

Multi-agent reinforcement learning for wall modeling in LES of flow over periodic hills

By D. Zhou[†], M. P. Whitmore, K. P. Griffin AND H. J. Bae[†]

We develop a wall model for large-eddy simulation (LES) that takes into account various pressure-gradient effects using multi-agent reinforcement learning (MARL). The model is trained using low-Reynolds-number flow over periodic hills with agents distributed on the wall along the computational grid points. The model utilizes a wall eddy-viscosity formulation as the boundary condition, which is shown to provide better predictions of the mean velocity field, rather than the typical wall-shear stress formulation. Each agent receives states based on local instantaneous flow quantities at an off-wall location, computes a reward based on the estimated wall-shear stress, and provides an action to update the wall eddy viscosity at each time step. The trained wall model is validated in wall-modeled LES (WMLES) of flow over periodic hills at higher Reynolds numbers, and the results show the effectiveness of the model on flow with pressure gradients. The analysis of the trained model indicates that the model is capable of distinguishing between the various pressure gradient regimes present in the flow.

1. Introduction

LES is an essential technology for the simulation of turbulent flows. The basic premise of LES is that energy-containing and dynamically important eddies must be resolved everywhere in the domain. This requirement is hard to meet in the near-wall region, as the stress-producing eddies become progressively smaller. The cost involved in resolving the near-wall region has motivated the development of WMLES, which utilizes a wall model to account for the effect of the energetic near-wall eddies. Because of such characteristics, wall modeling has been accepted as the next step to enable the increased use of high-fidelity LES in realistic engineering and geophysical applications.

Recently, Bae & Koumoutsakos (2022) demonstrated the efficacy of MARL as a development tool for wall models in zero-pressure-gradient flat-plate flow. In that work, a series of learning agents are distributed along the computational grid points, with each agent receiving local states and rewards then providing local actions at each time step. The MARL model is able to achieve these results by training in situ on moderate-Reynolds-number flows with a reward function only based on the mean wall-shear stress. Therefore, the trained models do not require any higher-fidelity simulation data and are data efficient, unlike supervised learning. Furthermore, MARL can develop novel models that are optimized to accurately reproduce the quantities of interest by discovering dominant patterns in flow physics rather than using *a-priori* parameterization.

In the present study, we extend the methodology of Bae & Koumoutsakos (2022) for pressure-gradient flows and apply it to the flow over periodic hills, training on low-Reynolds-number cases and testing on higher-Reynolds-number flows. Our objective is to develop a wall model for LES based on MARL that is robust to pressure-gradient

[†] Graduate Aerospace Laboratories, California Institute of Technology

Case	Mesh size ($N_x \times N_y \times N_z$)	x_{sep}/H	x_{rea}/H
RLWM	$128 \times 64 \times 64$	-0.03	3.91
RLWM, coarse mesh	$64 \times 32 \times 32$	0.36	4.16
RLWM, fine mesh	$128 \times 128 \times 72$	0.21	3.85
EQWM	$128 \times 64 \times 64$	0.57	3.05
DNS (Krank <i>et al.</i> 2018)	$896 \times 448 \times 448$	0.20	4.51
WRLES (Gloerfelt & Cinnella 2019)	$512 \times 256 \times 256$	-0.11	4.31

TABLE 1. Simulation cases in comparison to reference data, including mesh size, mean separation location x_{sep} , and mean reattachment location x_{rea} at $Re_H = 10595$.

effects in a data-efficient way. The remainder of this report is organized as follows. In Section 2, the setup of the flow simulation and reinforcement learning are introduced. In addition, an investigation of two candidate wall boundary conditions is presented. In Section 3, tests of the wall model are presented for flow over periodic hills at various Reynolds numbers. Finally, conclusions are drawn in Section 4.

2. Methodology

2.1. Flow simulation

For the flow solver, we utilize a finite volume, unstructured-mesh incompressible LES code (You *et al.* 2008). The subgrid-scale (SGS) stress is modeled using the dynamic Smagorinsky model (Germano *et al.* 1991; Lilly 1992).

