underestimation of the peak values. To sum up, the MANA model has a remarkably better performance in predicting the resolved fluctuating velocities and the resolved Reynolds shear stress than the ILES and the traditional eddy-viscosity models.

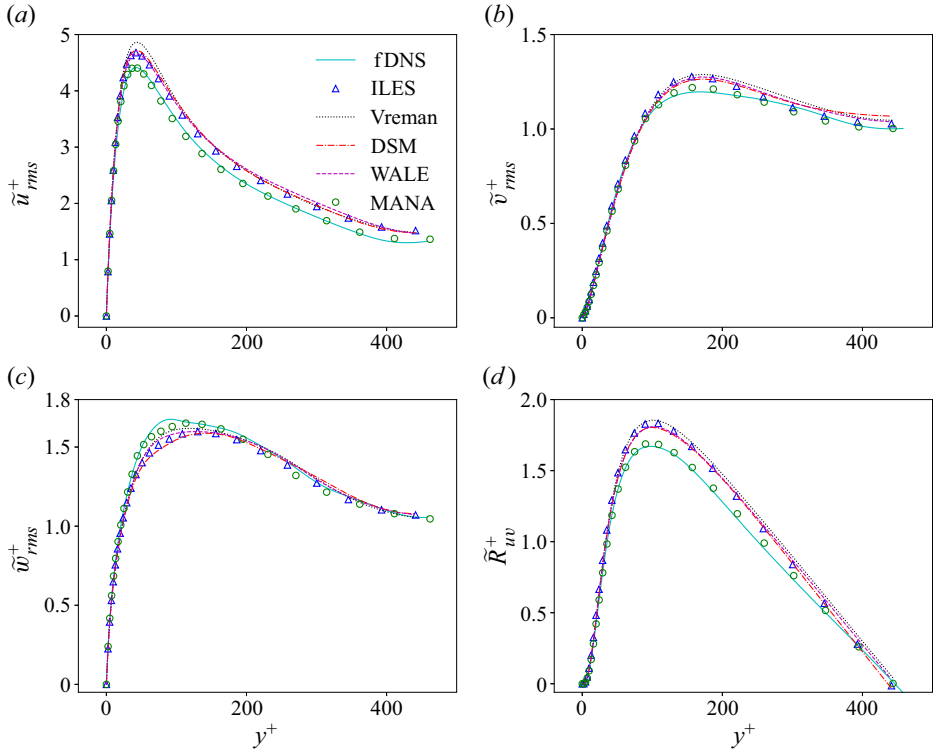


Figure 21. The normalised r.m.s. values of the resolved fluctuating velocities and the normalised resolved Reynolds shear stress along the wall-normal direction in the fDNS and different LES models in R2M30. (*a*) The streamwise resolved fluctuating velocities $\tilde{u}^+_{rms}$, (*b*) the wall-normal resolved fluctuating velocities $\tilde{v}^+_{rms}$, (*c*) the spanwise resolved fluctuating velocities $\tilde{w}^+_{rms}$ and (*d*) the resolved Reynolds shear stress $\tilde{R}^+_{uv}$.

The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along the wall-normal direction for the fDNS and different LES models in R2M30 are depicted in figure 22. It is found that the ILES and the traditional eddy-viscosity models have a good performance in predicting $\tilde{T}^+_{rms}$ and $\tilde{R}^+_{uT}$. However, the MANA model has a slightly better performance than the ILES and the traditional eddy-viscosity models, and the $\tilde{T}^+_{rms}$ and $\tilde{R}^+_{uT}$ profiles of the MANA model almost collapse to those of the fDNS result.

The streamwise–spanwise averages of the normalised SGS fluxes of the kinetic energy and the temperature variance $\langle \Pi^+_\tau \rangle_{xz}$ and $\langle \Pi^+_Q \rangle_{xz}$ along the wall-normal direction for the fDNS, ILES, Vreman, DSM, WALE and MANA models in R2M30 are depicted in figure 23. It is found that the Vreman and WALE models strongly underestimate the mean SGS fluxes of the kinetic energy and the temperature variance, while a remarkable overestimation appears in DSM. The MANA model has a significantly better performance in estimating the mean SGS fluxes of the kinetic energy and the temperature variance than the traditional eddy-viscosity models.

The details of the grid settings of the DNS and LES in case R3M10 are listed in table 6. The LES grid resolutions have the streamwise and spanwise filtered sizes $n_x = 12$ and $n_z = 4$ compared with the DNS grid resolution; therefore, the fDNS result is obtained by taking the top-hat filter on the DNS data with the streamwise and spanwise filtered sizes

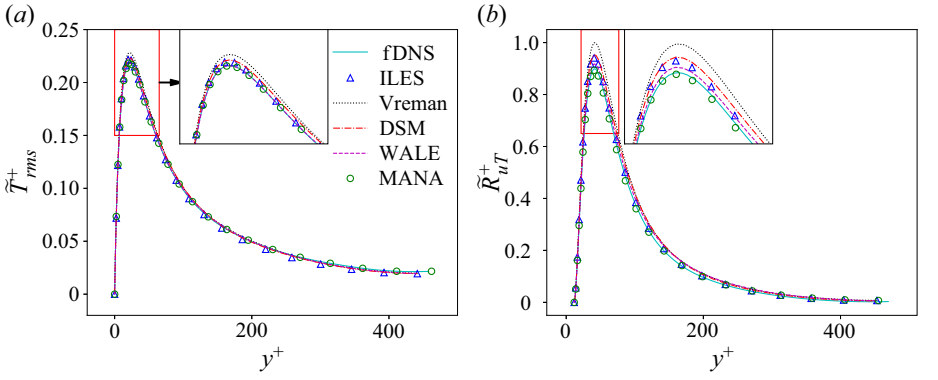


Figure 22. (*a*) The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and (*b*) the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along the wall-normal direction for the fDNS and different LES models in R2M30.

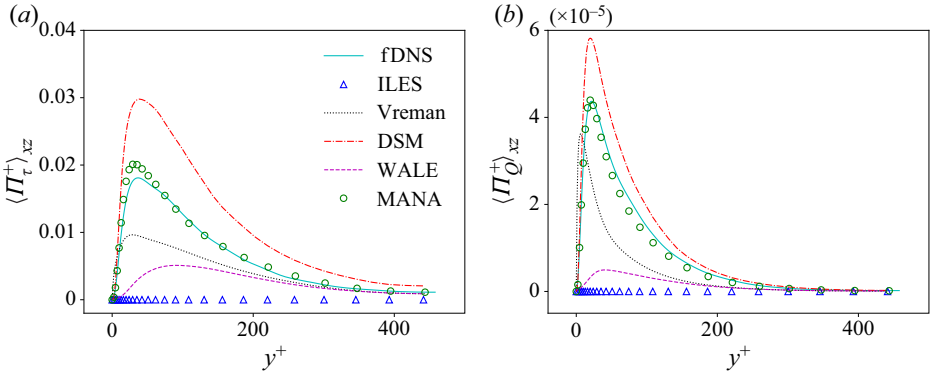


Figure 23. The streamwise–spanwise averages of the normalised SGS fluxes of the kinetic energy and the temperature variance (*a*) $\langle \Pi^+_\tau \rangle_{xz}$ and (*b*) $\langle \Pi^+_Q \rangle_{xz}$ along the wall-normal direction for the fDNS, ILES, Vreman, DSM, WALE and MANA models in R2M30.

	$N_x \times N_y \times N_z$	Δx^+	Δy^+_w	Δz^+
DNS	$768 \times 385 \times 256$	7.02	0.21	7.02
ILES	$64 \times 65 \times 64$	78.4	1.62	26.1
Vreman	$64 \times 65 \times 64$	77.7	1.61	25.9
DSM	$64 \times 65 \times 64$	78.4	1.62	26.1
WALE	$64 \times 65 \times 64$	77.5	1.60	25.8
MANA	$64 \times 65 \times 64$	81.2	1.68	27.0

Table 6. The details of the grid settings of the DNS, ILES, Vreman model, DSM, WALE models and the MANA model in the compressible channel flow case R3M10 with $Re = 7000$ and $M = 1.0$.

