Flow (2025), **5** E40
doi:10.1017/flo.2025.10032

FLOW CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Large language model driven development of turbulence models

Zhongxin Yang[1], Yuanwei Bin[2,3], Yipeng Shi[1] and Xiang I. A. Yang[4]

[1]College of Engineering, Peking University, Beijing, PR China
[2]Ningbo Institute of Digital Twin, Eastern Institute of Technology, Ningbo, Zhejiang, PR China
[3]Shenzhen Tenfong Technology Co., Ltd., Shenzhen, Guangdong, PR China
[4]Mechanical Engineering, Pennsylvania State University, State College, PA, USA
Corresponding author: Xiang I. A. Yang; Email: xzy48@psu.edu

Received: 2 May 2025; **Revised:** 23 September 2025; **Accepted:** 7 October 2025

Keywords: AI singularity; computational fluid dynamics; large language model; turbulence

Abstract

Artificial intelligence (AI) has achieved human-level performance in specialised tasks such as Go, image recognition and protein folding, raising the prospect of an AI singularity – where machines not only match, but surpass human reasoning. Here, we demonstrate a step towards this vision in the context of turbulence modelling. By treating a large language model (LLM), DeepSeek-R1, as an equal partner, we establish a closed-loop, iterative workflow in which the LLM proposes, refines and reasons about near-wall turbulence models under adverse pressure gradients (APGs), system rotation and surface roughness. Through multiple rounds of interaction involving long-chain reasoning and *a priori* and *a posteriori* evaluations, the LLM generates models that not only rediscover established strategies, but also synthesise new ones that outperform baseline wall models. Specifically, it recommends incorporating a material derivative to capture history effects in APG flows, modifying the law of the wall to account for system rotation and developing rough-wall models informed by surface statistics. In contrast to conventional data-driven turbulence modelling – often characterised by human-designed, black-box architectures – the models developed here are physically interpretable and grounded in clear reasoning.

Impact Statement

This work redefines the role of AI in turbulence modelling by engaging an LLM not as a tool, but as a collaborator in the scientific discovery process. Through a structured, iterative human–AI interaction, the LLM generates, reasons about and refines wall models for LES under complex non-equilibrium conditions – rediscovering known strategies and proposing new, physically interpretable formulations that outperform traditional models. The process leverages the LLM's comprehensive knowledge of the turbulence modelling literature – far exceeding that of any individual human. The results demonstrate, for the first time, that general-purpose LLMs can contribute meaningfully to open-ended, physics-based turbulence modelling tasks, suggesting a new paradigm in which AI actively advances core areas of fluid mechanics, rather than merely accelerating existing workflows.

1. Introduction

Artificial intelligence (AI) has achieved human-like performance in tasks once thought to require uniquely human intuition – mastering the game of Go (Silver *et al.*, 2017), predicting protein structures

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

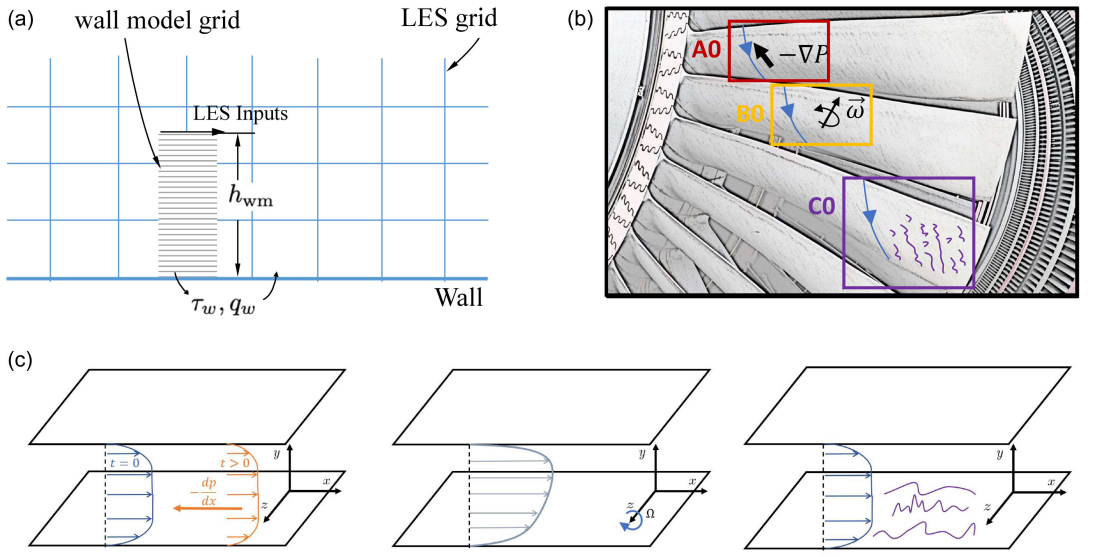


Figure 1. *(a) Schematic of wall-modelled LES (WMLES). A wall model predicts the wall fluxes – shear stress τ_w and heat flux q_w – based on LES-resolved flow quantities at a distance h_{wm} from the wall. (b) Blades in a turbine, illustrating flows subjected to non-equilibrium effects such as APGs (red box), system rotation (orange box) and surface roughness (purple box). (c) Model problems. From left to right: channel subjected to a suddenly imposed APG, channel with system rotation and channel with roughness on the bottom wall.*

(Jumper *et al.*, 2021) and driving through the streets (Yurtsever *et al.*, 2020). These advances have fuelled speculation about the advent of general artificial intelligence (GAI), an AI that can reason, adapt and solve problems across domains. A related concept gaining traction is the AI agent – a system capable of autonomous decision-making and iterative improvement in pursuit of a human-defined goal (Buehler, 2024; Ni and Buehler, 2024; Pandey *et al.*, 2025; Dong *et al.*, 2025). While most practical AI systems today remain narrow in scope, the emergence of large language models (LLMs) has narrowed the gap between domain-specific tools and general-purpose intelligence. These models, trained on vast corpora of human knowledge, can synthesise information, generate code and reason over complex topics (Chang *et al.*, 2024). Despite the hype surrounding GAI, compelling demonstrations of AI contributing new scientific insights – particularly in the physics – remain rare.

Among the most enduring grand challenges in physics is turbulence – a chaotic, multi-scale phenomenon that resists closed-form description and predictive modelling. Despite over a century of effort, turbulence modelling remains largely empirical, guided by human intuition, physical reasoning and hard-won insights from data (Meneveau and Katz, 2000; Piomelli and Balaras, 2002; Durbin, 2018). Yet, it underpins critical applications ranging from climate prediction (Alizadeh, 2022) and aerospace design (Mani and Dorgan, 2023) to wind and energy systems (Stevens and Meneveau, 2017). The gold standard of predictive fidelity, direct numerical simulation (DNS), is limited to canonical flows at modest Reynolds numbers due to its extreme computational cost (Yang and Griffin, 2021; Choi and Moin, 2012). Large-eddy simulation (LES), which resolves large-scale motions while modelling smaller ones, offers a more tractable alternative (Goc *et al.*, 2021, 2024) – but its cost remains prohibitive in high-Reynolds-number applications, especially near walls where turbulent eddies scale with their distance from the wall (Marusic and Monty, 2019). This makes wall modelling the pacing item for extending LES to realistic flows (Bose and Park, 2018; Larsson *et al.*, 2016).

