The Multiscale-based Data-Driven Subgrid-Scale Model for Large-Eddy Simulation of Wall-Bounded Turbulent Flow

Bahrul Jalaali, Kie Okabayashi

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

1. Introduction

Turbulent flows are ubiquitous in both natural and industrial fields. It has a critical role in various applications, such as aircraft, turbomachinery, and weather forecasting. Characterized by chaotic and unsteady motion, turbulence involves a broad range of vortices across different scales. The use of numerical simulations with reliable computational methods serves as valuable tools for fundamental research in turbulent flow. A direct numerical simulation (DNS) resolves all turbulence scales, but its computational cost is prohibitive for practical use due to its immense computational resources on fine mesh. Among existing modeling approaches, large-eddy simulation (LES) has become a widely used compromise between accuracy and computational efficiency. In LES, large-scale vortices are directly resolved using the numerical scheme, while a turbulence model is introduced to account for the small-scale vortices. The scales that are explicitly resolved are referred to as grid-scale (GS), whereas the small-scale are known as subgrid-scale (SGS) or residual stress (τ_{ij}), with the latter represented by SGS model. Many SGS models have been proposed such as the Smagorinsky (SMAG) model which is one of the most typical SGS models. However, these conventional models, which rely on linear eddy viscosity assumptions, often struggle to accurately capture the complex features of SGS.

Recently, advances in computational resource and machine learning have enabled the development of data-driven SGS models using deep neural networks (DNN). The data-driven SGS model aims to predict τ_{ij} without relying on explicit mathematical and physical assumptions. Early efforts, such as Gamahara & Hattori[1], used multilayer perceptrons (MLP) trained on filtered DNS (fDNS) data to predict τ_{ij}, while later Liu et al.[2] demonstrated an improved performance using convolutional neural networks (CNN) due to their ability on capturing spatial features. Despite this potential of data-driven SGS models, capturing the multiscale dynamics of turbulent fields remains challenging, particularly due to the nature of turbulent energy cascade where kinetic energy is progressively transferred from large to smaller vortices. To address this, we propose a multiscale CNN-based SGS model (MSC-SGS) which incorporates features across multiple spatial resolutions ranging from large to small vortices and evaluate its effectiveness on predicting τ_{ij}.

2. Methodology

The problem setting of this study is wall-bounded turbulent channel flow between two parallel flat plates driven by a constant pressure gradient with Reynolds number (Re) of 180. LES computation is conducted by solving the spatially filtered forms of the continuity and Navier–Stokes equations expressed in Eqs. (1) and (2).

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

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \overline{u_\iota}\,\overline{u_j}}{\partial x_j} = -\frac{\partial \bar{p}}{\partial x_i} + \frac{1}{\mathrm{Re}}\frac{\partial}{\partial x_j}\left(-\tau_{ij} + 2\bar{D}_{ij}\right) \quad (2)$$

The overbar denotes the filtering operation 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 GS rate-of-strain tensor. Here, $\bar{u}$ and $\bar{p}$ denotes the velocity and pressure, respectively. The SGS components, $\tau_{ij} = \overline{u_\iota u_j} - \overline{u_\iota}\,\overline{u_j}$, are unclosed term and need to be modeled. Here, the data-driven SGS model is proposed to predict τ_{ij}. A schematic diagram of data-driven SGS framework is depicted in Fig. 1. The data-driven SGS model is trained in a supervised manner using fDNS dataset that comprises a set of input

data X_{tr} and label data $\left(\tau_{ij}^{fDNS}\right)$. The training step (*a priori* test) aims to establish a functional relation between the input and output variables, thus yielding a nonlinear regression $\tau_{ij}^{p} = \mathcal{F}(X_{tr}; w)$, where w indicates the weight of the neural network and τ_{ij}^{p} is the predicted τ_{ij}. The training process aims to determine the optimum weight such that $w = \text{argmin}_w \left(\mathcal{L}\left(\tau_{ij}^{p}, \tau_{ij}^{fDNS}\right)\right)$, where $\mathcal{L}$ denotes the loss function. Subsequently, the trained data-driven SGS model is implemented in the CFD framework of actual LES computation to assess its performance (*a posteriori* test).