For training the wall model, a flow configuration that (i) has widely available wall-shear stress profiles for several Reynolds numbers and (ii) does not require tuning of the inlet profile or other boundary conditions is chosen. The flow over periodically arranged hills in a channel (Mellen *et al.* 2000) has well-defined boundary conditions, can be computed at affordable costs, and constitutes a canonical smooth-body separation and reattachment, which is representative of many aerodynamic applications. Furthermore, the flow does not require reconfiguration of the inlet boundary condition for different grid resolutions and wall models, which may be necessary for nonperiodic flows. This configuration has become a popular benchmark test case for validating computational fluid dynamics codes, since numerous experimental and high-fidelity numerical references exist (e.g., Krank *et al.* 2018; Gloerfelt & Cinnella 2019) that provide extensive data on a wide range of Reynolds numbers covering $700 \leq Re_H \leq 37000$, where $Re_H = U_b H / \nu$, H is the hill height, and U_b is the bulk velocity at the top of the hill. The periodic-hill configuration has the dimensions of $9H \times 3.035H \times 4.5H$ in the streamwise (x), vertical (y), and spanwise (z) directions, respectively. In the simulations of the present study, periodic boundary conditions are applied in streamwise and spanwise boundaries, and the equilibrium stress-balance wall model (EQWM; Kawai & Larsson 2012) is employed at the top wall. To maintain constant bulk velocity in time, the flow is driven by time-varying body force (Balakumar *et al.* 2014). A set of three meshes with increasing resolution are used in the present study. More details of the grids utilized are listed in Table 1. The grid is evenly spaced in the z direction and is slightly non-uniform in both the x and y directions. A CFL number of 1 is used for all simulations.

2.2. Reinforcement learning architecture

The MARL architecture of wall-model training in the study is based on Bae & Koumoutsakos (2022). During model training, the agents distributed above the wall receive states based on local instantaneous flow information and a reward based on the estimated wall-shear stress, then provide local actions to update the wall boundary condition at each time step. The agents infer a single optimized policy through their repeated interactions with the flow field to maximize its cumulative long-term reward. To utilize MARL as a model development tool for wall models, an open-source reinforcement learning (RL) toolbox smarties (Novati & Koumoutsakos 2019) is coupled with the flow solver. The coupling is validated using the same training configuration as the preliminary study of Bae & Koumoutsakos (2022).

2.3. Boundary conditions for wall modeling

Previous studies show that how the wall-shear stress is applied affects the mean quantities (Balakumar *et al.* 2014). We test two formulations of the boundary condition, namely the wall-shear stress and the wall eddy-viscosity formulations, to determine the appropriate action for the wall model. Two simulations for the flow over periodic hills at $Re_H = 10595$ are conducted using the same mesh ($N_x \times N_y \times N_z = 128 \times 64 \times 64$) but with the two different boundary conditions. For the wall-shear stress formulation, the wall-shear stress τ_w is set to the mean wall-shear stress from WRLES (Gloerfelt & Cinnella 2019). For the wall eddy-viscosity formulation, the wall-shear stress is calculated as $\tau_w = (\nu + \nu_{t,w})(\partial u_s / \partial n)_w$, where u_s is the wall-parallel velocity, n denotes the wall-normal direction pointing toward the interior, and the subscript w denotes quantities evaluated at the wall. The wall eddy viscosity $\nu_{t,w}(x)$ is obtained from the interpolated mean-flow field and mean skin friction of WRLES (Gloerfelt & Cinnella 2019). Similar boundary conditions for wall modeling have been used in the past (Bae & Lozano-Durán 2021).

The mean skin friction coefficients of the two simulations are compared in Figure 1(a). The skin friction coefficient is defined as $C_f = \langle \tau_w \rangle / (0.5\rho U_b^2)$, where ρ is the density and $\langle \cdot \rangle$ denotes the averaging in both time and spanwise direction. The C_f of the wall-shear stress formulation is equal to that from WRLES (Gloerfelt & Cinnella 2019) by design, whereas the wall eddy-viscosity formulation deviates slightly, with some differences observed near the hill peak and on the leeward side of the hill. Despite the similar predictions of C_f , the velocity fields of the two simulations are different [Figure 1(b)]. The size of the separation bubble behind the hill is largely underpredicted in the wall-shear stress formulation, whereas the prediction from the wall eddy-viscosity formulation agrees reasonably well with the WRLES results (see Figure 2).

The underprediction of the separation bubble in the wall-shear stress formulation can be explained by the directional inconsistency of the u_s at the wall-adjacent cell and τ_w . When the direction of the τ_w is different from u_s at the wall-adjacent cell, the forces acting on the control volume of the wall-adjacent cell will push the u_s in the incorrect direction in a positive feedback loop, leading to an inaccurate prediction of the separation bubble. We find that the wall eddy-viscosity formulation is a more robust method for WMLES of such flows and will utilize it for the remainder of the paper.