Figure 1a schematically illustrates wall modelling in the context of LES. The LES grid in the near-wall region typically scales with the outer layer rather than the local eddies, leaving the wall layer unresolved. A wall model is therefore used to reconstruct the near-wall turbulence and predict wall fluxes, such as

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

the wall shear stress τ_w and heat flux q_w, based on LES-resolved flow quantities at a distance h_{wm} from the wall – often referred to as the LES/wall-model matching location. The most widely used wall model is the equilibrium wall model (EWM) (Schumann, 1975; Kawai and Larsson, 2012; Yang *et al.*, 2017), which assumes the law of the wall (LoW) holds between the wall and h_{wm} locally and instantaneously. However, the LoW is valid only under equilibrium conditions for the mean flow. Consequently, the EWM falls short when applied to the non-equilibrium boundary layers. Figure 1b highlights a representative case in which near-wall turbulence is affected by non-equilibrium effects, including APGs, system rotation and surface roughness. Efforts have been made to address these non-equilibrium effects in the contexts of wall models. In the following, we highlight a few of these developments. Park and Moin (2014) proposed a dynamic non-equilibrium wall model that accounts for the temporal lag between outer-layer changes and the wall shear stress response. Yang *et al.* (2015) developed an integral wall model based on the momentum integral equation, enabling improved representation of mean velocity profiles. Bose and Moin (2014) introduced a dynamic slip boundary condition that allows the wall shear stress to adapt to the large-scale structures resolved by LES. Building on this concept, Bae *et al.* (2019) formulated a dynamic slip wall model with a self-consistent treatment of the slip length. Fowler *et al.* (2022) introduced a Lagrangian relaxation wall model that does not impose the LoW instantaneously, but instead allows the modelled stress to evolve towards equilibrium over time, thereby better capturing non-equilibrium effects such as pressure gradients. A more comprehensive review of recent progress in wall modelling is available from Fowler *et al.* (2022), Bose and Park (2018) and Yang *et al.* (2024*b*), and is not repeated here for brevity. While these advances have significantly expanded the applicability of LES, the process of model development remains human-driven – rooted in hypothesis generation, expert intuition and iterative refinement. In this context, near-wall turbulence not only remains a persistent bottleneck, but also serves as an ideal testbed for exploring whether AI can drive physical model development.

Recent years have seen a surge of interest in applying AI to turbulence modelling. In particular, machine learning (ML) tools have been used to develop turbulence closures by training on data from high-fidelity simulations and experiments (Duraisamy *et al.*, 2019; Pandey *et al.*, 2020; Shan *et al.*, 2023; Bin *et al.*, 2022, 2023). These efforts have yielded subgrid-scale models for LES (Maulik *et al.*, 2019; Cheng *et al.*, 2022; Xie *et al.*, 2020), Reynolds stress closures for Reynolds-averaged Navier–Stokes (RANS) (Ling *et al.*, 2016; Parish and Duraisamy, 2016; Wang *et al.*, 2017; Bin *et al.*, 2024*a*, *b*; Wu *et al.*, 2025) and wall models for LES (Yang *et al.*, 2019; Bae and Koumoutsakos, 2022; Vadrot *et al.*, 2023*b*; Zhou *et al.*, 2021; Ma and Lozano-Durán, 2025). Given the focus of this work on near-wall turbulence modelling, we briefly review several notable data-driven wall modelling efforts. Yang *et al.* (2019) proposed a predictive wall model based on supervised training of neural networks. The model learns a direct mapping between resolved flow quantities at the matching location and wall fluxes, with the training guided by the LoW. Zhou *et al.* (2021) adopted a similar supervised learning approach and demonstrated its effectiveness in LES of periodic hill flows, accurately capturing flow separation and reattachment. Bae and Koumoutsakos (2022) introduced a multi-agent reinforcement learning (RL) framework, in which local agents infer wall stress through trial-and-error interactions with the LES environment, with rewards based on physical performance metrics such as velocity field accuracy. Vadrot *et al.* (2023*b*) extended this RL approach and demonstrated that the trained model can recover the logarithmic law of the wall across a range of Reynolds numbers. Lozano-Durán and Bae (2023) proposed a machine learning wall model based on the principle of composability. Rather than training on complex geometries directly, the model is trained on simple 'building-block' flows – canonical configurations such as turbulent channel and Couette flows – with the assumption that more complex boundary layer behaviours can be composed from these elementary patterns. Additional reviews of recent developments in data-driven wall modelling can be found from Vadrot *et al.* (2023*a*) and are not repeated here for brevity. Despite these advances, the development process remains fundamentally human-guided. Researchers define the modelling objectives, curate training datasets, select architectures and tune loss functions. The resulting models are typically evaluated on predefined benchmarks. This raises a broader

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

question: can AI be tasked with open-ended problem solving in fluids engineering, where the objective is to iteratively generate and refine physically meaningful models?

In this study, we present a new paradigm in which a general-purpose LLM is tasked with solving problems in turbulence modelling. Unlike conventional applications where AI systems are trained to fit existing data within predefined architectures, we employ DeepSeek-R1, an open-weight LLM, as an autonomous agent operating within a closed-loop modelling framework. The LLM is prompted to generate strategies for near-wall turbulence modelling, which are then assessed both *a priori* and *a posteriori*. Based on performance feedback, the LLM revises its models through multiple iterations – mirroring the hypothesis generation, testing and refinement cycles traditionally performed by human researchers. This approach departs from standard one-shot benchmark-driven machine learning workflows such as reported by Pröhl *et al.* (2024) and Jiang *et al.* (2025). The present framework enables the model to engage in long-chain reasoning and to propose modelling strategies that exhibit both interpretability and performance. We note that the choice of DeepSeek-R1 here over more widely used commercial models, such as ChatGPT or Claude, is motivated by two key considerations. First, DeepSeek-R1 is open-source, which enables full control over prompt design, reproducibility and model deployment. Second, recent benchmarks suggest that DeepSeek-R1 exhibits strong reasoning capabilities, particularly in physics and mathematics domains (Gao *et al.*, 2025).

This article focuses on three longstanding challenges in near-wall turbulence modelling: the effects of APGs, system rotation and surface roughness. These physical mechanisms often appear simultaneously in engineering applications such as turbomachinery and aerospace flows. However, modelling their combined effects directly in complex configurations – such as those illustrated in Figure 1b – is ill-advised. The intertwined influence of multiple non-equilibrium effects makes it difficult to assess modelling errors. Here, we adopt a classical strategy in turbulence research: decomposing a complex modelling problem into a set of well-defined model problems. As sketched in Figure 1c, we study each effect individually in the context of periodic channel flows, which provide controlled environments for isolating the impact of APG, rotation or roughness. This serves multiple purposes. First, it enables scientific clarity – allowing us to assess whether the AI can identify the dominant mechanisms associated with each effect. Second, it facilitates validation against available high-fidelity data from DNS, which exists for these canonical flows – at least for mean velocity profiles and wall stresses (Yang *et al.*, 2020*b*; Chen *et al.*, 2023; Huang and Yang, 2021; Yang *et al.*, 2023; Nair *et al.*, 2024). Third, it mirrors the historical development of turbulence models, where insight is gained from simplified configurations before generalisation to more complex flow scenarios. For each model problem, DeepSeek-R1 is prompted to develop near-wall models. The proposed models are then evaluated in canonical LES settings and their performance is used to iteratively refine subsequent formulations. In doing so, we shall see that the AI demonstrates an emerging capability for open-ended problem solving in fluids engineering.

The rest of the paper is organised as follows. In § 2, the methodology is described. Results are presented in § 3, followed by concluding remarks in § 4.