$n_x = 12$ and $n_z = 4$. Moreover, the streamwise and spanwise filtered sizes in R3M10 are also different from those in the training process.

The normalised r.m.s. values of the resolved fluctuating velocities $\tilde{u}^+_{i,rms}$ and the normalised resolved Reynolds shear stress $\tilde{R}^+_{uv}$ along the wall-normal direction for the

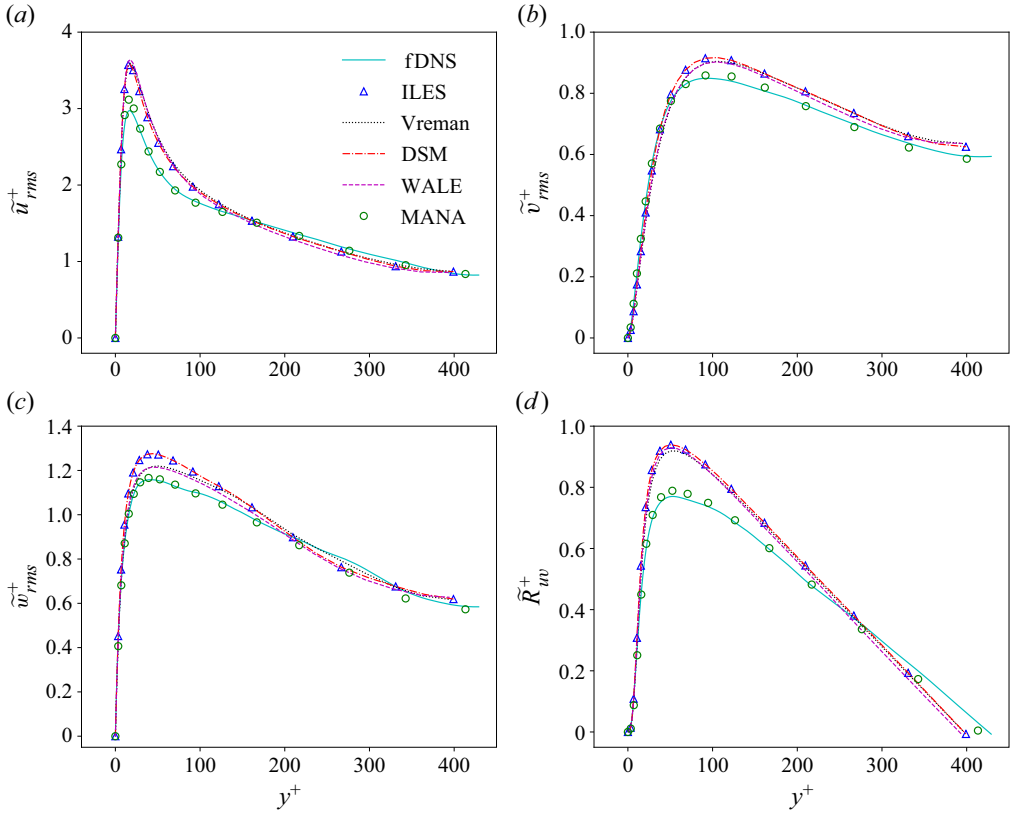


Figure 24. The normalised r.m.s. values of the resolved fluctuating velocities and the normalised resolved Reynolds shear stress along the wall-normal direction for the fDNS and different LES models in R3M10. (*a*) The streamwise resolved fluctuating velocities $\tilde{u}^+_{rms}$, (*b*) the wall-normal resolved fluctuating velocities $\tilde{v}^+_{rms}$, (*c*) the spanwise resolved fluctuating velocities $\tilde{w}^+_{rms}$ and (*d*) the resolved Reynolds shear stress $\tilde{R}^+_{uv}$.

fDNS and different LES models in R3M10 are shown in figure 24. It is found that the $\tilde{u}^+_{i,rms}$ and $\tilde{R}^+_{uv}$ profiles of the ILES and the traditional eddy-viscosity models exhibit significantly larger peak values than those of the fDNS result, while the MANA model provides accurate estimations, except for slightly larger peak values in the buffer layer. According to the above observations, it is concluded that the proposed MANA model offers significant advantages over the ILES and the traditional eddy-viscosity models in predicting the second-order statistics of velocity variables, especially under extremely sparse grid resolution.

Apart from the performance in predicting the second-order statistics of the velocity variables, the ability in estimating the second-order statistics of thermodynamic variables is also tested for different LES models. The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along the wall-normal direction for the fDNS and different LES models in R3M10 are shown in figure 25. It is found that the $\tilde{T}^+_{rms}$ and $\tilde{R}^+_{uT}$ profiles for the ILES and the traditional eddy-viscosity models have much larger peak values than those of the fDNS result. However, the MANA model has a better performance in predicting the $\tilde{T}^+_{rms}$ and

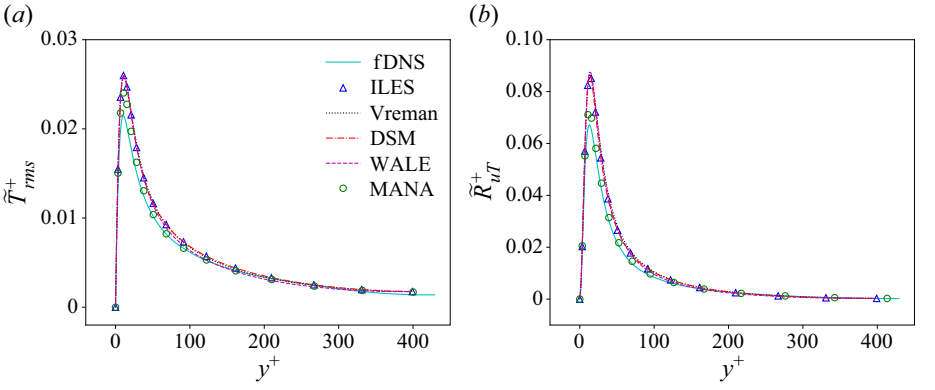


Figure 25. (*a*) The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and (*b*) the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along the wall-normal direction for the fDNS and different LES models in R3M10.

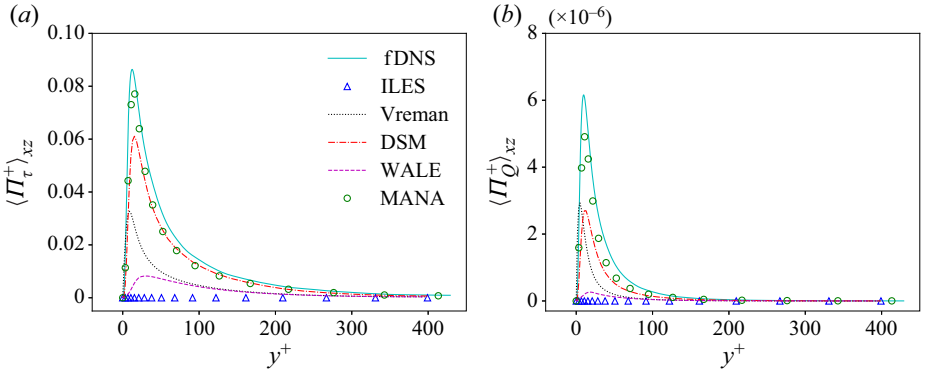


Figure 26. The streamwise–spanwise averages of the normalised SGS fluxes of the kinetic energy and the temperature variance (*a*) $\langle \Pi^+_\tau \rangle_{xz}$ and (*b*) $\langle \Pi^+_Q \rangle_{xz}$ along the wall-normal direction for the fDNS, ILES, Vreman, DSM, WALE and MANA models in R3M10.

$\tilde{R}^+_{uT}$ profiles, where the deviations of the peak values from the fDNS result are strongly reduced in the MANA model.

The streamwise–spanwise averages of the normalised SGS fluxes of the kinetic energy and the temperature variance $\langle \Pi^+_\tau \rangle_{xz}$ and $\langle \Pi^+_Q \rangle_{xz}$ along the wall-normal direction for the fDNS, ILES, Vreman, DSM, WALE and MANA models in R3M10 are depicted in figure 26. Remarkable underestimations of the mean SGS fluxes of the kinetic energy and the temperature variance appear in the LES with the traditional eddy-viscosity models, while the MANA model has a much better performance in predicting $\langle \Pi^+_\tau \rangle_{xz}$ and $\langle \Pi^+_Q \rangle_{xz}$.