2.4. Training of the wall model

The model training is carried out based on the LES of periodic-hill channel flow at $Re_H = 10595$ using the baseline mesh ($N_x \times N_y \times N_z = 128 \times 64 \times 64$). A total of 512 agents are uniformly distributed along the bottom wall, and the wall-normal locations of

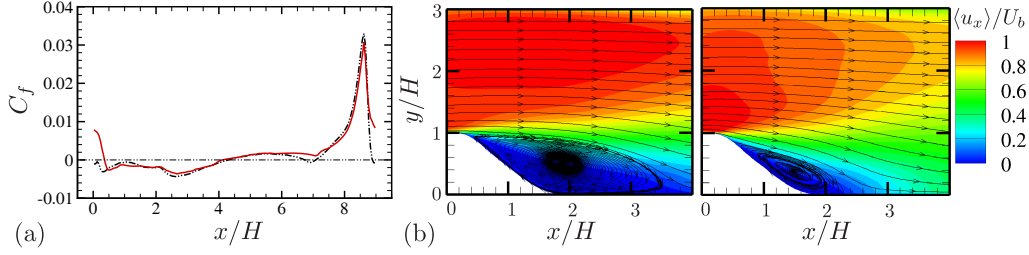


FIGURE 1. (a) Mean skin friction coefficient along the bottom wall at $Re_H = 10595$; —, LES using wall eddy-viscosity formulation; ---, LES using wall-shear stress formulation. (b) Contours of the mean streamwise velocity and the streamlines at $Re_H = 10595$: (left) LES using wall eddy-viscosity formulation and (right) LES using wall-shear stress formulation.

the agents h_m are randomly selected between $0.01H$ and $0.09H$ at each agent location. For each agent, we set three local flow quantities as states: (i) the instantaneous wall-parallel velocity u_s ; (ii) the wall-normal location of the agent $y_n = h_m$; and (iii) the turbulence strain rate $S_{12} = (\partial u_s / \partial n + \partial u_n / \partial s) / 2$, where s denotes the wall-parallel direction pointing toward the positive x -direction. Moreover, to increase the applicability of the wall model for a wide range of flow parameters, the states are nondimensionalized using kinematic viscosity ν and the instantaneous composite friction velocity $u_{\tau p} = (u_{\tau}^2 + u_p^2)^{1/2}$, where $u_p = |(\nu/\rho)(\partial p_w / \partial s)|^{1/3}$, p_w is the pressure on the bottom wall, and u_{τ} is the friction velocity based on the instantaneous modeled wall-shear stress. Nondimensionalized quantities are denoted by superscript $*$. Each agent acts to adjust the local wall eddy viscosity $\nu_{t,w}$ at each time step through a multiplication factor $\nu_{t,w}(t_{i+1}) = a\nu_{t,w}(t_i)$, where $a \in [1 - \alpha\Delta TU_b/\Delta x, 1 + \alpha\Delta TU_b/\Delta x]$, ΔT is the action time step, Δx is the grid size in x direction, and α is a constant that is selected to be 0.001. The local wall-shear stress can be calculated by the formula $\tau_w = \nu(1 + \nu_{t,w}^*)(\partial u_s / \partial n)_w$. The reward r is calculated based on $r(t_i) = [|\tau_w^{ref} - \tau_w(t_i)| - |\tau_w^{ref} - \tau_w(t_{i-1})|] / \tau_{w,rms}^{ref}$ at each location, where τ_w^{ref} and $\tau_{w,rms}^{ref}$ are the mean and root-mean-square wall-shear stress from reference simulation, respectively. The reward is proportional to the improvement in the modeled wall-shear stress compared to the one obtained in the previous time step, and an extra reward of 0.1 is added when the modeled τ_w is within 10% of the reference value.

During model training, each iteration is initialized with the normalized wall eddy viscosity $\nu_{t,w}^*$ that is randomly selected from $(0, 6]$. To generate the initial condition for training, the simulation is started from a flow field generated by the EQWM and run with the given initial $\nu_{t,w}$ for 20 flow-through times (FTTs) to remove numerical artifacts. Each iteration of the model training simulation is conducted for 5 FTTs, and to increase the data efficiency, the $\nu_{t,w}$ is updated every 100 time steps.

The model policy is learned based on a neural network with two hidden layers of 128 units each with a soft sign activation function. The parameters of the neural network are identical to those used in Bae & Koumoutsakos (2022). The model training is advanced for 1.6 million policy gradient steps.

