


Artificial neural network-based spatial gradient models for large-eddy simulation of turbulence

Cite as: AIP Advances **11**, 055216 (2021); <https://doi.org/10.1063/5.0053590>

Submitted: 09 April 2021 • Accepted: 28 April 2021 • Published Online: 12 May 2021

 Yunpeng Wang (王云朋), Zelong Yuan (袁泽龙),  Chenyue Xie (谢晨月), et al.



View Online



Export Citation



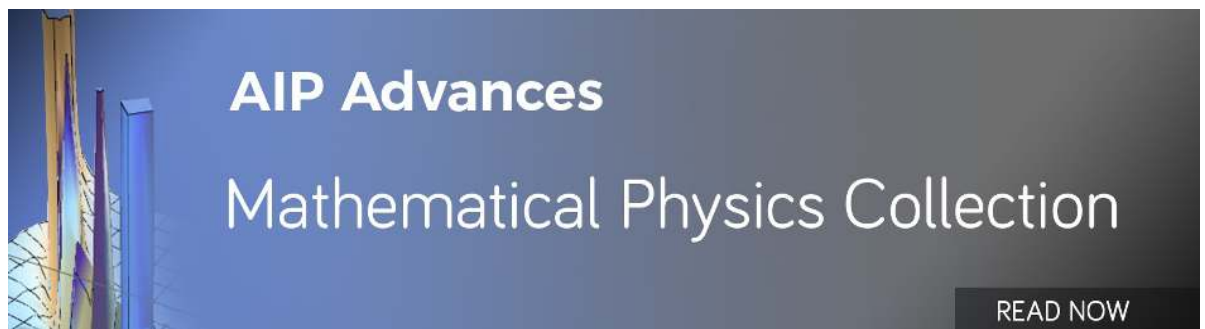
CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Deconvolutional artificial neural network models for large eddy simulation of turbulence](#)
Physics of Fluids **32**, 115106 (2020); <https://doi.org/10.1063/5.0027146>

[Artificial neural network-based nonlinear algebraic models for large eddy simulation of turbulence](#)
Physics of Fluids **32**, 115101 (2020); <https://doi.org/10.1063/5.0025138>

[A dynamic spatial gradient model for the subgrid closure in large-eddy simulation of turbulence](#)
Physics of Fluids **33**, 075119 (2021); <https://doi.org/10.1063/5.0058542>



AIP Advances
Mathematical Physics Collection

READ NOW

Artificial neural network-based spatial gradient models for large-eddy simulation of turbulence

Cite as: AIP Advances 11, 055216 (2021); doi: 10.1063/5.0053590

Submitted: 9 April 2021 • Accepted: 28 April 2021 •

Published Online: 12 May 2021



View Online



Export Citation



CrossMark

Yunpeng Wang (王云朋),^{1,2}  Zelong Yuan (袁泽龙),^{1,2} Chenyue Xie (谢晨月),^{1,2} 
and Jianchun Wang (王建春)^{1,2,a)} 

AFFILIATIONS

¹Guangdong Provincial Key Laboratory of Turbulence Research and Applications, Center for Complex Flows and Soft Matter Research, Department of Mechanics and Aerospace Engineering, Southern University of Science and Technology, Shenzhen 518055, People's Republic of China

²Guangdong-Hong Kong-Macao Joint Laboratory for Data-Driven Fluid Mechanics and Engineering Applications, Southern University of Science and Technology, Shenzhen 518055, People's Republic of China

^{a)} Author to whom correspondence should be addressed: wangjc@sustech.edu.cn

ABSTRACT

The subgrid-scale stress (SGS) of large-eddy simulation (LES) is modeled by artificial neural network-based spatial gradient models (ANN-SGMs). The velocity gradients at neighboring stencil locations are incorporated to improve the accuracy of the SGS stress. The consideration of the gradient terms in the stencil locations is in a semi-explicit form so that the deployed artificial neural network (ANN) can be considerably simplified. This leads to a much higher LES efficiency compared with previous “black-box” models while still retaining the level of accuracy in the *a priori* test. The correlation coefficients of the ANN-SGMs can be larger than 0.98 for the filter width in the inertial range. With the current formulation, the significances of the individual modeling terms are transparent, giving clear guidance to the potential condensation of the model, which further improves the LES efficiency. The computational cost of the current ANN-SGM method is found to be two orders lower than previous “black-box” models. In the *a posteriori* test, the ANN-SGM framework predicts more accurately the flow field compared with the traditional LES models. Both the flow statistics and the instantaneous field are accurately recovered. Finally, we show that the current model can be adapted to different filter widths with sufficient accuracy. These results demonstrate the advantage and great potential of the ANN-SGM framework as an attractive solution to the closure problem in large-eddy simulation of turbulence.

© 2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0053590>

I. INTRODUCTION

Accurate prediction of turbulent flow is crucial in both the research and engineering communities as most industrial flows turn out to be turbulent. As is well known, characterizing turbulence in complete detail by direct numerical simulation (DNS) is often impractical at high Reynolds number due to the large range of spatial and temporal scales to be resolved. Hence, averaged solutions become the major alternatives, mainly including the Reynolds-averaged Navier–Stokes (RANS) method and the large-eddy simulation (LES).¹ The RANS method only calculates the ensemble average of the flow properties, leaving the unclosed Reynolds stress modeled either empirically or through theoretical arguments.² In contrast, LES, which was originally pioneered for meteorological applications,^{3–5} directly represents a portion of the large-scale motions. The

small-scale dynamics in LES is modeled so that the computational cost is kept at manageable levels.⁶

Similar to the modeling of the Reynolds stress in RANS, the reconstruction of the subgrid-scale (SGS) stress is a central problem in the LES community.^{7–14} Extensive SGS models exist in the literature, among which the Smagorinsky model (SM),^{3,4} the dynamic Smagorinsky model (DSM),^{8,9,15,16} the scale-similarity model (SSM),^{17,18} the dynamic mixed model (DMM),^{19–22} the velocity gradient model (VGM),²³ Implicit-LES (ILES),^{24–28} higher-order gradient models,²⁹ nonlinear-viscosity-based LES,³⁰ non-local eddy viscosity approach,³¹ and many others^{32–40} constitute the most important ones.

Recently, increasing attention is attracted to machine-learning-based SGS models due to the development of modern computers and the availability of the large amounts of experimental and DNS

data. The earliest contribution, to our knowledge, is from Sarghini *et al.*⁴¹ who, in their study on channel flow, adopted an artificial neural network (ANN) to predict the Smagorinsky coefficient in Bardina's scale similarity model. Interestingly, what follows the work of Sarghini *et al.*⁴¹ is a long gap of more than a decade with very few works emerging on machine-learning-based turbulence closures until the quite recent studies.^{42–45} Thereafter, the literature has seen an enormous boom in this field.^{46–62}

With artificial neural networks, like that employed by Sarghini *et al.*,⁴¹ most researchers take velocity gradients (shear rates) as inputs, and it turns out to be effective.^{63–65} Gamahara and Hattori⁵⁷ predicted the SGS stress in turbulent channel flow using the artificial neural network without any assumptions on the form of the SGS term. With the inputs constructed using the filtered gradients from DNS data, the predicted SGS stress exhibits a comparable accuracy to the gradient model.²³ Vollant *et al.*⁵⁸ applied a similar artificial neural network in conjunction with an optimal estimation theory to predict the SGS flux of a passive scalar. To account for the spatial influence, Xie *et al.*^{66,67} proposed a series of ANN-based methods incorporating the first- and second-order gradients at neighboring stencil locations to achieve higher correlation coefficients. The reader is also referred to the summary by Duraisamy *et al.*² and the references therein for other recent progress. Despite these continuous efforts in the modeling of the SGS stress, it is important to note that highly accurate SGS stress can be unstable under perturbation.⁶⁸ While this approach may fail for more complex flows, it is sufficient for the current homogeneous isotropic case. It was shown that a hyper-viscosity can be implemented to keep the stability of SGS model without significantly affecting the accuracy in LES of isotropic turbulence.⁶⁷

To the best of our knowledge, most existing works in ANN-based SGS models make no assumptions on the form of the SGS term (i.e., “black-box” type). Consequently, the mathematical structure of the SGS stress is implicitly hidden in the “black-box” of the artificial neural network, rendering it difficult to generalize the model or to disentangle the specific role of the gradient terms. Moreover, the number of required neurons can be enormous when spatial stencil effects are considered,^{66,67,69} and over-fitting can exist when the velocity gradients at stencil locations are directly fed to the artificial neural network (cf. Fig. 2 of Ref. 66). Very recently, Xie *et al.*⁷⁰ proposed an ANN-based nonlinear algebraic model (ANN-NAM) that incorporates the Galilean invariance into the SGS stress. In their “semi-explicit” approach, the invariants of the velocity gradient tensors are taken as the input of the ANN and the nonlinear terms are the tensor bases of the local gradients. We note that spatial locations are not considered in their work and the accuracy of SGS stress predicted by the ANN-NAM framework is very close to that of the velocity gradient model (VGM). This also motivates the current work to examine in detail the influence of the spatial stencil locations.

In view of these limitations, we shall seek a semi-explicit formulation for the SGS stress using the artificial neural network, with particular emphasis on the individual contributions from the neighboring stencil locations. The dimensional consistency and physical constraint of the Galilean invariance give the model more robustness and stability. While accuracy is maintained at a comparable level, the computational efficiency is significantly increased compared with the “black-box” model. The training data are

generated from the filtered DNS (fDNS) data of incompressible isotropic turbulence at a grid resolution of 1024^3 in a cubic box of $(2\pi)^3$ with periodic boundary conditions. The dimensionless invariants of the velocity gradient tensor^{33,70,71} at the local position are taken as the training inputs. Since the velocity gradients are incorporated after the final hidden layer, the artificial neural network is considerably simplified compared with previous works,^{66,67} leading to a much faster LES. Furthermore, with the current method, the influence of each constitutive term can be clearly visualized in contrast to the previous models of “black-box” type. As we shall see, this leads to an even simpler version of the artificial neural network-based spatial gradient models (ANN-SGMs). Previous works^{66,67,69,72–74} show that the flow dynamics at scales between $\Delta/2$ and 2Δ play crucial roles in the transport of kinetic energy at filter width Δ . Hence, the stencil locations in the current work are tacitly taken at these scales.

Detailed comparisons between the present ANN-SGM method and existing SGS models are performed using both *a priori* and *a posteriori* tests. In the *a priori* test, the present models give much larger correlation coefficients than traditional gradient models. As the number of considered neighboring locations increases, the accuracy of the predicted SGS stress goes up. In the *a posteriori* test, we choose the most simplified model as it turns out to be amply adequate based on the *a priori* observation. Both the flow statistics and the instantaneous field are examined, invariably demonstrating the great potential of the present ANN-based SGS model.

The paper is organized as follows: Sec. II presents the governing equations of LES and the detailed structure of the artificial neural network-based spatial gradient models. Section III gives the statistics of the DNS database for incompressible turbulence and the parameterization of the numerical scheme. Sections IV and V highlight the key results in the *a priori* and the *a posteriori* tests, respectively. Section VI summarizes the paper and comments on the future perspectives.

II. GOVERNING EQUATIONS OF LARGE-EDDY SIMULATION AND ARTIFICIAL NEURAL NETWORK METHODS

In the current section, the governing equations for incompressible turbulence and large-eddy simulation are discussed, followed by a detailed description of the artificial neural network-based spatial gradient models (ANN-SGMs). The hyperparameters of the machine-learning process are also introduced.

A. Governing equations of incompressible turbulence and large-eddy simulation

The conservations of mass and momentum for incompressible turbulence are governed by^{1,14,75}

$$\frac{\partial u_i}{\partial x_i} = 0, \quad (1)$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial u_i u_j}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} + \mathcal{F}_i, \quad (2)$$

with u_i being the i th component of velocity, ν being the kinematic viscosity, p being the pressure divided by the constant density, and \mathcal{F} being a large-scale forcing to the fluid momentum. Here, the

summation convention is used. In addition, the Taylor microscale Reynolds number Re_λ is given by⁷⁵

$$Re_\lambda = \frac{u^{rms}\lambda}{\sqrt{3}\nu}, \quad (3)$$

where $u^{rms} = \sqrt{\langle u_i u_i \rangle}$ stands for the root mean square (rms) velocity and $\langle \cdot \rangle$ denotes a spatial average over the whole periodic domain D . The Taylor length scale λ is defined by

$$\lambda = \sqrt{\frac{5\nu}{\varepsilon}} u^{rms}, \quad (4)$$

where ε is the dissipation rate given by $\varepsilon = 2\nu \langle S_{ij} S_{ij} \rangle$ with $S_{ij} = \frac{1}{2}(\partial u_i \partial x_j + \partial u_j \partial x_i)$ being the strain rate tensor. Finally, the Kolmogorov length scale η and the integral length scale L_I are defined, respectively, by^{1,75}

$$\eta = \left(\frac{\nu^3}{\varepsilon} \right)^{1/4}, \quad (5)$$

$$L_I = \frac{3\pi}{2(u^{rms})^2} \int_0^\infty \frac{E(k)}{k} dk, \quad (6)$$

with the energy spectrum $E(k)$ defined as $\int_0^\infty E(k) dk = (u^{rms})^2 2$.

In LES, the governing equations are obtained through a filtering operation,

$$\bar{f}(\mathbf{x}) = \int_D f(\mathbf{x} - \mathbf{r}) G(\mathbf{r}, \mathbf{x}; \Delta) d\mathbf{r}, \quad (7)$$

where G is the filter kernel, Δ is the filter width, and D denotes the overall domain.^{6,10,76} In the current work, filter G is assumed to commute with the spatial derivatives (i.e., homogeneous filter¹). Applying Eq. (7) to Eq. (1) and Eq. (2), we have^{6,7,10,77}

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0, \quad (8)$$

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{\partial \bar{p}}{\partial x_i} - \frac{\partial \tau_{ij}}{\partial x_j} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} + \bar{\mathcal{F}}_i, \quad (9)$$

where the information of the filter is implicitly contained in the unclosed term τ_{ij} , recognized as the subgrid-scale (SGS) stress, defined by

$$\tau_{ij} = \bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j. \quad (10)$$

The central problem in LES boils down to the representation of the SGS term using the filtered (resolved) variables. Subsequently, the filtered equations can be solved just like the original Navier–Stokes equations using any suitable numerical algorithm. However, unlike the viscous stress, the constitutive equation for the SGS stress is not obvious since the SGS stress is not a true “stress.” It is formed completely due to nonlinear interactions of the fluctuating field. Consequently, the success of the traditional models (e.g., the Smagorinsky model³), which mimics the linear Newton’s Law for viscous stress, is quite limited. For enhanced predictions of the mean flow, many modifications to the Smagorinsky model were attempted. The dynamic Smagorinsky model (DSM) exploits the Germano identity¹⁵ and reproduces the overall energy exchange

between the resolved and the subgrid scales. The dynamic mixed model (DMM),¹⁸ which combines the scale-similarity model^{17,18} with an eddy viscosity part, yields a higher correlation coefficient. However, these models tend to overpredict the energy spectrum at larger scales while underestimates it at smaller scales due excessive SGS dissipation. These observations suggest an inadequacy in the accuracy of the SGS stress. In other words, these models can have similar net effects on the mean flow like the true SGS stress; however, the prediction of the SGS term per se and other flow properties may not be adequate, begging more accurate SGS models. In view of these limitations, we resort to the machine-learning technique to reconstruct the SGS stress.

