



Title	A priori Test of Data-Driven Subgrid-Scale Model with Few-shot Learning for Large-Eddy Simulation of Turbulent Channel Flow
Author(s)	Jalaali, Bahrul; Okabayashi, Kie
Citation	サイバーメディアHPCジャーナル. 2026, 16, p. 138-141
Version Type	VoR
URL	https://doi.org/10.18910/106256
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

A priori Test of Data-Driven Subgrid-Scale Model with Few-shot Learning for Large-Eddy Simulation of Turbulent Channel Flow

Bahrul Jalaali, Kie Okabayashi

Department of Mechanical Engineering, Graduate School of Engineering, The University of Osaka

1. Introduction

The use of numerical simulations is essential for investigations in turbulent flow. Among of them, large-eddy simulation (LES) has become widely adopted because it is considered cost-effective as well as more accurate than other coarse-grained models: In LES, the large vortices are explicitly resolved whereas the small vortices are modelled using a subgrid-scale (SGS) model. The SGS model has long been represented by mathematical models and has achieved a certain degree of success. Recently, there has been a rapid surge in the development of SGS models using deep neural networks (DNN), hereafter referred to as data-driven SGS models. Unlike mathematical models, these have the advantage of not involving any artificial assumptions or approximations. However, the data-driven SGS model often faces a generalization issue, where it refers to the ability of a data-driven SGS model to perform well on new “unseen” data, which are not fed in the training phase.

Several studies have been proposed to address this issue. Guan et al.¹ introduced a fine-tuning transfer learning (TL) technique to train data-driven SGS models on homogeneous turbulence. In their approach, DNN model that has been pre-trained using low Reynolds number (Re) dataset was re-trained using a dataset at higher Re, which is more complex and challenging condition than lower Re case. By using this scheme, the data-driven SGS model was able to predict higher-Re case with reduced training time. Although this model can predict higher-Re case, the capability to retain ability by predicting lower-Re remain questioned. Kim et al.² proposed a fusion DNN approach that combines channel flow and cylinder flow datasets on training phase to improve predictions across different domains.

This approach exhibited a good prediction for SGS quantities in channel and cylinder flow. Nevertheless, incorporating data from both domains during training may not be regarded as generalization. In addition, for a data-driven SGS model to be generalized and applicable to various flow fields, a large amount of training data is required. Generating such massive high-fidelity turbulence datasets is impractical due to the immense computational cost. These challenges are frequently highlighted in data-driven turbulence modeling³, making their application in engineering problems difficult.

To overcome the issues, this study introduces a data-driven SGS model with a TL algorithm to reduce the need for extensive re-training or large datasets for new flow condition. In this study, the analysis object is turbulent channel flow between two parallel flat plates, illustrated in Fig. 1. It is important to note that common data-driven model with TL is often experienced catastrophic forgetting where the model retrained using only new dataset may lose its ability to predict the previous knowledge. Therefore, the aim of this study is to predict high-Re case while maintaining accuracy to predict low-Re case. Here, we extended the multiscale convolutional neural network SGS model (MSC-SGS) introduced by Jalaali & Okabayashi^{4,5}, by incorporating a novel few-shot transfer learning (FSL) algorithm, hereinafter regarded to as MSC-FSL. The MSC-FSL model was re-trained on higher-Re case with merely a few portions of dataset. Despite that, the MSC-FSL model was able to maintain accurate prediction across both low- and high-Re case.

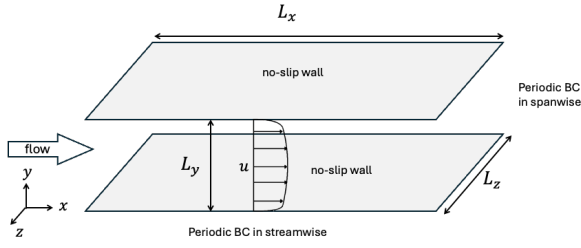


Figure 1. Channel flow domain with its boundary condition (BC).