3. Results

3.1. Testing for flow over periodic hills at $Re_H = 10595$

To evaluate the performance of the trained wall model (RLWM), three simulations for the periodic-hill channel flow at $Re_H = 10595$ are carried out using meshes with different

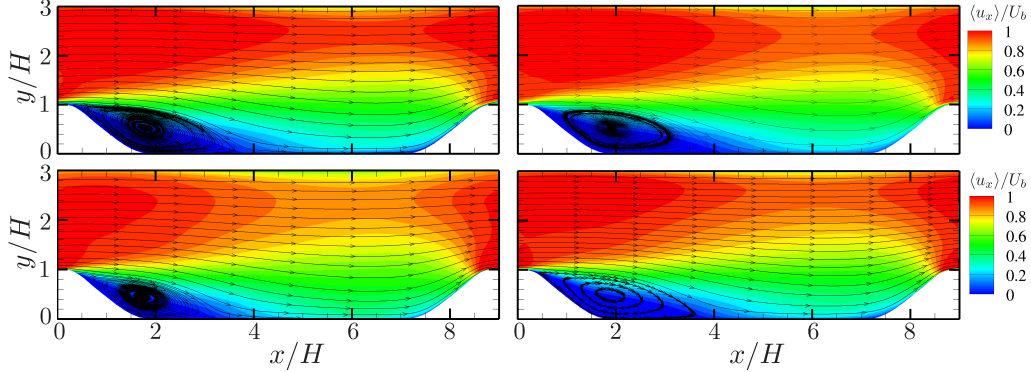


FIGURE 2. Contours of the velocity in x direction and the streamlines at $Re_H = 10595$: (top left) RLWM; (top right) RLWM, coarse mesh; (bottom left) EQWM; and (bottom right) WRLES (Gloerfelt & Cinnella 2019).

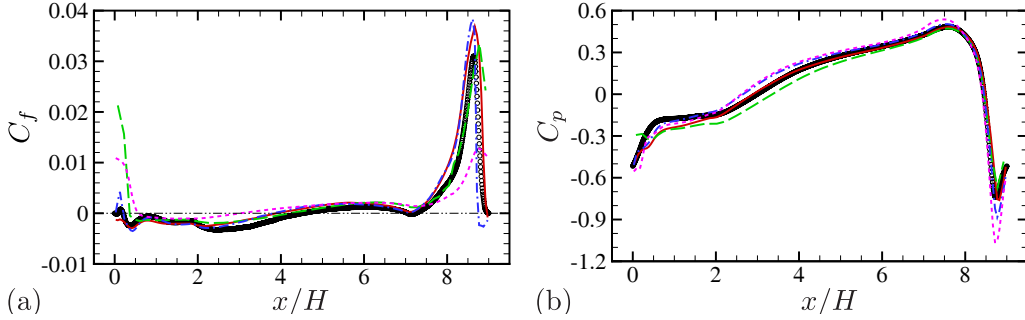


FIGURE 3. (a) Mean skin friction coefficient and (b) mean pressure coefficient along the bottom wall at $Re_H = 10595$. Lines indicate —, RLWM; ---, RLWM, coarse mesh; - · - ·, RLWM, fine mesh; - · - · - ·, EQWM; \circ , DNS (Krank *et al.* 2018).

resolutions. The details of the simulation cases are listed in Table 1. The number of agents on the bottom wall is consistent with the number of mesh cells on the wall. The wall-normal matching location of the agents was chosen to be within the first off-wall cell. The wall eddy viscosity $\nu_{t,w}$ is updated based on the model action at every time step. All simulations are run for about 60 FTTs after transients. The flow statistics of all simulations are averaged over spanwise direction and time. For comparison, an LES using EQWM is conducted and results from two high-fidelity reference databases for this flow (Krank *et al.* 2018; Gloerfelt & Cinnella 2019) are included.

Figure 2 shows the contours of the mean velocity in x direction and the mean-flow streamlines. The flow separates on the leeward side of the hill due to a strong adverse pressure gradient (APG), and a shear layer is generated near the top of the hill. The flow reattaches in the middle section of the channel, and as the flow approaches the windward side of the downstream hill, it is subjected to a strong favorable pressure gradient (FPG) and accelerates rapidly. The simulations with RLWM successfully capture the separation bubble on the leeward side of the hill and yield more accurate results than the EQWM (see Table 1 for quantitative comparison).

