

Subgrid-scale model for large-eddy simulation of isotropic turbulent flows using an artificial neural network

Zhideng Zhou^{a,b}, Guowei He^{a,b}, Shizhao Wang^{a,b}, Guodong Jin^{a,b,*}

^aThe State Key Laboratory of Nonlinear Mechanics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China

^bSchool of Engineering Science, University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:

Received 7 June 2019

Revised 12 August 2019

Accepted 9 October 2019

Available online 10 October 2019

Keywords:

Machine learning

Artificial neural network

Subgrid-scale model

Large-eddy simulation

Isotropic turbulent flows

ABSTRACT

An artificial neural network (ANN) is used to establish the relation between the resolved-scale flow field and the subgrid-scale (SGS) stress tensor, to develop a new SGS model for large-eddy simulation (LES) of isotropic turbulent flows. The data required for training and testing of the ANN are provided by performing filtering operations on the flow fields from direct numerical simulations (DNSs) of isotropic turbulent flows. We use the velocity gradient tensor together with filter width as input features and the SGS stress tensor as the output labels for training the ANN. In the *a priori* test of the trained ANN model, the SGS stress tensors obtained from the ANN model and the DNS data are compared by computing the correlation coefficient and the relative error of the energy transfer rate. The correlation coefficients are mostly larger than 0.9, and the ANN model can accurately predict the energy transfer rate at different Reynolds numbers and filter widths, showing significant improvement over the conventional models, for example the gradient model, the Smagorinsky model and its dynamic version. A real LES using the trained ANN model is performed as the *a posteriori* validation. The energy spectrum computed by the improved ANN model is compared with several SGS models. The Lagrangian statistics of fluid particle pairs obtained from the improved ANN model almost approach those from the filtered DNS, better than the results from the Smagorinsky model and dynamic Smagorinsky model.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Large-eddy simulation (LES) has emerged as an important numerical tool for turbulent flows. Compared with direct numerical simulation (DNS), LES resolves the turbulent flow fields at a coarser grid resolution and much lower computational cost. Additionally, the unsteady flow structures in high-Reynolds-number turbulent flows can be resolved using LES. In LES of incompressible turbulent flows, the resolved-scale flow fields are directly obtained by integrating the filtered Navier-Stokes equations, while the effects of the nonresolved subgrid-scale (SGS) motions are represented by an SGS stress tensor, which is usually modeled by an SGS model. It is a central issue to model the SGS stress tensor using the resolved-scale flow field in LES.

Currently, a variety of SGS models have been proposed to close the filtered Navier-Stokes equations [1,2]. These SGS models include the classic Smagorinsky model [3], the similarity model [4,5], the gradient model [6], the explicit nonlinear model [7], the mixed model [8], and the dynamic versions of these models [9–15]. Since

there is no backward scatter of SGS energy into the resolved scales in the Smagorinsky model, this model is usually too dissipative for the turbulent kinetic energy. Recently, Yang and Wang [16] have studied the topology of the SGS model using a proposed method based on the Euler angle and Euler axis of the relative rotation between the eigenframes of the SGS stress tensor and the resolved strain rate tensor, and they have pointed out that the topology of the Smagorinsky model deviates drastically from the DNS results. In contrast, the similarity model is insufficiently dissipative because of giving too much backward scatter of SGS energy into the resolved scales. The gradient model, based on the Taylor expansion, is accurate in estimating the local SGS stress tensor when the grid size is small, but this model becomes unstable with increasing grid size due to the incorrect prediction of energy transfer. Modulated gradient models have been proposed by controlling the energy transfer from the subgrid scales to the resolved scales, and they are limited to special conditions [17,18]. The dynamic versions of the aforementioned models and the mixed model can overcome the limitation of energy dissipation and give more accurate results in most cases, but these models are accompanied by high computational cost. Therefore, it is still desirable to develop a new SGS model for LES.

* Corresponding author.

E-mail address: gdjin@lnm.imech.ac.cn (G. Jin).

It is a great challenge to develop a new SGS model of high-fidelity. A feasible solution is to directly establish closure models for SGS stress based on existing DNS data. Due to the rapid growth in computing power, the computing scale for DNS is increasing quickly, and we can easily obtain large amounts of DNS data of turbulent flows at different Reynolds numbers. However, it is rather difficult to search for valuable information from such big data and establish reasonable SGS models. To solve these kinds of problems, machine learning has become a powerful tool due to improvements in computer hardware and algorithms. As a rapidly developing subject, machine learning has been applied to a wide range of fields, such as speech recognition [19], image recognition [20], natural language processing [21], and autonomous driving [22]. In this paper, we use an artificial neural network (ANN), a machine learning method, to develop a new SGS model based on DNS data.

The machine learning methods have been increasingly used to solve turbulent flow problems [23–28]. Duraisamy et al. [29] applied the ANN and Gaussian process regression methods to reconstruct improved functional forms of turbulence and transition modeling. Ling et al. [30] presented a deep neural network for Reynolds-averaged Navier-Stokes (RANS) turbulence modeling on an invariant tensor basis [31], which ensures Galilean invariance of the predicted Reynolds stress. Wang et al. [32] applied a random forests algorithm to learn the discrepancies between the Reynolds stresses from traditional RANS models and DNS and to predict Reynolds stress discrepancies in new flows using the obtained functions. Recently, Wu et al. [33] presented a comprehensive framework for physics-based implicit treatment to model Reynolds stress and successfully predicted the mean velocity field. In addition, there have been many applications of machine learning algorithms to quantify and eliminate the uncertainties of RANS models [34–37] and evaluate the extrapolation of Reynolds stress discrepancies between different flows [38–40]. For the application of machine learning in LES, Sarghini et al. [41] used a multilayer feedforward neural network as an SGS model to identify and reproduce the nonlinear behavior of turbulent flows. Vollant et al. [42] formulated new strategies based on an ANN to develop SGS models, and the LES results were very close to the filtered DNS (FDNS) results. Maulik and San [43] proposed a blind deconvolution ANN that performed well in *a priori* testing of Kraichnan, Kolmogorov and compressible stratified turbulent flows. They demonstrated that ANN is the optimal map for convolution and deconvolution of coarse-grained flow fields to close the two-dimensional Navier-Stokes equations [44]. Besides, Gamahara and Hattori [45] performed top-hat filtering on the DNS of turbulent channel flow to produce the flow data and used a single-hidden-layer ANN to find a new SGS model for the LES of turbulent channel flow. The data used for training is the resolved-scale flow field and SGS stress at one instant and a fixed filter width. Then they applied the trained ANN model to predict the spatial patterns of SGS stress at different Reynolds numbers and perform an *a posteriori* test in a real LES. Wang et al. [46] established the data-driven SGS model for an LES of decaying isotropic turbulence using the random forests and ANN algorithm. The data was also provided by performing top-hat filtering on the DNS flow field, and the sampled snapshots with a fixed filter width were chosen for the training and test. Their *a priori* and *a posteriori* test showed that ANN has better performance than the random forests algorithm for the regression problem. According to the eddy-viscosity assumption and the mixing-length theory of the SGS model, the SGS stress tensor not only depends on the velocity gradient tensor at the resolved scales but also the filter width which is a characteristic measure of the mixing-length. The relation between the SGS stress tensor and the velocity gradient tensor together with the filter width in the context of machine learning is in need of further study.

The objective of this paper is to establish and evaluate the relation between the resolved-scale flow field and the SGS stress tensor using an ANN in incompressible, homogeneous and isotropic turbulent flows. One important issue is to find a data-driven SGS model that works well in turbulent flows with different filter widths. The data required for training and testing of the ANN are the resolved-scale flow fields with different filter widths, which are obtained by performing filtering operations on the flow fields from DNSs of isotropic turbulent flows. The prediction accuracy and generalization of the trained ANN model are evaluated using DNS data at various Reynolds numbers, filter widths and time steps. Then, the *a posteriori* validation is performed by coupling the trained ANN model with a real LES of isotropic turbulent flows.

The paper is organized as follows: The DNS and FDNS of isotropic turbulent flows are introduced in Section 2. The correlation coefficient and the relative error of the energy transfer rate between the SGS stress tensor from the SGS model and that from the DNS data are used as two metrics for evaluating the Smagorinsky model and the gradient model. Based on the results, a suitable filter function, input features and output labels are chosen to produce the data for training and testing of the ANN. In Section 3, the basic procedures of a single-hidden-layer feedforward ANN are described in detail and the FDNS data with different filter widths are used to train the ANN model. In Section 4, we use the trained ANN model to predict the SGS stress tensors at different Reynolds numbers, filter widths and time steps. The results of correlation coefficients, the relative errors of the energy transfer rate and spatial structures of the SGS stress tensor are presented as the *a priori* test of the ANN model. In the *a posteriori* validation, we couple the ANN model with a real LES and compare the Eulerian energy spectra and Lagrangian statistics of fluid particle pairs from several SGS models. Section 5 gives the conclusion of this research.

2. Data preparation

To obtain the data for ANN training and testing, we performed filtering operations on the flow fields from DNS of isotropic turbulent flows [47–49]. The filtering operations using various filter functions and filter widths are carried out in this section.

2.1. Direct numerical simulation

The Navier-Stokes equations for incompressible flows are

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{u} \times \boldsymbol{\omega} - \nabla \left(\frac{p}{\rho} + \frac{1}{2} \mathbf{u}^2 \right) + \nu \nabla^2 \mathbf{u} + \mathbf{f}(\mathbf{x}, t), \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (2)$$

where \mathbf{u} denotes the velocity field, $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ denotes the vorticity field, p is the pressure, ρ is the fluid density, ν is the kinematic viscosity, and $\mathbf{f}(\mathbf{x}, t)$ denotes the forcing term to inject energy and maintain the turbulent state.

DNS of homogeneous and isotropic turbulent flows was performed using a standard pseudo-spectral method in a periodic cubic flow domain with each edge length $L = 2\pi$. In Fourier space, Eqs. (1) and (2) can be represented as

$$\left(\frac{\partial}{\partial t} + \nu k^2 \right) \hat{\mathbf{u}}(\mathbf{k}, t) = \mathbf{P}(\mathbf{k}) \mathcal{F}(\mathbf{u} \times \boldsymbol{\omega}) + \hat{\mathbf{f}}(\mathbf{k}, t), \quad (3)$$

where $\hat{\mathbf{u}}(\mathbf{k}, t)$ denotes the Fourier coefficient or the fluid velocity in Fourier space and \mathcal{F} denotes the Fourier transformation. The projection tensor $\mathbf{P}(\mathbf{k}) = \delta_{ij} - k_i k_j / k^2$ ($i, j = 1, 2, 3$) projects $\mathcal{F}(\mathbf{u} \times \boldsymbol{\omega})$ onto the plane normal to the wavenumber vector \mathbf{k} and eliminates the pressure gradient term in Eq. (1). The turbulent flow is driven by the deterministic forcing term $\hat{\mathbf{f}}(\mathbf{k}, t)$, which is

Table 1

DNS parameters of isotropic turbulent flows in a statistically steady state.

Re_λ	Grid size dx	k_{\max}	N^3
128.78	0.0245	256/3	256 ³
205.51	0.0123	512/3	512 ³
302.04	0.00614	1024/3	1024 ³

nonzero with a wavenumber magnitude $|\mathbf{k}| \leq 2$. Using this deterministic forcing method, a stationary turbulence was generated by maintaining the constant total energy in each of the first two wavenumber shells, and the energy ratio between the two shells was consistent with the $k^{-5/3}$ scaling law.

The cubic flow domain of isotropic turbulent flows was uniformly discretized into N^3 grids. The wavenumber components in Fourier space were defined as $k_j = n_j(2\pi/L)$, where $n_j = -N/2, \dots, -1, 0, 1, \dots, N/2 - 1$ for $j = 1, 2, 3$. The maximum wavenumber was approximately $k_{\max} = N/3$. The spatial resolution was monitored by the value of $k_{\max}\eta$, where η is the Kolmogorov length scale. The value of $k_{\max}\eta$ should be larger than 1.0 so that the Kolmogorov scale of the flow is well resolved [50–52]. This value was always nearly 1.5 in our simulations. The Fourier coefficients of the flow velocity were advanced in time using a second-order Adams-Bashforth method for the nonlinear term and an exact integration for the linear viscous term. The time step was chosen to ensure that the Courant-Friedrichs-Lewy (CFL) number was 0.5 or less for numerical stability and accuracy.

Table 1 presents the DNS parameters of isotropic turbulent flows in a statistically steady state, and the Taylor microscale Reynolds number $Re_\lambda = u'\lambda/\nu = 128.78, 205.51$ and 302.04 , where $u' = \sqrt{\langle u_i u_i \rangle}/3$ is the root mean square (rms) of the turbulent fluctuation velocity, $\lambda = (15\nu u'^2/\epsilon)^{1/2}$ is the Taylor microscale, and ϵ is the energy dissipation rate.