Here, the dataset is obtained using DNS data of a turbulent channel flow where it has been validated by comparing with Kim et al.[3]. DNS data are subsequently filtered using box filter to separate the GS and SGS components, resulting in the fDNS data. To satisfy the Galilean and rotational invariant of turbulence model, the GS variable of $\bar{D}_{ij}$ is used as the input variable. Therefore, the input (X_{tr}) and label data τ_{ij}^{fDNS} can be defined in Eqs. (3) and (4).

$$X_{tr} = \{x \in R^{6\times N_x\times N_y\times N_z} : x = \bar{D}_{ij}\} \quad (3)$$

$$\tau_{ij}^{fDNS} = \{\tau_{ij}^{fDNS} \in R^{6\times N_x\times N_y\times N_z}\} \quad (4)$$

As illustrated in Fig. 2, the MSC-SGS model incorporates multiscale representation of turbulent fields. The input features are separated into a quarter, half, and full scale which corresponds to large-, intermediate-, and full-scale vortices by using low-pass filter operation. Next, these inputs are sequentially encoded and then concatenated to form a comprehensive multiscale representation. To account for three-dimensional spatial interactions in turbulent flows, the MSC-SGS model utilizes 3D convolutional kernels with a uniform size of 3 across each spatial direction. For a more profound explanation of the MSC-SGS model and its training procedure, readers are referred to the study by Jalaali & Okabayashi[4].

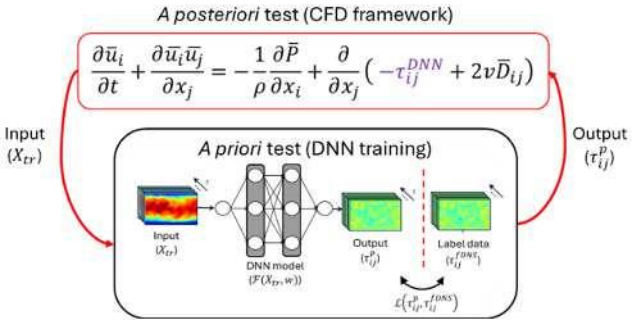

Figure 1. Schematic diagram of data-driven SGS framework.

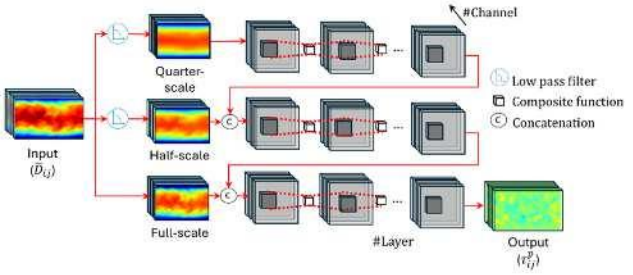

Figure 2. Schematic illustration of the MSC-SGS model network structure.

3. Result and discussion

In *a priori* test, the correlation coefficient (CC) expressed in Eq. (5) is introduced to evaluate the performance of the data-driven SGS model where $\langle . \rangle$ denotes the ensemble average in the streamwise–spanwise (x-z) direction and in time.

$$CC = \frac{\langle\left(\tau_{ij}^{fDNS} - \langle\tau_{ij}^{fDNS}\rangle\right)\left(\tau_{ij}^{p} - \langle\tau_{ij}^{p}\rangle\right)\rangle}{\sqrt{\langle\left(\tau_{ij}^{fDNS} - \langle\tau_{ij}^{fDNS}\rangle\right)^2\rangle}\sqrt{\langle\left(\tau_{ij}^{p} - \langle\tau_{ij}^{p}\rangle\right)^2\rangle}} \quad (5)$$

Figure 3 shows the correlation coefficient (CC) as a function of wall-normal distance (y^+), averaged over all components of τ_{ij}. The MSC-SGS model consistently yields high CC values, particularly in the region $5 < y^+ < 30$, where shear-dominated turbulence and SGS dynamics are prominent. A reduction is observed near the wall ($y^+ < 5$) due to diminished flow variability, while high correlations are observed for $y^+ > 30$. Compared to previous studies of Gamahara & Hattori[1], Bose & Roy[5], and Park & Choi[6], the MSC-SGS model demonstrates higher overall CC value, underscoring the advantage of incorporating multiscale information.

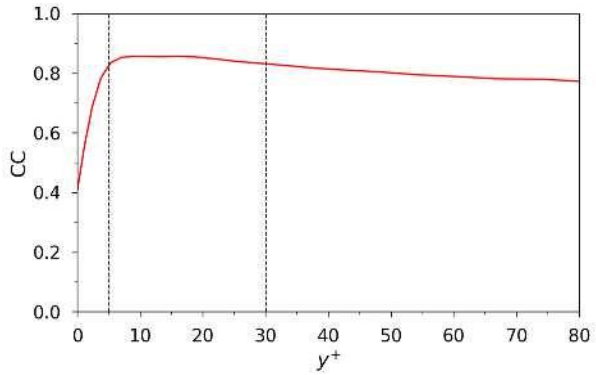

Figure 3. CC averaged in streamwise–spanwise (x-z) direction and in time between τ_{ij}^{p} and τ_{ij}^{fDNS}.