2. Methodology

2.1. Large language model

Our AI capabilities are driven by DeepSeek-R1, leveraging the computational power and infrastructure of the SiliconCloud API for seamless integration and performance. DeepSeek-R1 is an LLM released in January 2025 by DeepSeek AI (Guo *et al.*, 2025). It is an open-weight, transformer-based model designed to support long-form reasoning with a focus on handling complex, multi-step problems through iterative token generation and context management. The model is built on a dense decoder-only transformer architecture with 61 layers and approximately 671 billion parameters (Guo *et al.*, 2025), making it one of the largest openly accessible models in terms of parameter count. It supports a 32 000-token context window, facilitating the processing of long documents and maintaining coherence across extended dialogues or analytical tasks. The model was pretrained on a diverse mixture of high-quality data sources,

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

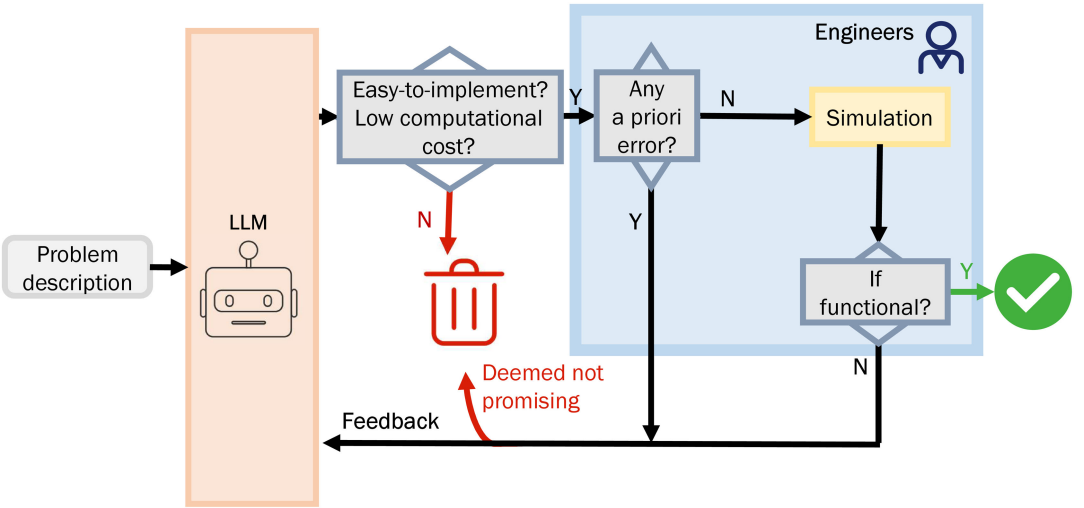


Figure 2. *A flow diagram representing the interaction between the engineers and LLM.*

including curated web documents, source code from multiple programming languages and scholarly articles from various scientific domains, which enhances its capability in both general and technical contexts. This was followed by extensive instruction tuning using supervised fine-tuning and reinforcement learning from human feedback (RLHF) to align its outputs with user intentions while preserving reasoning depth. The reasoning capabilities of Deepseek-R1 have been benchmarked against state-of-the-art models across scientific question-answering tasks, maths word problems and theorem proving. It consistently showing competitive performance, often matching or exceeding the performance of other leading models in complex reasoning scenarios.

Figure 2 illustrates the collaborative workflow between the engineers and the LLM, outlining the iterative process of proposal, evaluation and refinement of candidate models. When interacting with the LLM, we play the role of a fluids engineer, responsible for defining the modelling objectives and evaluating candidate solutions. The LLM, in turn, acts as the modeller, tasked with generating modelling strategies and outlining potential implementation pathways. The interaction begins by presenting the LLM with a flow configuration and the associated modelling challenge. It is then prompted to propose candidate solutions, often followed by clarification prompts to elicit additional details or justification needed for evaluation and practical implementation. In contrast to benchmark-style evaluations, which typically involve a single round of interaction focused on verifying the correctness of LLM responses, our process seeks to mimic the communication between a turbulence modeller in academia and a fluids engineer in industry. This communication is inherently iterative, involving multiple rounds of feedback and refinement. Here, multiple rounds of interaction are critical: they provide opportunities for both the user and the LLM to correct omissions or misunderstandings. For example, the instructions we initially provide may be incomplete or the models proposed by the LLM may contain errors. These shortcomings are often corrected through subsequent interactions, without explicit acknowledgment from either party. Once a potentially viable solution is provided, we proceed to formal evaluation. Two types of evaluation are employed. The first is based on feasibility: some solutions may require unavailable training data or demand impractical levels of manual coding, and such proposals are discarded. The second concerns accuracy, assessed through both *a priori* evaluations – focused on consistency with empirical knowledge – and *a posteriori* assessment via implementation and testing in CFD simulations. Each candidate model must be rigorously evaluated through high-fidelity numerical simulation before it can be considered a viable solution. Even if a model fails the *a posteriori* test, the simulation results still offer valuable performance feedback. The feedback is subsequently used to guide the LLM in refining the candidate

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

model. The iterative process terminates once a candidate model achieves satisfactory accuracy in the target flow configurations. Further details regarding the iterative interaction and evaluation process are provided in § 3.

2.2. Computational fluid dynamics

We solve the filtered incompressible Navier–Stokes equations, written in index notation with summation over repeated indices:

$$\partial_j \tilde{u}_j = 0, \tag{2.1}$$

$$\partial_t \tilde{u}_i + \partial_j \left(\tilde{u}_i \tilde{u}_j\right) = -\partial_i \left(\frac{\tilde{p}}{\rho}\right) + \nu \partial_j \partial_j \tilde{u}_i - \partial_j \tau_{ij} - 2\epsilon_{ijk}\Omega_j \tilde{u}_k, \tag{2.2}$$

where $\tilde{u}_i$ and $\tilde{p}$ are the resolved velocity and pressure, respectively; ρ is the fluid density, ν is the kinematic viscosity and Ω_j denotes system rotation. The subgrid-scale (SGS) stress tensor τ_{ij} is modelled according to the Boussinesq hypothesis:

$$\tau_{ij} = -2\nu_t \tilde{S}_{ij} + \frac{1}{3}\tau_{kk}\delta_{ij}, \tag{2.3}$$

where $\tilde{S}_{ij} = \frac{1}{2}\left(\partial_j \tilde{u}_i + \partial_i \tilde{u}_j\right)$ is the resolved rate-of-strain tensor and δ_{ij} is the Kronecker delta. The eddy viscosity ν_t is computed using the Vreman model, whose details could be found from Vreman (2004). Near-wall turbulence is not resolved and is modelled using a wall model. Two types are considered. The first is the EWM (Schumann, 1975; Piomelli *et al.*, 1989; Kawai and Larsson, 2012; Yang *et al.*, 2017), which imposes the LoW between the wall and the LES/wall-model matching location h_{wm}:

$$\tau_{w,x}/\rho = -\left[\frac{U_{||}^+}{f(y^+)}\right]^2 \frac{\tilde{u}_{LES}}{U_{||}}, \quad \tau_{w,z}/\rho = -\left[\frac{U_{||}^+}{f(y^+)}\right]^2 \frac{\tilde{w}_{LES}}{U_{||}}, \tag{2.4}$$

where the superscript + denotes normalisation by wall units, $\mathbf{U}_{||}$ is the wall-parallel velocity, $\tilde{u}_{LES}$ and $\tilde{w}_{LES}$ are the LES velocity in the streamwise and the transverse directions at the LES/wall-model matching location. In all of our WMLESs, $h_{wm}/h \approx 0.2$ and $f(y^+)$ is Spalding's LoW. The algebraic EWM in (2.4) has an ODE counterpart:

$$\frac{\mathrm{d}}{\mathrm{d}y}\left[(\nu + \nu_{t,wm})\frac{\mathrm{d}\tilde{u}_{||}}{\mathrm{d}y}\right] = 0, \tag{2.5}$$

where

$$\nu_t = \left[\kappa y(1 - \exp(-y^+/A^+))\right]^2 \left|\frac{\mathrm{d}\tilde{u}_{||}}{\mathrm{d}y}\right|, \quad A^+ = 26 \tag{2.6}$$

is the wall model eddy viscosity. The models in both (2.4) and (2.5) are equivalent. They both impose the LoW between the matching location and the wall. In the logarithmic region, the LoW is

$$U^+ = \frac{1}{\kappa}\ln y^+ + B, \tag{2.7}$$

with von Kármán constant $\kappa \approx 0.41$ and intercept $B \approx 5.2$ (Pope, 2001). For rough-wall flows, a roughness function ΔU^+ is subtracted to reflect the downward shift in the velocity profile:

$$U^+ = \frac{1}{\kappa}\ln y^+ + B - \Delta U^+. \tag{2.8}$$

In addition to the EWM, the second type of wall model includes those proposed from DeepSeek-R1. These LLM-driven closures are presented in § 3. While the details of the models differ, their implementations are rather alike, and we place the LES/wall-model matching location at $h_{wm}/h \approx 0.2$ as well.