2.2. Filtered DNS and resolved-scale flow fields

The velocity field after performing filtering operations is given by the convolution

$$\bar{\mathbf{u}}(\mathbf{x}) = \int G(\mathbf{r})\mathbf{u}(\mathbf{x} - \mathbf{r})d\mathbf{r}, \quad (4)$$

where G is the filter function. The integration is carried out over the whole flow domain. The filtered Navier-Stokes equations for incompressible turbulent flows are

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial (\bar{u}_i \bar{u}_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_k \partial x_k} - \frac{\partial \tau_{ij}}{\partial x_j} + f_i, \quad (5)$$

$$\frac{\partial \bar{u}_j}{\partial x_j} = 0, \quad (6)$$

where the SGS stress tensor is defined as

$$\tau_{ij} = \bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j, \quad (7)$$

The SGS stress tensor is unknown and is usually modeled with an SGS model. In this work, we use Eqs. (4) and (7) to compute the filtered flow field and SGS stress tensor by filtering the DNS data of isotropic turbulent flows in an *a priori* study. In an *a posteriori* study, the flow field is resolved using Eqs. (5)–(7) with different SGS models to represent the SGS stress tensor.

The three types of filter functions listed in Table 2 are used to perform filtering operations on the DNS data. $\Delta = \pi/k_c$ denotes the filter width, where k_c is the cutoff wavenumber and it is selected to be $k_c/k_{\max} = 0.03125, 0.0625, 0.125, 0.25$ and 0.5 to vary the filter width. The parameters corresponding to different cutoff wavenumbers are shown in Table 3. $N_x \times N_y \times N_z = 64 \times 64 \times 64$

Table 2

Different filter functions used in this work.

Filter function	FDNS velocity $\bar{\mathbf{u}}(\mathbf{x}, t) = \sum_{ \mathbf{k} =1}^{k_{\max}} \hat{\mathbf{u}}(\mathbf{k}, t)e^{i\mathbf{k}\cdot\mathbf{x}}$
Sharp spectral filter	$\hat{\mathbf{u}}(\mathbf{k}, t) = \hat{\mathbf{u}}(\mathbf{k}, t) \cdot H(k_c - \mathbf{k}), \Delta = \pi/k_c$
Gaussian filter	$\hat{\mathbf{u}}(\mathbf{k}, t) = \hat{\mathbf{u}}(\mathbf{k}, t) \cdot \exp\left(-\frac{ \mathbf{k} ^2 \Delta^2}{24}\right)$
Box filter	$\hat{\mathbf{u}}(\mathbf{k}, t) = \hat{\mathbf{u}}(\mathbf{k}, t) \cdot \frac{\sin(\mathbf{k} \Delta/2)}{ \mathbf{k} \Delta/2}$

Table 3

The cutoff wavenumber and filter width Δ at different Re_λ .

Re_λ (DNS)	k_c/k_{\max}	Δ
128.78 (256 ³)	0.125	0.2945
	0.25	0.1473
	0.5	0.0736
	0.0625	0.2945
205.51 (512 ³)	0.125	0.1473
	0.25	0.0736
	0.5	0.0368
	0.03125	0.2945
302.04 (1024 ³)	0.0625	0.1473
	0.125	0.0736
	0.25	0.0368
	0.5	0.0184

denotes the number of output grid points from the FDNS flow field. For each case corresponding to different Reynolds numbers and filter widths, the FDNS flow fields at five different time steps are extracted for ANN training and testing, and these time steps are labeled as $t_i (i = 1, \dots, 5)$.

2.3. Subgrid-scale model and correlation coefficient

We choose the commonly-used Smagorinsky model and the gradient model to estimate the SGS stress tensor. The SGS stress tensors obtained from the two SGS models and the DNS data are compared. The correlation coefficient between the SGS stress tensor computed by the SGS model and that from the DNS data is used as a metric to select suitable input features for the ANN training in the next section.

The Smagorinsky model [3] is expressed as

$$\tau_{ij} = -2(C_S \Delta)^2 |\bar{S}| \bar{S}_{ij} + (1/3)\tau_{kk} \delta_{ij}, \quad (8)$$

$$|\bar{S}| = (2\bar{S}_{ij} \bar{S}_{ij})^{1/2}, \quad (9)$$

where $\bar{S}_{ij} = \frac{1}{2}(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i})$ denotes the strain rate tensor, $C_S = \frac{1}{\pi}(\frac{2}{3C_K})^{3/4}$ denotes the Smagorinsky coefficient [53]. In this work, we set Kolmogorov constant $C_K = 2.0$, so that $C_S \approx 0.14$.

The gradient model has been proposed by Clark et al. since 1979 [6]. Based on a Taylor series expansion of the filtered velocity field, the gradient model is expressed as

$$\tau_{ij} = \sum_{k=1}^3 \frac{\Delta^2}{12} \frac{\partial \bar{u}_i}{\partial x_k} \frac{\partial \bar{u}_j}{\partial x_k}. \quad (10)$$

The performance of the SGS models is evaluated by two physical quantities. One is the correlation coefficient between the SGS stress tensor computed by the SGS model and that from the DNS data. The other is the relative error of the energy transfer rate between that computed by the SGS model and the DNS data.

The correlation coefficient between the SGS stress tensor computed by the SGS model and that from the DNS data is expressed

Table 4

The correlation coefficients between the SGS stress tensors obtained from the SGS models and those from the DNS data with different filter functions at $Re_\lambda = 128.78$.

k_c/k_{max}	SGS model	filter	τ_{xx}	τ_{xy}	τ_{xz}	τ_{yy}	τ_{yz}	τ_{zz}
0.125	Smagorinsky model	Sharp	0.05282	0.10336	0.08979	0.07656	0.09703	0.05026
		Box	0.12737	0.27656	0.24705	0.16409	0.25868	0.11469
		Gaussian	0.14153	0.31008	0.27705	0.18233	0.29134	0.12607
	Gradient model	Sharp	0.43892	0.41919	0.39670	0.45129	0.41605	0.43542
		Box	0.86863	0.86234	0.86155	0.86176	0.86240	0.86290
		Gaussian	0.93740	0.94453	0.94516	0.93238	0.94476	0.93435
0.25	Smagorinsky model	Sharp	0.05058	0.05818	0.05976	0.05778	0.06082	0.03811
		Box	0.13684	0.24233	0.22500	0.16558	0.23192	0.11554
		Gaussian	0.14408	0.26200	0.24280	0.17506	0.25009	0.12163
	Gradient model	Sharp	0.43379	0.42209	0.41028	0.45350	0.42470	0.43449
		Box	0.95819	0.95150	0.95166	0.95621	0.95189	0.95563
		Gaussian	0.97669	0.97683	0.97715	0.97526	0.97686	0.97549

as [45,46]

$$C_{ij}(M, X) = \frac{\langle (\tau_{ij}^{(M)} - \langle \tau_{ij}^{(M)} \rangle) (\tau_{ij}^{(X)} - \langle \tau_{ij}^{(X)} \rangle) \rangle}{\sqrt{\langle (\tau_{ij}^{(M)} - \langle \tau_{ij}^{(M)} \rangle)^2 \rangle} \sqrt{\langle (\tau_{ij}^{(X)} - \langle \tau_{ij}^{(X)} \rangle)^2 \rangle}}, \quad (11)$$

where M and X denote the SGS stress computed by the SGS model and the DNS data, respectively, and $\langle \cdot \rangle$ denotes the average over the flow domain and time.

The rate of energy transfer from the resolved scales to the sub-grid scales, also called the SGS energy dissipation rate [53], is defined as

$$D_\tau = -\tau_{ij} \bar{S}_{ij}. \quad (12)$$

Table 4 shows the correlation coefficients between the SGS stress tensors computed by the SGS models according to Eqs. (8)–(10) and those from the DNS data according to Eq. (7). For the three filter functions, the correlation coefficients between the SGS stress tensors computed by the Smagorinsky model and the DNS data are mostly less than 0.3. This finding demonstrates that the Smagorinsky model has little correlation with the real SGS stress tensor from DNS data [6,54].

For the gradient model, the correlation coefficients are very small, approximately 0.4~0.5, when a sharp spectral filter is used. The correlation coefficients are larger than 0.85 for different filter widths when a box or Gaussian filter is used. This result is because the sharp spectral filter is sharp in spectral space but decidedly nonlocal in physical space. The velocity gradient at a specific position is globally influenced by the filtering operation in physical space and cannot be used to model the local SGS stress tensor. In contrast, the box filter is local in physical space, but this filter is nonlocal in spectral space [53] and causes intense oscillations on the filtered energy spectrum, as shown in Fig. 1. Obviously, the box filter is not effective at filtering the energy at high wavenumbers. Among the three filters, only the Gaussian filter is reasonably compact in both physical and spectral space. Therefore, we choose the Gaussian filter to obtain the FDNS data for ANN training and testing. Furthermore, we choose the velocity gradient tensor and filter width as the ANN input features and the SGS stress tensor as the ANN output labels to train the SGS model, as described in the next section.

Table 5 shows the correlation coefficients between the SGS stress tensor computed by the gradient model and that computed by Eq. (7) using the DNS data and Gaussian filter. The correlation coefficients at different Reynolds numbers and filter widths are mostly larger than 0.9 and show a slight decrease with increasing filter width. Those results mean that the gradient model physically describes the relation between the SGS stress tensor and velocity gradient tensor at the resolved scales over a wide range of filter widths.

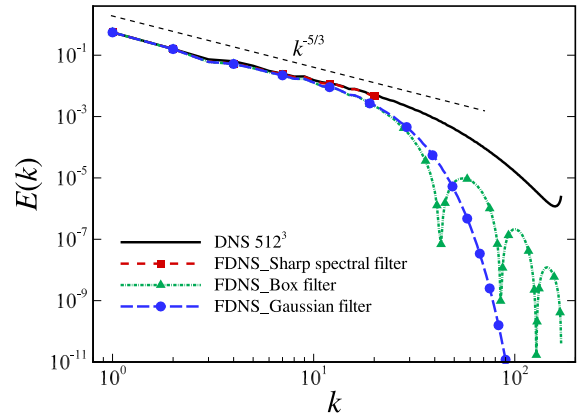


Fig. 1. Energy spectra of the FDNS ($k_c/k_{max} = 0.125$) with three filter functions: dashed line, sharp spectral filter; dash-dotted line, box filter; long-dashed line, Gaussian filter.

Table 6 shows the energy transfer rates obtained from the gradient model and the DNS data and the relative error between them. $D_{\tau,X}$ and $D_{\tau,M}$ are the energy transfer rates obtained from DNS and the modeled SGS stress tensor, respectively,

$$D_{\tau,X} = \frac{1}{N_g} \sum_{n=1}^{N_g} -\tau_{ij}^{(X)}(n) \bar{S}_{ij}(n), \quad (13)$$

$$D_{\tau,M} = \frac{1}{N_g} \sum_{n=1}^{N_g} -\tau_{ij}^{(M)}(n) \bar{S}_{ij}(n), \quad (14)$$

where N_g denotes the number of grid points in the FDNS flow field and $N_g = 64 \times 64 \times 64$ for all the filter widths and flows at different Reynolds numbers.

The relative error of the SGS energy transfer rate between the modeled and DNS stress tensors is defined as

$$E_D = \frac{(D_{\tau,M} - D_{\tau,X})}{D_{\tau,X}} \times 100\%. \quad (15)$$

In Table 6, the relative error E_D monotonically increases with the filter width. This result means that although the correlation coefficient approaches 1.0, the value of the SGS stress tensor is underestimated due to the coefficient used in the gradient model.

Therefore, it is necessary to establish the relation between the velocity gradient tensor together with the filter width and the SGS stress tensor using the ANN method.

Table 5

The correlation coefficients between the SGS stress tensors obtained from the gradient model and those from the DNS data with a Gaussian filter.

Re_λ	Δ	τ_{xx}	τ_{xy}	τ_{xz}	τ_{yy}	τ_{yz}	τ_{zz}
128.78	0.2945	0.93740	0.94453	0.94516	0.93238	0.94476	0.93435
	0.1473	0.97669	0.97683	0.97715	0.97526	0.97686	0.97549
	0.0736	0.99456	0.99427	0.99434	0.99427	0.99417	0.99426
205.51	0.2945	0.89927	0.91747	0.91432	0.89270	0.91390	0.89910
	0.1473	0.94192	0.94788	0.94841	0.94025	0.94678	0.94227
	0.0736	0.97740	0.97726	0.97769	0.97702	0.97690	0.97748
302.04	0.0368	0.99440	0.99401	0.99409	0.99439	0.99400	0.99447
	0.2945	0.90447	0.91856	0.90786	0.87630	0.89712	0.86402
	0.1473	0.91602	0.92859	0.92574	0.91070	0.92267	0.90696
	0.0736	0.95474	0.95483	0.95470	0.95044	0.95270	0.95018
	0.0368	0.98030	0.97986	0.97969	0.98099	0.98000	0.98146
	0.0184	0.99572	0.99520	0.99531	0.99554	0.99514	0.99579

Table 6

The energy transfer rates generated by the gradient model and the DNS stress tensor using a Gaussian filter and the relative error between them.