In a word, the newly proposed MANA model exhibits a much better performance in predicting second-order flow statistics and the mean SGS fluxes of the kinetic energy and the temperature variance than the ILES and the traditional eddy-viscosity models in the compressible turbulent channel flow with two untrained Reynolds and Mach numbers and at different grid resolutions than that in the training process. Therefore, the above observations indicate that the MANA model offers significant advantages over the ILES

	$N_x \times N_y \times N_z$	Δx^+	Δy_w^+	Δz^+
DNS	$384 \times 193 \times 128$	6.23	0.26	6.23
ILES	$64 \times 65 \times 64$	32.0	0.66	10.7
Vreman	$64 \times 65 \times 64$	32.7	0.68	10.9
DSM	$64 \times 65 \times 64$	31.9	0.66	10.6
WALE	$64 \times 65 \times 64$	32.8	0.68	10.9
MANA	$64 \times 65 \times 64$	37.3	0.77	12.4

Table 7. The details of the grid settings of the DNS, ILES, Vreman model, DSM, WALE models and MANA model in the compressible channel flow case R1M02 with $Re = 3000$ and $M = 0.2$.

and the traditional eddy-viscosity models in compressible turbulent channel flows with untrained Reynolds numbers, Mach numbers and grid resolutions.

4.3. *Application to compressible turbulent channel flows with extremely low Mach number*

When the Mach number is extremely low, the statistics of the velocity fields of the compressible turbulent channel flows are similar to those of incompressible turbulent channel flows. Therefore, the low-Mach-number compressible turbulent channel flow with the Reynolds number $Re = 3000$ and the Mach number $M = 0.2$, named 'R1M02' is also tested to measure the performance of the MANA model in nearly incompressible turbulent channel flows.

The details of the grid settings of the DNS and LES with different SGS models in case R1M02 are listed in table 7. It is found that the LES grid resolutions have the streamwise and spanwise filtered sizes $n_x = 6$ and $n_z = 2$ compared with the DNS grid resolution. Therefore, the fDNS result is obtained by utilising the top-hat filter on the DNS data with the streamwise and spanwise filtered sizes $n_x = 6$ and $n_z = 2$, and these filter sizes are different from those in the training process.

The normalised r.m.s. values of the resolved fluctuating velocities and the normalised resolved Reynolds shear stress along the wall-normal direction in the fDNS and different LES models in R1M02 are shown in figure 27. It is found that the ILES and the traditional eddy-viscosity models exhibit significantly larger peak values of $\tilde{u}_{rms}^+$ than that of the fDNS result, while they underestimate the peak values of $\tilde{v}_{rms}^+$ and $\tilde{w}_{rms}^+$ in the buffer layer. However, the $\tilde{u}_{rms}^+$, $\tilde{v}_{rms}^+$, $\tilde{w}_{rms}^+$ and $\tilde{R}_{uv}^+$ values predicted by the MANA model collapse well onto those of the fDNS result. Therefore, it is shown that the MANA model can give much more accurate predictions of the resolved fluctuating velocities and the resolved Reynolds shear stress than the ILES and the traditional eddy-viscosity models.

The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}_{rms}^+$ and the normalised resolved streamwise Reynolds heat flux $\tilde{R}_{uT}^+$ along the wall-normal direction for the fDNS and different LES models in R1M02 are depicted in figure 28. It is found that $\tilde{T}_{rms}^+$ of the fDNS result is very small, indicating that the temperature fluctuation is very weak. The ILES and the traditional eddy-viscosity models overestimate $\tilde{T}_{rms}^+$ and $\tilde{R}_{uT}^+$, while the MANA model has a much better performance in predicting $\tilde{T}_{rms}^+$ and $\tilde{R}_{uT}^+$.

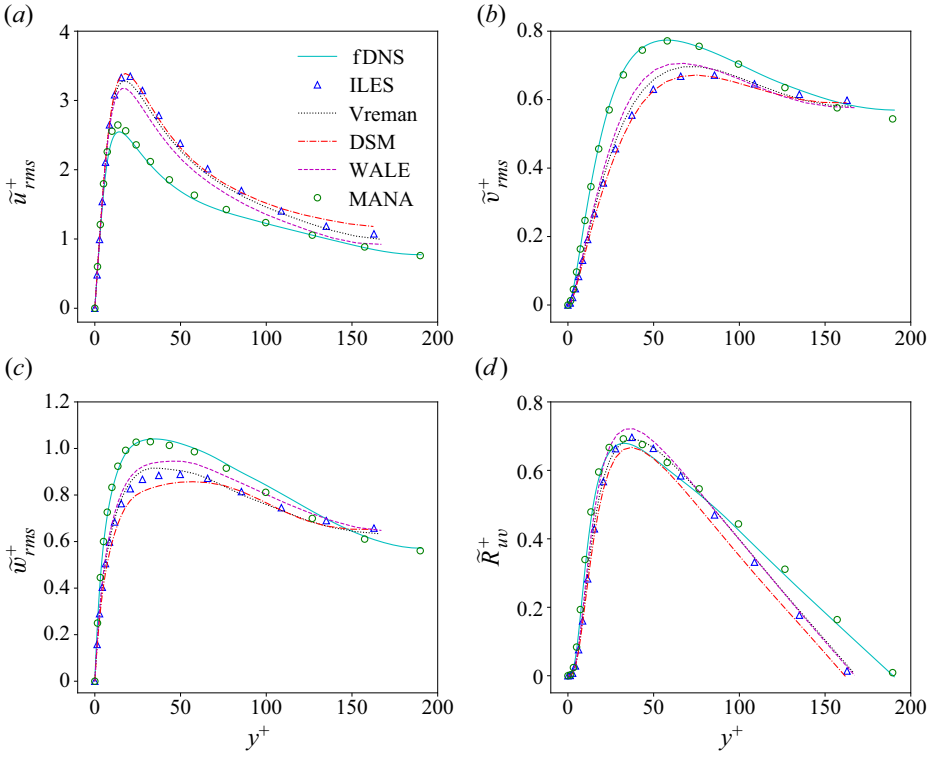


Figure 27. The normalised r.m.s. values of the resolved fluctuating velocities and the normalised resolved Reynolds shear stress along the wall-normal direction in the fDNS and different LES models in R1M02. (*a*) The streamwise resolved fluctuating velocities $\tilde{u}^+_{rms}$, (*b*) the wall-normal resolved fluctuating velocities $\tilde{v}^+_{rms}$, (*c*) the spanwise resolved fluctuating velocities $\tilde{w}^+_{rms}$ and (*d*) the resolved Reynolds shear stress $\tilde{R}^+_{uv}$.

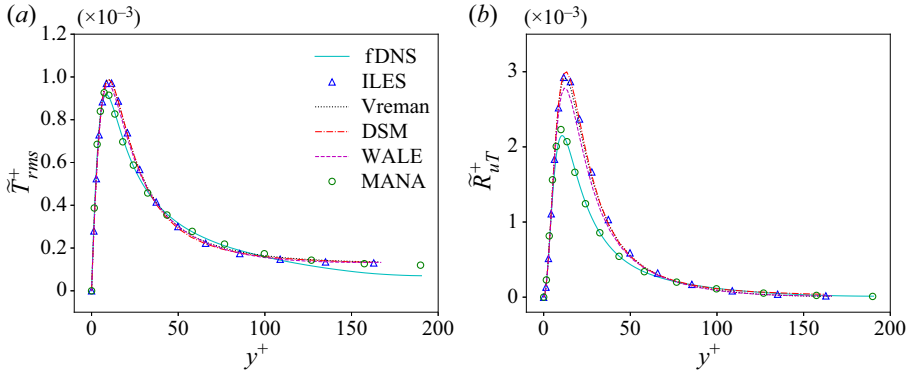


Figure 28. (*a*) The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and (*b*) the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along the wall-normal direction for the fDNS and different LES models in R1M02.