We focus on the model problems in Figure 1c. The flow configuration is a periodic channel. The flow is periodic in the streamwise and spanwise directions. The wall boundary condition is supplied

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

Table 1. *Details of WMLESs. The friction Reynolds number is defined as $Re_\tau = u_\tau h/\nu$, where u_τ is the friction velocity, h is the channel half-height and ν is the kinematic viscosity. For cases involving a suddenly imposed APG, Re_τ is the baseline value, i.e. at the time instant when the APG is applied. The APG strength is characterised by $\Pi_0 = (h/\tau_{w,0})(dp/dx)$, where $\tau_{w,0}$ is the wall shear stress prior to APG application. For cases with system rotation, the dimensionless rotation number is defined as $Ro_\tau = 2h\Omega/u_\tau$, where Ω is the rotation rate. Surface roughness is parametrised using the non-dimensional equivalent sandgrain roughness height, $k_s^+ = k_s u_\tau/\nu$. Case labels use the abbreviations APG, ROT and RW to indicate the presence of APG, rotation and roughness, respectively. The prefix 'R' denotes the nominal Reynolds number and is followed by $Re_\tau/100$.*

Case	Re_τ	$L_x \times L_y \times L_z$	$n_x \times n_y \times n_z$	**Remark**
			APG channel	
R5APG1	544	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	APG channel, $\Pi_0 = 1$
R5APG10	544	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	APG channel, $\Pi_0 = 10$
R5APG100	544	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	APG channel, $\Pi_0 = 100$
R10APG10	1000	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	APG channel, $\Pi_0 = 10$
R10APG100	1000	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	APG channel, $\Pi_0 = 100$
			Rotating channel	
R2ROT10	180	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 10$
R2ROT22	180	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 22$
R2ROT40	180	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 40$
R2ROT80	180	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 80$
R2ROT120	180	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 120$
R4ROT10	360	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 10$
R4ROT20	360	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 20$
R4ROT32	360	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 32$
R4ROT44	360	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Rotating channel, $Ro_\tau = 44$
			Rough wall channel	
R60RW1	6000	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Sandgrain, $k_s^+ = 54.2$
R60RW2	6000	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Sandgrain, $k_s^+ = 64.4$
R60RW3	6000	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Grit-blasted, $k_s^+ = 24.5$
R60RW4	6000	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Grit-blasted, $k_s^+ = 27.2$
R60RW5	6000	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Grit-blasted, $k_s^+ = 20.1$
R60RW6	6000	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Grit-blasted, $k_s^+ = 16.3$
R25RW7	2490	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Truncated core, $k_s^+ = 286$
R50RW8	4970	$2\pi h \times 2h \times 2\pi h$	$64 \times 64 \times 64$	Multiscale LEGO-like, $k_s^+ = 646$

by a wall model. Three flow effects are studied. First, a fully developed plane channel subjected to a suddenly imposed APG. Second, a plane channel subjected to spanwise system rotation. Third, channel flow with roughness on the bottom wall. Reference data are available from Chen *et al.* (2023); Xia *et al.* (2016); Flack *et al.* (2016); Medjnoun *et al.* (2021); Womack *et al.* (2022); Flack and Schultz (2023). The computational grid is uniformly spaced in all directions following the standard practice (Bose and Moin, 2014; Park and Moin, 2014; Yang *et al.*, 2020*a*). Key simulation parameters, including domain size, Reynolds number and grid resolution, are listed in Table 1. The domain size, grid spacing and aspect ratios closely match those used in prior studies (Anderson *et al.*, 2018; Abkar *et al.*, 2016; Martinez-Tossas *et al.*, 2018; Yang *et al.*, 2024*a*), ensuring the credibility of performance evaluation and cross-study comparisons.

All simulations are performed using a finite-volume solver adapted for wall-modelled LES. The solver builds on the SUPES-cwm code base developed by Lv *et al.* (2021) and Gao and Lv (2025), which has been validated for channel flow calculations. It employs a fully conservative finite-volume formulation, nominal third-order spatial accuracy via characteristic-variable-based reconstruction and

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

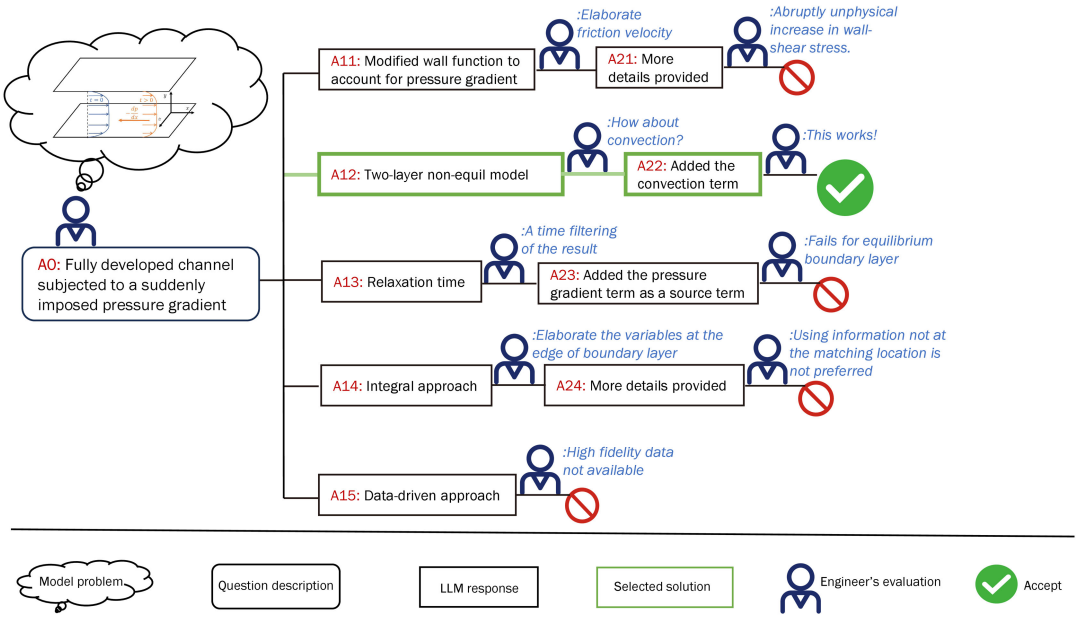


Figure 3. *Schematic overview of the interaction between the LLM and the user for the APG modelling problem.*

an explicit strong-stability-preserving Runge–Kutta scheme for time integration. Numerical fluxes are computed using a hybrid scheme blending central differencing and HLLC fluxes, with dissipation controlled via a blending factor. Further details of the code's numerics could be found from Lv *et al.* (2021) and Gao and Lv (2025) along with a validation, and are not repeated here for brevity.

3. Results

3.1. Adverse pressure gradient

The presence of a pressure gradient – whether favourable or adverse – modifies the velocity distribution and the shear stress within the boundary layer, leading to history effects (Bobke *et al.*, 2017; Chen *et al.*, 2023) and significant departures from the canonical LoW behaviour observed under zero-pressure-gradient conditions (Perry *et al.*, 2002; Volino, 2020; Pozuelo *et al.*, 2022). This deviation complicates near-wall turbulence modelling, which often relies on equilibrium assumptions.

In this subsection, we summarise our interactions with the LLM regarding this challenge, present the LLM-driven modelling solutions and assess its performance. The modelling challenge was presented to the LLM using the following prompt:

> *I am a fluid engineer. I use wall-modelled large-eddy simulation. I have the following flow problem: a fully developed channel flow with a suddenly imposed APG. The imposed APG is kept constant. You can imagine that the friction on the wall will decrease. The logarithmic law based wall model is not effective in simulating this flow. Please propose new models, taking into account simplicity and generality.*

An overview of the interaction between the LLM and the user is illustrated in Figure 3, with the initial prompt summarised and labelled as A0. In the following, we document the communication process, highlighting how the LLM refines its reasoning, iteratively breaks down the problem and ultimately converges on a viable solution. The communication here uses approximately 7000 tokens.