Re_λ	Δ	$D_{\tau,x}$	$D_{\tau,M}$	E_D
128.78	0.2945	0.15248	0.12097	-20.66%
	0.1473	0.09484	0.08871	-6.46%
	0.0736	0.04119	0.04070	-1.20%
205.51	0.2945	0.16938	0.10516	-37.92%
	0.1473	0.13604	0.10909	-19.81%
	0.0736	0.08276	0.07749	-6.37%
302.04	0.0368	0.03563	0.03519	-1.24%
	0.2945	0.17752	0.09286	-47.69%
	0.1473	0.15864	0.10525	-33.66%
	0.0736	0.12446	0.10361	-16.76%
	0.0368	0.07808	0.07399	-5.25%
	0.0184	0.03000	0.02972	-0.94%

3. Artificial neural network

3.1. Structure of the ANN

We use a single-hidden-layer feedforward ANN to establish the relation between the resolved-scale flow field and the SGS stress tensor. This ANN consists of an input layer, a hidden layer and an output layer.

The input layer is

$$\mathbf{X} = [x_1, x_2, \dots, x_{n_i}]^T, \quad (16)$$

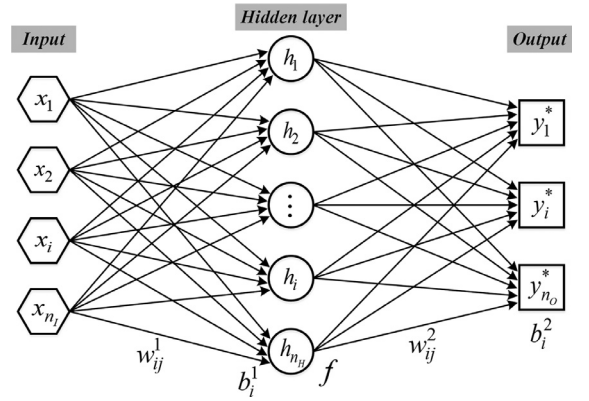
where x_i denotes the i th input feature, n_i is the number of neurons in the input layer. The matrix of weight and bias coefficient connecting the input layer and the hidden layer are

$$\mathbf{W}^1 = \begin{bmatrix} w_{11}^1 & w_{12}^1 & \dots & w_{1n_H}^1 \\ w_{21}^1 & w_{22}^1 & \dots & w_{2n_H}^1 \\ \vdots & \vdots & \ddots & \vdots \\ w_{n_i1}^1 & w_{n_i2}^1 & \dots & w_{n_in_H}^1 \end{bmatrix}, \quad \mathbf{B}^1 = \begin{bmatrix} b_1^1 \\ b_2^1 \\ \vdots \\ b_{n_H}^1 \end{bmatrix}, \quad (17)$$

where w_{ij}^1 ($i = 1, 2, \dots, n_i; j = 1, 2, \dots, n_H$) denotes the weight coefficient connecting the i th neuron in the hidden layer and the j th neuron in the input layer, b_i^1 denotes the bias coefficient for the i th neuron in the hidden layer, n_H is the number of neurons in the hidden layer. Initially, the weight coefficients are set to be random numbers from truncated normal distribution (0.0 mean and 0.1 standard deviation) and the bias coefficients are set to zero.

The output of the hidden layer is

$$\mathbf{H}^T = f(\mathbf{W}^1 \mathbf{X} + \mathbf{B}^1) = [h_1, h_2, \dots, h_{n_H}]^T, \quad h_i = f\left(\sum_{j=1}^{n_i} w_{ij}^1 x_j + b_i^1\right), \quad (18)$$

**Fig. 2.** Schematic diagram of the single-hidden-layer feedforward ANN.

where f denotes the activation function to carry out the nonlinear mapping of the ANN, and the superscript "T" denotes the transpose of matrix. The matrix of weight and bias coefficient connecting the hidden layer and the output layer are

$$\mathbf{W}^2 = \begin{bmatrix} w_{11}^2 & w_{12}^2 & \dots & w_{1n_o}^2 \\ w_{21}^2 & w_{22}^2 & \dots & w_{2n_o}^2 \\ \vdots & \vdots & \ddots & \vdots \\ w_{n_H1}^2 & w_{n_H2}^2 & \dots & w_{n_Hn_o}^2 \end{bmatrix}, \quad \mathbf{B}^2 = [b_1^2, b_2^2, \dots, b_{n_o}^2], \quad (19)$$

where w_{ij}^2 ($i = 1, 2, \dots, n_H; j = 1, 2, \dots, n_o$) denotes the weight coefficient connecting the i th neuron in the output layer and the j th neuron in the hidden layer, b_i^2 denotes the bias coefficient for the i th neuron in the output layer and n_o is the number of neurons in the output layer.

The output of the ANN is calculated by

$$\mathbf{Y}^* = \mathbf{W}^2 \mathbf{H}^T + \mathbf{B}^2 = [y_1^*, y_2^*, \dots, y_{n_o}^*]^T, \quad y_i^* = \sum_{j=1}^{n_H} w_{ij}^2 h_j + b_i^2. \quad (20)$$

Fig. 2 shows a schematic diagram of the single-hidden-layer feedforward ANN.

In this paper, the nine components of the velocity gradient tensor and the filter width are used as input features ($n_i = 10$) in the ANN,

$$\mathbf{X} = \left[\Delta, \frac{\partial \bar{u}_i}{\partial x_j}, \Delta \right]^T \quad (i, j = 1, 2, 3). \quad (21)$$

Table 7
The training data for the ANN model.

Re_λ	k_c/k_{\max}	Δ	$N_x \times N_y \times N_z$	Time step
128.78	0.125	0.2945	64 × 64 × 64	t_1
	0.25	0.1473		
	0.5	0.0736		
302.04	0.125	0.0736		
	0.25	0.0368		
	0.5	0.0184		

The output labels of the ANN are the six components of the SGS stress tensor ($n_O = 6$),

$$\mathbf{Y} = [\tau_{xx}, \tau_{xy}, \tau_{xz}, \tau_{yy}, \tau_{yz}, \tau_{zz}]. \quad (22)$$

The activation function used in this paper is the rectified linear unit (ReLU) [55],

$$f(x) = \begin{cases} 0, & \text{if } x < 0, \\ x, & \text{if } x \geq 0. \end{cases} \quad (23)$$

The loss function can be expressed as

$$E_{SGS} = \frac{1}{N_s} \sum_{i=1}^{N_s} |\mathbf{Y}_i - \mathbf{Y}_i^*|^2 + \frac{\lambda_0}{2N_s} \|\mathbf{w}\|_2^2, \quad (24)$$

where N_s is the number of sample points, w is the weight coefficient, and λ_0 is the regularization rate, which is set to 0.0001. The first term in Eq. (24) is the mean square error between the ANN output \mathbf{Y}^* and the labeled output \mathbf{Y} calculated from DNS data. The second term is an L2 regularization term that is included to avoid overfitting.

We use the error backpropagation (BP) scheme [56] implemented with TensorFlow [57] to train the ANN by optimizing the weight and bias coefficients to minimize the loss function E_{SGS} . The BP scheme is an iterative method. The main procedures of BP-ANN training are as follows:

- (1) Training data is provided to the input layer, the data signal is propagated forward layer by layer, and the result in the output layer can be computed according to Eqs. (16)–(20).
- (2) Compute the mean square error between the ANN output and the labeled output. Then, the loss function in the output layer is determined according to Eq. (24).
- (3) Propagate the loss function backward from the output layer to the hidden layer and input layer. Adjust the weight and bias coefficients using the gradient descent algorithm,

$$v = v + \Delta v, \quad \Delta v = -\eta' \frac{\partial E_{SGS}}{\partial v}, \quad (25)$$

where v denotes one of the weight and bias coefficients in the ANN and $\eta' \in (0, 1)$ denotes the learning rate, which is set to 0.2 initially and decays exponentially with the training step.

- (4) Repeat the above procedures as the loop iteration until the stop condition is met.

3.2. ANN training

For training the ANN model, we select the FDNS data at two Reynolds numbers, $Re_\lambda = 128.78$ and 302.04. Three different filter widths are used for each Reynolds number. The details of the dataset are listed in Table 7. The total number of sample points in the dataset is $64 \times 64 \times 64 \times 6$, which is then divided randomly into two parts. Three-quarters of the sample points are used as the training dataset to adjust the weight and bias coefficients to minimize the loss function, while the remaining quarter of the sample points are used as the validation dataset. In the training, we use

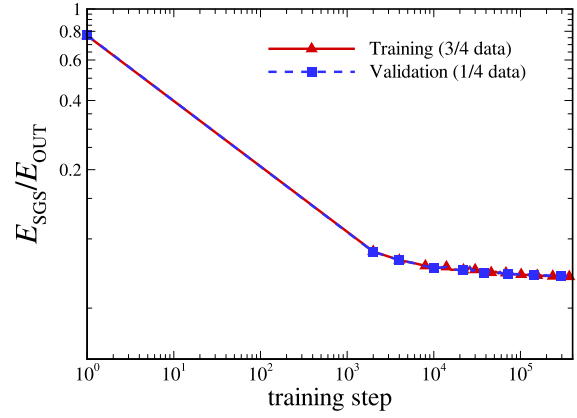


Fig. 3. Variation in the normalized error with the training step; the number of neurons in the hidden layer is $n_H = 300$.

Table 8
Three comparative cases with different n_H for the ANN model.

	Re_λ	k_c/k_{\max}	Time step	$N_x \times N_y \times N_z$	n_H
Case 1	128.78	0.125	t_3		50/100/
Case 2	205.51	0.0625	t_1	64 × 64 × 64	200/300/
Case 3	302.04	0.5	t_5		400/500, etc.

the mini-batch gradient descent (MBGD) algorithm and set batch size to 1024, which means 1024 sample points are randomly chosen to adjust the weight and bias coefficients. At every training step, the instantaneously adjusted ANN model based on the training dataset is applied to the validation dataset, and the error between the SGS stress from the ANN and DNS data is calculated to evaluate the generalization of the ANN model.

Fig. 3 shows the variations in the normalized errors E_{SGS}/E_{OUT} in the training dataset and validation dataset with the training step. E_{OUT} denotes the mean square of the labeled output \mathbf{Y} over the whole domain,

$$E_{OUT} = \frac{1}{N_s} \sum_{n=1}^{N_s} \mathbf{Y}^2. \quad (26)$$

Initially, the weight coefficients are randomly set, and the bias coefficients are set to zero; therefore, the normalized errors are large. With increasing training step, the errors decrease as the weight and bias coefficients are adjusted according to the gradient descent algorithm and then tend to approach steady values after the rapid decrease. The errors in the training and validation datasets exhibit similar variations, which shows that the ANN model has been successfully trained.

4. Results and discussion

After training the ANN model, we input the new flow data at different Reynolds numbers, time steps and filter widths into the trained model. The SGS stress tensor predicted by the ANN model and that from DNS data are compared to evaluate the prediction accuracy and generalization of the ANN model.

4.1. Effects of the number of neurons in the hidden layer

We obtain distinct ANN models with different numbers of neurons, n_H , in the hidden layer. Then, we apply all the ANN models to the same data and compare the correlation coefficients and the relative errors of the energy transfer rate predicted by various ANN models as comparative cases. The detailed information for three cases is listed in Table 8. In Cases 1 and 3, we choose data at different time steps from those of the training data. In Case 2, we

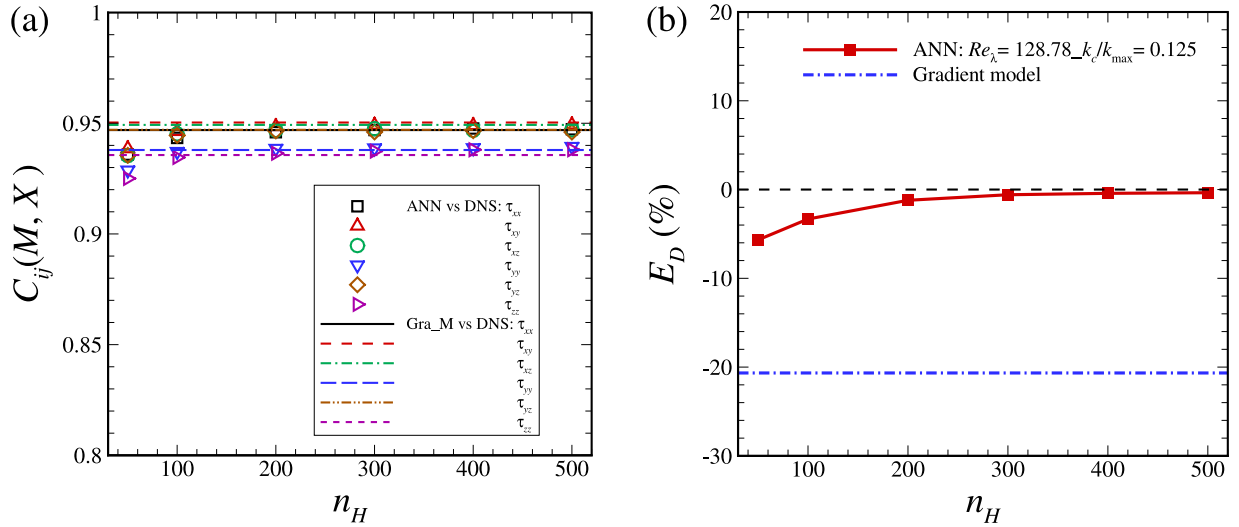


Fig. 4. Variations in the correlation coefficients and the relative error of the energy transfer rate with the number of neurons in the hidden layer in Case 1.