According to the above observations, it is found that the MANA model can also have a significantly better performance in predicting second-order flow statistics than the ILES and the traditional eddy-viscosity models in very low-Mach-number compressible turbulent channel flows.

	$N_x \times N_y \times N_z$	Δx^+	Δy_w^+	Δz^+
DNS	$10\,000 \times 90 \times 320$	5.77	0.58	5.26
ILES	$1000 \times 90 \times 80$	54.4	0.54	19.8
Vreman	$1000 \times 90 \times 80$	54.6	0.55	19.9
DSM	$1000 \times 90 \times 80$	51.8	0.52	18.9
WALE	$1000 \times 90 \times 80$	53.5	0.54	19.5
MANA	$1000 \times 90 \times 80$	56.5	0.57	20.6

Table 8. The details of the grid settings of the DNS, ILES, Vreman model, DSM, WALE models and MANA model in the supersonic adiabatic turbulent boundary layer.

4.4. *Application to the compressible zero-pressure-gradient flat-plate boundary layer*

In order to verify the performance of the MANA model in more complicated compressible wall-bounded flows, the application to the compressible zero-pressure-gradient flat-plate boundary layer is also demonstrated. The spatially evolving zero-pressure-gradient flat-plate boundary layer consists of laminar, transitional and fully turbulent regions, which flow is much more complex than the turbulent channel flows.

The spatially developing supersonic adiabatic flat-plate boundary layer with $Re = 635\,000$ and $M = 2.25$ (Pirozzoli *et al.* 2004) is simulated to check the performance of the MANA model in the compressible turbulent boundary layer. A schematic of the supersonic adiabatic transitional and turbulent boundary layer is shown in figure 29. The spatially evolving supersonic adiabatic transitional and turbulent boundary layer is numerically simulated with the following boundary conditions: the inflow and outflow boundary conditions, a wall boundary condition, an upper far-field boundary condition and a periodic boundary condition in the spanwise direction. To be specific, a time-independent laminar compressible boundary-layer similarity solution is applied at the inflow boundary. A region of wall blowing and suction is implemented to induce the laminar-to-turbulent transition. The form of blowing and suction is the same as that in Pirozzoli *et al.* (2004) except for the magnitude of the amplitude. In order to simulate a natural transition, an amplitude of 0.02 is applied in this case (Yu *et al.* 2022). For the outflow boundary condition, all the flow fields are extrapolated from the interior points to the outflow boundary points except the pressure in the subsonic region of the boundary layer. Moreover, the pressure in the subsonic region is set equal to the value of the first grid point where the flow is supersonic (Pirozzoli *et al.* 2004). Furthermore, the no-slip condition is applied for the wall boundary, and the non-reflecting boundary condition is imposed for the upper boundary (Pirozzoli *et al.* 2004). The computational domain size is $L_x \times L_y \times L_z = 6 \times 0.3 \times 0.175$ in the streamwise, wall-normal and spanwise directions, respectively, normalised by 1 inch (Pirozzoli *et al.* 2004; Yu *et al.* 2022). It is found in Pirozzoli *et al.* (2004) that at the streamwise location $x/L_x = 0.8$ (i.e. $x = 8.8$ in Pirozzoli *et al.* 2004), the flow reaches to the fully turbulent state. Therefore, the performance of the newly proposed MANA model in predicting the flow statistics of the supersonic turbulent boundary layer are tested at the streamwise location $x/L_x = 0.8$.

The details of the grid settings of the DNS and LES in the supersonic adiabatic turbulent boundary layer are listed in table 8. The LES grid resolutions have the streamwise and spanwise filtered sizes $n_x = 10$ and $n_z = 4$ compared with the DNS grid resolution. Therefore, the fDNS result is calculated by taking the top-hat filter on the DNS data with the streamwise and spanwise filtered sizes $n_x = 10$ and $n_z = 4$.

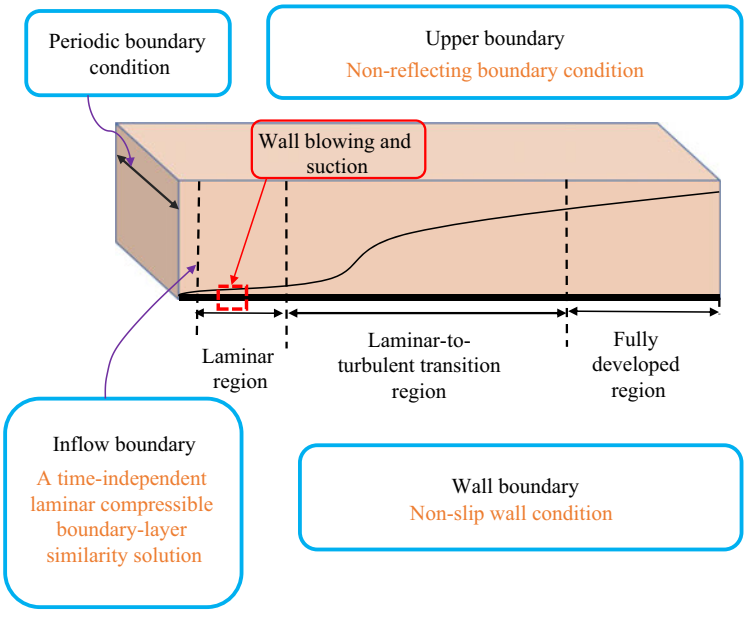


Figure 29. A schematic of the supersonic adiabatic transitional and turbulent boundary layer.

The normalised r.m.s. values of the resolved fluctuating velocities and the normalised resolved Reynolds shear stress along the wall-normal direction at streamwise location $x/L_x = 0.8$ in the fDNS and different LES models in the supersonic adiabatic turbulent boundary layer are shown in figure 30. It is found that the profiles of $\tilde{u}^+_{rms}$, $\tilde{v}^+_{rms}$, $\tilde{w}^+_{rms}$ and $\tilde{R}^+_{uv}$ predicted by the ILES and the traditional eddy-viscosity models show much larger peak values than those of the fDNS result. However, the MANA model accurately predicts the profiles of $\tilde{w}^+_{rms}$ and $\tilde{R}^+_{uv}$, and slightly overestimates the values of $\tilde{u}^+_{rms}$ and $\tilde{v}^+_{rms}$. According to the above observations, it is shown that the MANA model has a better performance in predicting the velocity statistics than the ILES and the traditional eddy-viscosity models in the supersonic adiabatic turbulent boundary layer.

The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along the wall-normal direction at streamwise location $x/L_x = 0.8$ for the fDNS and different LES models in the supersonic adiabatic turbulent boundary layer are shown in figure 31. It is found that the magnitudes of $\tilde{T}^+_{rms}$ and $\tilde{R}^+_{uT}$ predicted by the ILES and the traditional eddy-viscosity models are much larger than those of the fDNS result, while the MANA model only slightly overestimates the magnitudes of $\tilde{T}^+_{rms}$ and $\tilde{R}^+_{uT}$. Therefore, the MANA model gives much more accurate prediction of the thermodynamic statistics than the ILES and the traditional eddy-viscosity models in the supersonic adiabatic turbulent boundary layer.

To sum up, it is found that the proposed MANA model can also exhibit a much better performance in predicting the flow statistics in the supersonic adiabatic turbulent boundary layer than the ILES and the traditional eddy-viscosity models.

5. Summary and conclusions

In this paper, ANN-based nonlinear algebraic models are proposed for both the SGS stress and SGS heat flux in the LES of compressible wall-bounded turbulence.