The LLM proposes five potential modelling strategies, labelled A11 to A15 in Figure 3. Here, the four less successful strategies are discussed first, followed by the more promising one. The first strategy,

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

A11, modifies the wall function. The LLM introduces the concept of an adjusted friction velocity, on which we prompt it to elaborate. The model takes the following form:

$$\tau_w = \tau_{w,0}\left[1 + \alpha\, \mathrm{d}P^+\right], \tag{3.1}$$

where $\tau_{w,0}$ is the prediction of the EWM, $\mathrm{d}P^+$ is the dimensionless pressure gradient and α is a model constant. Although the intermediate model forms are not shown here for brevity, we note that the initial model form was dimensionally inconsistent. Nonetheless, this error was corrected in subsequent responses without explicit prompting – an encouraging behaviour akin to human model development. Upon implementation, we observe an abrupt and unphysical change in wall shear stress upon the application of the APG, regardless of the choice of α. Due to this fundamental misrepresentation of the flow physics, this strategy is discarded. The third strategy, A13, introduces a relaxation time scale to the wall stress evolution:

$$\tau_w^{n+1} = \tau_w^n + \Delta t \frac{\tau_{w,0} - \tau_w^n}{T_r}, \tag{3.2}$$

where τ_w^n and τ_w^{n+1} denote the wall shear stress at time steps n and $n+1$, Δt is the time step size, $\tau_{w,0}$ is the EWM prediction, and T_r is a relaxation time scale. Analysis shows that this model acts as a low-pass temporal filter for wall shear stress, but does not effectively capture pressure gradient effects. When this limitation is communicated to the LLM, it reformulates the model by adding a source term accounting for the APG:

$$\partial_t \tau_w = \frac{\tau_{w,0} - \tau_w}{T_r} + \gamma u_\tau \partial_x P, \tag{3.3}$$

where γ is a model constant. *A posteriori* tests show that the model suffers from the same issue as that in (3.1) and is therefore discarded as well. The fourth strategy, A14, is an integral model:

$$\partial_t \theta + U_e \partial_x \theta = \frac{\tau_w}{\rho U_e^2} - \frac{\theta}{U_e}\partial_x U_e. \tag{3.4}$$

Upon requesting clarification of the variables in the equation, we find that the model requires information at the edge of the boundary layer, which is generally not available in WMLES and is not desirable for practical implementation. The fifth strategy, A15, suggests a data-driven modelling approach. However, the lack of high-fidelity training data renders this strategy infeasible within the scope of the current study. The second strategy, A12, ultimately yields a viable solution. Upon follow-up discussions regarding the treatment of convective terms, the LLM refines the model from A12 to A22. The model is based on the thin boundary layer approximation, thus assuming $\partial_y P = 0$ and neglecting the wall-normal velocity. The final model takes the following form:

$$\begin{aligned} \frac{\mathrm{D}u_{wm}}{\mathrm{D}t} &= -\frac{1}{\rho}\frac{\partial P_{LES}}{\partial x} + \frac{\partial}{\partial y}\left[(\nu + \nu_{t,0})\frac{\partial u_{wm}}{\partial y}\right], \\ \frac{\mathrm{D}w_{wm}}{\mathrm{D}t} &= -\frac{1}{\rho}\frac{\partial P_{LES}}{\partial z} + \frac{\partial}{\partial y}\left[(\nu + \nu_{t,0})\frac{\partial w_{wm}}{\partial y}\right]. \end{aligned} \tag{3.5}$$

Here, u_{wm} and w_{wm} are the velocity components in the x and z directions, respectively; ν is the molecular viscosity and $\nu_{t,0}$ is the eddy viscosity, modelled in the same way as in the EWM. The model solves the velocity profiles between the LES/WM matching location and the wall in both wall-parallel directions, x and z, according to (3.5). The pressure gradients in both the x and z directions are obtained directly from the LES. Note that the pressure gradient and the material derivative terms require no closure, although storing the velocity profiles is necessary to compute the unsteady term. The rationale here is that history effects are largely captured by the material derivative and the pressure gradient, and no further modelling is required. In addition to the reasoning and the model formulation, the LLM also supplies an initial Python implementation, which – although not directly portable to our CFD code – provides a useful starting point. It is worth noting that some candidate modifications – for instance,

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

Table 2. *Time to incipient separation (normalised by $h/U_{c,0}$). Relative errors compared to DNS are also listed.*

Case	DNS	EWM	LLM-A
R5APG1	228.4	264.5 (+15.8 %)	210.6 (−7.79 %)
R5APG10	22.90	30.71 (+34.1 %)	19.30 (−15.7 %)
R5APG100	0.7644	3.086 (+303.7 %)	0.6790 (−11.2 %)
R10APG10	29.15	38.14 (+30.8 %)	26.06 (−10.6 %)
R10APG100	1.393	3.743 (+168.7 %)	1.148 (−17.6 %)

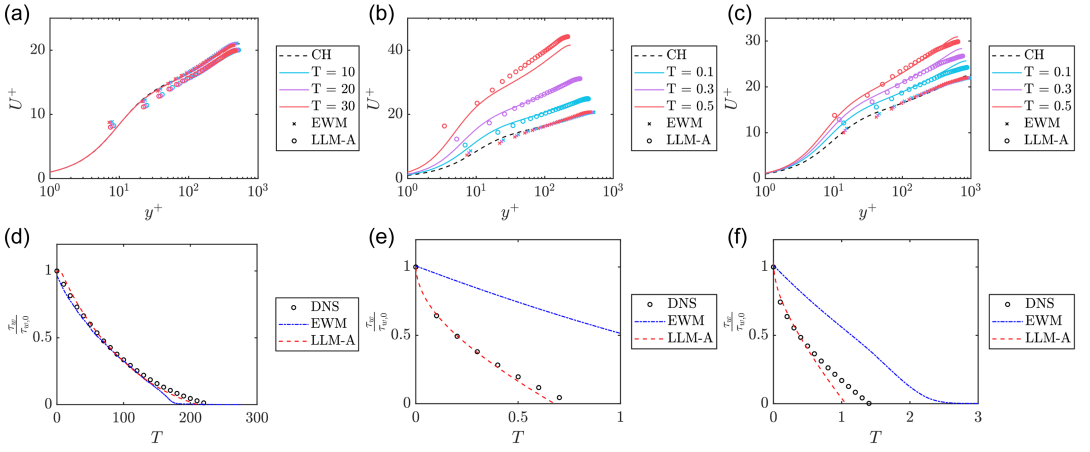


Figure 4. *(a–c) Inner-scaled mean velocity profiles following the imposition of an APG: (a) R5APG1; (b) R5APG100; (c) R10APG100. Time T is normalised by $h/U_{c,0}$, where $U_{c,0}$ is the channel centreline velocity at $t = 0$. (d–f) Evolution of the wall shear stress: (d) R5APG1; (e) R5APG100; (f) R10APG100. The dashed line corresponds to the results for the EWM. DNS reference data are shown in colour and predictions from the model in (3.5) are labelled 'LLM-A'.*

enhancing A11 through the introduction of a relaxation time, A13 – could, in principle, address deficiencies such as unphysical abrupt increases in subsequent iterations. However, during the second round of dialogue, a viable formulation, A22, emerged. Consequently, the iterative process was terminated.

We proceed to the *a posteriori* test. Figure 4 presents the results; the computational set-up has already been summarised in Table 1. Figures 4a–4c show inner-scaled mean velocity profiles at various time instants following the application of the APG, under mild (panel a) and strong (panels b, c) APG conditions at two Reynolds numbers. As expected, the EWM performs adequately under mild APG, but deteriorates under strong APG, regardless of Reynolds number. In contrast, the model in (3.5) accurately tracks the evolution of the mean flow. Figures 4d–4f show the evolution of wall shear stress over time for the three cases shown in Figures 4a–4c. Again, while the EWM is reasonably accurate under mild APG (panel d), it fails to capture the rapid decline in wall shear stress observed in the strong APG cases (panels e, f). By comparison, the DeepSeek model in (3.5) consistently produces more accurate wall shear stress predictions across all three cases. Table 2 further quantifies the time to incipient separation for each case. It is evident that the new model substantially reduces the prediction error compared with the baseline EWM.