The predictions of C_f and the mean pressure coefficient C_p are shown in Figure 3. The mean pressure coefficient is defined as $C_p = (\langle p_w \rangle - \langle p_{ref} \rangle) / (0.5\rho U_b^2)$, where the pressure at $x/H = 0$ on the top wall is chosen as reference pressure p_{ref} (Krank *et al.* 2018).

Regarding C_f , the results from RLWM simulations are in reasonable agreement with the direct numerical simulation (DNS) data, with large deviations found only near the top of the hill where the skin friction reaches its maximum value and then rapidly decreases to a negative value. As the mesh resolution increases, it yields a more accurate prediction of skin friction on the leeward side, but the maximum value of the skin friction on the windward side of the hill still is overpredicted. However, the results are better than the EQWM simulation, which largely underpredicts the skin friction on the windward side and overpredicts the skin friction on the leeward side. Furthermore, the mean locations of the separation and reattachment points (listed in Table 1) are better predicted by the RLWM, consistent with the streamline shown in Figure 2. Additionally, note that the separation bubble size changes as the mesh resolution increases or decreases from the baseline, and more details about the velocity field will be discussed later in this section.

All simulations capture the qualitative trend of the mean C_p on the bottom wall including the APG and FPG regimes, but relatively larger deviations among the simulation cases are visible near the top of the hill ($x/H \geq 8.5$ or $x/H \leq 0.5$) where the pressure sees a sudden change from strong FPG to strong APG and the flow separation emerges. Overall, the RLWM provides more accurate predictions of C_f and C_p than the EQWM.

Quantitative comparison of mean velocity and Reynolds stress components at five streamwise locations ($x/H = 0.05, 2, 4, 6,$ and 8) are shown in Figure 4. The results of the mean velocity profiles and Reynolds stress profiles from the RLWM simulations agree reasonably well with each other and with the reference DNS data. The discrepancies are visible for the regions near channel center. Note that the prediction of the velocity field depends not only on the wall boundary conditions but also on the SGS model, and the varying performance of the SGS model in different meshes may contribute to the inconsistency of those three RLWM simulations. Furthermore, the prediction from EQWM simulation is less ideal, particularly for the mean streamwise velocity $\langle u_x \rangle / U_b$.

To better understand the mechanism of the trained model, we examine the state-action maps, which are the probability density functions (PDFs) of the likelihood that the model takes a particular action conditioned on the occurrence of positive rewards. Figure 5 shows the maps based on instantaneous states and actions at three streamwise positions $x/H = 0.025, 2,$ and 8.3 which are located at the top of the hill, within the separation bubble, and the windward side of the hill, respectively. Overall, the action contour lines for increasing and decreasing $\nu_{t,w}$ are well separated with respect to the mean quantities obtained from the WRLES simulation (Gloerfelt & Cinnella 2019), which illustrates that the model is able to distinguish flow states and provide appropriate actions. Some minor overlaps of the contour lines not clearly separated by the mean relation from WRLES are visible at the top of the hill [see Figures 5(a,d)] where the APG is strong. This implies that the capability of the trained model to robustly adjust the wall-shear stress through appropriate actions is less ideal within this regime.

3.2. Testing for flow over periodic hills at higher Reynolds numbers

In this section, the RLWM is applied to a WMLES of periodic-hill channel flow at $Re_H = 19000$ and 37000 . The simulations are conducted by using the baseline mesh ($128 \times 64 \times 64$) and the coarse mesh ($64 \times 32 \times 32$), and the implementation of RLWM is similar to the simulations at $Re_H = 10595$. All simulations are run for about 60 FTTs after transients. The results from EQWM with baseline mesh and the WRLES (Gloerfelt & Cinnella 2019) are included for comparison.

The mean skin friction coefficients along the bottom wall at $Re_H = 19000$ and 37000 are shown in Figure 6(a,b). The distributions of C_f at higher Reynolds numbers have