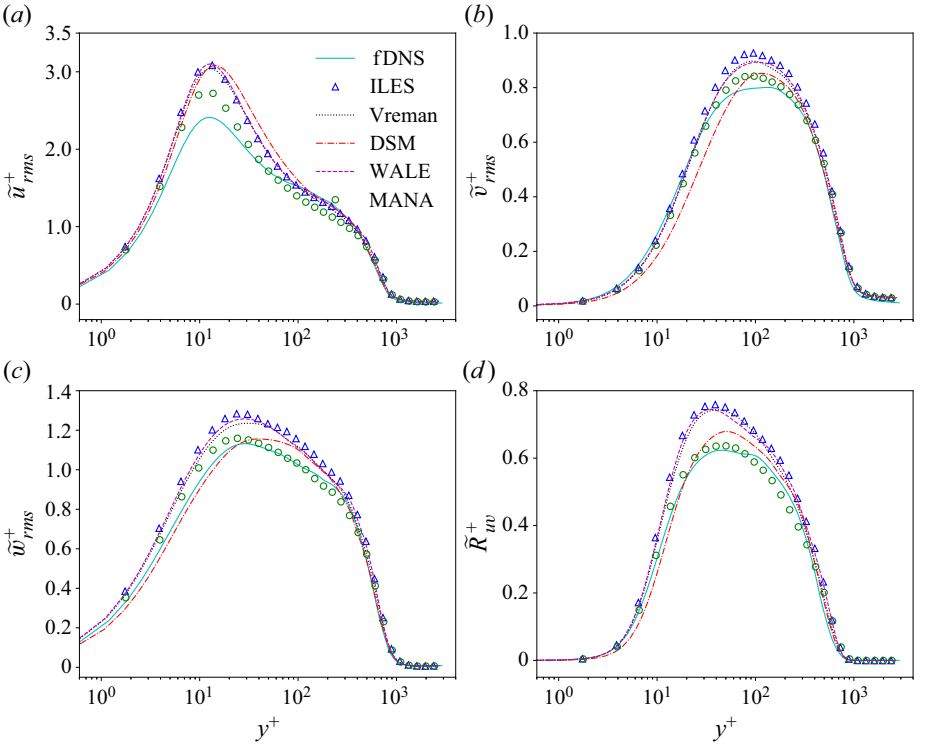


Figure 30. The normalised r.m.s. values of the resolved fluctuating velocities and the normalised resolved Reynolds shear stress along the wall-normal direction at streamwise location $x/L_x = 0.8$ in the fDNS and different LES models in the supersonic adiabatic turbulent boundary layer. (*a*) The streamwise resolved fluctuating velocities $\tilde{u}^+_{rms}$, (*b*) the wall-normal resolved fluctuating velocities $\tilde{v}^+_{rms}$, (*c*) the spanwise resolved fluctuating velocities $\tilde{w}^+_{rms}$ and (*d*) the resolved Reynolds shear stress $\tilde{R}^+_{uv}$.

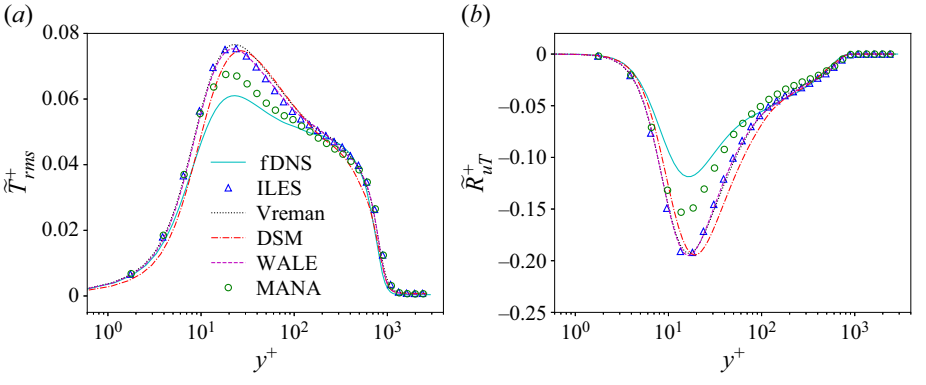


Figure 31. (*a*) The normalised r.m.s. values of the resolved fluctuating temperature $\tilde{T}^+_{rms}$ and (*b*) the normalised resolved streamwise Reynolds heat flux $\tilde{R}^+_{uT}$ along wall-normal direction at streamwise location $x/L_x = 0.8$ for the fDNS and different LES models in the supersonic adiabatic turbulent boundary layer.

The input quantities and the tensor bases of the proposed ANN-based nonlinear algebraic models are significantly modified from the original nonlinear algebraic model proposed by Lund & Novikov (1992) and Wang *et al.* (2007), which gives rise to the drastically

improved performance compared with the ILES and the traditional eddy-viscosity models (DSM, Vreman and WALE) in the LES of the compressible turbulent channel flow. Furthermore, the modified ANN-based nonlinear algebraic models can be applied to accurately predict the flow statistics in turbulent channel flows with untrained Reynolds numbers, Mach numbers and grid resolutions as well as the supersonic turbulent boundary layer. More importantly, the modified ANN-based nonlinear algebraic models are more computationally efficient than the DSM.

With the increasing popularity of the ANN method, it is straightforward to use the ANN method to generate proper dimensionless model coefficients of the original nonlinear algebraic model proposed by Lund & Novikov (1992) and Wang *et al.* (2007). However, it is found that the original ANN-based nonlinear algebraic model (OANA model) has low correlation coefficients and large relative errors compared with the fDNS result in the *a priori* tests, although it is slightly better than the traditional eddy-viscosity models (DSM, Vreman and WALE). Therefore, we innovatively propose a modified ANN-based nonlinear algebraic model (MANA model). The local grid widths along the streamwise, wall-normal and spanwise directions are used to normalise the corresponding gradients of the flow variables, and this MANA model exhibits a much better performance than the OANA model and the traditional eddy-viscosity models (DSM, Vreman and WALE) in the *a priori* tests, where the correlation coefficients are all larger than 0.91, and the relative errors are smaller than 0.4. Furthermore, the MANA model can accurately predict the mean SGS fluxes of the kinetic energy and the temperature variance, which is a significant advantage over the traditional eddy-viscosity models. It is worth noting that the information of the local grid resolution is inherently embedded in the input quantities and the tensor bases of the MANA model, therefore it is expected that the proposed MANA model can show good performance in compressible turbulent channel flows with untrained grid resolutions.

Furthermore, the performance of the proposed MANA model in predicting the flow statistics and the mean SGS fluxes of the kinetic energy and the temperature variance is also examined in *a posteriori* tests. It is confirmed that the MANA model can give more accurate prediction of the second-order flow statistics and the mean SGS fluxes of the kinetic energy and the temperature variance than the ILES and the traditional eddy-viscosity models (DSM, Vreman and WALE) in compressible turbulent channel flows with untrained Reynolds numbers, Mach numbers and grid resolutions. Moreover, the MANA model can give much more accurate predictions of the flow statistics in the supersonic turbulent boundary layer. The inverse transfer of the kinetic energy and the temperature variance cannot be estimated in the traditional eddy-viscosity models, while the MANA model can well predict the direct and inverse transfer of the kinetic energy and the temperature variance. Moreover, the MANA model is computationally cheaper than the DSM.

In summary, the newly proposed modified ANN-based nonlinear algebraic model in this paper has been verified to be an effective SGS model in LES of compressible wall-bounded turbulence. It is noted that many ANN-based models proposed previously (including those of Park & Choi (2021), Xie *et al.* (2019*c*) and so on) directly constructed the ANN relations between the filtered quantities and the unclosed SGS terms. Such 'black box' models cannot have a good performance in untrained cases. On the contrary, the newly proposed MANA model combines plenty of prior physical knowledge, and therefore can give accurate predictions in a variety of untrained cases. However, it should also be emphasised that the MANA model still needs to be further tested and improved in LES of