3.2. System rotation

System rotation profoundly modifies the dynamics of wall-bounded turbulent flows and arises in a wide range of applications, including turbomachinery, atmospheric flows and rotating devices. A special

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

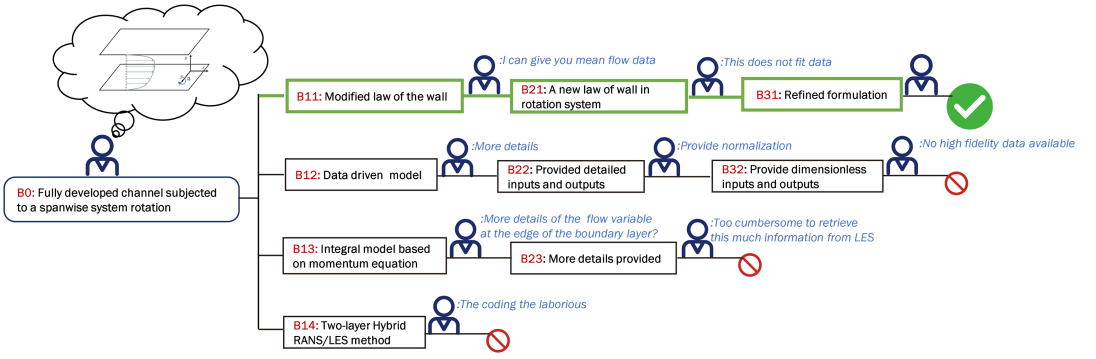


Figure 5. *Schematic overview of the interaction between the LLM and the user for the spanwise rotation modelling problem.*

case that has received much attention is when the flow is subjected to spanwise system rotation which rotates along the z axis with the Coriolis force parallel to $x-y$ plane, where the Coriolis force induces a wall-normal pressure gradient that leads to the formation of a 'pressure side' and a 'suction side', with turbulence suppressed and enhanced, respectively (Johnston *et al.*, 1972; Xia *et al.*, 2016). This redistribution of turbulence intensity results in substantial deviations from the classical law-of-the-wall behaviour: on the pressure side, the mean velocity exhibits a near-linear profile instead of the canonical logarithmic scaling (Yang *et al.*, 2020*b*; Brethouwer, 2017). Consequently, the EWM no longer applies.

In this section, we engage the LLM to address this challenge. The modelling task was presented to DeepSeek-R1 using the following prompt:

> *I am a fluid engineer using wall-modelled large-eddy simulation. Turbulent flow in a rotating system has long been a challenging problem. I am currently considering a simplified case: a fully developed plane channel flow between two infinitely large plates, where x, y and z denote the streamwise, wall-normal and spanwise directions, respectively. A uniform volume force is applied along the x direction and the system rotates about the z axis at an angular velocity Ω. Are there any existing wall models that can handle this type of flow?*

An overview of the interaction between the LLM and the user is illustrated in Figure 5.

Before attempting a solution, DeepSeek-R1 undertakes an extended chain of reasoning. It first analyses the mean flow by invoking and simplifying the RANS equations, concluding that spanwise mean velocities or streamwise vortices may emerge due to Coriolis–Reynolds stress interactions. It correctly concludes that spanwise system rotation modifies turbulence anisotropically: turbulence is suppressed on the stable (suction) side and enhanced on the unstable (pressure) side. Additionally, the altered Reynolds stresses redistribute momentum, fundamentally modifying the mean flow structure. DeepSeek-R1 then analyses the limitations of existing wall models. It reasons that the Coriolis force disrupts the equilibrium state assumed in traditional EWMs. Rotation selectively suppresses or enhances turbulent fluctuations and conventional models are unable to capture such anisotropic effects. Due to this extended reasoning chain, including the discussion on modelling strategies, the conversation uses approximately 21 000 tokens.

Following this analysis, DeepSeek-R1 proposes several possible solutions. The overarching principles are similar to those observed in the APG problem. A data-driven approach (labelled B12 in Figure 5) is proposed but ultimately discarded due to the lack of high-fidelity training data. An integral formulation, B13, is also suggested but rejected because it requires information at the boundary-layer edge. What we find particularly interesting in this interaction is that the LLM identifies transport-equation-based models with rotation corrections, B14, of which we were not previously aware and which could potentially form the basis of an effective two-layer model, but is not selected due to the significant amount of coding required. The selected solution, B11, involves modifying the law of the wall. Initially, DeepSeek-R1

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

Table 3. *Bulk velocity predictions normalised by u_τ. Relative errors compared with DNS are also listed.*

Case	DNS	EWM	LLM-B
R2ROT10	17.62	21.16 (+20.1 %)	17.86 (+1.36 %)
R2ROT22	22.25	25.91 (+16.4 %)	23.89 (+7.37 %)
R2ROT80	45.56	39.10 (−14.2 %)	47.83 (+4.98 %)
R2ROT120	55.40	39.60 (−28.5 %)	52.03 (−6.08 %)
R4ROT10	20.25	24.69 (+21.9 %)	21.66 (+6.96 %)
R4ROT20	24.86	28.88 (+16.2 %)	25.57 (+2.86 %)
R4ROT32	31.24	33.41 (+6.95 %)	30.93 (−0.99 %)
R4ROT44	37.84	37.72 (−0.32 %)	37.44 (−1.06 %)

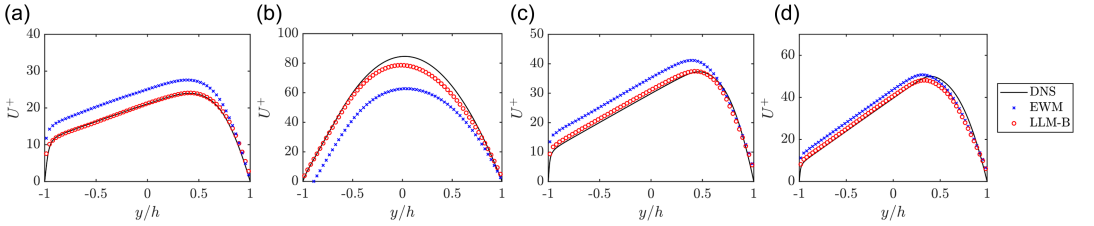


Figure 6. *Mean velocity profiles in spanwise rotating channels: (a) R2ROT10; (b) R2ROT120; (c) R4ROT20; (d) R4ROT32. Label 'LLM-B' corresponds to the model in (3.7).*

recommends altering the logarithmic law. Upon examining the mean flow data, DeepSeek-R1 proposes a modified law of the wall:

$$U^+ = 2Ro^+ y^+ + (16.5 + 60.6Ro^+)/(1 + 46Ro^+), \tag{3.6}$$

where Ro^+ is the rotation number based on wall units. Inverting the mean flow scaling in (3.7) yields a near-wall model:

$$\tau_w/\rho = \left[\frac{U_{LES}}{2Ro^+ y^+ + (16.5 + 60.6Ro^+)/(1 + 46Ro^+)}\right]^2. \tag{3.7}$$

Note that, due to the use of u_τ for non-dimensionalisation, evaluating the right-hand side requires knowledge of the wall shear stress, making (3.7) an implicit equation for τ_w. Nonetheless, one can avoid an iterative solution by using the wall shear stress from the previous time step to evaluate the right-hand side, following the practice of Yang *et al.* (2017), Yang *et al.* (2015), among others.

We test the model in (3.7) within the WMLES framework for plane channel flows subjected to various levels of spanwise rotation. Figure 6 presents the mean velocity profiles for four representative cases. For comparison, we include results from the EWM and reference DNS data. The model in (3.7) consistently outperforms the baseline EWM and accurately captures the mean flow behaviour in spanwise rotating channels. Additional results are provided in Table 3, listing the predicted bulk velocities normalised by u_τ. Note that a pressure gradient is imposed, the friction Reynolds number is fixed and the bulk velocity emerges as a prediction of the model. Across all cases, the proposed model significantly improves bulk velocity predictions relative to the EWM over a broad range of rotation numbers.