B. The structure of the artificial neural network-based spatial gradient models

As discussed, most existing works^{41,57,58,64,65} take the velocity gradients as the inputs of the artificial neural network. Very recently, Xie *et al.*^{66,67,69} further improved the previous models by considering the velocity gradients of different orders at neighboring spatial locations near the filter width (cf. Fig. 2). However, since no forms are assumed for the SGS stress. The constitutive relation is implicitly hidden in the network, rendering it difficult to unravel the exact contributions of the neighboring locations. Considering dimensional consistency, the SGS stress can be written in the most general form as

$$\tau_{ij}^{model}|_{l,m,n} = \sum_{\substack{p_1, p_2, p_3 = -N_{p,q} \\ q_1, q_2, q_3 = 0}}^{N_{p,q}} \sum_{\substack{I_1, I_2, I_3 = 0 \\ J_1, J_2, J_3 = 0}}^{\substack{p_1, p_2, p_3 \\ q_1, q_2, q_3}} g_{I_1, J_1, I_2, J_2, I_3, J_3}^{I_1+I_2+I_3, J_1+J_2+J_3} \times \frac{\partial^{I_1+I_2+I_3} \bar{u}_i}{\partial x_1^{I_1} \partial x_2^{I_2} \partial x_3^{I_3}} \Big|_{n+p_3}^{l+p_1} \frac{\partial^{J_1+J_2+J_3} \bar{u}_j}{\partial x_1^{J_1} \partial x_2^{J_2} \partial x_3^{J_3}} \Big|_{n+q_3}^{m+q_2} \cdot \quad (11)$$

Here, (l, m, n) represent the local grid point and each component of (p_1, p_2, p_3) and (q_1, q_2, q_3) can take the values between $\pm N_{p,q}$ (e.g., 0, $\pm 1, \pm 2, \pm 3 \dots$), which represent different neighboring stencil locations. When only the local gradients are considered, a rigorous form exists using Taylor expansion.²⁹ In the current treatment, which considers also the information at multiple spatial locations, Eq. (11) is an assumption based on the dimensional consistency of the SGS stress. In general cases, the numerical construction of first-order derivatives requires less efforts than higher-order derivatives, which also induces further numerical errors at large LES grid spacings. Besides, the previous “black-box” model has shown that incorporating first-order gradients at multiple stencil locations can have higher SGS accuracy than including higher-order derivatives at the local grid point.^{64,67} Meanwhile, higher-order gradients can be, in principle, constructed using lower-order gradients among the spatial stencil locations. Hence, we will only consider the first-order velocity gradients, yielding

$$\tau_{ij}^{model}|_{l,m,n} = \sum_{\substack{p_1, p_2, p_3 = -N_{p,q} \\ q_1, q_2, q_3 = 0}}^{N_{p,q}} \sum_{k=1}^3 g^{(p_1, p_2, p_3; q_1, q_2, q_3)} \times \Delta^2 \frac{\partial \bar{u}_i}{\partial x_k} \Big|_{l+p_1, m+p_2, n+p_3} \frac{\partial \bar{u}_j}{\partial x_k} \Big|_{l+q_1, m+q_2, n+q_3}. \quad (12)$$

In the current work, we only choose the stencil locations along the x , y and z coordinate directions, as shown in Fig. 1. As a result, the SGS stress reduces to

$$\tau_{ij}^{model}|_{l,m,n} = \sum_{\substack{N_{pq} \\ p_1, p_2, p_3 = -N_{pq} \\ q_1, q_2, q_3 = -N_{pq} \\ p_1 p_2 = p_2 p_3 = p_1 p_3 = 0 \\ q_1 q_2 = q_2 q_3 = q_1 q_3 = 0}} \sum_{k=1}^3 g^{(p_1, p_2, p_3; q_1, q_2, q_3)} \times \Delta^2 \frac{\partial \bar{u}_i}{\partial x_k} |_{l+p_1, m+p_2, n+p_3} \frac{\partial \bar{u}_j}{\partial x_k} |_{l+q_1, m+q_2, n+q_3}. \quad (13)$$

For $N_{pq} = 1$, there are $M = 1 + 6N_{pq} = 7$ stencil locations, and the model will be abbreviated as ANNSGM-7-49 (since there are $M^2 = 49$ possible combinations of the gradient products on the right-hand side of the equation). Likewise, for $N_{pq} = 2$, $M = 13$, and the model will be abbreviated as ANNSGM-13-169 . . . , and so on. In general, the model is abbreviated as ANNSGM- M - M^2 . For simplicity, hereafter, we shall denote (l, m, n) , (p_1, p_2, p_3) , and (q_1, q_2, q_3) by \mathbf{L} , \mathbf{s}_1 , and \mathbf{s}_2 , respectively. We shall also drop

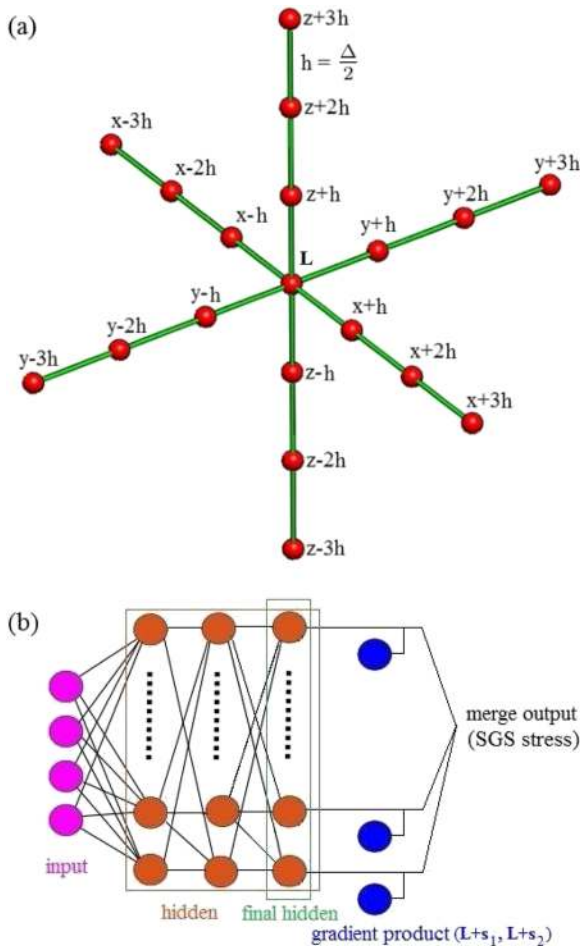


FIG. 1. Schematic of the ANN-SGM: (a) spatial locations in consideration and (b) the structure of the neural network.

the summation symbol Σ due to the summation convention. Consequently, the SGS stress can be written as

$$\tau_{ij}^{model} = g^{s_1 s_2} \Delta^2 \left(\frac{\partial \bar{u}_i}{\partial x_k} \Big|_{\mathbf{L}+s_1} \right) \left(\frac{\partial \bar{u}_j}{\partial x_k} \Big|_{\mathbf{L}+s_2} \right). \quad (14)$$

Inspired by previous works,^{33,42,78} the model coefficients $g^{s_1 s_2}$ are assumed to be functions of $\Lambda_n^{L^*}$ ($n = 1, 2, 3, 4, 5$), which are the invariants of the local velocity gradient tensor, defined as

$$\Lambda_1^{L^*} = \text{Tr}(\bar{\mathbf{S}}^2), \quad \Lambda_2^{L^*} = \text{Tr}(\bar{\mathbf{\Omega}}^2), \quad \Lambda_3^{L^*} = \text{Tr}(\bar{\mathbf{S}}^3), \quad (15)$$

$$\Lambda_4^{L^*} = \text{Tr}(\bar{\mathbf{\Omega}}^2 \bar{\mathbf{S}}), \quad \Lambda_5^{L^*} = \text{Tr}(\bar{\mathbf{\Omega}}^2 \bar{\mathbf{S}}^2),$$

with the filtered strain and rotation rates given by

$$\bar{\mathbf{S}}_{ij} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right), \quad \bar{\mathbf{\Omega}}_{ij} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} - \frac{\partial \bar{u}_j}{\partial x_i} \right). \quad (16)$$

Following previous works,^{33,71} we further reduce the five invariants to four dimensionless variables by rescaling, yielding

$$\Lambda_1^L = \frac{\text{Tr}(\bar{\mathbf{\Omega}}^2)}{\text{Tr}(\bar{\mathbf{S}}^2)}, \quad \Lambda_2^L = \frac{\text{Tr}(\bar{\mathbf{S}}^3)}{\text{Tr}(\bar{\mathbf{S}}^2)^{3/2}}, \quad \Lambda_3^L = \frac{\text{Tr}(\bar{\mathbf{\Omega}}^2 \bar{\mathbf{S}})}{\text{Tr}(\bar{\mathbf{S}}^2)^{1/2} \text{Tr}(\bar{\mathbf{\Omega}}^2)}, \quad (17)$$

$$\Lambda_4^L = \frac{\text{Tr}(\bar{\mathbf{\Omega}}^2 \bar{\mathbf{S}}^2)}{\text{Tr}(\bar{\mathbf{S}}^2) \text{Tr}(\bar{\mathbf{\Omega}}^2)}.$$

In turn, Eq. (14) can be recast as

$$\tau_{ij}^{model} = g^{s_1 s_2} (\Lambda_1^L \Lambda_2^L \Lambda_3^L \Lambda_4^L) \Delta^2 \left(\frac{\partial \bar{u}_i}{\partial x_k} \Big|_{\mathbf{L}+s_1} \right) \left(\frac{\partial \bar{u}_j}{\partial x_k} \Big|_{\mathbf{L}+s_2} \right). \quad (18)$$

Clearly, the building blocks of Eq. (18) explicitly involve the velocity gradients at different stencil locations. As discussed above, both \mathbf{s}_1 and \mathbf{s}_2 represent seven stencil locations (including the local point) if one neighboring point is chosen in each direction ($x-$, $x+$, $y-$, $y+$, $z-$, $z+$). Consequently, $7^2 = 49$ terms appear in the summation on the right-hand side of the equation, and 49 dimensionless coefficients ($g^{s_1 s_2}$) are needed. These coefficients ($g^{s_1 s_2}$) are only functions of the local dimensionless invariants. Based on our test (not reported), including the invariants at the neighboring stencil locations, i.e., incorporating Λ_s at the seven stencil locations in the expressions of the coefficients g (4×7 terms in total), makes little improvements, but including the velocity gradients at the neighboring stencil locations is indeed necessary for improving the accuracy of the SGS stress. Finally, the relations between $g^{s_1 s_2}$ and the invariants are obtained by the ANN. As we can see, the gradient terms are not in the hidden layers of the ANN and no parameters need to be trained for them. The only implicit part is in the dimensionless model coefficients $g^{s_1 s_2}$ generated from the ANN. This “semi-explicit” model needs only four inputs (the invariants) for the hidden layers to yield the model coefficients. As already noted, Eq. (18) is mathematically and dimensionally consistent with the true SGS stress. Furthermore, the non-dimensional character of the coefficients inherently gives the model a wide applicability, making it easier to adapt to other filter widths. With these considerations, we aim at constructing a highly explicable and widely applicable model with desired accuracy.

We recall that Ling *et al.*⁴² constructed an ANN-based semi-explicit model for the anisotropic part of the Reynolds stress with

embedded invariance properties. Recently, Xie *et al.*⁷⁰ adopted a similar approach in the reconstruction of the SGS stress in LES. Both works show the importance of the embedded invariance properties. However, these works only exploit the information of the local point. In addition, the accuracy of the predicted SGS stress is not significantly better than the velocity gradient model (VGM).⁷⁰ As a new extension to these studies, the current work also incorporates the invariance property. Meanwhile, the contributions from the spatial stencil locations are examined in detail.

To obtain the SGS stress τ_{ij} , the dimensionless model coefficients $g^{s_{1s_2}}$ in Eq. (18) are determined by the artificial neural network shown in Fig. 1. The output of any layer l (except for the input and output layers) can be calculated as

$$X_i^l = \sigma \left(\sum_j W_{ij}^l X_j^{l-1} + b_i^l \right), \quad (19)$$

where W_{ij}^l is the weight from the j th neuron in layer $l-1$ to the i th neuron in layer l . b_i^l is the bias parameter for the i th neuron in layer l . W_{ij}^l are initialized using Glorot-uniform,⁷⁹ and b_i^l are initialized to zero. Here, σ is an activation function.^{80,81} In the current work, we choose the Leaky-ReLU function,⁴² given by

$$\sigma(a) = \begin{cases} a & \text{if } a > 0 \\ 0.2a & \text{if } a \leq 0 \end{cases}. \quad (20)$$

The present artificial neural network consists of four layers of neurons ($4 : 20 : 20 : N_o$) with the four neurons in the input layer receiving the invariants $\Lambda_1^L \dots \Lambda_4^L$, and the N_o neurons send out the coefficients $g^{s_{1s_2}}$ for each of the gradient products as per Eq. (18). Here, the number of coefficients depends on the number of considered spatial locations. These coefficients finally combine with the gradient products, as shown in Eq. (18), to match the target (i.e., the true SGS stress generated from the filtered DNS data).

In this work, we choose $N_{pq} = 1, 2, 3$ neighboring locations of the local grid point in each of the six directions, namely, $x-, x+, y-, y+, z-, z+$, as shown in Fig. 1. These locations are equally spaced at the distance of $\frac{\Delta}{2}$. The local grid point is denoted by \mathbf{L} , and the neighboring points are denoted by $x-h, x-2h, x-3h, x+h, x+2h \dots y-h, \dots z-h \dots$ as Fig. 1 shows. Throughout the study, the grid space h of LES is taken as $h = \Delta/2$ unless otherwise specified. As discussed, the models corresponding to $N_{pq} = 1, 2, 3$ will be denoted by ANNSGM-7-49, ANNSGM-13-169, and ANNSGM-19-361, respectively.