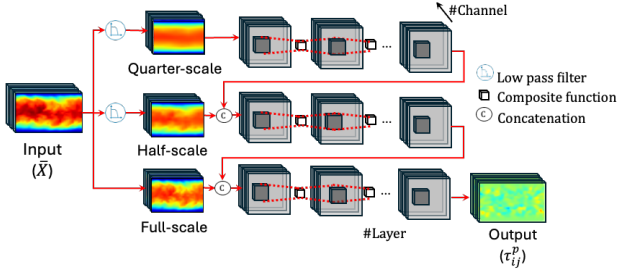


Figure 2. Schematic illustration of the MSC-SGS model network structure.

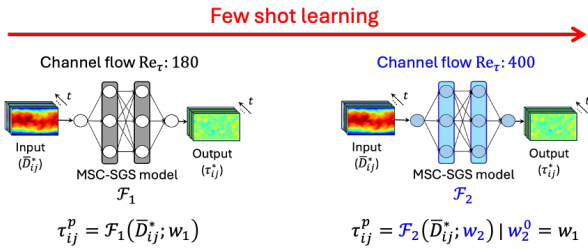


Figure 3. Schematic illustration of FSL algorithm. w_2^0 denotes the initial value of w_2 before retraining. $\bar{D}_{ij}^*$ is the normalized $\bar{D}_{ij}$.

2. Methodology

The problem setting is to develop a data-driven SGS model for LES that is capable of predicting both low- and high-Re. LES computation is conducted by solving the spatially filtered forms of the continuity and Navier–Stokes equations expressed in Eqs. (1) and (2).

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \quad (1)$$

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{\partial \bar{p}}{\partial x_i} + \frac{1}{\text{Re}_\tau} \frac{\partial}{\partial x_j} (-\tau_{ij} + 2\bar{D}_{ij}) \quad (2)$$

The overbar denotes the filtering operation, Re_τ is the friction Re which is specific flow parameter for channel flows, and $\bar{D}_{ij} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right)$, is the grid-scale rate-of-strain tensor. Here, $\bar{u}$ and $\bar{p}$ denotes the filtered velocity

and pressure, respectively. The SGS components, $\tau_{ij} = \overline{u_i u_j} - \bar{u}_i \bar{u}_j$, are unclosed term and need to be modeled. Here, the MSC-FSL model is proposed to predict τ_{ij} . The present study is restricted to an *a priori* test, where the predictive capability of MSC-FSL model is evaluated without its implementation in actual LES framework.

The MSC-FSL model is developed from the MSC-SGS model^{4,5}. The schematic of MSC-SGS model is depicted in Fig. 2. The MSC-SGS model was trained in a supervised learning. Predicted τ_{ij} (τ_{ij}^p) can be defined as $\tau_{ij}^p = \mathcal{F}(\bar{D}_{ij}; w)$, where $\mathcal{F}$ is the DNN model, $\bar{D}_{ij}$ is the input quantity, and w indicates the weight of the DNN. The training process aims to determine the optimum weight such that $w = \text{argmin}_w (\mathcal{L}(\tau_{ij}^p, \tau_{ij}^{\text{fDNS}}))$, where $\mathcal{L}$ denotes the loss function. For a more detailed explanation of the MSC-SGS model readers are referred to the study by Jalaali & Okabayashi^{4,5}.

Subsequently, the training procedure for MSC-FSL model is described where the schematic illustration of MSC-FSL framework is depicted in Fig. 3. The first step of training is to set the input quantity. Here the input and output normalization are employed; hence, the MSC-FSL can be formulated as:

$$\tau_{ij}^p / k_{\text{sgs}} = \mathcal{F}_2(\bar{D}_{ij} \delta / U_{\text{bulk}}, w),$$

where k_{sgs} is the SGS kinetic energy ($k_{\text{sgs}} = \frac{1}{2} (\overline{u_k u_k} - \bar{u}_k \bar{u}_k)$), δ denotes the channel half-width, and U_{bulk} is the bulk velocity of channel flow. This normalization is intended to represent the similar physical characteristic regardless of Re_τ .

Subsequently, the MSC-FSL will be trained with low-Re dataset, $\text{Re}_\tau = 180$, where this model is regarded to as $\mathcal{F}_1$. The training dataset is obtained from filtered direct numerical simulation (fDNS) data of a turbulent channel flow. During this training process, a total of 12,000 flow-field samples, each comprising paired input ($\bar{D}_{ij}^{\text{fDNS}} \delta / U_{\text{bulk}}$) and label data ($\tau_{ij}^{\text{fDNS}} / k_{\text{sgs}}^{\text{fDNS}}$), were used for training.