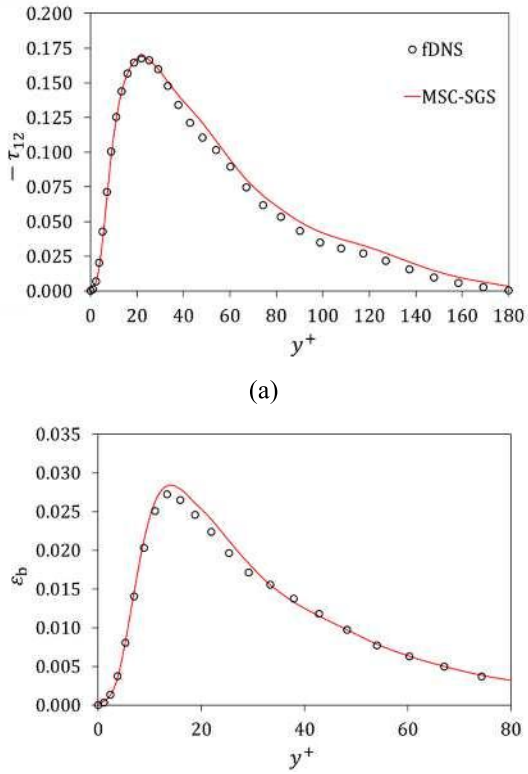

Figure 4. Wall-normal distribution of (a) τ_{12} and (b) SGS backscatter

Figure 4(a) shows wall-normal distribution of τ_{12}, the dominant shear stress component in wall-bounded flows. Fig. 4(b) presents the SGS backscatter $\left(\varepsilon_b = -\frac{1}{2}\langle \varepsilon_{sgs} - |\varepsilon_{sgs}| \rangle\right)$ where $\varepsilon_{sgs} = -\tau_{ij}\overline{D}_{ij}$ represents SGS backscatter (energy transfer from small to large fluctuations) and accurately capturing this term is important for turbulence modeling. From the results, the MSC-SGS model closely aligns fDNS results for both quantities. Nevertheless, as noted in Park & Choi[6] and Duraisamy[7], *a priori* result does not guarantee accuracy in *a posteriori* simulations, prompting further LES validation of the model.

In *a posteriori* test, the data-driven SGS model is implemented within an LES of turbulent channel flow, under conditions similar to those used in the DNS computation. The computational domain and grid resolution are $(L_x \times L_y \times L_z) = (2\pi\delta \times 2\delta \times \pi\delta)$ and $(n_x \times n_y \times n_z) = (32 \times 64 \times 32)$, respectively. For comparison, the conventional mathematics model of SMAG model with van Driest damping is also evaluated.

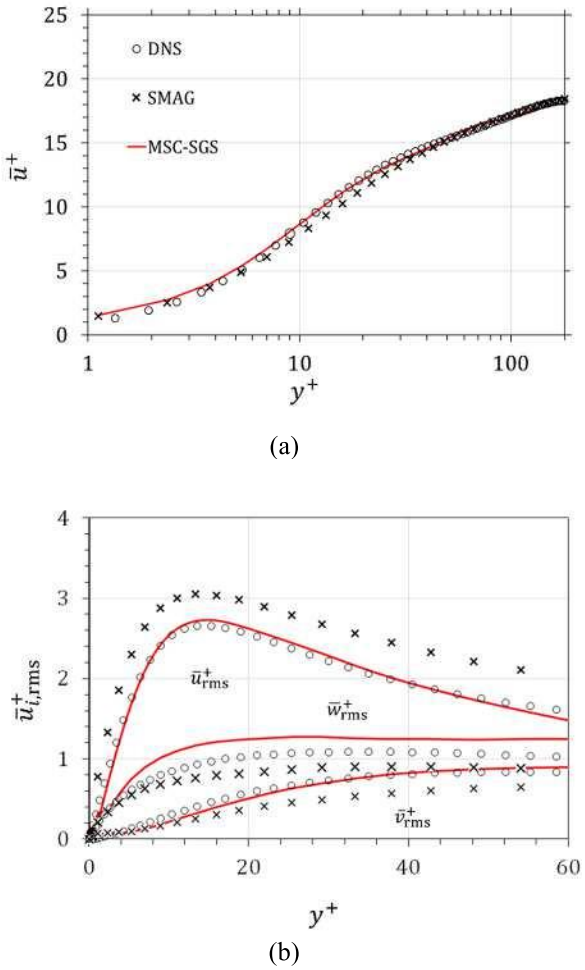

Figure 5. Wall-normal distribution of turbulence statistics of *a posteriori* result for (a) mean velocity and (b) root-mean-square velocity.