For the stability of the training process,^{43,44,58} the primary inputs ($\Lambda_1^L \dots \Lambda_4^L$) are normalized to the range $[0, 1]$ by $\hat{x} = \frac{x - \min\{x^{DNS}\}}{\max\{x^{DNS}\} - \min\{x^{DNS}\}}$, while the secondary inputs ($\frac{\partial u_i}{\partial x_k}|_{\mathbf{L}+s_1}, \frac{\partial u_i}{\partial x_k}|_{\mathbf{L}+s_2}$) and the outputs τ_{ij} are scaled to $O(1)$ by $\hat{x}_{ij} = \frac{x_{ij}}{x_{minmax}}$ with $x_{minmax} = \frac{1}{6} \sum_{i=1}^3 \sum_{j=1}^i \frac{\max\{x_{ij}^{DNS}\} - \min\{x_{ij}^{DNS}\}}{2}$. The mean squared error ($(\frac{1}{6} \sum_{i=1}^3 \sum_{j=1}^i (\tau_{ij}^{A_{predicted}} - \tau_{ij}^{A_{DNS}})^2)$) is finally minimized by the back-propagation method⁵⁷ to optimize the values of W_{ij}^l and b_i^l .

III. DNS DATABASE AND THE FILTERED DNS

In the current study, the DNS data are obtained from the direct numerical simulation of a forced incompressible isotropic turbulence at the Taylor Reynolds number $Re_\lambda \approx 250$ on a uniform grid

TABLE I. Numerical simulation parameters and one-point statistical quantities for 1024^3 grid resolution of incompressible turbulence.

Reso.	Re_λ	η/h_{DNS}	L_1/η	λ/η	ν	u^{rms}	ω^{rms}	ε
1024^3	259	1.01	233	31.7	0.001	2.29	26.3	0.69

of 1024^3 . The grid space is denoted by h_{DNS} . The numerical parameters and one-point statistics of the DNS are detailed in Table I.^{67,82} With this configuration, ηh_{DNS} is close to 1.01, giving a resolution of $k_{max}\eta \approx 2.11$ with $k_{max} = N/3 = 2\pi/3h_{DNS}$ being the maximum wavenumber. We note that a resolution of $k_{max}\eta \geq 2.1$ guarantees the convergence of the kinetic energy at all wavenumbers.^{75,83} The kinematic viscosity ν is taken to be 0.001. The root mean squared (rms) value of vorticity is defined as $\omega^{rms} = \sqrt{\langle \omega_i \omega_i \rangle}$.

The numerical simulations of the incompressible homogeneous isotropic turbulence are performed in a cubic box of $(2\pi)^3$ with periodic boundary conditions. A pseudospectral code is applied for the numerical simulations on a uniform grid.^{67,75} A second order two-step Adams–Bashforth scheme is used for time integration. Full dealiasing is implemented using the two-thirds rule.⁸⁴ The velocity field is forced by fixing the energy spectrum within the two lowest wavenumber shells with $f(k=1)/f(k=2) \approx 3.17$,⁸⁵ where $f(k)$ represents the component of the large-scale forcing \mathcal{F} at wavenumber k .

In the current machine-learning process, all the training data are generated through the filtered DNS field. The filtered variables and the corresponding SGS stress τ_{ij} are obtained using a Gaussian filter given by¹

$$G(r) = \left(\frac{6}{\pi \Delta^2} \right)^{1/2} \exp\left(-\frac{6r^2}{\Delta^2}\right), \quad (21)$$

where the filter width is $\Delta = nh_{DNS}$. Figure 2 illustrates the velocity spectrum of the DNS and the filtered DNS data at filter width $\Delta = 16h_{DNS}$, corresponding to $\Delta/\eta \approx 16$, which covers well the inertial subrange. In Fig. 2, the cutoff wavenumber is defined as $k_c = \pi/\Delta$. It is found that 5% of the turbulent kinetic energy resides in the SGS field at $\Delta/h_{DNS} = 16$. In the *a priori* analysis, the training data are

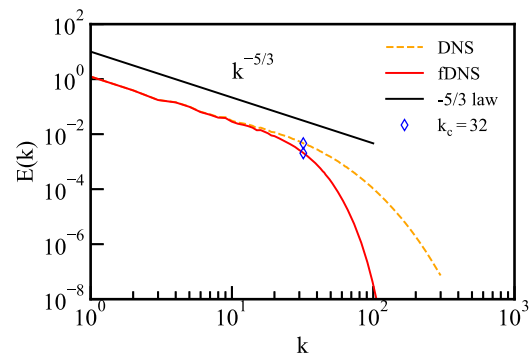


FIG. 2. Velocity spectrum of direct numerical simulation of a forced incompressible isotropic turbulence with a uniform grid of 1024^3 . Diamond represents the cutoff wavenumber $k_c = 32$ with a filter width of $\Delta = 16h_{DNS}$.

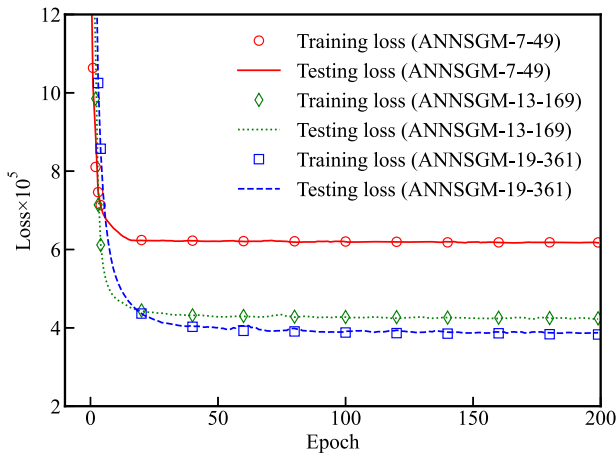


FIG. 3. Learning curves of the proposed models for the unclosed SGS stress.

invariably generated at this filter width. However, the results turn out to have wide applicability to other filter widths as we shall see later.

IV. TRAINING OF THE ANN-SGMs AND A SIMPLIFIED MODEL

In the training of the ANN, 4×64^3 of the 1024^3 filtered data points at filter width $\Delta = 16h_{DNS}$ are randomly selected for ANN training. 70% of the data are used for training and the rest for testing. The network is trained by the Adam algorithm to update W_{ij}^l and b_i^{ls6} for 200 iterations with the batch size being 4000. The training and testing losses for the three proposed ANN-SGMs are shown in Fig. 3. As can be seen, 200 iterations are sufficient for convergence. Since the learning curves of the training and testing sets almost merge, the trainings are free from over-fitting. Furthermore, the impact of the hyperparameters (i.e., the size of the artificial neural network: the number of layers and the number of neurons in each layer) on the ANN has been analyzed, and the network structure 4 : 20 : 20 : N_o turns out to be adequate.

The *a priori* accuracy of the ANN-SGM framework is examined in detail in Appendix A. Overall, the ANN-SGMs yield very high (close to 0.99) correlation coefficients for the SGS stress, which is much higher than that of the traditional velocity gradient model. With increasing number of considered stencil locations,

the accuracy exhibits a saturation (cf. Appendix A). Seven stencil locations turn out to be adequate for the current analysis.

Taking the ANNSGM-7-49 model as an example, we examine the dimensionless coefficients $g^{s_1s_2}$ through their mean values in Table II. As observed, the mean values vary across multiple orders of magnitude. Apparently, the major contributions, bold-faced in Table II, are from the terms associated with the local grid point (i.e., the elements in the first row and the first column) and the self-associated neighboring terms (i.e., the elements in the diagonal line). The rest of the terms are at least one order smaller than the “leading order” (g^{LL}) and “second order” ($g^{Lx-h}, \dots, g^{x-hL}, \dots$) terms. The standard deviations of the dimensionless coefficients $\sqrt{((g^{s_1s_2} - \langle g^{s_1s_2} \rangle)^2)}$ normalized by the mean values ($\langle g^{s_1s_2} \rangle$) are shown in Table III. As observed, the coefficients of the dominant terms are more narrowly distributed (i.e., the deviations are two orders smaller than the mean values). For the “less important” terms, the mean values and their deviations are “closer,” indicating a wider distribution. In light of these observations, it is natural to consider only these dominant terms so as to potentially condense the model without a large sacrifice of accuracy.

In this attempt, we artificially set the coefficients of the less important terms to zero. Taking the ANNSGM-7-49 model, for instance, the simplified version, abbreviated as ANNSGM-7-19, contains $1 + 6 \times 3 = 19$ terms (cf. Table II). Comparisons between the ANNSGM-7-19 and ANNSGM-7-49 models are given in Appendixes B and C in terms of the *a priori* and *a posteriori* accuracies, respectively. As it turns out, the ANNSGM-7-19 model gives almost the same performance as the ANNSGM-7-49 model. In this case, we will only examine in detail the capability of the ANNSGM-7-19 model in the *a posteriori* LES.

To close the current section, we examine the average coefficients of the ANNSGM-7-19 model in Table IV. The most important observation is that the distribution of coefficients is more uniform compared with the ANNSGM-7-49 model. This should be expected due to the isotropic nature of the flow. However, this trend is not so significant for the ANNSGM-7-49 model (cf. Table II) presumably due to the numerical noise induced by the large number of stencil terms. In this sense, the ANNSGM-7-19 model is also more robust than the ANNSGM-7-49 model. It is also worth noting that the symmetry of the SGS stress is not incorporated in Eq. (18). However, the symmetry of the model coefficients in Table IV, to a large extent, guarantees the symmetry of the SGS stress. In this case, we will take the ANNSGM-7-19 model as an example to examine the ANN-SGM framework in LES.

TABLE II. Mean values of the coefficients $g^{s_1s_2}$ for the ANNSGM-7-49 model at filter width $\Delta = 16h_{DNS}$. Boldface denotes the values of the dominant terms.

$\langle g^{s_1s_2} \rangle$	$s_1 = L$	$s_1 = x - h$	$s_1 = x + h$	$s_1 = y - h$	$s_1 = y + h$	$s_1 = z - h$	$s_1 = z + h$
$s_2 = L$	$+1.9 \times 10^{-1}$	-1.4×10^{-2}	-1.4×10^{-2}	-1.5×10^{-2}	-1.5×10^{-2}	-2.7×10^{-2}	-2.7×10^{-2}
$s_2 = x - h$	-2.1×10^{-2}	$+1.6 \times 10^{-2}$	-5.0×10^{-4}	$+6.0 \times 10^{-4}$	-1.3×10^{-6}	$+3.2 \times 10^{-3}$	$+3.4 \times 10^{-3}$
$s_2 = x + h$	-2.1×10^{-2}	-4.6×10^{-4}	$+1.6 \times 10^{-2}$	-4.0×10^{-4}	-1.9×10^{-4}	$+4.0 \times 10^{-3}$	$+4.0 \times 10^{-3}$
$s_2 = y - h$	-1.9×10^{-2}	-1.9×10^{-4}	-4.0×10^{-4}	$+1.6 \times 10^{-2}$	-3.8×10^{-4}	$+1.3 \times 10^{-3}$	$+1.5 \times 10^{-3}$
$s_2 = y + h$	-2.1×10^{-2}	-6.0×10^{-5}	-5.8×10^{-5}	-1.6×10^{-4}	$+1.6 \times 10^{-2}$	$+2.5 \times 10^{-3}$	$+2.7 \times 10^{-3}$
$s_2 = z - h$	-1.4×10^{-2}	-5.5×10^{-4}	-3.8×10^{-4}	-2.3×10^{-4}	-2.1×10^{-4}	$+1.6 \times 10^{-2}$	-5.9×10^{-4}
$s_2 = z + h$	-1.3×10^{-2}	-4.6×10^{-4}	-4.2×10^{-4}	-2.3×10^{-4}	-3.2×10^{-4}	-5.5×10^{-4}	$+1.5 \times 10^{-2}$

TABLE III. Standard deviations of $g^{s_1 s_2}$ normalized by mean values for the ANNSGM-7-49 model at filter width $\Delta = 16h_{DNS}$. Boldface denotes the values of the dominant terms.

$\frac{\sqrt{\langle (g^{s_1 s_2} - \langle g^{s_1 s_2} \rangle)^2 \rangle}}{\langle g^{s_1 s_2} \rangle}$	$s_1 = L$	$s_1 = x - h$	$s_1 = x + h$	$s_1 = y - h$	$s_1 = y + h$	$s_1 = z - h$	$s_1 = z + h$
$s_2 = L$	$+1.8 \times 10^{-2}$	-1.7×10^{-2}	-2.6×10^{-2}	-3.4×10^{-2}	-1.5×10^{-2}	-2.2×10^{-2}	-2.2×10^{-2}
$s_2 = x - h$	-3.7×10^{-2}	$+1.9 \times 10^{-2}$	-4.4×10^{-1}	$+4.5 \times 10^{-1}$	-1.5×10^2	$+1.2 \times 10^{-1}$	$+1.2 \times 10^{-1}$
$s_2 = x + h$	-3.4×10^{-2}	-4.3×10^{-1}	$+2.3 \times 10^{-2}$	-1.0×10^{-1}	-1.1×10^{-1}	$+0.9 \times 10^{-1}$	$+0.8 \times 10^{-1}$
$s_2 = y - h$	-2.8×10^{-2}	-5.2×10^{-1}	-3.3×10^{-1}	$+2.6 \times 10^{-2}$	-5.7×10^{-1}	$+2.5 \times 10^{-1}$	$+2.2 \times 10^{-1}$
$s_2 = y + h$	-2.2×10^{-2}	-5.7×10^{-1}	-5.2×10^{-1}	-7.5×10^{-1}	$+1.6 \times 10^{-2}$	$+0.8 \times 10^{-1}$	$+1.1 \times 10^{-1}$
$s_2 = z - h$	-2.6×10^{-2}	-1.6×10^{-1}	-2.2×10^{-1}	-2.7×10^{-1}	-3.2×10^{-1}	$+2.0 \times 10^{-2}$	-5.8×10^{-1}
$s_2 = z + h$	-2.7×10^{-2}	-1.6×10^{-1}	-3.6×10^{-1}	-5.2×10^{-1}	-1.1×10^{-1}	-6.2×10^{-1}	$+2.5 \times 10^{-2}$

TABLE IV. Spatially averaged value for the coefficients $g^{s_1 s_2}$ of the ANNSGM-7-19 model at filter width $\Delta = 16h_{DNS}$. Boldface denotes the non-zero values, corresponding to the dominant terms of the ANNSGM-7-49 model in Tables II and III.