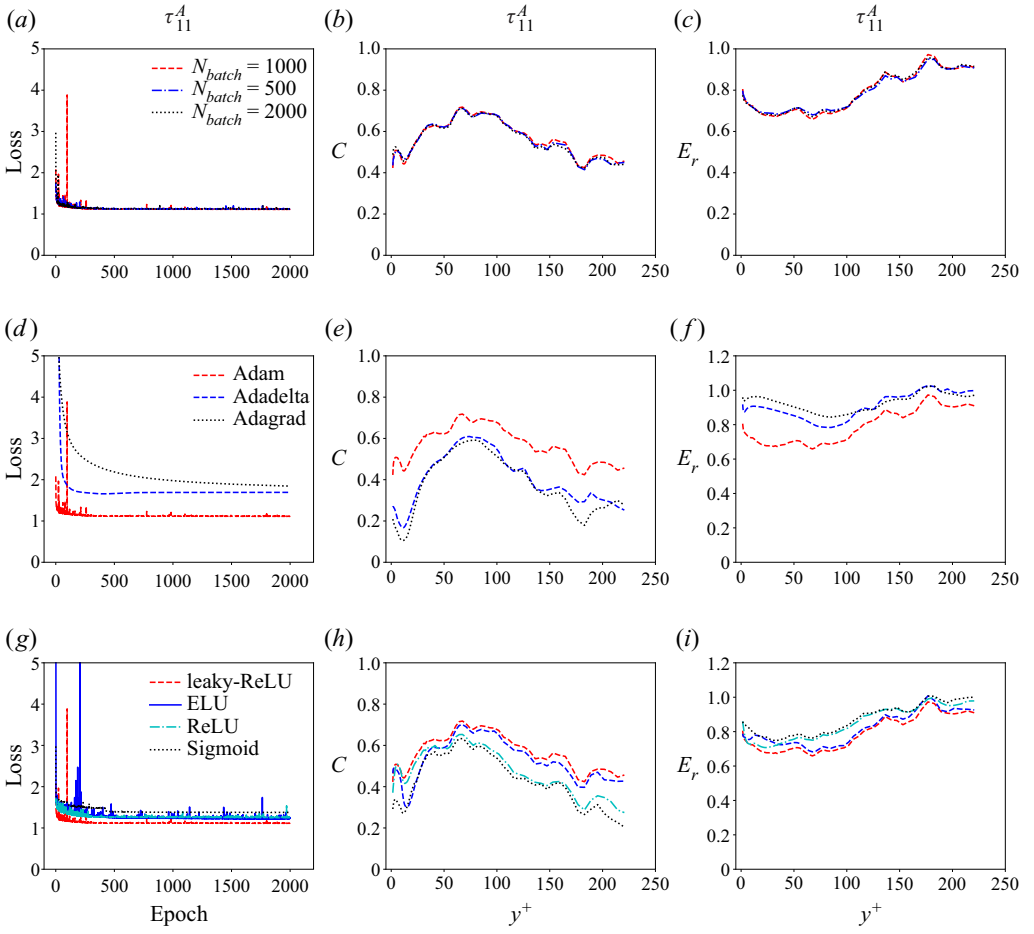


Figure 32. The learning curves of the validation loss, the correlation coefficients C and the relative errors E_r in the testing set with different mini-batch sizes N_{batch} (*a–c*), different optimisers (*d–f*) and different activation functions (*g–i*) in the OANA model. The number of neurons of the hidden layers is fixed to $N_h = 20$.

wall-bounded turbulent flows with higher Reynolds numbers, higher Mach numbers and more complex geometries.

Supplementary materials. Supplementary materials are available at https://doi.org/10.1017/jfm.2023.179.

Funding. This work was supported by the NSFC Basic Science Center Program (grant no. 11988102), by the National Natural Science Foundation of China (NSFC grant nos. 91952104, 92052301, 12172161 and 91752201), by the Technology and Innovation Commission of Shenzhen Municipality (grant nos. KQTD20180411143441009 and JCYJ20170412151759222) and by the Department of Science and Technology of Guangdong Province (grant no. 2019B21203001). This work was also supported by the Center for Computational Science and Engineering of Southern University of Science and Technology.

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

Author ORCIDs.
Dehao Xu https://orcid.org/0000-0003-2442-4150;
Jianchun Wang https://orcid.org/0000-0001-5101-7791;
Changping Yu https://orcid.org/0000-0002-2126-1344.

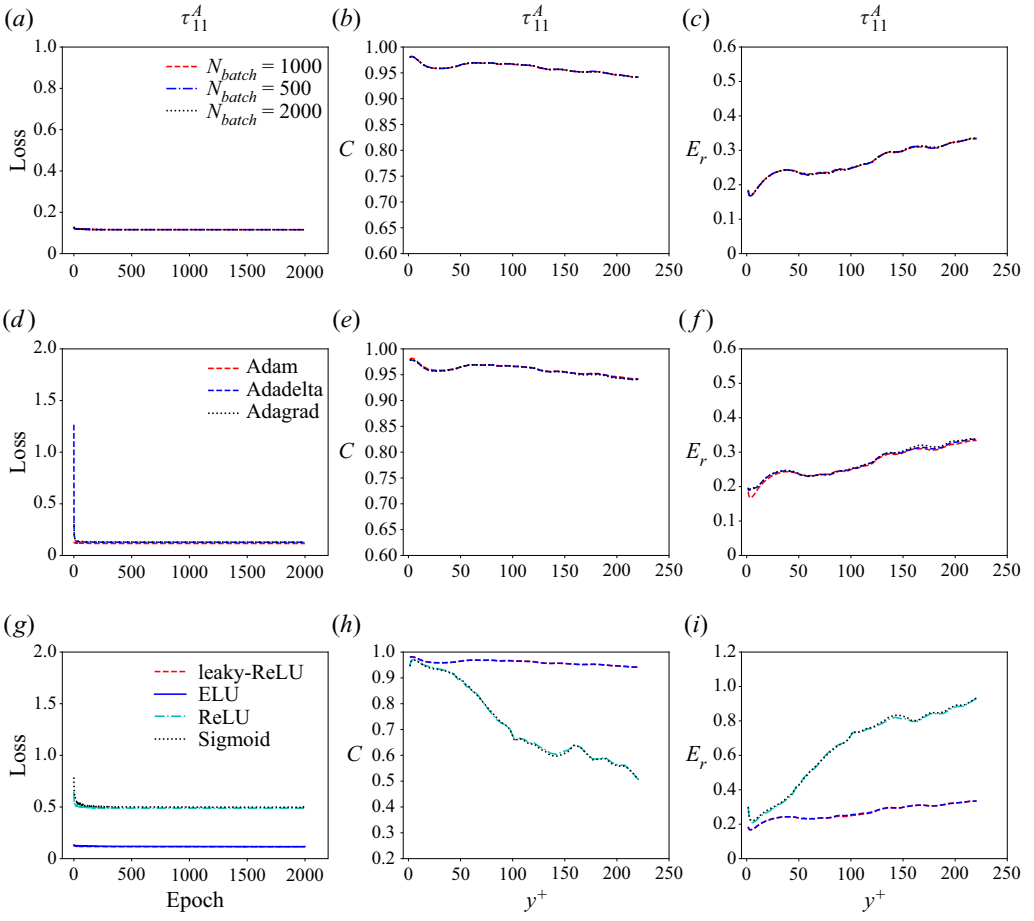


Figure 33. The learning curves of the validation loss, the correlation coefficients C and the relative errors E_r in the testing set with different mini-batch sizes N_{batch} (*a–c*), different optimisers (*d–f*) and different activation functions (*g–i*) in the MANA model. The number of neurons of the hidden layers is fixed to $N_h = 20$.

Appendix. The influence of the mini-batch size N_{batch}, different optimisers and activation functions in *a priori* test

In this appendix, the influences of the size of the mini-batch N_{batch}, different optimisers and activation functions on the performance of the OANA model and MAMA model in *a priori* tests are examined to determine the best choices of N_{batch}, optimiser and activation function. It is noted that only the performance of the ANN of τ_{11}^A is shown in this appendix; the results in other SGS unclosed components are similar to that of τ_{11}^A and are omitted for brevity.

The learning curves of the validation loss, the correlation coefficients C and the relative errors E_r in the testing set with different mini-batch sizes N_{batch}, different optimisers and different activation functions in the OANA model and MANA model are shown in figures 32 and 33 respectively. The number of neurons of the hidden layers is fixed to $N_h = 20$.

It is shown in figures 32 and 33(*a–c*) that, when the mini-batch size N_{batch} is changed from 1000 to 500 or 2000, the learning curves of the validation loss, the correlation

coefficients C and the relative errors E_r in the testing set are almost the same as the results of $N_{batch} = 1000$, which indicates that the mini-batch size N_{batch} has a negligible influence on the performance of the OANA model and MANA model in the range $500 \leq N_{batch} \leq 2000$. Therefore, the mini-batch size N_{batch} is fixed to $N_{batch} = 1000$ in all ANN training processes in this paper, which is consistent with the N_{batch} setting in many previous studies (Xie *et al.* 2019*b*, 2020*b*; Yu *et al.* 2022).