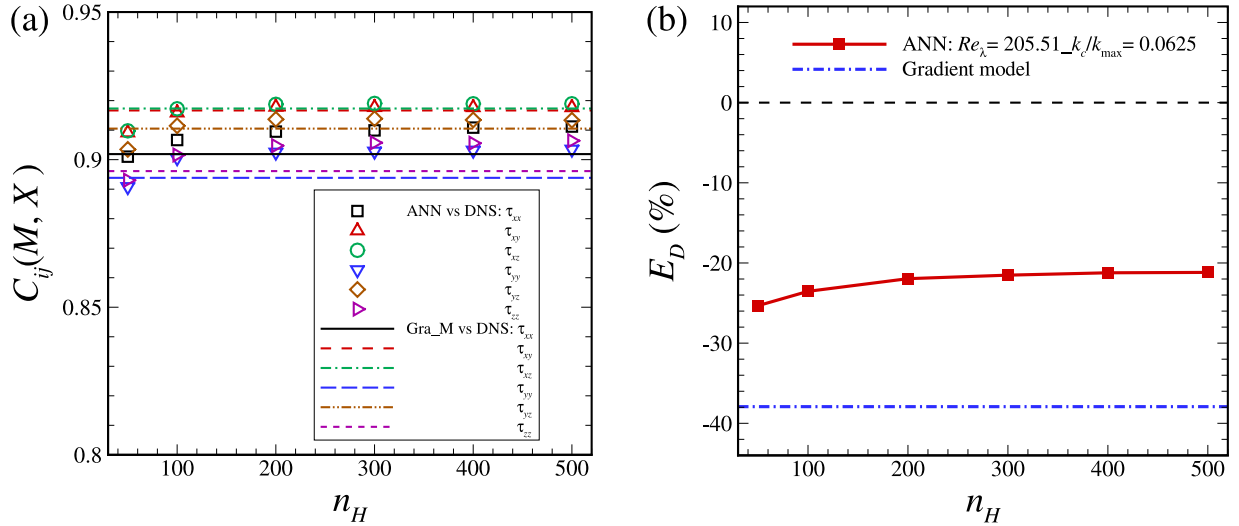


Fig. 5. Variations in the correlation coefficients and the relative error of the energy transfer rate with the number of neurons in the hidden layer in Case 2.

choose data at a different Reynolds number from that of the training data. According to the comparative cases, we can determine a proper number of neurons in the hidden layer for the ANN model.

Figs. 4–6 show the variations in the correlation coefficients and the relative errors of the energy transfer rate computed by the ANN model and the gradient model with the number of neurons. For the above three cases, the correlation coefficients increase monotonously with the number of neurons for all six components of the SGS stress tensor and reach stable values after $n_H \geq 300$ [Figs. 4(a)–6(a)]. Meanwhile, the relative errors computed by the ANN model significantly change with the number of neurons in the hidden layer. When k_c/k_{\max} is small, i.e., the filter width is large, the relative error of the energy transfer rate computed by the gradient model is quite large, while the ANN model can reduce the relative errors of the energy transfer rate at different Reynolds numbers and filter widths. As shown in Figs. 4(b)–6(b), the relative errors of the energy transfer rate computed by the ANN model are large when $n_H = 50$. Then, these errors gradually decrease and approach zero with increasing n_H . When $n_H \geq 300$, the relative errors approach steady values. In Case 1, the relative error reduces to zero after $n_H \geq 300$. In Case 2, although the relative error remains at approximately -20% , this error has been significantly reduced

compared with that of the gradient model. The latter has a relative error as large as -40% . In Case 3, the filter width is very small, so the relative error of the gradient model is very small. The relative error of the ANN model is comparable to that of the gradient model when $n_H = 200$.

Taking the above three cases into account, we select $n_H = 300$ in this paper to obtain an ANN model for subsequent prediction of the SGS stress tensor.

4.2. A priori test of the ANN model

After selecting $n_H = 300$, we applied the trained ANN model to predict the SGS stress tensor at different Reynolds numbers, filter widths and time steps. The parameters of FDNS data used for a priori test of the ANN model are listed in Table 9.

Fig. 7 shows the correlation coefficients and the relative error of the energy transfer rate at five time steps when $Re_\lambda = 128.78$ and $k_c/k_{\max} = 0.125$. For the isotropic turbulent flow fields at different time steps, steady values are maintained for the correlation coefficients and relative error of the energy transfer rate between the SGS stress tensor predicted by the ANN and that from DNS data. This finding demonstrates that the trained ANN model accu-

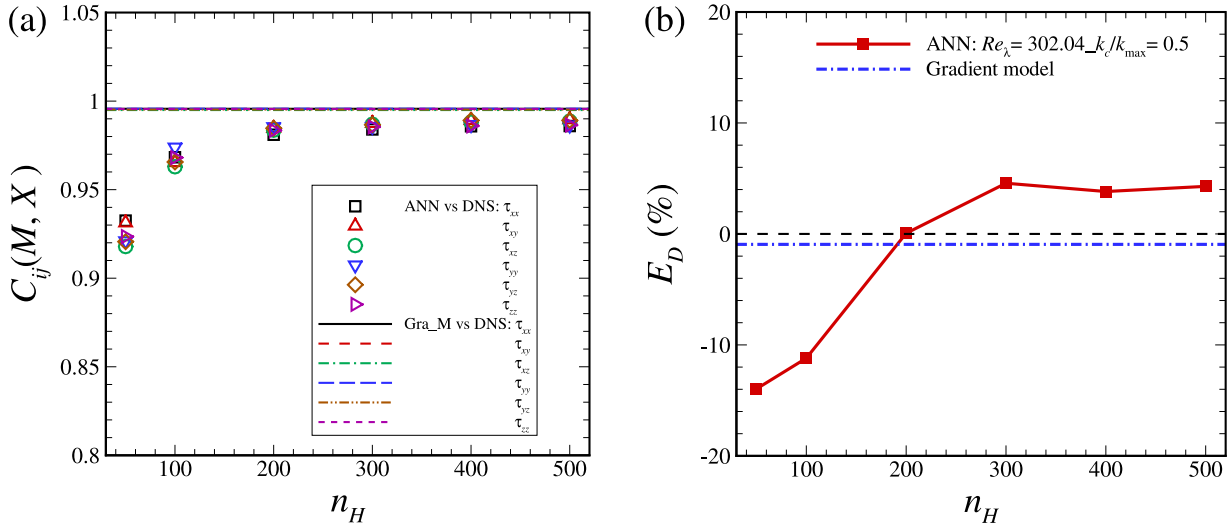


Fig. 6. Variations in the correlation coefficients and the relative error of the energy transfer rate with the number of neurons in the hidden layer in Case 3.

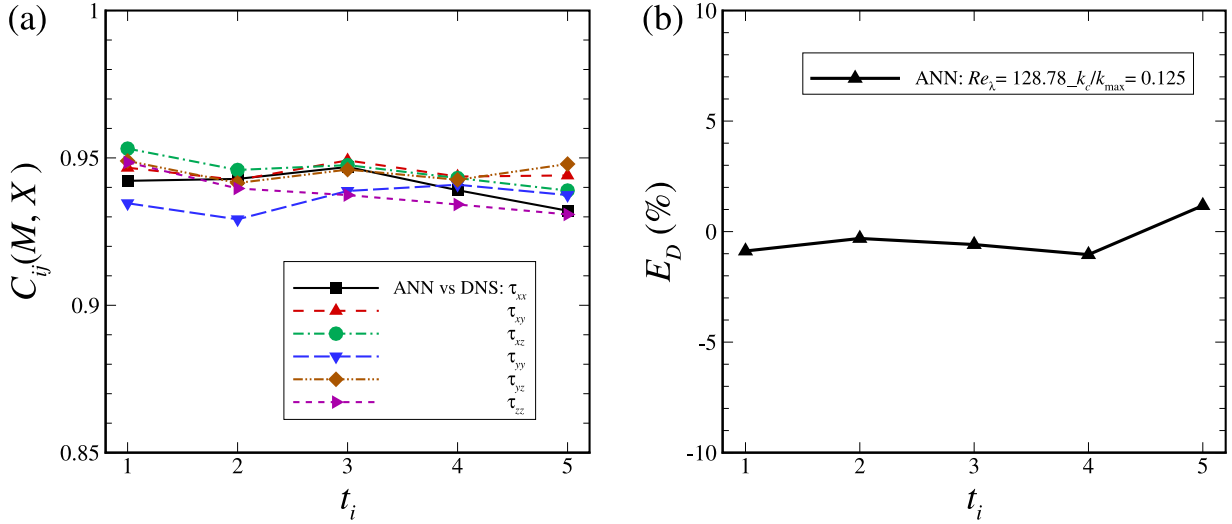


Fig. 7. The correlation coefficients and relative error of the energy transfer rate predicted by the ANN model at five time steps when $Re_\lambda = 128.78$ and $k_c/k_{max} = 0.125$.

Table 9
The parameters of FDNS data used for *a priori* test of the ANN model.

Re_λ	Δ	$N_x \times N_y \times N_z$	Remarks
128.78	0.2945, 0.1473, 0.0736		Average over 5 time steps
205.51	0.2945, 0.1473, 0.0736, 0.0368	$64 \times 64 \times 64$	
302.04	0.2945, 0.1473, 0.0736, 0.0368, 0.0184		

rately describes the relation between the velocity gradient tensor and SGS stress tensor at different time steps. Therefore, we can obtain the statistical results of ANN model by averaging the results at different time steps, shown as Table 10.

Figs. 8–10 show the variations in the correlation coefficients and the relative errors of the energy transfer rate with the filter width computed by the ANN model, the gradient model and the Smagorinsky model. Most of the correlation coefficients determined by the ANN model are larger than 0.9 and close to those by the gradient model, which both are obviously better than the Smagorinsky model [Figs. 8(a)–10(a)]. These high correlation coefficients indicate that the ANN model accurately predicts the spatial patterns of the SGS stress tensor at different Reynolds numbers and filter widths. In Figs. 8(b)–10(b), the gradient model underesti-

mates the energy transfer rate with increasing filter width and the Smagorinsky model significantly overestimates the energy transfer rate. When the filter width is large, the ANN model partially reduces the relative error of the energy transfer rate compared to the gradient model (the maximum absolute value decreases from -47.69% to -33.28% at $Re_\lambda = 302.04$, from -37.92% to -21.14% at $Re_\lambda = 205.51$, from -20.66% to -0.33% at $Re_\lambda = 128.78$, shown as Tables 6 and 10). When the filter width is small, the relative errors computed by the ANN model are sometimes a little larger than zero, but still close to the gradient model. In all, the ANN model significantly improves the prediction of the energy transfer rate at different Reynolds numbers and filter widths.

In order to assess the importance of filter width in the input features, two different ANNs are trained to compare with the present ANN model. First, we choose $[\frac{\partial u_i}{\partial x_j}]$ as input features and $[\tau_{ij}]$ as output labels. The dataset is the same as the present ANN model, listed in Table 7. The normalized errors in the training and validation datasets can hardly be reduced with the increasing training step, demonstrating the failure of this training. Actually, this failure is easy to understand that we attempt to train an ANN model similar to the gradient model using FDNS data with different filter widths, but not taking the filter width into consideration.