$\langle g^{s_1 s_2} \rangle$	$s_1 = L$	$s_1 = x - h$	$s_1 = x + h$	$s_1 = y - h$	$s_1 = y + h$	$s_1 = z - h$	$s_1 = z + h$
$s_2 = L$	$+1.8 \times 10^{-1}$	-1.6×10^{-2}	-1.6×10^{-2}	-1.6×10^{-2}	-1.6×10^{-2}	-1.7×10^{-2}	-1.6×10^{-2}
$s_2 = x - h$	-1.7×10^{-2}	$+1.7 \times 10^{-2}$	0	0	0	0	0
$s_2 = x + h$	-1.6×10^{-2}	0	$+1.6 \times 10^{-2}$	0	0	0	0
$s_2 = y - h$	-1.6×10^{-2}	0	0	$+1.6 \times 10^{-2}$	0	0	0
$s_2 = y + h$	-1.6×10^{-2}	0	0	0	$+1.6 \times 10^{-2}$	0	0
$s_2 = z - h$	-1.6×10^{-2}	0	0	0	0	$+1.7 \times 10^{-2}$	0
$s_2 = z + h$	-1.6×10^{-2}	0	0	0	0	0	$+1.6 \times 10^{-2}$

V. A POSTERIORI TEST OF THE ANN-SGM

In the *a posteriori* test, LESs are performed at $Re_\lambda \approx 250$ using the same time integration scheme as in DNS. Meanwhile, three grid resolutions are adopted, namely, 128^3 , 64^3 , and 32^3 , corresponding to filter widths $\Delta = 16h_{DNS}$, $\Delta = 32h_{DNS}$, and $\Delta = 64h_{DNS}$, respectively. Here, the different filter widths are chosen to test the range of applicability of the proposed ANN-SGM. However, the key results and analyses are only for the $\Delta = 16h_{DNS}$ case, for which the inertial subrange is well covered.¹ We investigate and validate the performance of the ANNSGM-7-19 model using the filtered DNS result as benchmarks. Meanwhile, other popular LES models are introduced for comparison, including implicit LES (ILES), the dynamic Smagorinsky model (DSM), and the dynamic mixed model (DMM). The grid size of LES is invariably taken at $h = \Delta/2$.

The implicit-LES (ILES) is based on a numerical viscosity without explicit modeling for the SGS stress.^{24–28} To avoid the instability often encountered in gradient-type SGS models, a dissipative numerical method is introduced with higher numerical dissipation at larger wavenumbers, namely,⁶⁷

$$\bar{\mathbf{u}} = \mathbf{u} - \Delta t \cdot C_0^I \left(\frac{\mathbf{k}}{k_0} \right)^n \bar{\mathbf{u}}, \quad k_0 = N/3, \quad (22)$$

where $n = 4$, $C_0^I = 3$, N is the number of LES grids, and k_0 is the largest wavenumber in LES due to the two-thirds dealiasing rule. This configuration is invariably used for all models in the *a posteriori* test for consistency. It was shown that the numerical dissipation is necessary to generate the flux of kinetic energy to smaller scales, which cannot be resolved at the LES grids.⁶⁷ The grid scale is chosen to be smaller than the filter width to ensure that the numerical

dissipation does not significantly affect the high accuracy of SGS models.

The spirit of the dynamic Smagorinsky model (DSM) is the assumption that the SGS stress at different filter widths in the inertia subrange is rigorously similar in an eddy viscosity form.¹ By exploiting the Germano identity,⁹ the DSM model can be written as^{6,8,10,16}

$$\tau_{ij}^A = \tau_{ij} - \frac{\delta_{ij}}{3} \tau_{kk} = -2C_s^2 \Delta^2 |\bar{S}| \bar{S}_{ij}, \quad (23)$$

with Δ being the primary filter width, $|\bar{S}| = (2\bar{S}_{ij}\bar{S}_{ij})^{1/2}$ being the characteristic filtered rate of strain, and C_s^2 calculated dynamically as

$$C_s^2 = \frac{\langle \mathcal{L}_{ij} \mathcal{M}_{ij} \rangle}{\langle \mathcal{M}_{kl} \mathcal{M}_{kl} \rangle}, \quad (24)$$

with $\mathcal{L}_{ij} = \widetilde{\widetilde{u_i u_j}} - \widetilde{u_i} \widetilde{u_j} \alpha_{ij} = -2\Delta^2 |\bar{S}| \bar{S}_{ij} \beta_{ij} = -2\widetilde{\Delta^2} |\widetilde{S}| \widetilde{S}_{ij}$, and $\mathcal{M}_{ij} = \beta_{ij} - \widetilde{\alpha}_{ij}$. Here, following the convention,¹ an overbar denotes the filtering operation at scale Δ , a tilde represents a coarser filtering ($\widetilde{\Delta} = 2\Delta$), and a tilde over an overbar denotes a double filtering.

The dynamic mixed model (DMM)^{17,19} is a combination of the scale-similarity model and the eddy-viscosity model.¹ This model guarantees sufficient dissipation with the eddy-viscosity part while retaining the high correlation coefficient of scale similarity part. To briefly describe the dynamic procedure in the DMM model, we first write the Germano identity, given by⁹

$$L_{ij}^A = T_{ij}^A - \widetilde{\tau}_{ij}^A, \quad (25)$$

with $T_{ij}^A = T_{ij} - \frac{\delta_{ij}}{3} T_{kk}$, where $T_{ij} = \widetilde{\widetilde{u_i u_j}} - \widetilde{u_i} \widetilde{u_j}$ is the SGS stress at filter width $\widetilde{\Delta}$, and $L_{ij}^A = L_{ij} - \frac{\delta_{ij}}{3} L_{kk}$ with $L_{ij} = \widetilde{u_i u_j} - \widetilde{u_i} \widetilde{u_j}$ representing

the resolved stress.¹ A square error is obtained by replacing τ_{ij}^A and T_{ij}^A with the modeled stresses τ_{ij}^{Amodel} and T_{ij}^{Amodel} , yielding

$$E_{model} = \langle (L_{ij}^A - L_{ij}^{Amodel})^2 \rangle, \quad (26)$$

where $L_{ij}^{Amodel} = T_{ij}^{Amodel} - \widehat{\tau_{ij}^{Amodel}}$. By assuming the coefficients are scale invariant and minimizing the right-hand side of Eq. (26), we obtain a set of equations for the coefficients, including the SGS stress terms at scales Δ and $\widehat{\Delta}$ ^{18,22,87}

$$\tau_{ij}^{Amodel} = C_1 h_{1ij}^A + C_2 h_{2ij}^A, \quad (27)$$

$$T_{ij}^{Amodel} = C_1 H_{1ij}^A + C_2 H_{2ij}^A, \quad (28)$$

where $h_{1ij}^A = -2\Delta^2 |\widehat{S}_{ij}| \widehat{S}_{ij}$, $h_{2ij}^A = h_{2ij} - \frac{\delta_{ij}}{3} h_{2kk}$, $h_{2ij} = \widehat{\overline{u_i u_j}} - \widetilde{\overline{u_i u_j}}$, $H_{1ij}^A = -2\widehat{\Delta}^2 |\widehat{\widetilde{S}}_{ij}| \widehat{\widetilde{S}}_{ij}$, $H_{2ij}^A = H_{2ij} - \frac{\delta_{ij}}{3} H_{2kk}$, and $H_{2ij} = \widehat{(\overline{\widetilde{u_i u_j}})} - \widetilde{(\overline{\widetilde{u_i u_j}})}$ with the right arrow denoting the filtering at filter width 4Δ . These yield

$$C_1 = \frac{\langle N_{ij}^2 \rangle \langle L_{ij}^A M_{ij} \rangle - \langle M_{ij} N_{ij} \rangle \langle L_{ij}^A N_{ij} \rangle}{\langle N_{ij}^2 \rangle \langle M_{ij}^2 \rangle - \langle M_{ij} N_{ij} \rangle^2}, \quad (29)$$

$$C_2 = \frac{\langle M_{ij} \rangle \langle L_{ij}^A N_{ij} \rangle - \langle M_{ij} N_{ij} \rangle \langle L_{ij}^A M_{ij} \rangle}{\langle N_{ij}^2 \rangle \langle M_{ij}^2 \rangle - \langle M_{ij} N_{ij} \rangle^2}, \quad (30)$$

where $M_{ij} = H_{1ij}^A - \widehat{h_{1ij}^A}$ and $N_{ij} = H_{2ij}^A - \widehat{h_{2ij}^A}$

All the LESs are performed on a grid resolution of $h = \Delta/2$ with filter widths $\Delta = 16, 32, 64h_{DNS}$. The value $FGR = 2$ ($FGR = \Delta/h$), which is the ratio of the filter width to the grid space, ensures that the differences between the LES and the filtered DNS are mainly due to the modeling errors of the SGS terms.⁸⁸⁻⁹⁵ The ratio of the time

steps in LES and DNS is kept at $\Delta t_{LES}/\Delta t_{DNS} = 10$ for all the LESs, where Δt_{LES} and Δt_{DNS} are the time steps for LES and DNS, respectively. In this research, the time discretization error of the LES is small and thus neglected.⁸⁸⁻⁹⁵ As we shall see, both the flow statistics and the instantaneous field can be accurately predicted by the current ANN-SGM in the *a posteriori* test.

Due to the stationary nature of the flow, all the results are time-averaged values, except for the result of the instantaneous field (Figs. 9, 10, 15, and 16).

A. A posteriori tests at filter width $\Delta = 16h_{DNS}$

Figure 4 displays the velocity spectrum of the ANNSGM-7-19 model together with the results from the DNS and the filtered DNS (fDNS). The results of the ILES, DSM, and DMM models are also shown in Fig. 4. As expected, the velocity spectrum of DNS has a long inertial regime with the $k^{-5/3}$ scaling.¹ The filtered DNS spectrum is lower than that of DNS especially at high wavenumbers after the cutoff value (cf. Fig. 2) since the small scale energy is filtered out. The error in the ILES result increases as k increases. As Fig. 4 illustrates, the DSM and DMM models tend to overestimate the spectrum at larger scales (small wavenumber) while they underestimate it at small scales (large wavenumber). On the other hand, the ANNSGM-7-19 model recovers nearly the whole range of velocity spectra except a small deviation from the filtered DNS result near the cutoff wavenumber. As for efficiency, the ratio of the computational cost of LES using different SGS models is found to be ILES (no-model): SM (Smagorinsky model): DSM: DMM: ANNSGM-7-19 = 0.4:0.72:1:1.5:4 (it should be noted that the ratio is 0:0.32:0.6:1.1:3.6 for the subgrid modeling cost only). For comparison, the ratio of the

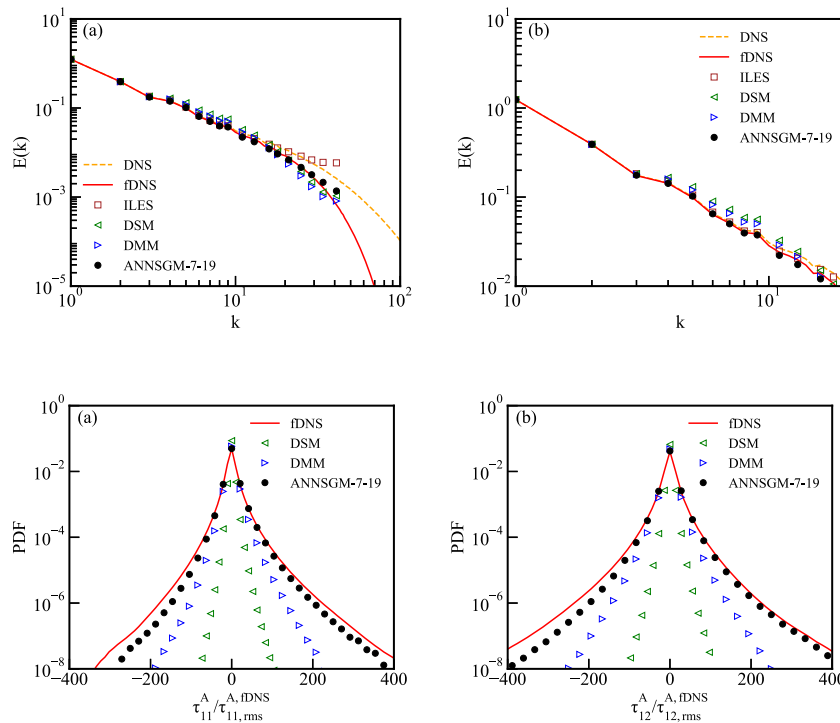


FIG. 4. Spectrum of velocity for LES at a grid resolution of 128^3 ($h = \Delta/2$) in the *a posteriori* test at filter width of $\Delta = 16h_{DNS}$ (a) at wavenumber $k < 100$ and (b) at wavenumber $k < 20$.

FIG. 5. PDF of normal and shear components of the SGS stress τ_{ij} in the *a posteriori* test at filter width $\Delta = 16h_{DNS}$: (a) normal component of SGS stress and (b) shear component of SGS stress.

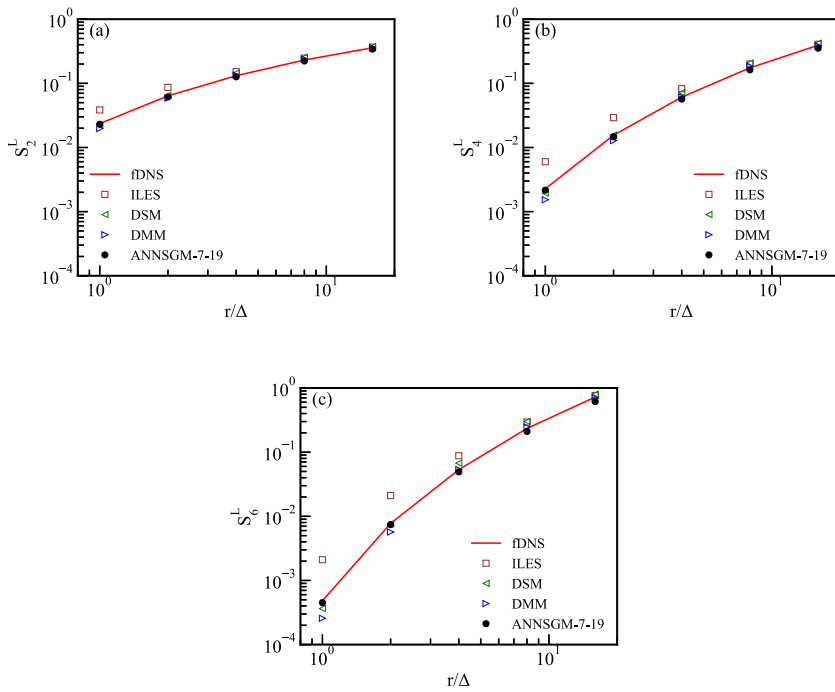


FIG. 6. Structure functions in the *a posteriori* analysis at filter width $\Delta = 16h_{DNS}$: (a) S_2^L , (b) S_4^L , and (c) S_6^L .

computational cost of LES using the current model to the previous “black-box” model⁶⁷ is 4:256, clearly demonstrating the efficiency of the current ANN-SGM framework considering the very small sacrifice of accuracy.