After the training of model $\mathcal{F}_1$ has converged, the TL procedure was performed by transferring the trained

parameters w_1 from $\mathcal{F}_1$ to new DNN model, regarded to as $\mathcal{F}_2$, which will be served as MSC-FSL model after retraining. These parameters w_1 were used to initialize the subsequent w_2 retraining process. Next, $\mathcal{F}_2$ was then retrained using datasets that consist of both $Re_\tau = 180$ and $Re_\tau = 400$ fDNS data, each comprising 600 flow-field samples. The use of this limited amount of training data constituted the key feature of FSL strategy. It is worth noting that FSL adopted in this study has a different characteristic than common FSL in computer science. The distinguishing characteristics of the novel FSL framework are described as follows. During retraining, only weight the parameters associated with the half-scale and quarter-scale branches were updated. In contrast, the weight parameters of the full-scale branch were frozen and retained the values of previous w_1 . This approach is based on the universal characteristic of a small scale, regardless of Re . Therefore, it is assumed that $w_1 = w_2$ for a full-scale branch. The retraining process was conducted for 400 epochs.

3. Result and discussion

For comparison, this study presents the τ_{ij} prediction using the baseline MSC-SGS model^{4,5}. Here, the MSC-SGS model did not employ normalization and FSL. Fig. 4 shows the wall-normal distributions of τ_{ij} , averaged over the $x - z$ directions for $Re_\tau = 180$. In general, both MSC-SGS and MSC-FSL models accurately predicted all six components of τ_{ij} , despite of overestimated τ_{12} in MSC-SGS model. It is important to note that one of the limitations of TL-based model is catastrophic forgetting, in which the model lost previous acquired knowledge after retrained using new dataset. In this study, the MSC-FSL model was able to retain its predictive accuracy for low- Re case, $Re_\tau = 180$, even after being retrained using new dataset which included $Re_\tau = 400$. This result indicates that the MSC-FSL model mitigated catastrophic forgetting and preserved the knowledge acquired from the low- Re case.

The wall-normal distribution of the averaged τ_{ij} at

$Re_\tau = 400$ is depicted in Fig. 5. Consistent with the results on low- Re case, the MSC-FSL model accurately predicted all six components of τ_{ij} . Nevertheless, the predicted τ_{12} showed a noticeable discrepancy as in higher- Re , vortices interactions were more intricate, resulting relatively more difficult to predict. However, more importantly, despite being retrained using only a few amounts of dataset in high- Re , the MSC-FSL model was relatively performed in satisfactory for predicting τ_{ij} in $Re_\tau = 400$. In contrast, the baseline MSC-SGS model failed to predict high- Re .

The combination of nondimensional normalization and the novel FSL therefore provided effective generalization across both lower- and higher- Re cases. The ability to predict higher- Re flows is particularly important because such flows exhibit a broader range of interacting turbulent scales.