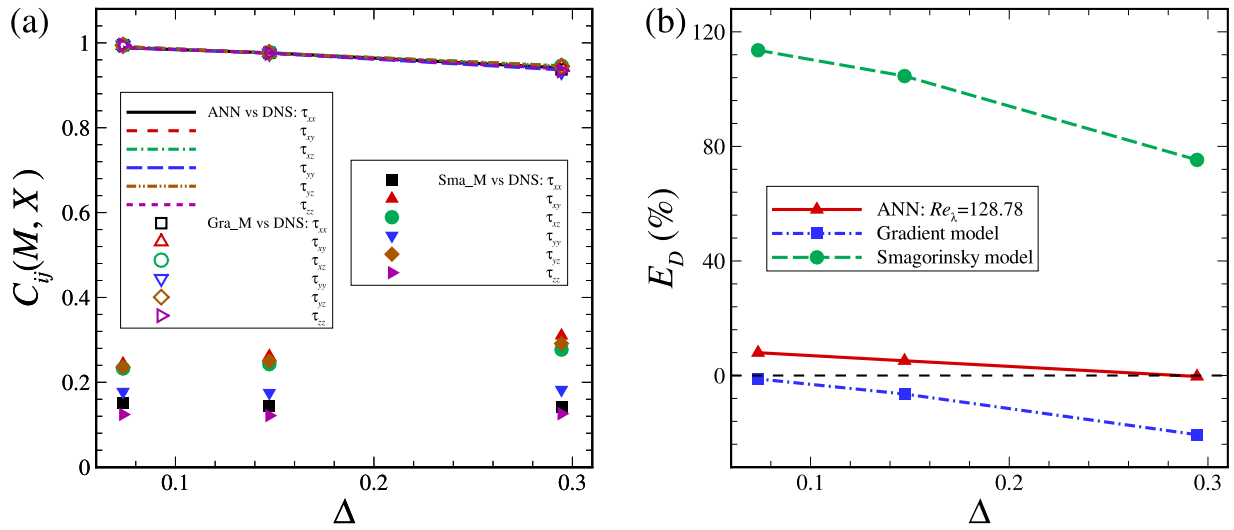


Fig. 8. Variations in the correlation coefficients and relative errors of the energy transfer rate with filter width computed by the ANN model, the gradient model and the Smagorinsky model at $Re_\lambda = 128.78$, where the filter width $\Delta = \pi/k_c$.

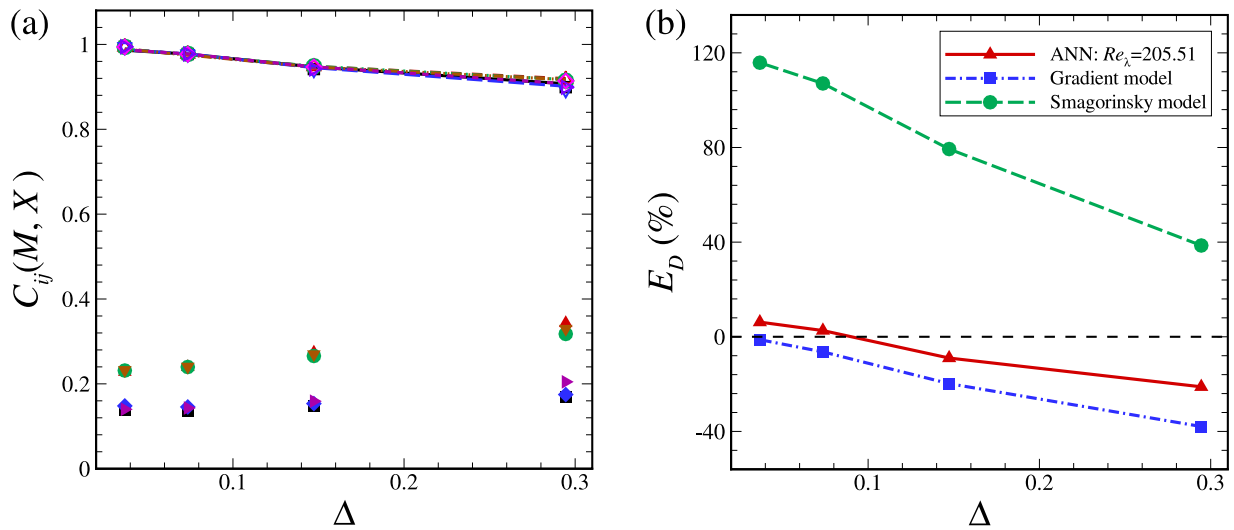


Fig. 9. Variations in the correlation coefficients and relative errors of the energy transfer rate with filter width computed by the ANN model, the gradient model and the Smagorinsky model at $Re_\lambda = 205.51$. The line legend in (a) is the same as that in Fig. 8(a).

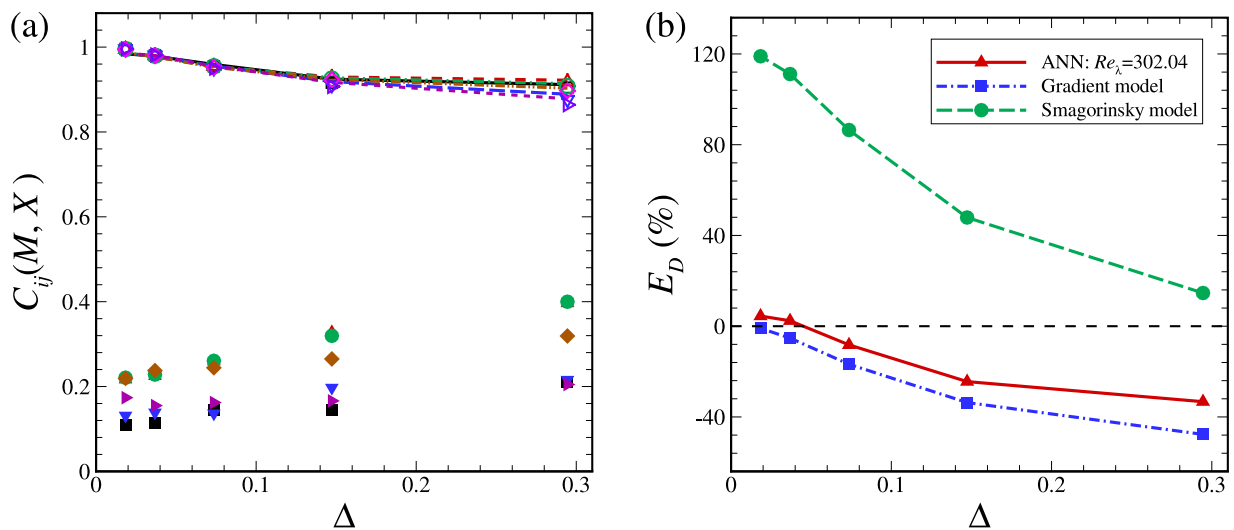


Fig. 10. Variations in the correlation coefficients and relative errors of the energy transfer rate with filter width computed by the ANN model, the gradient model and the Smagorinsky model at $Re_\lambda = 302.04$. The line legend in (a) is the same as that in Fig. 8(a).

Table 10

The correlation coefficients and relative errors of the energy transfer rate between the SGS stress tensors obtained from the ANN model and the DNS data.

Re_λ	Δ	τ_{xx}	τ_{xy}	τ_{xz}	τ_{yy}	τ_{yz}	τ_{zz}	E_D
128.78	0.2945	0.94117	0.94533	0.94617	0.93642	0.94547	0.93877	-0.33%
	0.1473	0.97667	0.97596	0.97623	0.97562	0.97586	0.97580	5.15%
	0.0736	0.98844	0.99108	0.99104	0.98892	0.99077	0.98873	7.96%
205.51	0.2945	0.90722	0.91935	0.91720	0.90147	0.91710	0.90744	-21.14%
	0.1473	0.94687	0.94804	0.94844	0.94530	0.94673	0.94713	-8.98%
	0.0736	0.97716	0.97587	0.97618	0.97743	0.97506	0.97757	2.63%
302.04	0.0368	0.98640	0.98927	0.98905	0.98737	0.98868	0.98731	6.19%
	0.2945	0.91157	0.92197	0.91390	0.88895	0.90304	0.87811	-33.28%
	0.1473	0.92450	0.92968	0.92764	0.91871	0.92367	0.91692	-24.42%
	0.0736	0.95914	0.95451	0.95433	0.95520	0.95233	0.95522	-8.23%
	0.0368	0.97856	0.97678	0.97739	0.98002	0.97749	0.98012	2.39%
	0.0184	0.98418	0.98729	0.98705	0.98467	0.98619	0.98572	4.47%

Table 11

The training data with a fixed filter width $\Delta_0 = 0.1473$.

Re_λ	$N_x \times N_y \times N_z$	Time step
128.78		
205.51	64 × 64 × 64	t_1 and t_5
302.04		

Second, we still choose $[\frac{\partial \bar{u}_i}{\partial x_j}]$ as input features and $[\tau_{ij}]$ as output labels. But the dataset is selected from the FDNS data with a fixed filter width $\Delta_0 = 0.1473$ at different Reynolds numbers, listed in Table 11. This ANN is successfully trained, named as “ANN (Δ_0)”. Then we perform the *a priori* test using the ANN (Δ_0) model.

Fig. 11 shows the comparison of the *a priori* results by the present ANN model and the ANN (Δ_0) model. The correlation coefficients determined by the ANN (Δ_0) model are mostly larger than 0.9, close to the present ANN model. Only when $\Delta < \Delta_0$, the ANN (Δ_0) model slightly underestimates the correlation coefficient. As for the energy transfer rate, the ANN (Δ_0) model accurately predicts the values only if $\Delta = \Delta_0$. If $\Delta < \Delta_0$, the ANN (Δ_0) model significantly overestimates the relative error; if $\Delta > \Delta_0$, the ANN (Δ_0) model significantly underestimates the relative error. The main reason is that the coefficient of SGS model related to filter width ($\propto \Delta^2$) is not used as an input feature, but included in the trained ANN (Δ_0) model. So the predicted SGS stress could be proportional to the real SGS stress, and the ratio approximates Δ_0^2/Δ^2 , shown in Fig. 11(b), (d) and (f).

If we select the FDNS data with another fixed filter width Δ'_0 as the training dataset, the ANN model can be successfully trained. It can also be expected that this ANN model cannot accurately predicts the SGS stress tensor with $\Delta \neq \Delta'_0$. Thus, it is significant to include the filter width as input feature in the training of ANN model.

4.3. Spatial structures of the SGS stress tensor

To intuitively compare the SGS stress tensors obtained from the ANN model and DNS data, we plot the contours of the components of the SGS stress tensor. Fig. 12 shows the three-dimensional contours of a diagonal component τ_{xx} and an off-diagonal component τ_{xy} computed respectively from the DNS data, the ANN model, the gradient model and the Smagorinsky model at $Re_\lambda = 205.51$ and $k_c/k_{max} = 0.125$. It is obvious that the ANN model and the gradient model accurately reconstructs the spatial structures of the SGS stress tensor, the Smagorinsky model has little correlation with the real SGS stress from the DNS data [6,46]. Besides, the ANN model is more accurate than the gradient model. The latter underesti-

mates the magnitude of the stress tensor, leading to a lower transfer rate of turbulent kinetic energy.

Fig. 13 shows the joint probability density functions (PDFs) of two components τ_{xx} and τ_{xy} of the SGS stress tensor from the ANN model and the real SGS stress tensor from the DNS data. The root mean square of the components of the SGS stress tensor computed by the DNS data is used to normalize the components,

$$\bar{\tau}_{ij,DNS} = \sqrt{\frac{1}{N_g} \sum_{n=1}^{N_g} \tau_{ij,DNS}^2} \quad (i, j = 1, 2, 3). \quad (27)$$

For isotropic turbulent flow, the joint PDFs of the three diagonal components are almost identical. Additionally, the joint PDFs of the three off-diagonal components are very similar. If the correlation coefficients between the SGS stress tensor from the ANN model and that from the DNS data approach 1.0, the main axis of the iso-value lines of the joint PDFs is a line with unit slope, as shown by the dash-dotted line in Fig. 13.

The correlation coefficients of the six components are larger than 0.9. The iso-values of the joint PDFs are located near the line with unit slope, indicating that the ANN model accurately predicts the spatial structures of the SGS stress tensor.

4.4. A posteriori validation of the ANN model

In this subsection, we couple the proposed ANN model with a real LES by introducing the SGS stress tensor predicted by the ANN into the filtered Navier-Stokes equation shown in Eq. (5). We compute and compare the Eulerian energy spectra and Lagrangian statistics of fluid particle pairs from LES with the ANN model and LES with the conventional spectral eddy viscosity model (refer to our previous work [48,49]), the Smagorinsky model, the dynamic Smagorinsky model and the FDNS. As proposed by Germano et al. [9] and Lilly [58], the dynamic Smagorinsky model defines two filters: a grid filter with a width of Δ and a test filter with a width of $\tilde{\Delta}$ (here we use $\tilde{\Delta} = 2\Delta$). The modeled SGS stress is given by $\tau_{ij} = -2C_D \Delta^2 |\tilde{S}| \tilde{S}_{ij}$, where the Smagorinsky coefficient $C_D = \frac{1}{2} (L_{ij} M_{ij} / M_{ij}^2)$ is determined dynamically using the least squares approach, $L_{ij} = \widetilde{\widetilde{u}_i u_j} - \widetilde{\widetilde{u}_i} \widetilde{\widetilde{u}_j}$, $M_{ij} = \widetilde{\widetilde{S}^2} |\tilde{S}| \tilde{S}_{ij} - \Delta^2 \widetilde{\widetilde{S}^2} |\tilde{S}| \tilde{S}_{ij}$.