We now compare the probability density functions (PDFs) of anisotropic SGS stress $\tau_{ij}^A (= \tau_{ij} - \frac{\delta_{ij}}{3}\tau_{kk})$ for the ANNSGM-7-19 model with other SGS models in Fig. 5. The SGS stress is normalized by the root mean square (rms) value of the true SGS stress obtained

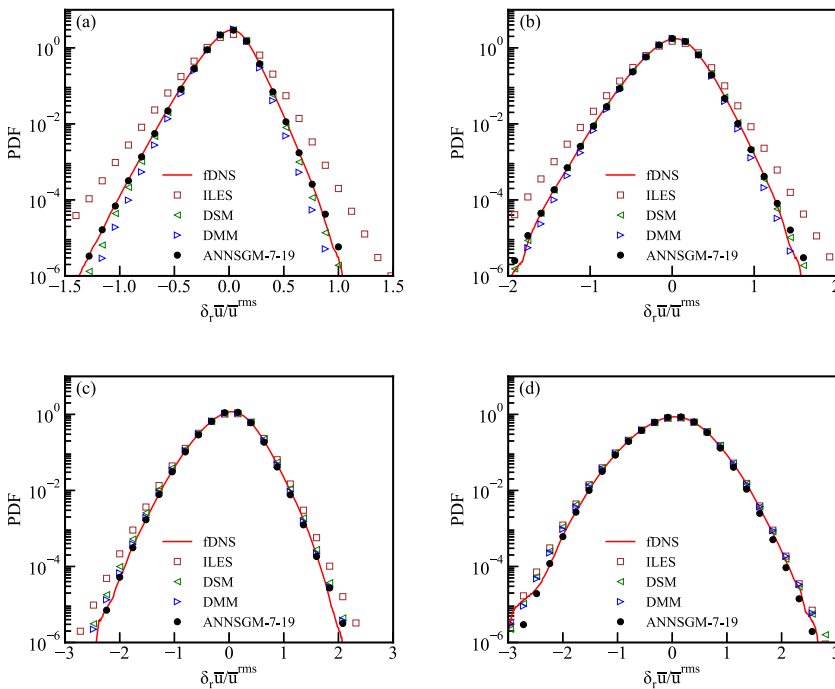


FIG. 7. PDF of the normalized velocity increment in the *a posteriori* test at filter width $\Delta = 16h_{DNS}$: (a) $r = \Delta$, (b) $r = 2\Delta$, (c) $r = 4\Delta$, and (d) $r = 8\Delta$.

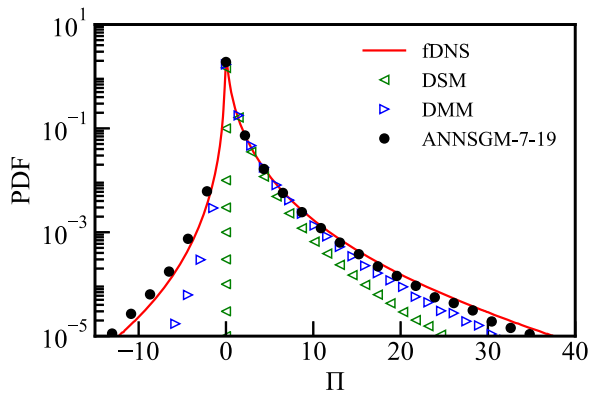


FIG. 8. PDF of the SGS flux in the *a posteriori* test at filter width $\Delta = 16h_{DNS}$.

from the filtered DNS field, calculated as $\tau_{ijrms}^{A fDNS} = \sqrt{\langle (\tau_{ij}^{A fDNS})^2 \rangle}$. Again, $\langle \cdot \rangle$ denotes a spatial average. Here, only one component of the normal stress and one component of the shear stress are presented due to the isotropic character of the flow. The PDFs of τ_{ij}^A are almost symmetric for isotropic turbulence. Clearly, the ANNSGM-7-19 model predicts best the SGS stress, while both the DSM and DMM models predict a narrower PDF. Meanwhile, we also observe

that the DMM model performs better in the prediction of the SGS stress than the DSM model.

Next, we analyze the spatial structures of the LES flow field using the structure functions and PDFs of the velocity increment. The longitudinal structure function of velocity can be written as

$$S_n^L(r) \equiv \left\langle \left| \frac{\delta_r \bar{u}}{\bar{u}^{rms}} \right|^n \right\rangle, \quad (31)$$

where $\delta_r \bar{u} = [\bar{\mathbf{u}}(\mathbf{x} + \mathbf{r}) - \bar{\mathbf{u}}(\mathbf{x})] \cdot \hat{\mathbf{r}}$ is the longitudinal increment of the velocity at the separation \mathbf{r} with $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$. Here, the velocity increment is normalized by the rms velocity \bar{u}^{rms} . For isotropic flow, we can arbitrarily choose the direction and take \mathbf{r} in the x direction for simplicity. As Fig. 6 illustrates, the structure functions at large separations can be accurately predicted by all the models. This is not surprising due to the nature of LES. As the separation \mathbf{r}/Δ decreases, ILES does a poor job, predicting a value higher than the filtered DNS result, whereas the structure functions from the DSM and DMM models are lower than the filtered DNS result at small separation and their errors go up as n increases. In contrast, the ANNSGM-7-19 model can accurately predict structure functions at the whole range of scales.

The PDFs of the normalized velocity increment are shown in Fig. 7. As Fig. 7 depicts, all the models have reasonable agreement with the filtered DNS result at large \mathbf{r} . At smaller \mathbf{r} , the DSM, DMM, and ILES models show some discrepancies, consistent with the trend

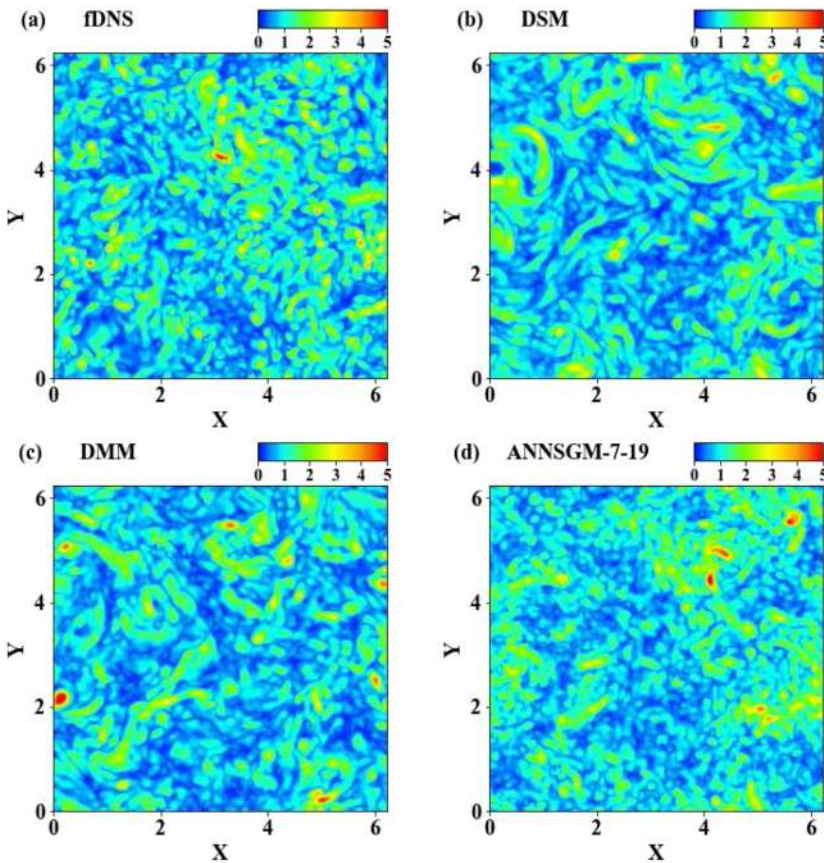


FIG. 9. The instantaneous field of normalized vorticity in the *a posteriori* test at filter width $\Delta = 16h_{DNS}$: (a) fDNS, (b) DSM, (c) DMM, and (d) ANNSGM-7-19.

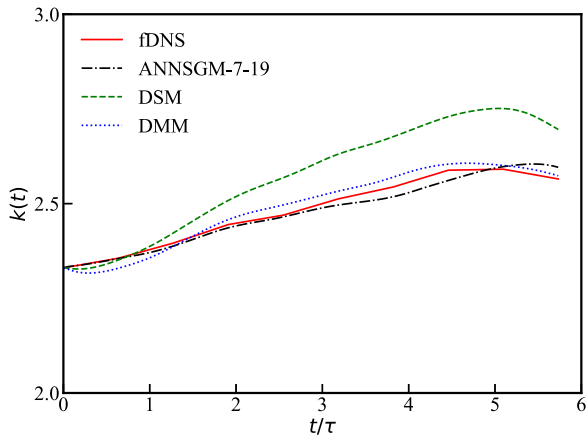


FIG. 10. Time evolution of the kinetic energy in the *a posteriori* test at filter width $\Delta = 16h_{DNS}$.

in Fig. 6. Again, the ANNSGM-7-19 model agrees well with the filtered DNS result at both small and large separations.

In the study of turbulence, a very important and intriguing phenomenon is the energy cascade. In this regard, we examine the SGS flux, defined by¹

$$\Pi = -\tau_{ij}\bar{\delta}_{ij}, \quad (32)$$

which constitutes a direct representation of the rate of energy transfer to the residual field (small scale) from the resolved field (large scale). We note that this transport of energy is entirely due to nonlinear interactions, and it can be negative (i.e., backscatter). The PDFs of the SGS flux are displayed in Fig. 8. As shown, the ANNSGM-7-19

model predicts accurately the energy transfer between the resolved and the SGS fields. The DSM predicts a strictly positive SGS flux, thus no backscatter. This is not surprising due to the formulation of the model [cf. Eq. (24)]. The DMM model, on the other hand, does predict a backscatter. However, the backscatter is much weaker in intensity than the filtered DNS.

In Fig. 9, we show the instantaneous vorticity field normalized by the rms value at arbitrarily selected xy cross sections of the domain in order to scrutinize the coherent structure of the flow field.⁶⁷ Here, all the models are initialized to the same condition and run for $t/\tau = 5.34$ with τ being the large-eddy turnover time, defined by $\tau = L_I/u^{rms}$. As shown in Fig. 9, the ANNSGM-7-19 model simulates reasonably well the vorticity field with most of the flow structures recovered. On the other hand, the predicted vorticity fields by the DSM and DMM models are quite different compared with the filtered DNS. We also note that the result of DSM is very close to that of the DMM due to the similarity in their formulations (they both have a “Smagorinsky part”). The evolution of the kinetic energy $k(t)$ with time is illustrated in Fig. 10. As can be seen, the evolution of the kinetic energy predicted by the ANNSGM-7-19 model is the closest to the filtered DNS result.

B. A posteriori tests at larger filter widths $\Delta = 32h_{DNS}$ and $\Delta = 64h_{DNS}$

To test the applicability of the current ANN method at filter widths larger than the training one (i.e., $\Delta = 16h_{DNS}$), we rerun the *a posteriori* tests at filter width $\Delta = 32h_{DNS}$ and $\Delta = 64h_{DNS}$ to examine some of the major results in Subsection V A. Now, the corresponding grids applied in the LESs are 64^3 and 32^3 , respectively.

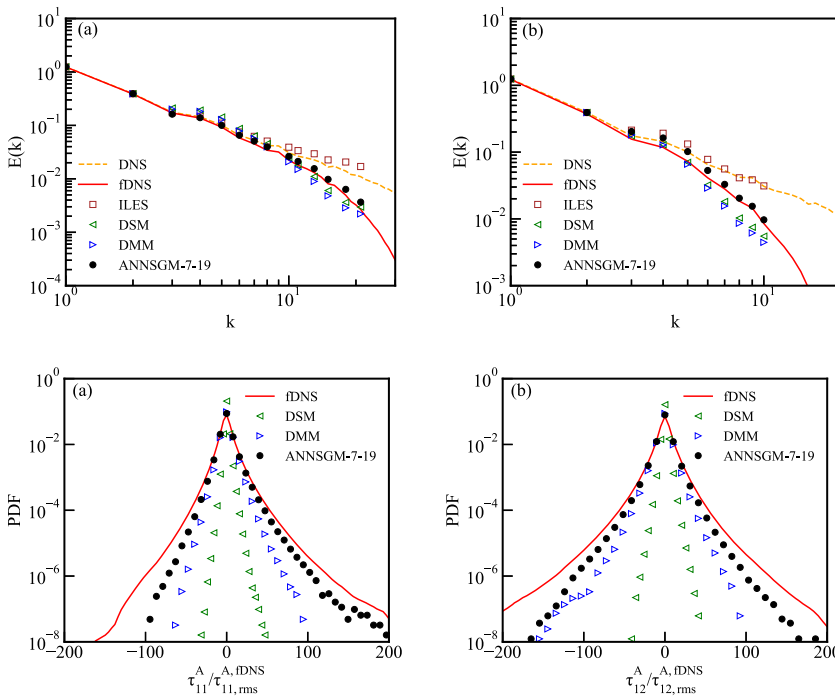


FIG. 11. Spectrum of velocity for LES in the *a posteriori* test at larger filter widths: (a) $\Delta = 32h_{DNS}$ and (b) $\Delta = 64h_{DNS}$.

FIG. 12. PDF of normal and shear components of the SGS stress τ_{ij} in the *a posteriori* test at filter width $\Delta = 32h_{DNS}$: (a) normal component of SGS stress and (b) shear component of SGS stress.