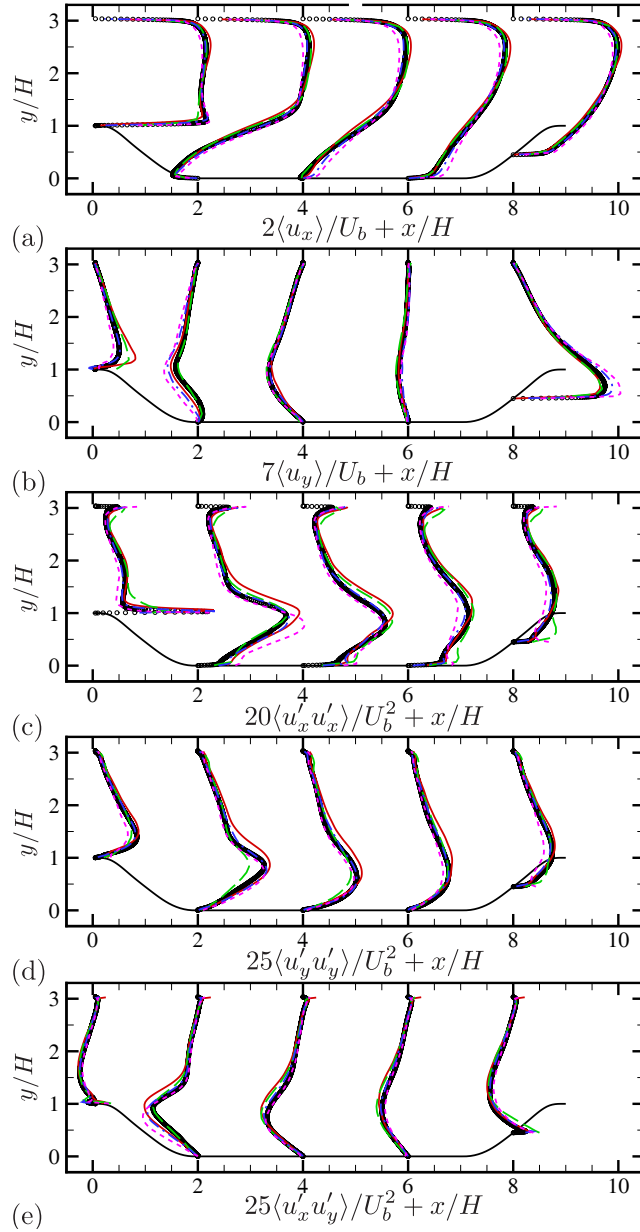


FIGURE 4. Mean velocity and Reynolds stress components at $Re_H = 10595$: (a) streamwise velocity, (b) vertical velocity, (c) streamwise Reynolds stress, (d) vertical Reynolds stress, and (e) Reynolds shear stress. Lines indicate —, RLWM; - - -, RLWM, coarse mesh; - · - ·, RLWM, fine mesh; · · · ·, EQWM; ○, DNS (Krank *et al.* 2018).

a similar shape as the one shown in Figure 3(a). As the Reynolds number increases, the maximum value of the C_f on the windward side of the hill decreases. The results from the two RLWM simulations are in reasonable agreement with the WRLES results. Regarding the separation point, the predicted locations from the RLWM simulations

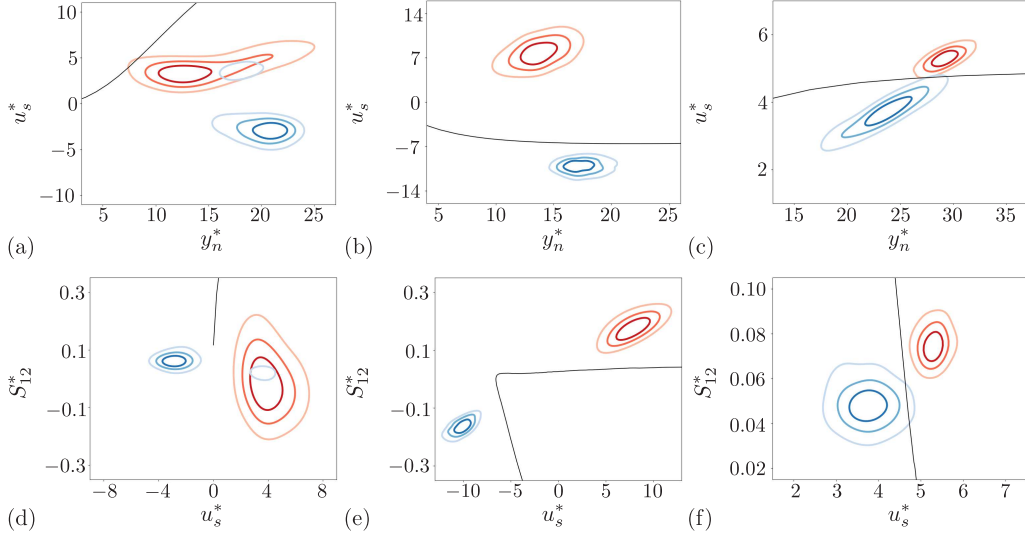


FIGURE 5. PDFs of states for RLWM conditioned to events with $r > 0.1$ and $a < 0.9996$ (blue) or $a > 1.0004$ (red) at (a,d) $x/H = 0.025$, (b,e) $x/H = 2$, and (c,f) $x/H = 8.3$ at $Re_H = 10595$: PDFs of states (a–c) u_s^* and y_n^* and (d–f) S_{12}^* and u_s^* . The black line is the mean relation of the states from WRLES (Gloerfelt & Cinnella 2019). Contour levels are 25, 50, and 75% of the maximum value.