We define the mean value and standard deviation of the energy transfer rate as

$$\mu_{DNS} = \frac{1}{N_g} \sum_{n=1}^{N_g} D_{\tau,DNS}, \quad (28)$$

$$\sigma_D = \sqrt{\frac{1}{N_g} \sum_{n=1}^{N_g} (D_{\tau,DNS} - \mu_{DNS})^2}. \quad (29)$$

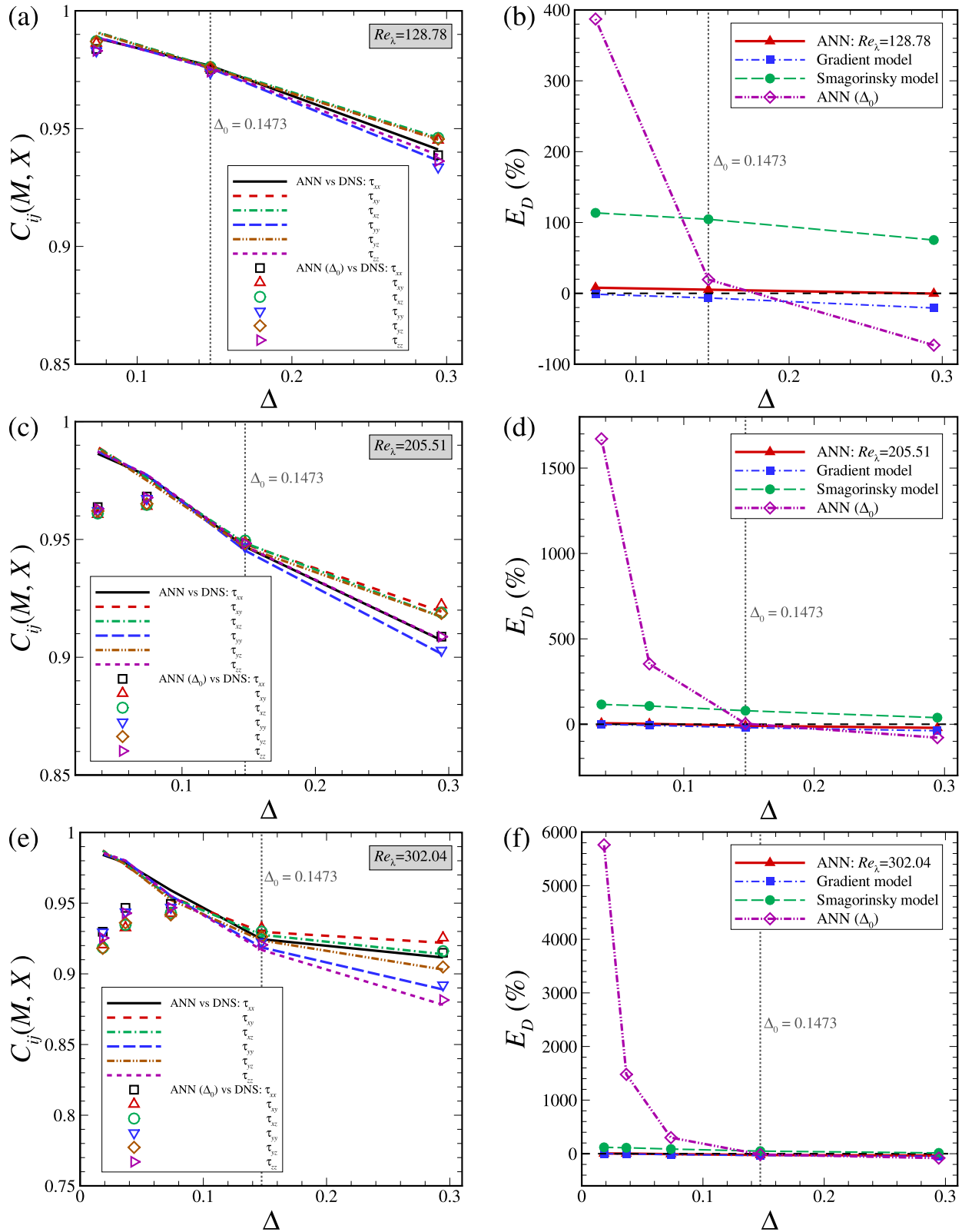


Fig. 11. Comparison of the *a priori* results by the present ANN model and the ANN (Δ_0) model at $Re_\lambda = 128.78$ [(a), (b)], 205.51 [(c), (d)] and 302.04 [(e), (f)].

Fig. 14 shows the PDFs of the energy transfer rate obtained from the FDNS and the ANN model. There is a negative rate of energy transfer in the left part of the figure for both the FDNS data and the ANN model. Since we define a turbulent energy trans-

fer rate from the resolved scales to the subgrid scales as positive, a negative energy transfer rate indicates energy backward scatter from the subgrid scales to the resolved scales. It is well known that this backward scatter of energy is a physical property, but it may lead to numerical instability of the SGS models.

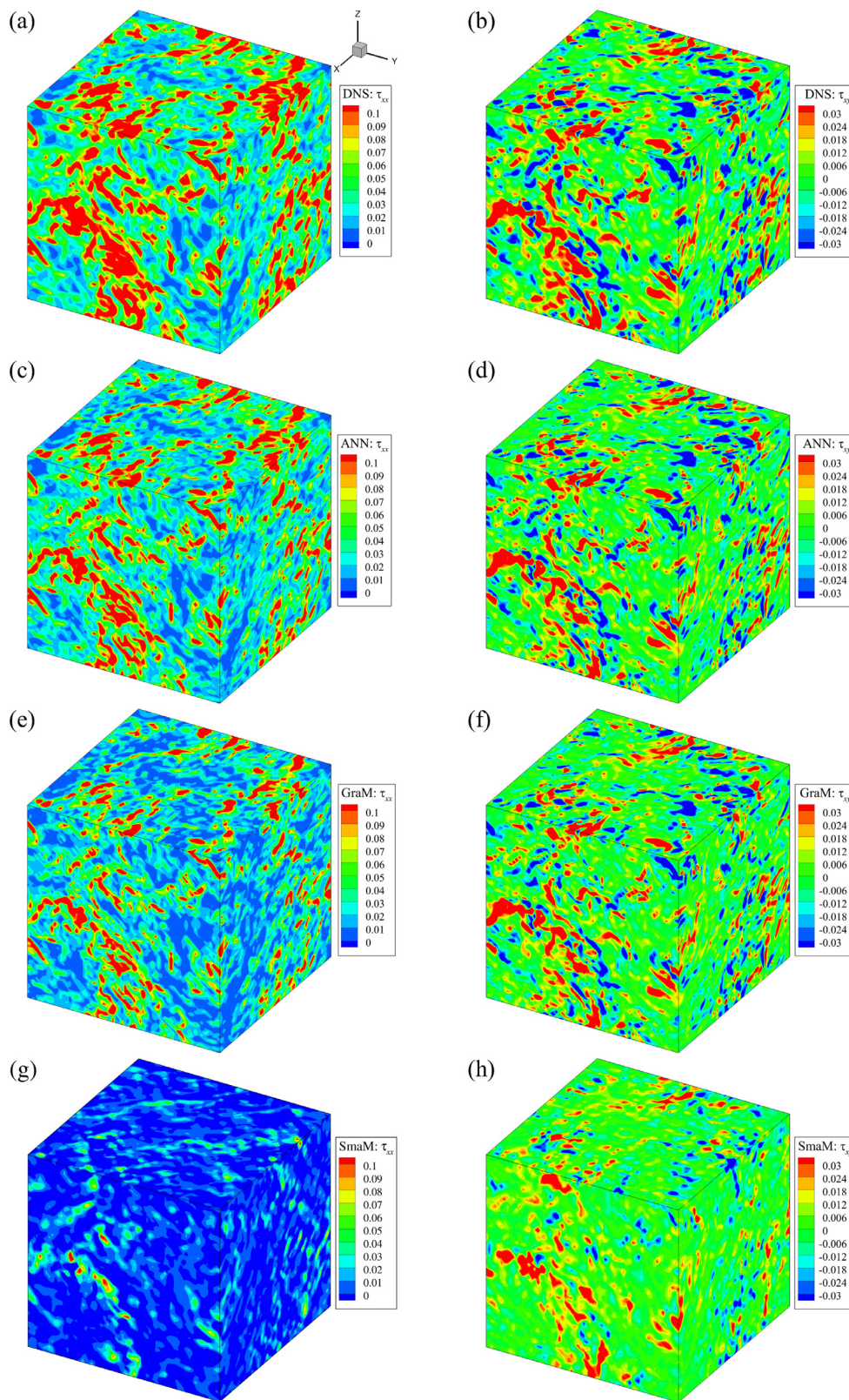


Fig. 12. Three-dimensional contours of the SGS stress components τ_{xx} and τ_{xy} obtained from the DNS data [(a), (b)], the ANN model [(c), (d)], the gradient model [(e), (f)] and the Smagorinsky model [(g), (h)] at $Re_\lambda = 205.51$ and $k_c/k_{max} = 0.125$.

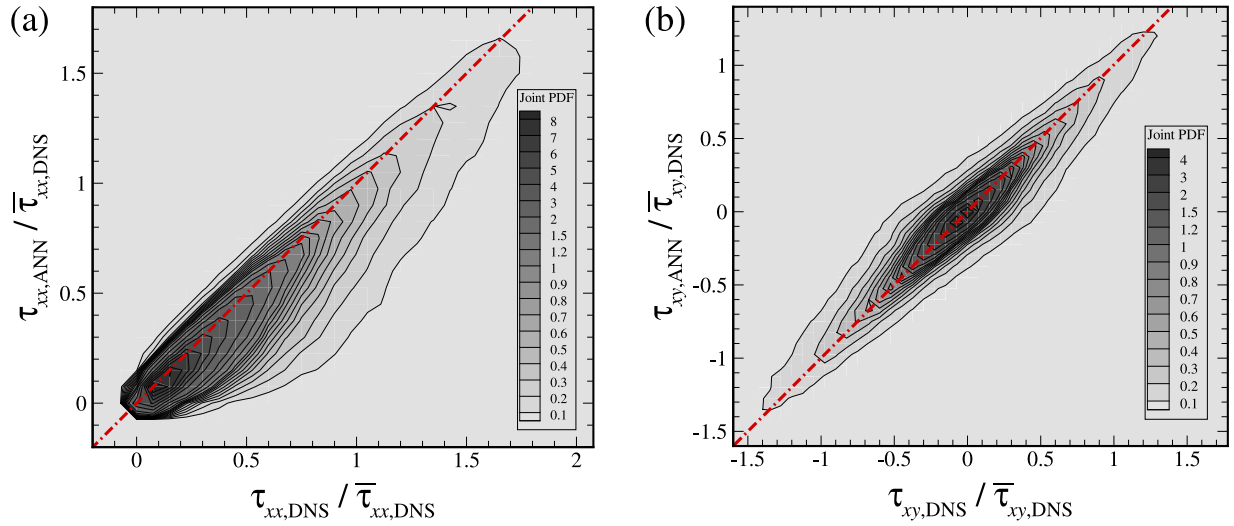


Fig. 13. The joint PDFs of the SGS stress components τ_{xx} and τ_{xy} obtained from the ANN model and the DNS data at $Re_\lambda = 205.51$ and $k_c/k_{max} = 0.125$. The slope of the dash-dotted line is 1.

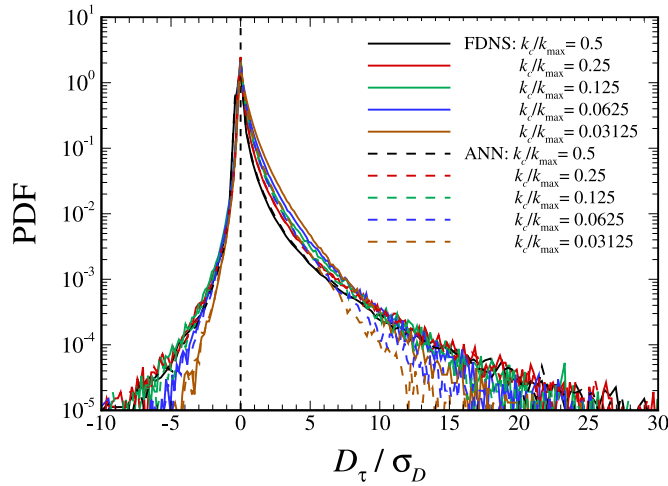


Fig. 14. The PDFs of the local energy transfer rate determined by the ANN model and the FDNS at different cutoff wavenumbers at $Re_\lambda = 302.04$.