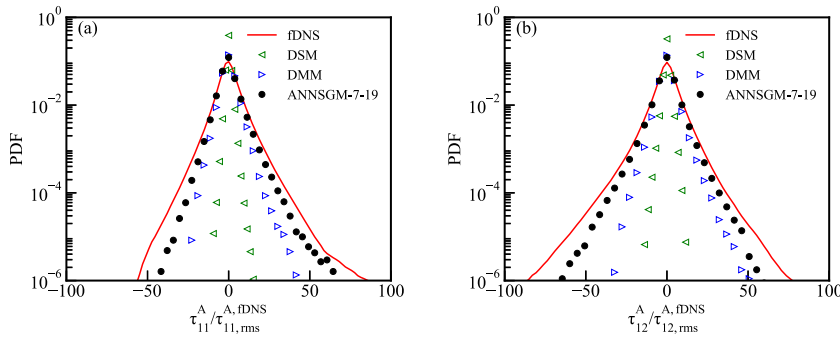


FIG. 13. PDF of normal and shear components of the SGS stress τ_{ij} in the *a posteriori* test at filter width $\Delta = 64h_{DNS}$: (a) normal component of SGS stress and (b) shear component of SGS stress.

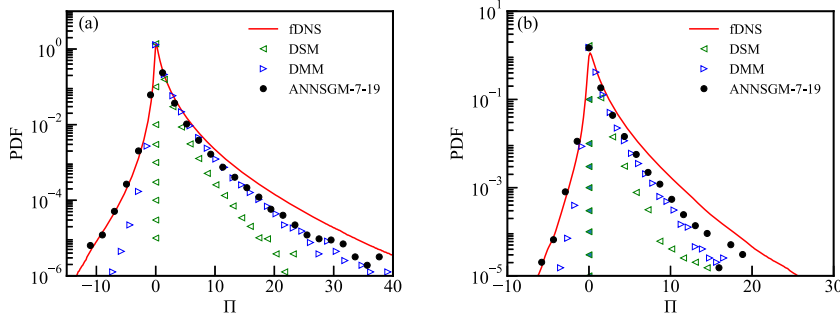


FIG. 14. PDF of the SGS flux in the *a posteriori* test at larger filter widths: (a) $\Delta = 32h_{DNS}$ and (b) $\Delta = 64h_{DNS}$.

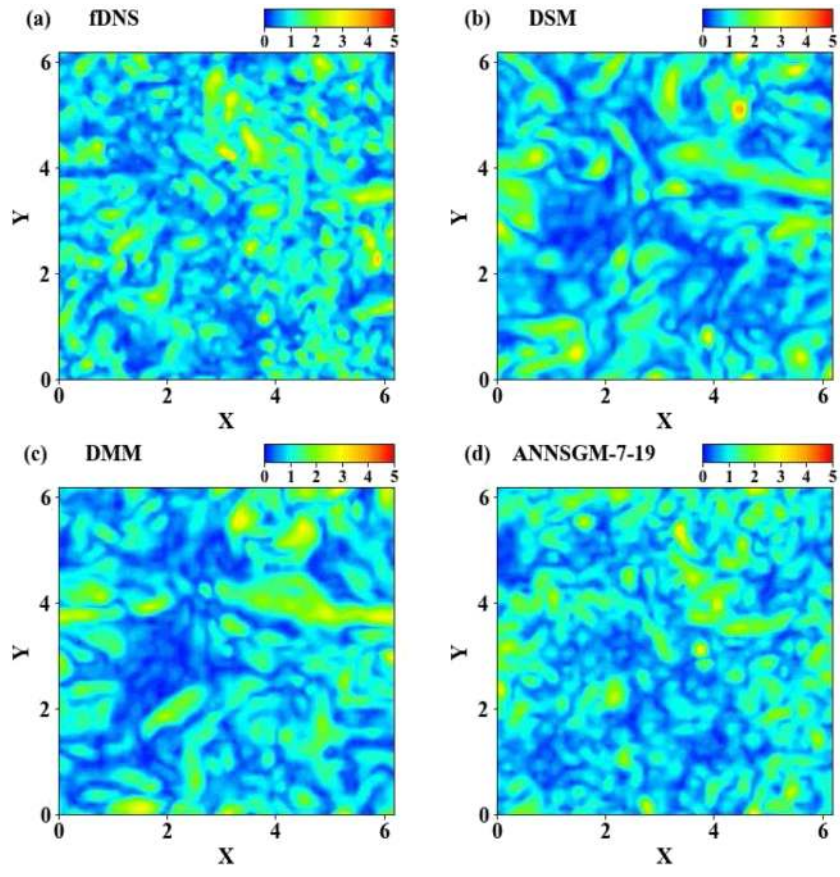


FIG. 15. The instantaneous field of normalized vorticity in the *a posteriori* test at filter width $\Delta = 32h_{DNS}$: (a) fDNS, (b) DSM, (c) DMM, and (d) ANNSGM-7-19.

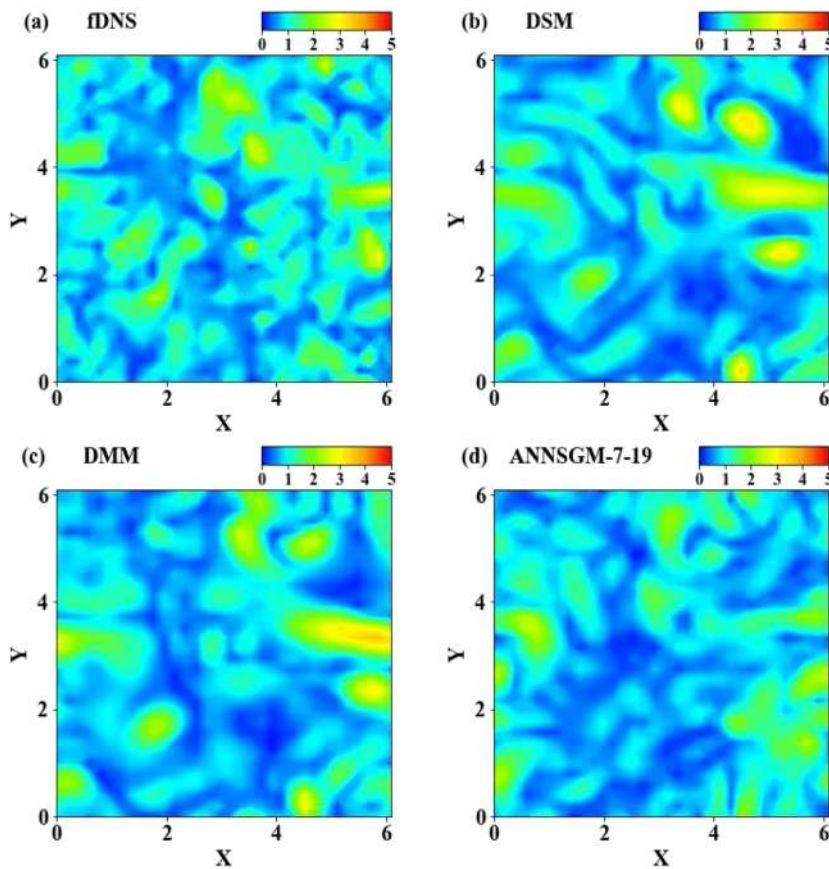


FIG. 16. The instantaneous field of normalized vorticity in the *a posteriori* test at filter width $\Delta = 64h_{DNS}$: (a) fDNS, (b) DSM, (c) DMM, and (d) ANNSGM-7-19.

We apply the same weights and biases obtained from the training at $\Delta = 16h_{DNS}$ due to the assumed independence of the model coefficients g^{s1s2} on filter width as per Eq. (18). In other words, the dependence of the SGS stress on the filter width is completely attributed to the Δ^2 term. This turns out to be not a bold assumption as both the $\Delta = 32h_{DNS}$ and the $\Delta = 64h_{DNS}$ cases yield satisfactory results.

In the numerical configuration, we apply the same type of time advancing scheme as in the $\Delta = 16h_{DNS}$ case. The predictions of the ILES, DSM, and DMM models are also reproduced. Shown in Fig. 11 are the velocity spectra at $\Delta = 32h_{DNS}$ and $\Delta = 64h_{DNS}$. At $\Delta = 32h_{DNS}$, the result of the ANNSGM-7-19 model almost merges with the filtered DNS result. For the $\Delta = 64h_{DNS}$ case, the ANNSGM-7-19 model still gives the most satisfactory prediction among the tested models despite some discrepancies in the middle range.

The PDFs of the SGS stress and SGS flux at $\Delta = 32h_{DNS}$ and $\Delta = 64h_{DNS}$ are displayed in Figs. 12–14. Again, the SGS stress is normalized by the root mean square (rms) value of the true SGS stress similar to Fig. 5. Overall, the result from the ANNSGM-7-19 model is the closest to the filtered DNS result in all these comparisons. Interestingly, the ANNSGM-7-19 model accurately predicts the backscatter in a consistent manner (cf. Fig. 14), whereas the positive branches of the PDFs of the SGS flux are not perfectly recovered at coarser grids. It is worth noting that even though the positive energy transfer is underestimated, numerical instability is

not encountered due to the applied dissipative numerical scheme, as shown in Eq. (22).

The snapshots of the vorticity fields for the $\Delta = 32h_{DNS}$ and $\Delta = 64h_{DNS}$ cases are given in Figs. 15 and 16 at the same instant

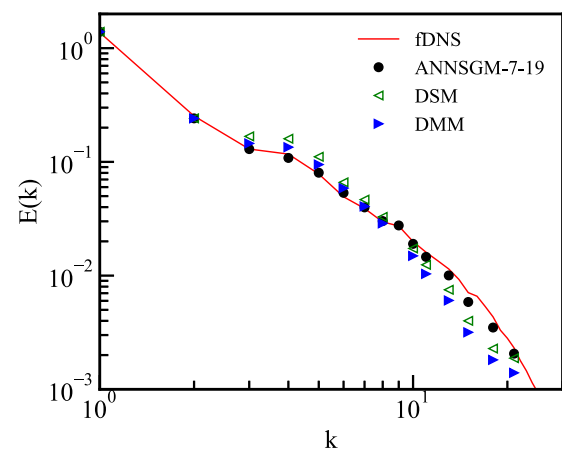


FIG. 17. Spectrum of velocity in the LES for the ANNSGM-7-19 model in the *a posteriori* test at a different forcing, $f(k=1)/f(k=2) \approx 5.79$. Here, the filter width is $\Delta = 32h_{DNS}$.

(i.e., $t/\tau = 5.34$). To compare with the $\Delta = 16h_{DNS}$ case, we take the same color bar limits as in Fig. 9. As Figs. 15 and 16 suggest, the performance of the ANNSGM-7-19 model is quite promising especially for the $\Delta = 32h_{DNS}$ case, for which the differences between the ANNSGM-7-19 model and the other two models (i.e., DSM and DMM) are even larger than the $\Delta = 16h_{DNS}$ case. In addition, the maximum vorticity level is found to drop as the grids become coarser (cf. Fig. 9). In fact, this should be expected since larger vorticity generally exists in smaller vortexes, causing a larger shearing.

Finally, we test the performance of the ANNSGM-7-19 model under a different forcing ratio of $f(k=1)/f(k=2) = 5.79$ in Fig. 17. Here, the Taylor Reynolds number Re_λ is ~ 269 . As observed, the current ANNSGM-7-19 model still gives a quite satisfying prediction, demonstrating its applicability at different flow parameters.

VI. CONCLUSIONS

In this study, the subgrid-scale (SGS) stress of incompressible isotropic turbulence is reconstructed by artificial neural network-based spatial gradient models (ANN-SGMs). The invariants of the velocity gradient tensor are chosen as inputs of the neural network. The dimensionless model coefficients are obtained from the final hidden layer. The influence of the neighboring stencil locations is explicitly incorporated after the final hidden layer in an effort to unravel the intriguing nonlinear structure of the SGS stress.

In the *a priori* analysis, we show that the current model can give quite high correlation coefficients and low relative errors. Inspired by the trends of the model coefficients g^{s1s2} (Table II), we constructed a simplified model (ANNSGM-7-19), which is used to test the *a posteriori* performance. In the *a posteriori* analysis, the spectra and statistical properties of the velocity field as well as the instantaneous vorticity structures predicted by the ANNSGM-7-19 model are comprehensively explored, and good agreements are achieved when compared against the filtered DNS result at grid resolution 128^3 (i.e., $\Delta = 16h_{DNS}$). More importantly, the *a posteriori* efficiency of the current ANN-SGM framework is found to be two orders higher than the previous models of the “black-box” type.

Furthermore, we have tested the applicability of the ANNSGM-7-19 model at larger filter widths $\Delta = 32h_{DNS}$ and $\Delta = 64h_{DNS}$ with grids 64^3 and 32^3 , respectively. Again, the ANNSGM-7-19 model gives satisfactory performance and reasonable agreements are achieved with the filtered DNS result. These observations consistently demonstrate the merits of the current ANN-SGM framework as a promising tool in the reconstruction of SGS stress for large-eddy simulations.

We note that, due to the uniform grids applied, the current work is not intended for wall bounded turbulence or body fitted coordinate systems. In such problems, the applied grid is generally non-uniform. Existing work⁹⁶ has shown that the aspect ratio of the computational mesh can have tangible effects on the flow even in the isotropic case. Consequently, the aspect ratio of the local grid as well as the distance from the wall would have to be considered in the inputs of the ANN so as to potentially build a satisfying SGS model.

Certainly, to construct ANN-based LES models with higher accuracy and wider applications, more knowledge on the physical mechanisms behind LES should be incorporated. In any case, the current semi-explicit approach provides a practical and robust alternative to improve the accuracy of the SGS stress. More importantly,

it sheds light on the potential mathematical forms of the SGS term, inspiring more rigorous theoretical considerations in future works.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (NSFC Grants Nos. 91952104, 92052301, and 91752201), the National Numerical Windtunnel Project (No. NNW2019ZT1-A04), the Shenzhen Science and Technology Program (Grant Nos. KQTD20180411143441009 and JCYJ20170412151759222), and the Department of Science and Technology of Guangdong Province (Grant Nos. 2019B21203001 and 2020B1212030001). This research was also funded by the Center for Computational Science and Engineering of the Southern University of Science and Technology.