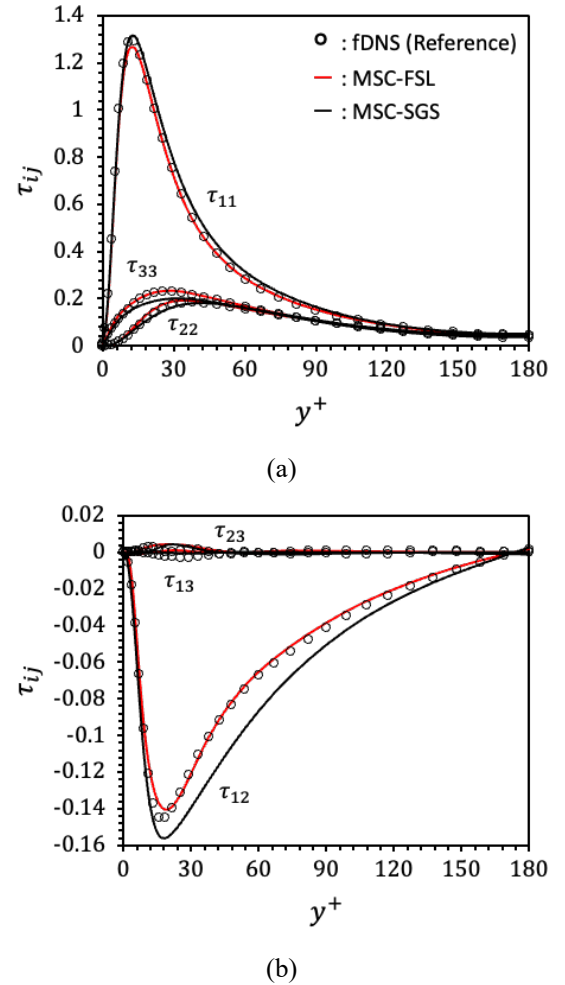
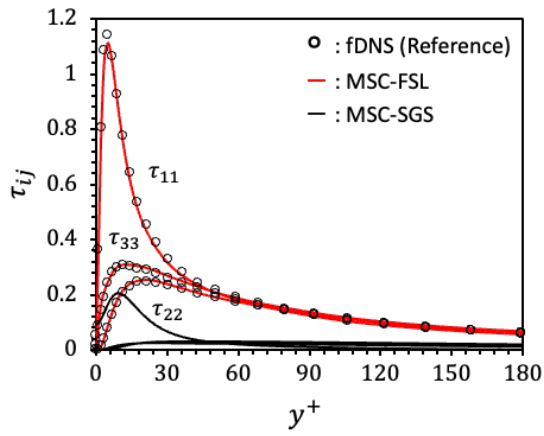
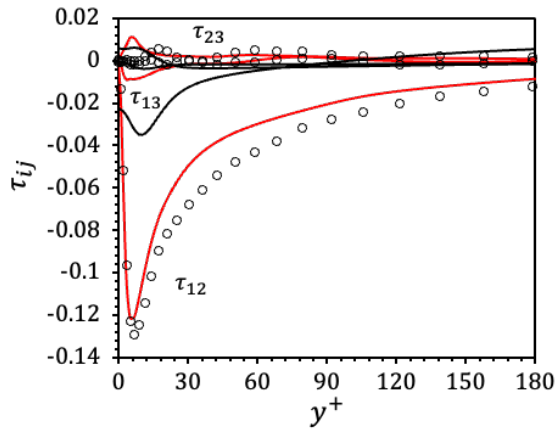


Figure 4. Wall-normal distribution of τ_{ij} for $Re_\tau = 180$ for (a) isotropic and (b) anisotropic component.



(a)



(b)

Figure 5. Wall-normal distribution of τ_{ij} for $Re_\tau = 400$ for (a) isotropic and (b) anisotropic component.

4. Conclusion and future direction

The MSC-FSL model was developed by incorporating FSL into a baseline MSC-SGS model⁴. This model exhibited a generalization capability by predicting SGS quantities across different Reynolds number cases which is an objective that remains challenging for common data-driven SGS model. Moreover, this model was also successfully mitigated the catastrophic forgetting, a common issue in TL models. As the next step, implementation into actual LES framework will be conducted.

Acknowledgements

This work was partly supported by the Research Proposal-based Use of the Project for Nurturing Student Competing with the World at the Large-Scale Computer

System-D3 Center, The University of Osaka. This study was financially supported by JSPS KAKENHI grant No. JP25K07584 and Indonesian Education Scholarship Program (BPI), administered by the Center for Higher Education Funding and Assessment (PPAPT), Ministry of Higher Education, Science, and Technology of the Republic of Indonesia, and the Indonesia Endowment Fund for Education (LPDP), Ministry of Finance of the Republic of Indonesia, BPI No. 202205080366.

Reference:

- (1) Y. Guan, A., et al. *Journal of Computational Physics*. (2022).
- (2) M. Kim., et al. *J. Fluid Mech.* (2024).
- (3) K. Duraisamy., et al. *Annu. Rev. Fluid Mech.* (2019).
- (4) Jalaali, B., Okabayashi, K. *Journal of Turbulence*. (2026).
- (5) Jalaali, B., Okabayashi, K. *Cybermedia HPC Journal*. (2025).