To ensure numerical stability, we use a clipping procedure to fix all backward scatter energy as zero [59]. The ANN model is simply improved with

$$\tau_{ij,ANN} = \begin{cases} \tau_{ij,ANN}, & \text{if } D_{\tau,ANN} \geq 0, \\ 0, & \text{if } D_{\tau,ANN} < 0, \end{cases} \quad (30)$$

where $D_{\tau,ANN}$ is the local energy transfer rate at a grid point in the physical flow domain. Thus, the backward scatter of energy to the resolved scales is eliminated and the improved ANN model can only dissipate energy toward the subgrid scales.

The energy spectrum of the Fourier mode is defined as

$$E(\mathbf{k}, t) = \frac{1}{2} \langle \hat{\mathbf{u}}^*(\mathbf{k}, t) \cdot \hat{\mathbf{u}}(\mathbf{k}, t) \rangle, \quad (31)$$

where $\hat{\mathbf{u}}^*(-\mathbf{k}, t) = \hat{\mathbf{u}}(\mathbf{k}, t)$ denotes the conjugate symmetry of complex vector.

Fig. 15 compares the energy spectra obtained from LES coupled with different SGS models and DNS. The Smagorinsky model predicts excessive dissipation, and the resulting energy spectrum is lower than those obtained from the spectral eddy viscosity model, the dynamic Smagorinsky model and DNS, especially near the cutoff wavenumber. Thanks to the fitting of C_S to Kolmogorov con-

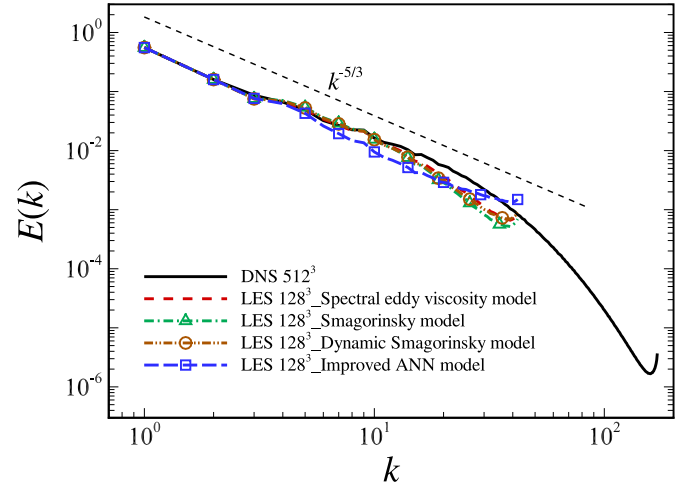


Fig. 15. Energy spectra of the isotropic turbulent flow fields obtained from LES coupled with the improved ANN model, spectral eddy viscosity model, Smagorinsky model and dynamic Smagorinsky model.

stant, the deviation of energy spectrum between the Smagorinsky model and DNS is small. The energy spectrum obtained from the improved ANN model exhibits slight underprediction in the inertial subrange.

To further evaluate the performance of the ANN model, we calculate the Lagrangian statistics of fluid particle pairs from LES with the improved ANN model and compare the results with those from the FDNS and LES with the spectral eddy viscosity model, Smagorinsky model and the dynamic Smagorinsky model. For a particle pair, the separation distance and its variance are defined as

$$R(\mathbf{r}, t_0 | \tau) = \sqrt{\mathbf{R}(\mathbf{r}, t_0 | \tau) \cdot \mathbf{R}(\mathbf{r}, t_0 | \tau)}, \quad (32)$$

$$\mathbf{R}(\mathbf{r}, t_0 | \tau) = \mathbf{X}(\mathbf{x}_0, t_0 | t_0 + \tau) - \mathbf{X}(\mathbf{x}_0 + \mathbf{r}, t_0 | t_0 + \tau), \quad (33)$$

$$\begin{aligned} \sigma_2^2(r, \tau) &= \langle [R(\mathbf{r}, t_0 | \tau) - \langle R(\mathbf{r}, t_0 | \tau) \rangle]^2 \rangle \\ &= \langle \mathbf{R}(\mathbf{r}, t_0 | \tau) \cdot \mathbf{R}(\mathbf{r}, t_0 | \tau) \rangle - \langle R(\mathbf{r}, t_0 | \tau) \rangle^2. \end{aligned} \quad (34)$$

where $\mathbf{R}(\mathbf{r}, t_0 | \tau)$ denotes the separation vector of the particle pair, \mathbf{r} is the initial separation vector, $r = |\mathbf{r}|$ is the initially prescribed

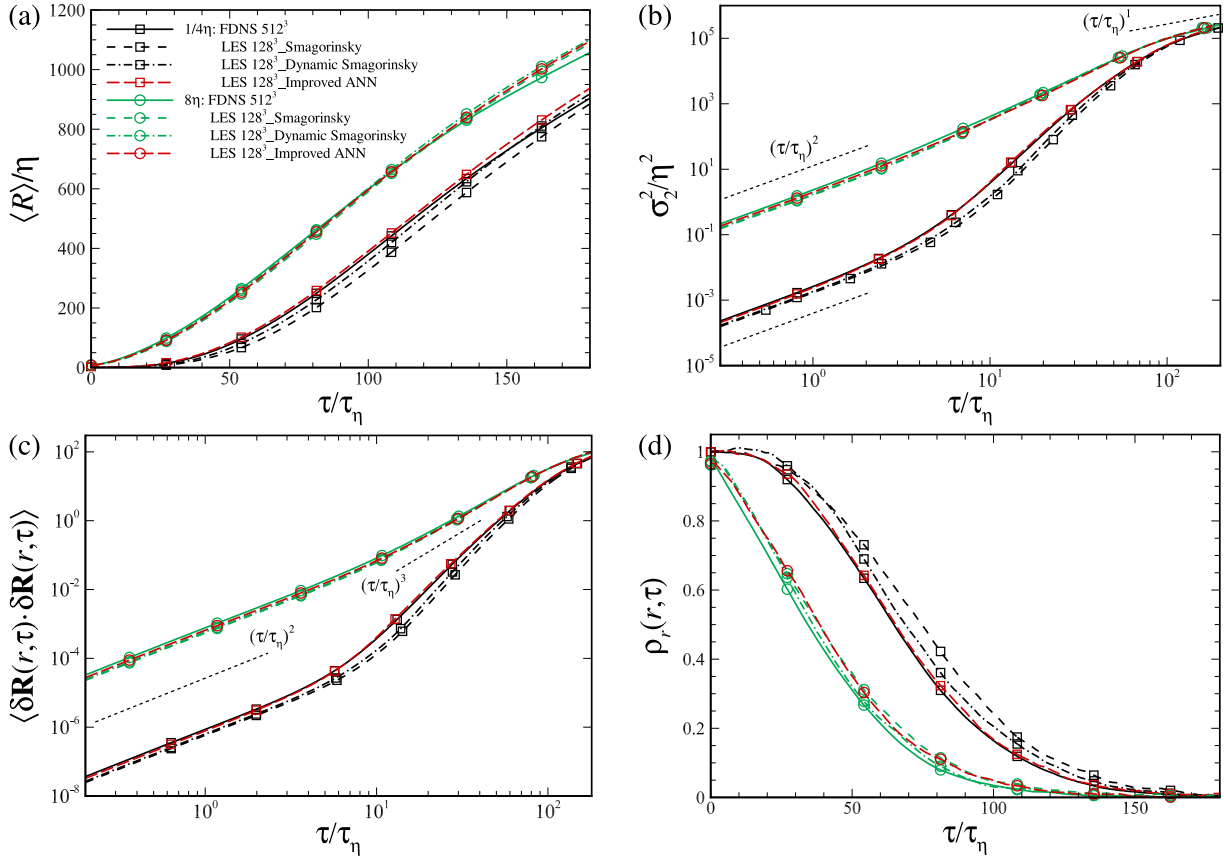


Fig. 16. Lagrangian statistics of fluid particle pairs in LES coupled with the improved ANN model, Smagorinsky model, dynamic Smagorinsky model and FDNS: (a) Mean of the separation distance; (b) Variance of the separation distance; (c) Relative dispersion; (d) One-time two-point Lagrangian velocity correlation function.

separation distance, τ is the time lag, and $\langle \rangle$ denotes the average condition of the particle pairs. The relative dispersion is defined as $\langle \delta \mathbf{R}(r, \tau) \cdot \delta \mathbf{R}(r, \tau) \rangle$, where $\delta \mathbf{R}(r, \tau) = \mathbf{R}(\mathbf{x}_0, t_0 | \tau) - \mathbf{r}$ is the separation vector increment. The one-time two-point Lagrangian velocity correlation function is defined as

$$\rho_r(r, \tau) = \frac{\langle \mathbf{V}(\mathbf{x}_0, t_0 | t_0 + \tau) \cdot \mathbf{V}(\mathbf{x}_0 + \mathbf{r}, t_0 | t_0 + \tau) \rangle}{\langle \mathbf{V}(\mathbf{x}_0, t_0) \cdot \mathbf{V}(\mathbf{x}_0, t_0) \rangle}, \quad (35)$$

where $\mathbf{V}(\mathbf{x}_0, t_0)$ denotes the Lagrangian velocity of the fluid particle located at \mathbf{x}_0 at the initial time t_0 and $\mathbf{V}(\mathbf{x}_0 + \mathbf{r}, t_0 | t_0 + \tau)$ denotes the Lagrangian velocity of that fluid particle at time $t_0 + \tau$.

Fig. 16 shows the mean and variance of the separation distance, the relative dispersion and the one-time two-point Lagrangian velocity correlation function of fluid particle pairs when $r = 1/4\eta$ and 8η ; the results from FDNS are obtained by using a sharp spectral filter, and the cutoff wavenumber is $k_c/k_{\max} = 0.25$. The FDNS can be regarded as an ideal LES with a grid resolution of 128^3 . Due to the SGS model error, it is difficult for the LES to accurately predict the Lagrangian statistics of fluid particle pairs. For the Smagorinsky model, the excessive dissipation reduces the velocity fluctuations at the resolved scales, causing the slower separation of fluid particle pairs and more correlated flow fields than FDNS. The dynamic Smagorinsky model determines the model coefficient dynamically, which reduces the excessive dissipation and gets better results. The results by the spectral eddy viscosity model coincide with the dynamic Smagorinsky model, which are not shown in the figure (see Ref. [48]). The Lagrangian statistics obtained from the improved ANN model almost approach those from FDNS, better than the results from the Smagorinsky model and dynamic Smagorinsky model. These observations imply that the improved ANN model can be used to predict the Lagrangian statistics.

5. Conclusion

We established the relation between the resolved-scale velocity gradient tensor and the SGS stress tensor in forcing isotropic turbulent flows in the framework of an ANN. The prediction accuracy and generalization of the trained ANN model were investigated by comparing the predicted SGS stress tensor with the DNS data, the gradient model and the Smagorinsky model. *A posteriori* validation was carried out by applying the ANN model to a real LES of isotropic turbulent flows.

We first obtained the resolved-scale flow field and SGS stress tensor from DNS data by performing filtering operations. A Gaussian filter with different filter widths was used in the FDNS. The nine components of the velocity gradient tensor together with the filter width were used as the input features and the six components of the SGS stress tensor were used as the output labels. Using the DNS data with each three filter widths at $Re_\lambda = 128.78$ and 302.04 , a single-hidden-layer feedforward ANN was applied to train the new SGS model for LES of isotropic turbulent flows.

After training the ANN model, the velocity gradient tensors and filter widths at different Reynolds numbers, filter widths and time steps in Table 3 were used as new inputs, and the predicted SGS stress tensors were used for the *a priori* test of the ANN model. The correlation coefficient and the relative error of the energy transfer rate between the SGS stress tensors from the ANN model and DNS data were used as two metrics for evaluating the performance of ANN model. The correlation coefficients determined by the ANN model were mostly larger than 0.9, much better than the Smagorinsky model. The relative errors of the energy transfer rate predicted by the ANN model exhibited significant improvement over those computed by the gradient model. Subsequently,

two more ANN models were trained to compare with the present ANN model, demonstrating the importance of filter width in the input features. In addition, we plotted the three-dimensional contours and joint PDFs of the SGS stress components, and these plots intuitively demonstrate that the ANN model accurately predicts the spatial structures of the SGS stress tensor.

We then performed *a posteriori* validation by coupling the improved ANN model with an LES of isotropic turbulent flows. The energy spectrum is computed by the improved ANN model and then compared with several SGS models. The Lagrangian statistics of fluid particle pairs, such as the mean and variance of the separation distance, relative dispersion and Lagrangian velocity correlations, obtained from the improved ANN model nearly approached the results from the FDNS and better than those from the Smagorinsky model and dynamic Smagorinsky model. We will study this ANN model in turbulent flows in more complex geometries in future work.