APPENDIX A: A PRIORI PERFORMANCE OF THE ANN-SGM FRAMEWORK

We assess the *a priori* accuracy of the ANN-SGM framework in this appendix by examining the correlation coefficient $C(H)$, the relative error $E_r(H)$ and the normalized root-mean-square value $D_N(H)$ defined, respectively, by

$$C(H) = \frac{\langle (H - \langle H \rangle)(H^{model} - \langle H^{model} \rangle) \rangle}{\langle ((H - \langle H \rangle)^2)(H^{model} - \langle H^{model} \rangle)^2 \rangle^{1/2}}, \quad (A1)$$

$$E_r(H) = \frac{\sqrt{\langle (H - H^{model})^2 \rangle}}{\sqrt{\langle H^2 \rangle}}, \quad (A2)$$

$$D_N(H) = \frac{\sqrt{\langle (H^{model} - \langle H^{model} \rangle)^2 \rangle}}{\sqrt{\langle (H - \langle H \rangle)^2 \rangle}}, \quad (A3)$$

where H stands for the real unclosed terms obtained from the filtered DNS field, H^{model} is the modeled term, and $\langle \cdot \rangle$ denotes the averaging over the entire spatial volume.

Table V shows the correlation coefficients, the relative errors, and the normalized root mean square value D_N of τ_{ij} for the different models in the testing set at filter width $\Delta/h_{DNS} = 16$. The traditional velocity gradient model (VGM),^{17,18} given by $\tau_{ij} = \frac{\Delta^2}{12} \frac{\partial \bar{u}_i}{\partial x_k} \frac{\partial \bar{u}_j}{\partial x_k}$, and the scale-similarity model (SSM), given by $\tau_{ij} = \bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j$, are also introduced here for comparison. As can be seen, the ANN-based models have the highest SGS accuracy. Meanwhile, the VGM model performs better than the SSM model. The correlation coefficients are about 0.985 for the ANNSGM-7-49 model and 0.99 for the ANNSGM-13-169 and ANNSGM-19-361 models. The normalized root mean square value D_N also illustrates the good performance of the ANN-based models as they are all very close to unity. With no surprise, the accuracy increases when more spatial information is included. However, the improvement from the ANNSGM-13-169 model to the ANNSGM-19-361 model is minimal, reflecting a saturation.

To illustrate the applicability of the ANN-SGM framework in the *a priori* test, we show the correlation coefficients and relative

TABLE V. Correlation coefficient (C), relative error (E_r), and normalized root mean square value (D^N) of τ_{ij} for the testing set with different models at filter width $\Delta = 16h_{DNS}$.

	τ_{11}	τ_{22}	τ_{33}	τ_{12}	τ_{13}	τ_{23}
C						
ANNSGM-7-49	0.985	0.985	0.984	0.981	0.981	0.982
ANNSGM-13-169	0.989	0.990	0.989	0.987	0.987	0.987
ANNSGM-19-361	0.991	0.991	0.991	0.988	0.988	0.988
VGM	0.945	0.944	0.945	0.946	0.948	0.947
SSM	0.920	0.919	0.921	0.907	0.910	0.909
E_r						
ANNSGM-7-49	0.135	0.134	0.135	0.197	0.194	0.192
ANNSGM-13-169	0.111	0.109	0.111	0.165	0.164	0.161
ANNSGM-19-361	0.106	0.105	0.105	0.157	0.155	0.155
VGM	0.368	0.375	0.369	0.333	0.329	0.332
SSM	0.538	0.545	0.537	0.541	0.536	0.542
D_N						
ANNSGM-7-49	0.972	0.971	0.981	1.019	1.000	1.019
ANNSGM-13-169	0.981	0.981	0.990	1.019	1.000	1.018
ANNSGM-19-361	0.981	0.981	0.981	1.000	1.000	1.000
VGM	0.757	0.757	0.769	0.870	0.857	0.854
SSM	0.512	0.500	0.511	0.566	0.569	0.562

TABLE VI. Correlation coefficient (C) and relative error (E_r) of τ_{ij} for the testing set with different models at filter widths $\Delta = 32h_{DNS}$ and $\Delta = 64h_{DNS}$.

	τ_{11}	τ_{22}	τ_{33}	τ_{12}	τ_{13}	τ_{23}
$C(\Delta = 32h_{DNS})$						
ANNSGM-7-49	0.948	0.947	0.949	0.944	0.943	0.942
VGM	0.901	0.898	0.902	0.912	0.915	0.910
SSM	0.873	0.875	0.873	0.866	0.868	0.867
$E_r(\Delta = 32h_{DNS})$						
ANNSGM-7-49	0.301	0.303	0.298	0.329	0.335	0.338
VGM	0.502	0.515	0.503	0.425	0.421	0.432
SSM	0.649	0.660	0.649	0.616	0.613	0.625
$C(\Delta = 64h_{DNS})$						
ANNSGM-7-49	0.915	0.922	0.919	0.923	0.921	0.915
VGM	0.864	0.866	0.868	0.893	0.894	0.883
SSM	0.827	0.845	0.831	0.857	0.859	0.843
$E_r(\Delta = 64h_{DNS})$						
ANNSGM-7-49	0.355	0.351	0.342	0.402	0.399	0.409
VGM	0.582	0.606	0.585	0.470	0.466	0.490
SSM	0.711	0.730	0.709	0.631	0.619	0.651

TABLE VII. Correlation coefficient (C), relative error (E_r), and normalized root mean square value (D^N) of τ_{ij} for the testing set of the ANNSGM-7-49 model and the ANNSGM-7-19 model at filter width $\Delta = 16h_{DNS}$.

	τ_{11}	τ_{22}	τ_{33}	τ_{12}	τ_{13}	τ_{23}
C						
ANNSGM-7-49	0.985	0.985	0.984	0.981	0.981	0.982
ANNSGM-7-19	0.985	0.985	0.984	0.981	0.981	0.982
E_r						
ANNSGM-7-49	0.135	0.134	0.135	0.197	0.194	0.192
ANNSGM-7-19	0.135	0.134	0.135	0.197	0.196	0.194
D_N						
ANNSGM-7-49	0.972	0.971	0.981	1.019	1.000	1.019
ANNSGM-7-19	0.972	0.971	0.981	1.019	1.018	1.018

errors for $\Delta/h_{DNS} = 32$ and 64 in [Table VI](#), using the model developed for $\Delta/h_{DNS} = 16$ for the seven-point stencil case. As can be seen, the *a priori* accuracy drops from the training accuracy, which is expected for a larger filter width. Nevertheless, the accuracies are still much higher than the VGM model as shown in [Table VI](#). Again, the results of the scale-similarity model (SSM) are shown for comparison. Similar to the $\Delta/h_{DNS} = 16$ case, the SSM model performs slightly worse than the VGM model when $\Delta/h_{DNS} = 16$ and 32 .

APPENDIX B: A PRIORI TEST OF THE ANNSGM-7-19 MODEL

In this appendix, we show the *a priori* performance of the ANNSGM-7-19 model, which contains only the dominant terms, as discussed in [Sec. IV](#). As shown in [Table VII](#), this simplified model exhibits almost the same accuracy except a negligible difference in the relative errors and the normalized rms values. The influence of this finding is twofold: it not only saves the computing efforts in the training process but also facilitates a faster LES.

APPENDIX C: A COMPARISON BETWEEN THE ANNSGM-7-49, ANNSGM-7-19, ANNSGM-1-1, AND VGM MODELS

The energy spectra and PDFs of the SGS flux are compared in this appendix among the ANNSGM-7-19, ANNSGM-7-49, ANNSGM-1-1, and VGM models. The ANNSGM-1-1 model corresponds to the case where only the local gradients are considered. Here, we adopt two different *FGRs* ($FGR = \Delta/h$), which are the ratios of the filter width to the grid space. As shown in [Figs. 18](#) and [19](#), at $FGR = 2$, the predictions of the ANNSGM-7-19 and the ANNSGM-7-49 models are very close to the filtered DNS result. The ANNSGM-1-1 and VGM models yield slightly worse predictions (the ANNSGM-1-1 model performs slightly better than the VGM model), but the differences among these models are not very large. However, the differences become much more pronounced at $FGR = 4$: While the result of the ANNSGM-7-19 model still agrees reasonably well with the filtered DNS result, the ANNSGM-1-1 model performs much worse, showing a strong sensitivity to the

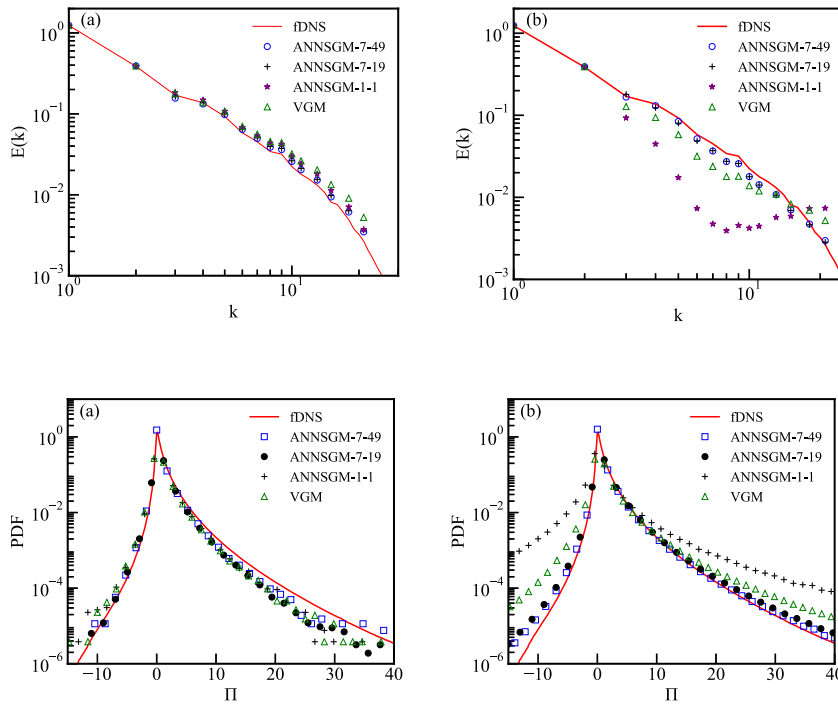


FIG. 18. Spectrum of velocity in the LES for the ANNSGM-7-49, ANNSGM-7-19, ANNSGM-1-1, and VGM models: (a) $FGR = 2$ and (b) $FGR = 4$.

FIG. 19. PDF of the SGS flux in the LES for the ANNSGM-7-49, ANNSGM-7-19, ANNSGM-1-1, and VGM models: (a) $FGR = 2$ and (b) $FGR = 4$.

applied grids (FGRs). Also worth noting is that the VGM model, unlike the ANNSGM-1-1 model, turns out to be less sensitive to the FGR value since its predictions are closer to the filtered DNS result than the ANNSGM-1-1 result at $FGR = 4$. Meanwhile, the differences between the ANNSGM-7-19 and ANNSGM-7-49 models are very little in both cases so that the ANNSGM-7-19 model is sufficient.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- S. B. Pope, *Turbulent Flows* (Cambridge University Press, 2000).
- K. Duraisamy, G. Iaccarino, and H. Xiao, "Turbulence modeling in the age of data," *Annu. Rev. Fluid Mech.* **51**, 537 (2019).
- J. Smagorinsky, "General circulation experiments with the primitive equations, I. The basic experiment," *Mon. Weather Rev.* **91**, 99 (1963).
- D. K. Lilly, "The representation of small-scale turbulence in numerical simulation experiments," in *Proceedings of the IBM Scientific Computing Symposium on Environmental Sciences* (IBM, 1967), p. 195.
- J. W. Deardorff, "A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers," *J. Fluid Mech.* **41**, 453 (1970).
- P. Sagaut, *Large Eddy Simulation for Incompressible Flows* (Berlin Heidelberg Springer, 2006).
- E. Garnier, N. Adams, and P. Sagaut, *Large Eddy Simulation for Compressible Flows* (Springer-Verlag, 2009).
- P. Moin, K. Squires, W. Cabot, and S. Lee, "A dynamic subgrid-scale model for compressible turbulence and scalar transport," *Phys. Fluids A* **3**, 2746 (1991).
- M. Germano, "Turbulence: The filtering approach," *J. Fluid Mech.* **238**, 325 (1992).

- C. Meneveau and J. Katz, "Scale-invariance and turbulence models for large-eddy simulation," *Annu. Rev. Fluid Mech.* **32**, 1 (2000).
- D. I. Pullin, "A vortex-based model for the subgrid flux of a passive scalar," *Phys. Fluids* **12**, 2311 (2000).
- S. Chen, Z. Xia, S. Pei, J. Wang, Y. Yang, Z. Xiao, and Y. Shi, "Reynolds-stress-constrained large-eddy simulation of wall-bounded turbulent flows," *J. Fluid Mech.* **703**, 1 (2012).
- P. A. Durbin, "Some recent developments in turbulence closure modeling," *Annu. Rev. Fluid Mech.* **50**, 77 (2018).
- M. Buzzicotti, M. Linkmann, H. Aluie, L. Biferale, J. Brasseur, and C. Meneveau, "Effect of filter type on the statistics of energy transfer between resolved and sub-filter scales from a-priori analysis of direct numerical simulations of isotropic turbulence," *J. Turbul.* **19**, 167 (2018).
- M. Germano, U. Piomelli, P. Moin, and W. H. Cabot, "A dynamic subgrid-scale eddy-viscosity model," *Phys. Fluids A* **3**, 1760–1765 (1991).
- D. K. Lilly, "A proposed modification of the Germano subgrid scale closure method," *Phys. Fluids A* **4**, 633 (1992).
- J. Bardina, J. H. Ferziger, and W. C. Reynolds, "Improved subgrid scale models for large eddy simulation," AIAA Paper No. 80-1357, 1980.
- S. Liu, C. Meneveau, and J. Katz, "On the properties of similarity subgrid-scale models as deduced from measurements in a turbulent jet," *J. Fluid Mech.* **275**, 83 (1994).
- C. G. Speziale, G. Erlebacher, T. A. Zang, and M. Y. Hussaini, "The subgrid-scale modeling of compressible turbulence," *Phys. Fluids* **31**, 940 (1988).
- G. Erlebacher, M. Y. Hussaini, C. G. Speziale, and T. A. Zang, "Toward the large-eddy simulation of compressible turbulent flows," *J. Fluid Mech.* **238**, 155 (1992).
- T. A. Zang, R. B. Dahlburg, and J. P. Dahlburg, "Direct and large-eddy simulations of three-dimensional compressible Navier Stokes turbulence," *Phys. Fluids A* **4**, 127 (1992).
- Y. Shi, Z. Xiao, and S. Chen, "Constrained subgrid-scale stress model for large eddy simulation," *Phys. Fluids* **20**, 011701 (2008).
- R. A. Clark, J. H. Ferziger, and W. C. Reynolds, "Evaluation of sub-grid scalar models using an accurately simulated turbulent flow," *J. Fluid Mech.* **91**, 1–16 (1979).