Then, the influence of different optimisers is also checked. The performances of the Adam algorithm (Kingma & Ba 2014), Adadelta algorithm (Zeiler 2012) and Adagrad algorithm (Duchi, Hazan & Singer 2011) in the OANA model and MANA model are compared in figures 32 and 33(*d–f*), respectively. It is found that the OANA model with the Adam algorithm (Kingma & Ba 2014) has the smallest validation loss, the largest correlation coefficient and the smallest relative error among the three optimisers. On the other hand, in the MANA model, the performance of the Adam algorithm (Kingma & Ba 2014) is slightly better than those of the Adadelta algorithm (Zeiler 2012) and Adagrad algorithm (Duchi *et al.* 2011). Therefore, the Adam algorithm (Kingma & Ba 2014) is used in all ANN training processes in this paper.

Finally, the influences of different activation functions in the OANA model and MANA model are tested in figures 32 and 33(*g–i*), respectively. It is found that the Sigmoid activation function has the worst performance, mainly due to the reason that the Sigmoid can result in the 'vanishing gradient problem': when the inputs have very high or very low values, there is almost no change to the prediction. This problem can lead to the network refusing to learn further, or being too slow to reach an accurate prediction (Goodfellow, Bengio & Courville 2016). Furthermore, the ReLU activation function also has poor performance, which can be ascribed to the 'dying ReLU problem': when the input is less than or equal to zero, the gradient of the function becomes zero and the network cannot perform backpropagation (Lu *et al.* 2020). It is noted that the exponential linear unit (ELU) (Clevert, Unterthiner & Hochreiter 2015) and leaky-ReLU activation function can remove the 'dying ReLU problem' of ReLU activation (Clevert *et al.* 2015; Xu *et al.* 2015). However, the ELU activation function has a larger computational cost and worse performance compared with those of the leaky-ReLU activation function. Therefore, the leaky-ReLU activation function is applied in all ANN structures in this paper.

References

Aluie, H. 2013 Scale decomposition in compressible turbulence. *Physica* D **247**, 54–65.

Aluie, H. & Eyink, G.L. 2009 Localness of energy cascade in hydrodynamic turbulence. II. Sharp spectral filter. *Phys. Fluids* **21**, 115108.

Anderson, R. & Meneveau, C. 1999 Effects of the similarity model in finite-difference LES of isotropic turbulence using a Lagrangian dynamic mixed model. *Flow Turbul. Combust.* **62**, 201–225.

Beck, A., Flad, D. & Munz, C.-D. 2019 Deep neural networks for data-driven LES closure models. *J. Comput. Phys.* **398**, 108910.

Bonett, D.G. & Wright, T.A. 2000 Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika* **65**, 23–28.

Chamecki, M., Meneveau, C. & Parlange, M.B. 2007 The local structure of atmospheric turbulence and its effect on the smagorinsky model for large eddy simulation. *J. Atmos. Sci.* **64**, 1941–1958.

Clark, R.A. 1979 Evaluation of sub-grid scalar models using an accurately simulated turbulent flow. *J. Fluid Mech.* **91**, 1–16.

Clevert, D., Unterthiner, T. & Hochreiter, S. 2015 Fast and accurate deep network learning by exponential linear units (ELUS). arXiv:1511.07289.

Coleman, G.N., Kim, J. & Moser, R.D. 1995 A numerical study of turbulent supersonic isothermal-wall channel flow. *J. Fluid Mech.* **305**, 159–183.