Acknowledgments

This work was supported by the [National Natural Science Foundation of China](#) (Nos. 11988102, 11772337 and 11572331), the [Science Challenge Program](#) (No. TZ2016001), the Strategic Priority Research Program, CAS (No. XDB22040104), and the Key Research Program of Frontier Sciences, CAS (No. QYZDJ-SSW-SYS002).

References

- [1] Lesieur M, Métais O. New trends in large-eddy simulations of turbulence. *Annu Rev Fluid Mech* 1996;28:45–82. doi:10.1146/annurev.fl.28.010196.000401.
- [2] Meneveau C, Katz J. Scale-invariance and turbulence models for large-eddy simulation. *Annu Rev Fluid Mech* 2000;32:1–32. doi:10.1146/annurev.fluid.32.1.1.
- [3] Smagorinsky J. General circulation experiments with the primitive equations: I. The basic experiment. *Mon Weather Rev* 1963;91:99–164. doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.
- [4] Bardina J, Ferziger JH, Reynolds WC. Improved subgrid scale models for large eddy simulation. In: *AIAA 13th fluid & plasma dynamics conference*, AIAA–80–1357; 1980. <https://doi.org/10.2514/6.1980-1357>.
- [5] Bardina J, Ferziger JH, Reynolds WC. Improved turbulence models based on large eddy simulation of homogeneous, incompressible, turbulent flows. Report No TF-19, Department of Mechanical Engineering, Stanford University; 1983.
- [6] Clark RA, Ferziger JH, Reynolds WC. Evaluation of subgrid-scale models using an accurately simulated turbulent flow. *J Fluid Mech* 1979;91:1–16. doi:10.1017/S002211207900001X.
- [7] Lund TS, Novikov EA. Parameterization of subgrid-scale stress by the velocity gradient tensor. In: *Annual research briefs (Center for Turbulence Research, Stanford University and NASA)*; 1992. p. 27–43.
- [8] Zang Y, Street RL, Koseff JR. A dynamic mixed subgrid-scale model and its application to turbulent recirculating flows. *Phys Fluids A* 1993;5:3186. doi:10.1063/1.858675.
- [9] Germano M, Piomelli U, Moin P, Cabot WH. A dynamic subgrid-scale eddy viscosity model. *Phys Fluids A* 1991;3:1760. doi:10.1063/1.857955.
- [10] Germano M. Turbulence: the filtering approach. *J Fluid Mech* 1992;238:325–36. doi:10.1017/S0022112092001733.
- [11] Vreman B, Geurts B, Kuerten H. Large eddy simulation of the temporal mixing layer using the Clark model. *Theor Comp Fluid Dyn* 1996;8:309–24.
- [12] Morinishi Y, Vasilyev OV. A recommended modification to the dynamic two-parameter mixed subgrid scale model for large eddy simulation of wall bounded turbulent flows. *Phys Fluids* 2001;13:3400. doi:10.1063/1.1404396.
- [13] Yang ZX, Cui GX, Zhang ZS, Xu CX. A modified nonlinear sub-grid scale model for large eddy simulation with application to rotating turbulent channel flows. *Phys Fluids* 2012;24:075113. doi:10.1063/1.4739063.
- [14] Yang ZX, Cui GX, Xu CX, Zhang ZS. Large eddy simulation of rotating turbulent channel flow with a new dynamic global-coefficient nonlinear subgrid stress model. *J Turbul* 2012;13:N48. doi:10.1080/14685248.2012.726996.
- [15] Ghaisas NS, Frankel SH. Dynamic gradient models for the sub-grid scale stress tensor and scalar flux vector in large eddy simulation. *J Turbul* 2015;17:30–50. doi:10.1080/14685248.2015.1083106.
- [16] Yang ZX, Wang B-C. On the topology of the eigenframe of the subgrid-scale stress tensor. *J Fluid Mech* 2016;798:598–627. doi:10.1017/jfm.2016.336.
- [17] Lu H, Porté-Agel F. A modulated gradient model for large-eddy simulation: application to a neutral atmospheric boundary layer. *Phys Fluids* 2010;22:015109. doi:10.1063/1.3291073.
- [18] Vollant A, Balarac G, Corre C. A dynamic regularized gradient model of the subgrid-scale stress tensor for large-eddy simulation. *Phys Fluids* 2016;28:025114. doi:10.1063/1.4941781.
- [19] Hinton G, Deng L, Yu D, Dahl GE, Mohamed A-R, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Process Mag* 2012;29:82–97. doi:10.1109/MSP.2012.2205597.
- [20] Krizhevsky A, Sutskever I, Hinton G. Imagenet classification with deep convolutional neural networks. *Proc Adv Neural Inf Process Syst* 2012;25:1090–8. doi:10.1145/3065386.
- [21] Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa P. Natural language processing (almost) from scratch. *J Mach Learn Res* 2011;12:2493–537.
- [22] Hadsell R, Sermanet P, Ben J, Erkan A, Scoffier M, Kavukcuoglu K, et al. Learning long-range vision for autonomous off-road driving. *J Field Robot* 2009;26:120–44. doi:10.1002/rob.20276.
- [23] Ma M, Lu J, Tryggvason G. Using statistical learning to close two-fluid multiphase flow equations for a simple bubbly system. *Phys Fluids* 2015;27:092101. doi:10.1063/1.4930004.
- [24] Tracey B, Duraisamy K, Alonso JJ. A machine learning strategy to assist turbulence model development. In: *53rd AIAA aerospace science meeting*; 2015. p. p.1287. doi:10.2514/6.2015-1287.
- [25] Parish EJ, Duraisamy K. A paradigm for data-driven predictive modeling using field inversion and machine learning. *J Comput Phys* 2016;305:758–74. doi:10.1016/j.jcp.2015.11.012.
- [26] Kutz JN. Deep learning in fluid dynamics. *J Fluid Mech* 2017;814:1–4. doi:10.1017/jfm.2016.803.
- [27] Duraisamy K, Iaccarino G, Xiao H. Turbulence modeling in the age of data. *Annu Rev Fluid Mech* 2019;51:357–77. doi:10.1146/annurev-fluid-010518-040547.
- [28] Brunton SL, Noack BR, Koumoutsakos P. Machine learning for fluid mechanics. *arXiv:190511075v1* 2019;.
- [29] Duraisamy K, Zhang ZJ, Singh AP. New approaches in turbulence and transition modeling using data-driven techniques. In: *53rd AIAA aerospace science meeting*; 2015. p. p.1284. doi:10.2514/6.2015-1284.
- [30] Ling J, Kurzawski A, Templeton J. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *J Fluid Mech* 2016;807:155–66. doi:10.1017/jfm.2016.615.
- [31] Ling J, Jones R, Templeton J. Machine learning strategies for systems with invariance properties. *J Comput Phys* 2016;318:22–35. doi:10.1016/j.jcp.2016.05.003.
- [32] Wang J-X, Wu J-L, Xiao H. Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data. *Phys Rev Fluids* 2017;2:034603. doi:10.1103/PhysRevFluids.2.034603.
- [33] Wu J-L, Xiao H, Paterson E. Physics-informed machine learning approach for augmenting turbulence models: a comprehensive framework. *Phys Rev Fluids* 2018;3:074602. doi:10.1103/PhysRevFluids.3.074602.
- [34] Tracey B, Duraisamy K, Alonso JJ. Application of supervised learning to quantify uncertainties in turbulence and combustion modeling. In: *51st AIAA aerospace science meeting*; 2013. p. p.0259. doi:10.2514/6.2013-259.
- [35] Singh AP, Duraisamy K. Using field inversion to quantify functional errors in turbulence closures. *Phys Fluids* 2016;28:045110. doi:10.1063/1.4947045.
- [36] Xiao H, Wu J-L, Wang J-X, Sun R, Roy CJ. Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier-Stokes simulations: a data-driven, physics-informed Bayesian approach. *J Comput Phys* 2016;324:115–36. doi:10.1016/j.jcp.2016.07.038.
- [37] Wu J-L, Sun R, Laizet S, Xiao H. Representation of stress tensor perturbations with application in machine-learning-assisted turbulence modeling. *Comput Methods Appl Mech Engrg* 2018;Inpress. doi:10.1016/j.cma.2018.09.010.
- [38] Ling J, Templeton J. Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty. *Phys Fluids* 2015;27:085103. doi:10.1063/1.4927765.
- [39] Wu J-L, Wang J-X, Xiao H. A Bayesian calibration-prediction method for reducing model-form uncertainties with application in RANS simulations. *Flow Turbul Combust* 2016;97:761–86. doi:10.1007/s10494-016-9725-6.
- [40] Wu J-L, Wang J-X, Xiao H, Ling J. A priori assessment of prediction confidence for data-driven turbulence modeling. *Flow Turbul Combust* 2017;99:25–46. doi:10.1007/s10494-017-9807-0.
- [41] Sarghini F, de Felice G, Santini S. Neural networks based subgrid scale modeling in large eddy simulations. *Comput Fluids* 2003;32:97–108. doi:10.1016/S0045-7930(01)00098-6.
- [42] Vollant A, Balarac G, Corre C. Subgrid-scale scalar flux modelling based on optimal estimation theory and machine-learning procedures. *J Turbul* 2017;18:854–78. doi:10.1080/14685248.2017.1334907.
- [43] Maulik R, San O. A neural network approach for the blind deconvolution of turbulent flows. *J Fluid Mech* 2017;831:151–81. doi:10.1017/jfm.2017.637.
- [44] Maulik R, San O, Rasheed A, Vedula P. Data-driven deconvolution for large eddy simulations of Kraichnan turbulence. *Phys Fluids* 2018;30:125109. doi:10.1063/1.5079582.
- [45] Gamahara M, Hattori Y. Searching for turbulence models by artificial neural network. *Phys Rev Fluids* 2017;2:054604. doi:10.1103/PhysRevFluids.2.054604.
- [46] Wang Z, Luo K, Li D, Tan JH, Fan JR. Investigations of data-driven closure for subgrid-scale stress in large-eddy simulation. *Phys Fluids* 2018;30:125101. doi:10.1063/1.5054835.
- [47] Jin GD, He G-W, Wang L-P. Large-eddy simulation of turbulent collision of heavy particles in isotropic turbulence. *Phys Fluids* 2010;22:055106. doi:10.1063/1.3425627.
- [48] Zhou ZD, Chen JC, Jin GD. Prediction of lagrangian dispersion of fluid particles in isotropic turbulent flows using large-eddy simulation method. *Acta Mech* 2017;228:3203–22. doi:10.1007/s00707-017-1877-5.
- [49] Zhou ZD, Wang SZ, Jin GD. A structural subgrid-scale model for relative dis-

- persion in large-eddy simulation of isotropic turbulent flows by coupling kinematic simulation with approximate deconvolution method. *Phys Fluids* 2018;30:105110. doi:[10.1063/1.5049731](https://doi.org/10.1063/1.5049731).
- [50] Yeung PK, Pope SB. An algorithm for tracking fluid particles in numerical simulations of homogeneous turbulence. *J Comput Phys* 1988;79:373–416. doi:[10.1016/0021-9991\(88\)90022-8](https://doi.org/10.1016/0021-9991(88)90022-8).
- [51] Eswaran V, Pope SB. An examination of forcing in direct numerical simulations of turbulence. *Comput Fluids* 1988;16:257–78. doi:[10.1016/0045-7930\(88\)90013-8](https://doi.org/10.1016/0045-7930(88)90013-8).
- [52] Yeung PK, Pope SB. Lagrangian statistics from direct numerical simulations of isotropic turbulence. *J Fluid Mech* 1989;207:531–86. doi:[10.1017/S0022112089002697](https://doi.org/10.1017/S0022112089002697).
- [53] Pope SB. *Turbulent flows*. Cambridge: Cambridge University Press; 2000. doi:[10.1017/CBO9780511840531](https://doi.org/10.1017/CBO9780511840531).
- [54] Liu S, Meneveau C, Courville A. On the properties of similarity subgrid-scale models as deduced from measurements in a turbulent jet. *J Fluid Mech* 1994;275:83–119. doi:[10.1017/S0022112094002296](https://doi.org/10.1017/S0022112094002296).
- [55] Goodfellow I, Bengio Y, Courville A. *Deep learning*. Cambridge, MA: MIT Press; 2016.
- [56] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986;323:533–6. doi:[10.1038/323533a0](https://doi.org/10.1038/323533a0).
- [57] Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. TensorFlow: a system for large-scale machine learning. In: *12th USENIX symposium on operating systems design and implementation*, 16; 2016. p. 265–83.
- [58] Lilly DK. A proposed modification of the Germano subgrid-scale closure method. *Phys Fluids A* 1992;4:633. doi:[10.1063/1.858280](https://doi.org/10.1063/1.858280).
- [59] Fang L, Bos WJ, Shao L, Bertoglio J-P. Time reversibility of Navier-Stokes turbulence and its implication for subgrid scale models. *J Turbul* 2012;13:N3. doi:[10.1080/14685248.2011.639777](https://doi.org/10.1080/14685248.2011.639777).