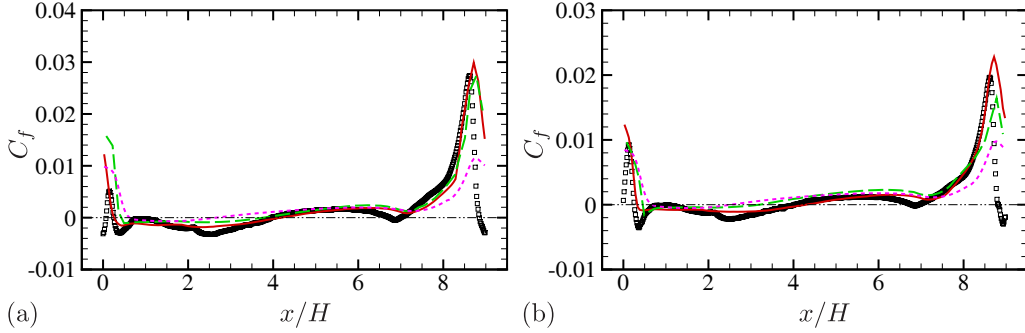


FIGURE 6. Mean skin friction coefficient along the bottom wall at (a) $Re_H = 19000$ and (b) $Re_H = 37000$. Lines indicate —, RLWM; - - -, RLWM, coarse mesh; ····, EQWM; □, WRLES (Gloerfelt & Cinnella 2019).

are further downstream than that of the WRLES. The predicted reattachment location from the RLWM simulation with baseline mesh is slightly downstream from that of the WRLES, but the prediction from the RLWM simulation with coarse mesh is further upstream. Compared to the EQWM simulation, where the separation bubble size is largely underpredicted and the maximum value of C_f on the windward side of the hill is much smaller than the value of WRLES, RLWM is more accurate.

Figure 7 shows the profiles of streamwise components of mean velocity and Reynolds stress at five streamwise stations ($x/H = 0.05, 2, 4, 6$, and 8) for $Re_H = 19000$ and 37000 . The deviations among the profiles from RLWM simulations and the WRLES profiles grow as the Reynolds number increases, albeit slightly, which implies the degradation of performance in the RLWM. Particularly for the Reynolds stress, the EQWM outperforms the RLWM even though the prediction of C_f is inaccurate, as shown in Figure 6.