- ²⁴J. P. Boris, F. F. Grinstein, E. S. Oran, and R. L. Kolbe, "New insights into large eddy simulation," *Fluid Dyn. Res.* **10**, 199–228 (1992).
- ²⁵E. Garnier, M. Mossi, P. Sagaut, P. Comte, and M. Deville, "On the use of shock-capturing schemes for large-eddy simulation," *J. Comput. Phys.* **153**, 273–311 (1999).
- ²⁶M. R. Visbal, P. E. Morgan, and D. P. Rizzetta, "An implicit LES approach based on high-order compact differencing and filtering schemes (Invited)," AIAA Paper No. 2003–4098, 2003.
- ²⁷N. A. Adams, S. Hickel, and S. Franz, "Implicit subgrid-scale modeling by adaptive deconvolution," *J. Comput. Phys.* **200**, 412 (2004).
- ²⁸F. F. Grinstein, L. G. Margolin, and W. J. Rider, *Implicit Large Eddy Simulation, Computing Turbulent Fluid Dynamics* (Cambridge University Press, 2007).
- ²⁹D. Carati, G. S. Winckelmans, and H. Jeanmart, "On the modelling of the subgrid-scale and filtered-scale stress tensors in large-eddy simulation," *J. Fluid Mech.* **441**, 119–138 (2001).
- ³⁰J. M. Taboada-Vázquez and R. Taboada-Vázquez, "Numerical behaviour of a new LES model with nonlinear viscosity," *J. Comput. Appl. Math.* **377**, 112868 (2020).
- ³¹P. Clark Di Leoni, T. A. Zaki, G. Karniadakis, and C. Meneveau, "Two-point stress-strain-rate correlation structure and non-local eddy viscosity in turbulent flows," *J. Fluid Mech.* **914**, A6 (2021).
- ³²R. D. Moser, S. W. Haering, and G. R. Yalla, "Statistical properties of subgrid-scale turbulence models," *Annu. Rev. Fluid Mech.* **53**, 255–286 (2021).
- ³³T. S. Lund and E. Novikov, *Parameterization of Subgrid-scale Stress by the Velocity Gradient Tensor* (Center for Turbulence Research Annual Research Brief Stanford University, 1992).
- ³⁴C. Meneveau, T. S. Lund, and P. Moin, "Search for subgrid scale parameterization by projection pursuit regression," in *Center for Turbulence Research Proceedings of Summer Program* (Stanford University, 1992).
- ³⁵V. C. Wong, "A proposed statistical-dynamic closure method for the linear and non-linear subgrid-scale stresses," *Phys. Fluids A* **4**, 1080 (1992).
- ³⁶B. Tao, J. Katz, and C. Meneveau, "Geometry and scale relationships in high Reynolds number turbulence determined from three-dimensional holographic velocimetry," *Phys. Fluids* **12**, 941 (2000).
- ³⁷V. M. Canuto and Y. Cheng, "Determination of the Smagorinsky–Lilly constant C_s ," *Phys. Fluids* **9**, 1368 (1997).
- ³⁸B. Kosović, "Subgrid-scale modelling for the large-eddy simulation of high-Reynolds-number boundary layers," *J. Fluid Mech.* **336**, 151 (1997).
- ³⁹B. Tao, J. Katz, and C. Meneveau, "Statistical geometry of subgrid-scale stresses determined from holographic particle image velocimetry measurements," *J. Fluid Mech.* **457**, 35 (2002).
- ⁴⁰B.-C. Wang and D. J. Bergstrom, "A dynamic nonlinear subgrid-scale stress model," *Phys. Fluids* **17**, 035109 (2005).
- ⁴¹F. Sarghini, G. de Felice, and S. Santini, "Neural networks based subgrid scale modeling in large eddy simulations," *Comput. Fluids* **32**, 97 (2003).
- ⁴²J. Ling, A. Kurzawski, and J. Templeton, "Reynolds averaged turbulence modelling using deep neural networks with embedded invariance," *J. Fluid Mech.* **807**, 155 (2016).
- ⁴³J. Ling and J. Templeton, "Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty," *Phys. Fluids* **27**, 085103 (2015).
- ⁴⁴H. Xiao, J.-L. Wu, J.-X. Wang, R. Sun, and C. J. Roy, "Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier–Stokes simulations: A data-driven, physics-informed Bayesian approach," *J. Comput. Phys.* **324**, 115 (2016).
- ⁴⁵J. Ling, R. Jones, and J. Templeton, "Machine learning strategies for systems with invariance properties," *J. Comput. Phys.* **318**, 22 (2016).
- ⁴⁶J.-X. Wang, J.-L. Wu, and H. Xiao, "Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data," *Phys. Rev. Fluids* **2**, 034603 (2017).
- ⁴⁷J. L. Wu, H. Xiao, and E. Paterson, "Physics-Informed Machine Learning Approach for Augmenting Turbulence Models: A Comprehensive Framework," *Phys. Rev. Fluids* **3**, 074602 (2018).
- ⁴⁸Z. Wang, K. Luo, D. Li, J. Tan, and J. Fan, "Investigations of data-driven closure for subgrid-scale stress in large-eddy simulation," *Phys. Fluids* **30**, 125101 (2018).
- ⁴⁹R. Maulik, O. San, A. Rasheed, and P. Vedula, "Data-driven deconvolution for large eddy simulations of Kraichnan turbulence," *Phys. Fluids* **30**, 125109 (2018).
- ⁵⁰L. Zhu, W. Zhang, J. Kou, and Y. Liu, "Machine learning methods for turbulence modeling in subsonic flows around airfoils," *Phys. Fluids* **31**, 015105 (2019).
- ⁵¹M. Yu, W.-X. Huang, and C.-X. Xu, "Data-driven construction of a reduced-order model for supersonic boundary layer transition," *J. Fluid Mech.* **874**, 1096–1114 (2019).
- ⁵²J. Wu, H. Xiao, R. Sun, and Q. Wang, "Reynolds averaged Navier Stokes equations with explicit data driven Reynolds stress closure can be ill conditioned," *J. Fluid Mech.* **869**, 553–586 (2019).
- ⁵³A. Beck, D. Flad, and C.-D. Munz, "Deep neural networks for data-driven turbulence models," *J. Comput. Phys.* **398**, 108910 (2019).
- ⁵⁴X. I. A. Yang, S. Zafar, J.-X. Wang, and H. Xiao, "Predictive large-eddy-simulation wall modeling via physics-informed neural networks," *Phys. Rev. Fluids* **4**, 034602 (2019).
- ⁵⁵C. Xie, J. Wang, H. Li, M. Wan, and S. Chen, "Artificial neural network mixed model for large eddy simulation of compressible isotropic turbulence," *Phys. Fluids* **31**, 085112 (2019).
- ⁵⁶J. Han, C. Ma, Z. Ma, and W. E., "Uniformly accurate machine learning-based hydrodynamic models for kinetic equations," *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21983–21991 (2019).
- ⁵⁷M. Gamahara and Y. Hattori, "Searching for turbulence models by artificial neural network," *Phys. Rev. Fluids* **2**(5), 054604 (2017).
- ⁵⁸A. Volland, G. Balarac, and C. Corre, "Subgrid-scale scalar flux modelling based on optimal estimation theory and machine-learning procedures," *J. Turbul.* **18**(9), 854 (2017).
- ⁵⁹C. Ma, J. Wang, and W. E., "Model reduction with memory and the machine learning of dynamical systems," *Commun. Comput. Phys.* **25**(4), 947–962 (2019).
- ⁶⁰R. Maulik, O. San, A. Rasheed, and P. Vedula, "Subgrid modelling for two-dimensional turbulence using neural networks," *J. Fluid Mech.* **858**, 122 (2019).
- ⁶¹J. Park and H. Choi, "Toward neural-network-based large eddy simulation: Application to turbulent channel flow," *J. Fluid Mech.* **914**, A16 (2021).
- ⁶²A. Subel, A. Chattopadhyay, Y. Guan, and P. Hassanzadeh, "Data-driven subgrid-scale modeling of forced Burgers turbulence using deep learning with generalization to higher Reynolds numbers via transfer learning," *Phys. Fluids* **33**, 031702 (2021).
- ⁶³Z. Zhou, G. He, S. Wang, and G. Jin, "Subgrid-scale model for large-eddy simulation of isotropic turbulent flows using an artificial neural network," *Comput. Fluids* **195**, 104319 (2019).
- ⁶⁴C. Xie, J. Wang, K. Li, and C. Ma, "Artificial neural network approach to large-eddy simulation of compressible isotropic turbulence," *Phys. Rev. E* **99**, 053113 (2019).
- ⁶⁵C. Xie, K. Li, C. Ma, and J. Wang, "Modeling subgrid-scale force and divergence of heat flux of compressible isotropic turbulence by artificial neural network," *Phys. Rev. Fluids* **4**, 104605 (2019).
- ⁶⁶C. Xie, J. Wang, H. Li, M. Wan, and S. Chen, "Spatial artificial neural network model for subgrid-scale stress and heat flux of compressible turbulence," *Theor. Appl. Mech. Lett.* **10**, 27 (2020).
- ⁶⁷C. Xie, J. Wang, and W. E., "Modeling subgrid-scale forces by spatial artificial neural networks in large eddy simulation of turbulence," *Phys. Rev. Fluids* **5**, 054606 (2020).
- ⁶⁸B. T. Nadiga and D. Livescu, "Instability of the perfect subgrid model in implicit-filtering large eddy simulation of geostrophic turbulence," *Phys. Rev. E* **75**, 046303 (2007).
- ⁶⁹C. Xie, J. Wang, H. Li, M. Wan, and S. Chen, "Spatially multi-scale artificial neural network model for large eddy simulation of compressible isotropic turbulence," *AIP Adv.* **10**, 015044 (2020).
- ⁷⁰C. Xie, Z. Yuan, and J. Wang, "Artificial neural network-based nonlinear algebraic models for large eddy simulation of turbulence," *Phys. Fluids* **32**, 115101 (2020).

- ⁷¹M. Chamecki, C. Meneveau, and M. B. Parlange, "The local structure of atmospheric turbulence and its effect on the Smagorinsky model for large eddy simulation," *J. Atmos. Sci.* **64**, 1941–1958 (2007).
- ⁷²Z. Yuan, C. Xie, and J. Wang, "Deconvolutional artificial neural network models for large eddy simulation of turbulence," *Phys. Fluids* **32**, 115106 (2020).
- ⁷³G. L. Eyink and H. Aluie, "Localness of energy cascade in hydrodynamic turbulence. I. Smooth coarse graining," *Phys. Fluids* **21**, 115107 (2009).
- ⁷⁴H. Aluie, "Compressible turbulence: The cascade and its locality," *Phys. Rev. Lett.* **106**, 174502 (2011).
- ⁷⁵T. Ishihara, T. Gotoh, and Y. Kaneda, "Study of high-Reynolds number isotropic turbulence by direct numerical simulation," *Annu. Rev. Fluid Mech.* **41**, 165–180 (2009).
- ⁷⁶A. Leonard, "Energy cascade in large-eddy simulations of turbulent fluid flows," *Adv. Geophys.* **18**, 237–248 (1974).
- ⁷⁷B. Vreman, B. Geurts, and H. Kuerten, "Subgrid-modeling in LES of compressible flow," *Appl. Sci. Res.* **54**, 191 (1995).
- ⁷⁸S. B. Pope, "A more general effective-viscosity hypothesis," *J. Fluid Mech.* **72**, 331–340 (1975).
- ⁷⁹X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 2010.
- ⁸⁰G. Zhang, B. Eddy Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecast.* **14**, 35 (1998).
- ⁸¹H. B. Demuth, M. H. Beale, O. D. Jess, and M. T. Hagan, *Neural Network Design* (Martin Hagan, Stillwater, OK, 2014).
- ⁸²J. Wang, M. Wan, S. Chen, and S. Chen, "Kinetic energy transfer in compressible isotropic turbulence," *J. Fluid Mech.* **841**, 581 (2018).
- ⁸³T. Ishihara, Y. Kaneda, M. Yokokawa, K. Itakura, and A. Uno, "Small-scale statistics in high-resolution direct numerical simulation of turbulence: Reynolds number dependence of one-point velocity gradient statistics," *J. Fluid Mech.* **592**, 335–366 (2007).
- ⁸⁴G. S. Patterson and S. A. Orszag, "Spectral calculations of isotropic turbulence: Efficient removal of aliasing interactions," *Phys. Fluids* **14**, 2538–2541 (1971).
- ⁸⁵J. Wang, L.-P. Wang, Z. Xiao, Y. Shi, and S. Chen, "A hybrid numerical simulation of isotropic compressible turbulence," *J. Comput. Phys.* **229**, 5257 (2010).
- ⁸⁶D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
- ⁸⁷C. Xie, J. Wang, H. Li, M. Wan, and S. Chen, "A modified optimal LES model for highly compressible isotropic turbulence," *Phys. Fluids* **30**, 065108 (2018).
- ⁸⁸S. Ghosal, "An analysis of numerical errors in large-eddy simulations of turbulence," *J. Comput. Phys.* **125**, 187–206 (1996).
- ⁸⁹J. Meyers, B. J. Geurts, and M. Baelmans, "Database analysis of errors in large-eddy simulation," *Phys. Fluids* **15**, 2740 (2003).
- ⁹⁰F. K. Chow and P. Moin, "A further study of numerical errors in large-eddy simulations," *J. Comput. Phys.* **184**, 366–380 (2003).
- ⁹¹O. J. Mcmillan and J. H. Ferziger, "Direct testing of subgrid scale models," *AIAA J.* **17**, 1340–1346 (1979).
- ⁹²S. B. Pope, "Ten questions concerning the large-eddy simulation of turbulent flows," *New J. Phys.* **6**, 35–59 (2004).
- ⁹³S. T. Bose, P. Moin, and D. You, "Grid-independent large-eddy simulation using explicit filtering," *Phys. Fluids* **22**, 105103 (2010).
- ⁹⁴S. Radhakrishnan and J. Bellan, "Explicit filtering to obtain grid-spacing-independent and discretization-order-independent large-eddy simulation of compressible single-phase flow," *J. Fluid Mech.* **697**, 399–435 (2012).
- ⁹⁵U. Piomelli, A. Rouhi, and B. J. Geurts, "A grid-independent length scale for large-eddy simulations," *J. Fluid Mech.* **766**, 499–527 (2015).
- ⁹⁶S. W. Haering, M. Lee, and R. D. Moser, "Resolution-induced anisotropy in large-eddy simulations," *Phys. Rev. Fluids.* **4**, 114605 (2019).