3.3. Surface roughness

Surface roughness significantly impacts boundary-layer flows, altering momentum transfer and introducing substantial drag penalties across a wide range of engineering applications, including ships, aircraft, turbomachinery and atmospheric boundary layers (Barlow and Coceal, 2008; Bons, 2010; Schultz, 2007). Predicting the influence of roughness has been a longstanding challenge due to the diversity of

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

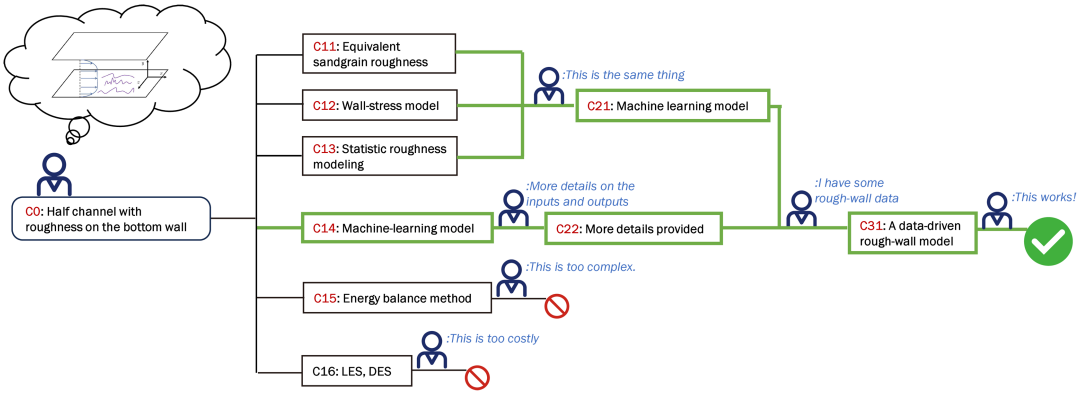


Figure 7. *Schematic overview of the interaction between the LLM and the user for the roughness modelling problem.*

roughness topographies and their complex interactions with near-wall turbulence (Chung *et al.*, 2021). The objective of rough-wall modelling is typically to predict quantities such as the equivalent sand-grain roughness height k_s or the roughness function ΔU^+, and significant advances have been made since the seminal work of Nikuradse (Nikuradse, 1950; Yang *et al.*, 2023; Colebrook, 1939; Moody, 1944; Flack and Schultz, 2014). However, incorporating these insights directly into wall modelling for LES remains limited (Bose and Park, 2018). Here, we engage DeepSeek-R1 to explore new modelling strategies for rough-wall effects in the context of WMLES. We initiated the discussion with DeepSeek-R1 by posing a general question:

> *I am a fluid engineer and I use wall-modelled large-eddy simulation. Modelling surface roughness has always been a difficult problem. What are the main approaches to rough-wall modelling? How can we accommodate surface roughness in wall models?*

DeepSeek-R1 has an awareness of the extensive literature on rough-wall turbulence and highlights several key points. First, it notes that the effect of surface roughness is often parametrised by a roughness function, which manifests as a downward shift of the log-law velocity profile relative to smooth-wall flows. Second, it recognises that explicit geometric resolution of roughness in CFD simulations can accurately capture roughness effects, but doing that can be computationally costly. Third, it identifies machine-learning approaches that relate near-wall flow information to wall shear stress.

The full conversation involves approximately 24 000 tokens and the portion relevant to modelling is summarised schematically in Figure 7. In this dialogue, DeepSeek-R1 identifies three key variables for rough-wall modelling: the equivalent sandgrain roughness k_s (labelled C11), wall shear stress τ_w (labelled C12), and parameterizations of the roughness geometry (labelled C13). When prompted on how to determine these quantities, the LLM recommends a data-driven approach, C14. The training data include drag measurements, roughness geometric statistics and equivalent sandgrain roughness, which are available from public roughness datasets (Yang *et al.*, 2023). Upon confirming the feasibility of this data-driven approach, DeepSeek-R1 provides detailed guidance on constructing a neural network, selecting input and output features, and performing the training. The resulting wall model takes the form:

$$\tau_w/\rho = \left[\frac{U_{\parallel}}{\ln(h_{wm}/k_s)/\kappa + A}\right]^2, \quad k_s = \mathrm{ANN}\,(\text{Parametrisation of Roughness Geometry}), \tag{3.8}$$

where the parametrisation of roughness geometry involves single-point statistics such as the peak-to-trough height, second, third and fourth-order moments of the roughness height, as well as combinations of these variables. We note that comparing the model in (3.8) to the EWM is inappropriate, as the EWM assumes a smooth flat plate. As a baseline for comparison, we adopt the rough-wall model proposed by

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

Table 4. *Predicted roughness function ΔU^+ and the roughness functions measured from experiments. Relative errors compared with experimental measurements are also listed.*

Case	Exp	Ref	LLM-C
R60RW1	6.26	8.80 (+40.6 %)	5.75 (−8.11 %)
R60RW2	6.67	9.42 (+41.2 %)	6.93 (+3.86 %)
R60RW3	4.43	5.59 (+26.1 %)	4.33 (−2.18 %)
R60RW4	4.69	5.87 (+25.1 %)	4.10 (−12.7 %)
R60RW5	3.89	4.94 (+26.9 %)	3.93 (+1.06 %)
R60RW6	3.30	4.23 (+28.3 %)	3.08 (−6.64 %)
R25RW7	10.6	13.3 (+25.3 %)	10.66 (+0.14 %)
R50RW8	12.7	15.5 (+21.4 %)	13.24 (+3.96 %)

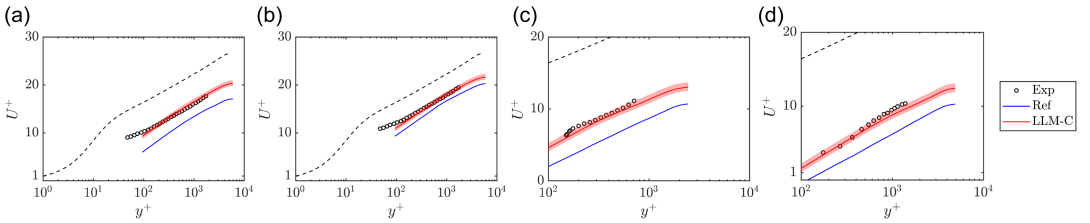


Figure 8. *Predicted mean velocity profiles for rough-wall flows: (a) R60RW1; (b) R60RW3; (c) R25RW7; (d) R50RW8. 'Ref' corresponds to the roughness model of Forooghi et al. (2017). 'LLM-C' corresponds to the model in (3.8). The dashed black line indicates the LoW. The shaded regions represents the uncertainty in the training data.*

Forooghi *et al.* (2017), which is otherwise identical to (3.8) except that k_s is given by

$$k_s/k_{rms} = 3.41(1 + \mathrm{SK})^{0.61}, \tag{3.9}$$

where k_{rms} is the root-mean-square of the roughness height and SK denotes the skewness.

We evaluate the performance of the proposed model within the WMLES framework. Here, the surface roughness is not explicitly resolved; instead, its effects on the flow are entirely modelled. Figure 8 presents comparisons of mean velocity profiles for several rough-wall cases. Results are compared against experimental reference data from Flack *et al.* (2016), Flack and Schultz (2023), Medjnoun *et al.* (2021), Womack *et al.* (2022), as well as against the baseline model. The roughness morphologies considered range from random Gaussian surfaces to regular truncated cones. We observe that the model proposed by DeepSeek-R1 consistently yields more accurate mean velocity predictions, particularly in the logarithmic layer. More quantitative comparisons are provided in Table 4, which lists the predicted roughness function ΔU^+ for all rough surfaces considered. The corresponding roughness morphologies are already detailed in Table 1. Across all cases, the present model outperforms the reference model of Forooghi *et al.* (2017), highlighting the effectiveness of the LLM-guided modelling approach.