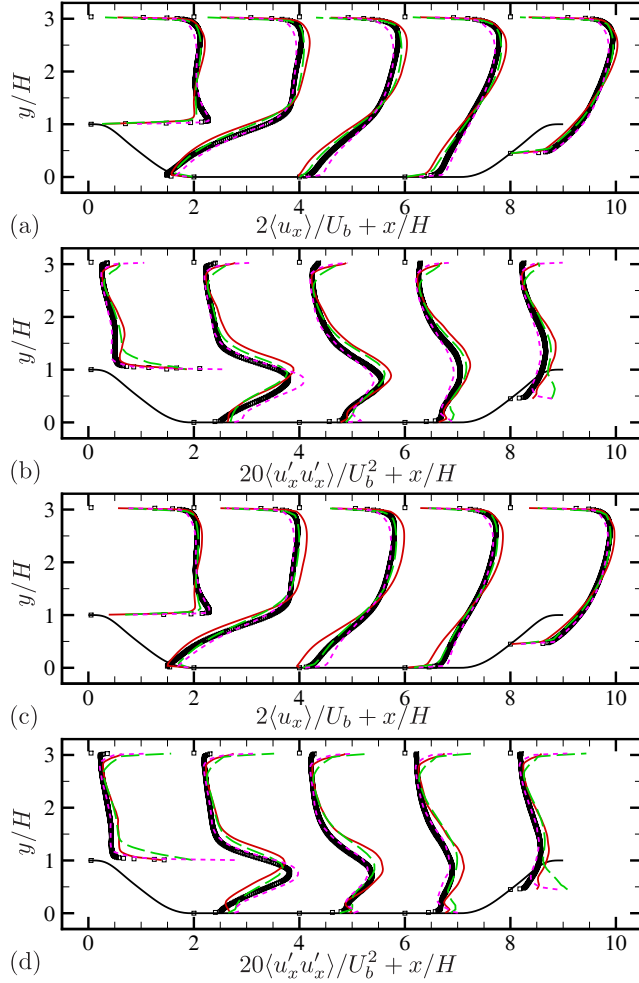


FIGURE 7. Mean velocity and Reynolds stress components at (a,b) $Re_H = 19000$ and (c,d) $Re_H = 37000$: (a,c) mean streamwise velocity and (b,d) streamwise Reynolds stress. Lines indicate —, RLWM; - - -, RLWM, coarse mesh; - - -, EQWM; \square , WRLES (Gloerfelt & Cinnella 2019).

Furthermore, at $Re_H = 19000$ the coarse-mesh RLWM simulation results agree better with the WRLES results than the baseline results. However, the performance of the SGS model affects the velocity field away from the wall more than the wall model, and this could be an indication that a better SGS model is necessary.

4. Conclusions

In this work, a wall model that can adapt to various pressure-gradient effects is developed for turbulent flow over periodic hills using multi-agent reinforcement learning. The model behaves as a control policy for wall eddy viscosity to predict the correct wall-shear stress. The optimized policy of the wall model is learned through the training process based on LES of low-Reynolds-number flow over periodic hills with cooperating agents using the recovery of the correct wall-shear stress as a reward.

The developed wall model is first validated in the LES of the periodic hill configuration at the same Reynolds number of model training. The wall model provides good predictions of mean wall-shear stress, mean wall pressure, and the mean velocity as well as Reynolds stress in the flow field. The test results also show that the developed model outperforms the EQWM. The performance of the developed model is further evaluated at $Re_H = 19000$ and 37000 . Good predictions are obtained for the mean wall-shear stress.

Overall, we have extended the MARL-based wall modeling paradigm and demonstrated that it is robust in flows with pressure gradients. Additional investigations into the generalizability of the developed wall model for different geometries and flow configurations need to be carried out, and the grid sensitivity of the model needs to be further analyzed.

Acknowledgments

The authors acknowledge the XSEDE computing resource from PSC under project PHY210119. The authors would like to thank Dr. Xiang Yang for their insightful remarks on the report.

REFERENCES

- BAE, H. J. & KOUMOUTSAKOS, P. 2022 Scientific multi-agent reinforcement learning for wall-models of turbulent flows. *Nat. Commun.* **13**, 1443.
- BAE, H. J. & LOZANO-DURÁN, A. 2021 Effect of wall boundary conditions on a wall-modeled large-eddy simulation in a finite-difference framework. *Fluids* **6**, 112.
- BALAKUMAR, P., PARK, G. & PIERCE, B. 2014 DNS, LES, and wall-modeled LES of separating flow over periodic hills. *Proceedings of the Summer Program, Center for Turbulence Research, Stanford University*, pp. 407–415.
- GERMANO, M., PIOMELLI, U., MOIN, P. & CABOT, W. H. 1991 A dynamic subgrid-scale eddy viscosity model. *Phys. Fluids* **3**, 1760–1765.
- GLOERFELT, X. & CINNELLA, P. 2019 Large eddy simulation requirements for the flow over periodic hills. *Flow Turbul. Combust.* **103**, 55–91.
- KAWAI, S. & LARSSON, J. 2012 Wall-modeling in large eddy simulation: Length scales, grid resolution, and accuracy. *Phys. Fluids* **24**, 015105.
- KRANK, B., KRONBICHLER, M. & WALL, W. A. 2018 Direct numerical simulation of flow over periodic hills up to $Re_H = 10,595$. *Flow Turbul. Combust.* **101**, 521–551.
- LILLY, D. K. 1992 A proposed modification of the Germano subgrid-scale closure method. *Phys. Fluids A* **4**, 633–635.
- MELLEN, C. P., FRÖHLICH, J. & RODI, W. 2000 Large eddy simulation of the flow over periodic hills. In *Proceedings of the 16th IMACS World Congress*, pp. 21–25.
- NOVATI, G. & KOUMOUTSAKOS, P. 2019 Remember and forget for experience replay. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 4851–4860.
- YOU, D., HAM, F. & MOIN, P. 2008 Discrete conservation principles in large-eddy simulation with application to separation control over an airfoil. *Phys. Fluids* **20**, 101515.