DA SILVA, C.B. & MÉTAIS, O. 2002 On the influence of coherent structures upon interscale interactions in turbulent plane jets. *J. Fluid Mech.* **473**, 103–145.
DEARDORFF, J.W. 1970 A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. *J. Fluid Mech.* **41**, 453–480.
DUAN, L., BEEKMAN, I. & MARTIN, M.P. 2010 Direct numerical simulation of hypersonic turbulent boundary layers. Part 2. Effect of wall temperature. *J. Fluid Mech.* **655**, 419–445.
DUBOIS, T., DOMARADZKI, J.A. & HONEIN, A. 2002 The subgrid scale estimation model applied to large eddy simulations of compressible turbulence. *Phys. Fluids* **14**, 1781–1801.
DUCHI, J., HAZAN, E. & SINGER, Y. 2011 Adaptive subgradient methods for online learning and stochastic optimization. *J. Machine Learning Res.* **12**, 2121–2159.
DURBIN, P.A. 2018 Some recent developments in turbulence closure modeling. *Annu. Rev. Fluid Mech.* **50**, 77–103.
EYINK, G.L. 2005 Locality of turbulent cascades. *Physica* D **207**, 91–116.
EYINK, G.L. & ALUIE, H. 2009 Localness of energy cascade in hydrodynamic turbulence. I. Smooth coarse graining. *Phys. Fluids* **21**, 115107.
FREZAT, H., BALARAC, G., LE SOMMER, J., FABLET, R. & LGUENSAT, R. 2021 Physical invariance in neural networks for subgrid-scale scalar flux modeling. *Phys. Rev. Fluids* **6**, 024607.
GAMAHARA, M. & HATTORI, Y. 2017 Searching for turbulence models by artificial neural network. *Phys. Rev. Fluids* **2**, 054604.
GARNIER, E., ADAMS, N. & SAGAUT, P. 2009 *Large Eddy Simulation for Compressible Flows*. Springer.
GERMANO, M., PIOMELLI, U., MOIN, P. & CABOT, W.H. 1991 A dynamic subgrid-scale eddy viscosity model. *Phys. Fluids* A **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.
GLOROT, X. & BENGIO, Y. 2010 Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research* (ed. Y.W. Teh & M. Titterington), vol. 9, pp. 249–256. PMLR.
GOODFELLOW, I., BENGIO, Y. & COURVILLE, A. 2016 *Deep Learning*. MIT Press.
HUNT, J.C.R., WRAY, A.A. & MOIN, P. 1988 Eddies, streams, and convergence zones in turbulent flows. In *Studying Turbulence Using Numerical Simulation Databases*, 2, pp. 193–208. Center for Turbulence Research, Ames Research Center.
JIANG, Z., XIAO, Z., SHI, Y. & CHEN, S. 2013 Constrained large-eddy simulation of wall-bounded compressible turbulent flows. *Phys. Fluids* **25**, 106102.
JIMÉNEZ, C., DUCROS, F., CUENOT, B. & BÉEDAT, B. 2001 Subgrid scale variance and dissipation of a scalar field in large eddy simulations. *Phys. Fluids* **13**, 1748–1754.
KAWAI, S. & LARSSON, J. 2013 Dynamic non-equilibrium wall-modeling for large eddy simulation at high Reynolds numbers. *Phys. Fluids* **25**, 015105.
KINGMA, D.P. & BA, J. 2014 Adam: a method for stochastic optimization. arXiv:1412.6980.
LAGHA, M., KIM, J., ELDREDGE, J.D. & ZHONG, X. 2011 A numerical study of compressible turbulent boundary layers. *Phys. Fluids* **23**, 015106.
LESIEUR, M. & METAIS, O. 1996 New trends in large-eddy simulations of turbulence. *Annu. Rev. Fluid Mech.* **28**, 45–82.
LILLY, D.K. 1992 A proposed modification of the germano subgrid-scale closure method. *Phys. Fluids* A **238**, 633–635.
LING, J., KURZAWSKI, A. & TEMPLETON, J.P. 2016 Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *J. Fluid Mech.* **807**, 155–166.
LU, L., SHIN, Y., SU, Y. & KARNIADAKIS, G. 2020 Dying relu and initialization: theory and numerical examples. *Commun. Comput. Phys.* **28**, 1671–1706.
LUND, T.S. & NOVIKOV, E.A. 1992 *Parameterization of Subgrid-Scale Stress By the Velocity Gradient Tensor*. Center for Turbulent Research, Stanford University, NASA.
MAEDER, T., ADAMS, N.A. & KLEISER, L. 2001 Direct simulation of turbulent supersonic boundary layers by an extended temporal approach. *J. Fluid Mech.* **429**, 187–216.
MARTIN, M.P., PIOMELLI, U. & CANDLER, G. 2000 Subgrid-scale models for compressible large-eddy simulations. *J. Theor. Comput. Fluid Dyn.* **13**, 361–376.
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.
MOIN, P., SQUIRES, K., CABOT, W. & LEE, S. 1991 A dynamic subgrid-scale model for compressible turbulence and scalar transport. *Phys. Fluids* A **3**, 2746.
MOSER, R.D., HAERING, S.W. & YALLA, G.R. 2021 Statistical properties of subgrid-scale turbulence models. *Annu. Rev. Fluid Mech.* **53**, 255–286.
NICOUD, F. & DUCROS, F. 1999 Subgrid-scale stress modelling based on the square of the velocity gradient tensor. *Flow Turbul. Combust.* **63**, 183–200.
NOLL, W. 1967 Representations of certain isotropic tensor functions. *Arch. Math.* **21**, 87–90.
PARK, J. & CHOI, H. 2021 Toward neural-network-based large eddy simulation: application to turbulent channel flow. *J. Fluid Mech.* **914**, A16.
PIOMELLI, U. 1993 High Reynolds number calculations using the dynamic subgrid-scale stress model. *Phys. Fluids* A **5**, 1484–1490.
PIOMELLI, U. 1999 Large-eddy simulation: achievements and challenges. *Prog. Aerosp. Sci.* **35**, 335–362.
PIROZZOLI, S. & BERNARDINI, M. 2011 Turbulence in supersonic boundary layers at moderate Reynolds number. *J. Fluid Mech.* **688**, 120–168.
PIROZZOLI, S., BERNARDINI, M. & GRASSO, F. 2008 Characterization of coherent vortical structures in a supersonic turbulent boundary layer. *J. Fluid Mech.* **613**, 205–231.
PIROZZOLI, S., BERNARDINI, M. & GRASSO, F. 2010 On the dynamical relevance of coherent vortical structures in turbulent boundary layers. *J. Fluid Mech.* **648**, 325–349.
PIROZZOLI, S., GRASSO, F. & GATSKI, T.B. 2004 Direct numerical simulation and analysis of a spatially evolving supersonic turbulent boundary layer at $M = 2.25$. *Phys. Fluids* **16**, 530–545.
POPE, S.B. 1975 A more general effective-viscosity hypothesis. *J. Fluid Mech.* **72**, 331–340.
RAGAB, S.A., SHEEN, S.C. & SREEDHAR, M. 1992 An investigation of finite-difference methods for large-eddy simulation of a mixing layer. *AIAA Paper* 92-0554.
SAGAUT, P. 2006 *Large Eddy Simulation for Incompressible Flows*. Springer.
SAYADI, T. & MOIN, P. 2012 large eddy simulation of controlled transition to turbulence. *Phys. Fluids* **24**, 114103.
SHU, C.-W. & OSHER, S. 1988 Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.* **77** (2), 439–471.
SMAGORINSKY, J. 1963 General circulation experiments with the primitive equations: I. The basic experiment. *Mon. Weath. Rev.* **91**, 99–164.
VOLLANT, V., BALARAC, G. & CORRE, C. 2017 Subgrid-scale scalar flux modelling based on optimal estimation theory and machine-learning procedures. *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.
WANG, Z., LUO, K., LI, D., TAN, J. & FAN, J. 2018*b* Investigations of data-driven closure for subgrid-scale stress in large-eddy simulation. *Phys. Fluids* **30**, 125101.
WANG, J., WAN, M., CHEN, S. & CHEN, S. 2018*a* Kinetic energy transfer in compressible isotropic turbulence. *J. Fluid Mech.* **841**, 581–613.
WANG, B.C., YIN, J., YEE, E. & BERGSTROM, D.J. 2007 A complete and irreducible dynamic SGS heat-flux modelling based on the strain rate tensor for large-eddy simulation of thermal convection. *Intl J. Heat Fluid Flow* **28**, 1227–1243.
WANG, Y., YUAN, Z., XIE, C. & WANG, J. 2021*a* Artificial neural network-based spatial gradient models for large-eddy simulation of turbulence. *AIP Adv.* **11**, 055216.
WANG, Y., YUAN, Z., XIE, C. & WANG, J. 2021*b* A dynamic spatial gradient model for the subgrid closure in large-eddy simulation of turbulence. *Phys. Fluids* **33**, 075119.
XIE, C., LI, K., MA, C. & WANG, J. 2019*a* 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., LI, K. & MA, C. 2019*c* Artificial neural network approach to large-eddy simulation of compressible isotropic turbulence. *Phys. Rev.* E **99**, 053113.
XIE, C., WANG, J., LI, H., WAN, M. & CHEN, S. 2019*b* Artificial neural network mixed model for large eddy simulation of compressible isotropic turbulence. *Phys. Fluids* **31**, 085112.
XIE, C., WANG, J. & WEINAN, E. 2020*a* Modeling subgrid-scale forces by spatial artificial neural networks in large eddy simulation of turbulence. *Phys. Rev. Fluids* **5**, 054606.
XIE, C., YUAN, Z. & WANG, J. 2020*b* Artificial neural network-based nonlinear algebraic models for large eddy simulation of turbulence. *Phys. Fluids* **32**, 115101.
XU, D., WANG, J. & CHEN, S. 2022*a* Skin-friction and heat-transfer decompositions in hypersonic transitional and turbulent boundary layers. *J. Fluid Mech.* **941**, A4.

XU, B., WANG, N., CHEN, T. & LI, M. 2015 Empirical evaluation of rectified activations in convolutional network. arXiv:1505.00853.

XU, D., WANG, J., WAN, M., YU, C., LI, X. & CHEN, S. 2021*a* Compressibility effect in hypersonic boundary layer with isothermal wall condition. *Phys. Rev. Fluids* **6**, 054609.

XU, D., WANG, J., WAN, M., YU, C., LI, X. & CHEN, S. 2021*b* Effect of wall temperature on the kinetic energy transfer in hypersonic turbulent boundary layer. *J. Fluid Mech.* **929**, A33.

XU, D., WANG, J., YU, C., LI, X. & CHEN, S. 2022*b* Contribution of flow topology to the kinetic energy flux in hypersonic turbulent boundary layer. *Phys. Fluids* **34**, 046103.

XU, D., WANG, J., YU, C., LI, X. & CHEN, S. 2022*c* Effect of compressibility on the small-scale structures in hypersonic turbulent boundary layer. *Phys. Fluids* **34**, 055121.

YANG, X., ZAFAR, S., WANG, J. & XIAO, H. 2019 Predictive large-eddy-simulation wall modeling via physics-informed neural networks. *Phys. Rev. Fluids* **4**, 034602.

YU, C., YUAN, Z., QI, H., WANG, J., LI, X. & CHEN, S. 2022 Kinetic-energy-flux-constrained model using an artificial neural network for large-eddy simulation of compressible wall-bounded turbulence. *J. Fluid Mech.* **932**, A23.

YUAN, Z., WANG, Y., XIE, C. & WANG, J. 2022 Dynamic nonlinear algebraic models with scale-similarity dynamic procedure for large-eddy simulation of turbulence. *Adv. Aerodyn.* **4**, 16.

YUAN, Z., XIE, C. & WANG, J. 2020 Deconvolutional artificial neural network models for large eddy simulation of turbulence. *Phys. Fluids* **32**, 115106.

ZANG, Y., STREET, R.L. & KOSEFF, J.R. 1993 A dynamic mixed subgrid-scale model and its application to turbulent recirculating flows. *Phys. Fluids* A **5**, 3186–3196.

ZEILER, M.D. 2012 Adadelta: an adaptive learning rate method. arXiv:1212.5701.

ZHAO, D. & ALUIE, H. 2018 Inviscid criterion for decomposing scales. *Phys. Rev. Fluids* **3**, 054603.

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.