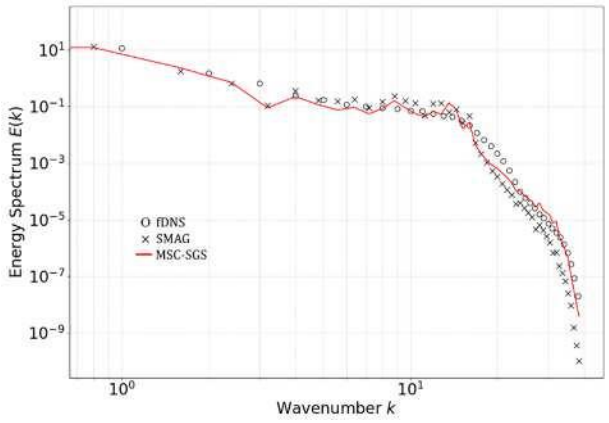

Figure 6. Energy spectrum of velocity fluctuations.

Figure 5(a) compares the mean velocity profiles of the MSC-SGS model, SMAG model, and DNS. The MSC-SGS model closely matches DNS mean velocity profile, while the SMAG model exhibits deviation. These discrepancies can be attributed to its reliance on a predefined eddy viscosity assumption. Fig. 5(b) shows the root-mean-square (rms) velocity statistics. In comparison with the DNS results, the SMAG model tends to overestimate $\bar{u}^{+}_{\mathrm{rms}}$, while underestimating both $\bar{v}^{+}_{\mathrm{rms}}$ and $\bar{w}^{+}_{\mathrm{rms}}$ which indicates inadequate energy redistribution. On the other hand, although the MSC-SGS model slightly overpredicted the DNS for the $\bar{w}^{+}_{\mathrm{rms}}$, it produces the $\bar{u}^{+}_{\mathrm{rms}}$ and $\bar{v}^{+}_{\mathrm{rms}}$ well owing to the appropriate energy redistribution.

Figure 6 presents the three-dimensional energy spectra of the velocity fluctuations. At low wavenumbers, both SMAG and MSC-SGS models agree well with fDNS, capturing large-scale structures. However, at high wavenumbers, the SMAG model underpredicts energy due to excessive dissipation. In contrast, the MSC-SGS model aligns closely with fDNS across all scales, indicating its ability to preserve fine-scale dynamics and accurately represent the energy cascade.

4. Conclusion and Future direction

In this study, the multiscale algorithm of the multiscale CNN-based SGS model (MSC-SGS) model was employed to predict the SGS residual stress τ_{ij}. The model encoded multiscale input representations obtained through low-pass filtering at various scales. *A priori* results showed that the MSC-SGS model consistently achieved a high correlation coefficient across all wall regions. It also revealed that the MSC-SGS model provided accurate predictions for shear stress τ_{12} and SGS backscatter. In *a posteriori* tests, LES computations based on the MSC-SGS model showed good agreement with DNS data in terms of turbulence statistics and energy spectra, compared to the conventional mathematical model.

Despite its advantages, the MSC-SGS model has certain limitations such as the generalization ability to predict flow other than training dataset. Future work will focus on enhancing the model's ability to generalize across diverse flow conditions beyond the training datasets, thereby improving its applicability to a wider range of turbulent flow scenarios.

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. JP22K03925 and a grant from Indonesian Education Scholarship (BPI), Center for Higher Education Funding and Assessment (PPAPT), and Indonesian Endowment Fund for Education (LPDP).

Reference:

(1) Gamahara, M., Hattori, Y. *Phys. Rev. Fluids*. (2017)
(2) Liu, B., et al. *AIP Advances*. (2022)
(3) Kim, J., et al. *J. Fluid Mech*. (1987)
(4) Jalaali, B., Okabayashi, K. *Physics of Fluids*. (2025)
(5) Bose, R., Roy, A.M. *Engineering Applications of Artificial Intelligence*. (2024)
(6) Park, J., Choi, H. *J. Fluid Mech*. (2021)
(7) Duraisamy, K. *Phys. Rev. Fluid*. (2021)