4. Concluding remarks

In this study, we explored the utility of LLMs for turbulence modelling by engaging DeepSeek-R1 in a closed-loop, iterative framework. Within this framework, DeepSeek-R1 engages a human engineer, and proposes and refines wall models to address three challenges in near-wall turbulence modelling: APGs, system rotation and surface roughness. Our results demonstrate that the LLM-driven models not only rival, but in many cases outperform, baseline wall models.

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

A key distinction between the present work and existing data-driven turbulence modelling efforts lies in the role played by the AI. Conventional data-driven turbulence models are in fact human-driven: researchers design network architectures, curate training data, and employ machine learning to optimise weights and biases. The resulting models often function as black boxes, offering limited interpretability or physical reasoning. In contrast, the models developed in this study were truly AI-driven. DeepSeek-R1 autonomously provided physical reasoning, logical model structures and complete modelling strategies, resulting in models that are transparent, interpretable and grounded in physical arguments. It should be emphasised that the *a priori* knowledge introduced does not predefine the model itself, but rather establishes the constraints of the problem – defining relevant variables, enforcing dimensional consistency and ensuring adherence to fundamental physical laws. Within this constrained framework, the LLM acts as a hypothesis-generation engine, enabling the systematic exploration of physically plausible model forms.

Equally important is the paradigm shift in how the AI is treated during the process. In most prior applications, AI tools are treated as subordinate assistants, tasked with executing narrowly defined objectives with minimal feedback loops from a human. Here, we treat the LLM as an equal partner, engaging in multiple rounds of iterative dialogue analogous to collaboration among human researchers. Recent studies in human–AI co-creation suggest that enabling reciprocal communication significantly enhances collaboration quality and creative outcomes (Rezwana & Maher, 2022). Our experience supports this view: through sustained interaction, DeepSeek-R1 is able to correct its own errors, refine incomplete formulations and meaningfully contribute to model development without requiring explicit correction at each step.

The results also highlight a limitation of conventional benchmark evaluations of LLMs, which typically assess models based on a single-response correctness (Pröhl *et al.*, 2024; Jiang *et al.*, 2025). As shown here, when allowed to participate in multi-turn, collaborative interactions, LLMs demonstrate capabilities for adaptive reasoning and creative problem-solving that more closely resemble scientific inquiry than simple information retrieval.

Despite these promising results, several limitations of the present investigation must be acknowledged. First, in real engineering applications, these three effects – pressure gradient, rotation and surface roughness – typically occur simultaneously. Current models fail in this regard. Second, the field of LLM development is advancing rapidly. Since the release of DeepSeek-R1, several newer models have demonstrated even stronger reasoning capabilities, making the present work an initial but already dated baseline. Second, while DeepSeek-R1 proposed models with strong performance and clear logic, it did not invent entirely novel modelling paradigms. The strategies it identified – including data-driven approaches, integral methods and modified laws of the wall – have precedents in the turbulence modelling literature. Nevertheless, it is important to recognise that individual researchers are typically familiar with only a subset of these methods. The ability of the LLM to synthesise diverse approaches and reason across multiple frameworks represents a significant augmentation of human expertise.

In all, our experiences point to a broader opportunity in applied fluid mechanics: LLMs can serve as powerful collaborators in engineering model developments. They offer a platform for brainstorming ideas, synthesising knowledge beyond the reach of any individual researcher and refining concepts through interactive reasoning.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/flo.2025.10032.

Data availability. Raw data are available from the author upon reasonable request.

Funding. Zhongxin Yang and Yipeng Shi acknowledge NSFC11988102 for financial support.

Competing interests. The authors declare no conflict of interest.

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

References

Abkar, M., Bae, H. J., & Moin, P. (2016). Minimum-dissipation scalar transport model for large-eddy simulation of turbulent flows. *Physical Review Fluids*, *1*(4), 041701.

Alizadeh, O. (2022). Advances and challenges in climate modeling. *Climatic Change*, *170*(1), 18.

Anderson, W., Yang, J., Shrestha, K., & Awasthi, A. (2018). Turbulent secondary flows in wall turbulence: Vortex forcing, scaling arguments, and similarity solution. *Environmental Fluid Mechanics*, *18*, 1351–1378.

Bae, H. J., & Koumoutsakos, P. (2022). Scientific multi-agent reinforcement learning for wall-models of turbulent flows. *Nature Communications*, *13*(1), 1443.

Bae, H. J., Lozano-Durán, A., Bose, S. T., & Moin, P. (2019). Dynamic slip wall model for large-eddy simulation. *Journal of Fluid Mechanics*, *859*, 400–432.

Barlow, J., & Coceal, O. (2008). *A review of urban roughness sublayer turbulence*. Technical report, Met Office.

Bin, Y., Chen, L., Huang, G., & Yang, X. I. A. (2022). Progressive, extrapolative machine learning for near-wall turbulence modeling. *Physical Review Fluids*, *7*(8), 084610.

Bin, Y., Hu, X., Li, J., Grauer, S. J., & Yang, X. I. A. (2024a). Constrained re-calibration of two-equation Reynolds-averaged Navier–Stokes models. *Theoretical and Applied Mechanics Letters*, *14*(2), 100503.

Bin, Y., Huang, G., Kunz, R., & Yang, X. I. A. (2024b). Constrained recalibration of Reynolds-averaged Navier–Stokes models. *AIAA Journal*, *62*(4), 1434–1446.

Bin, Y., Huang, G., & Yang, X. I. A. (2023). Data-enabled recalibration of the Spalart–Allmaras model. *AIAA Journal*, *61*(11), 4852–4863.

Bobke, A., Vinuesa, R., Örlü, R., & Schlatter, P. (2017). History effects and near equilibrium in adverse-pressure-gradient turbulent boundary layers. *Journal of Fluid Mechanics*, *820*, 667–692.

Bons, J. P. (2010). A review of surface roughness effects in gas turbines. *Journal of Turbomachinery-Transactions of the ASME*, *132*(2), 021004.

Bose, S. T., & Moin, P. (2014). A dynamic slip boundary condition for wall-modeled large-eddy simulation. *Physics of Fluids*, *26*(1), 015104.

Bose, S. T., & Park, G. I. (2018). Wall-modeled large-eddy simulation for complex turbulent flows. *Annual Review of Fluid Mechanics*, *50*(1), 535–561.

Brethouwer, G. (2017). Statistics and structure of spanwise rotating turbulent channel flow at moderate Reynolds numbers. *Journal of Fluid Mechanics*, *828*, 424–458.

Buehler, M. J. (2024). MechGPT, a language-based strategy for mechanics and materials modeling that connects knowledge across scales, disciplines, and modalities. *Applied Mechanics Reviews*, *76*(2), 021001.

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., Wang, Y., et al. (2024). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, *15*(3), 1–45.

Chen, P. E., Wu, W., Griffin, K. P., Shi, Y., & Yang, X. I. A. (2023). A universal velocity transformation for boundary layers with pressure gradients. *Journal of Fluid Mechanics*, *970*, A3.

Cheng, Y., Giometto, M. G., Kauffmann, P., Lin, L., Cao, C., Zupnick, C., Li, H., Li, Q., Huang, Y., Abernathey, R., Gentine, P. (2022). Deep learning for subgrid-scale turbulence modeling in large-eddy simulations of the convective atmospheric boundary layer. *Journal of Advances in Modeling Earth Systems*, *14*(5), e2021MS002847.

Choi, H., & Moin, P. (2012). Grid-point requirements for large eddy simulation: Chapman's estimates revisited. *Physics of Fluids*, *24*(1), 011702.

Chung, D., Hutchins, N., Schultz, M. P., & Flack, K. A. (2021). Predicting the drag of rough surfaces. *Annual Review of Fluid Mechanics*, *53*(1), 439–471.

Colebrook, C. F. (1939). Turbulent flow in pipes, with particular reference to the transition region between the smooth and rough pipe laws. *Journal of the Institution of Civil Engineers*, *11*(4), 133–156.

Dong, Z., Lu, Z., & Yang, Y. (2025). Fine-tuning a large language model for automating computational fluid dynamics simulations. *Theoretical and Applied Mechanics Letters*, *15*, 100594.

Duraisamy, K., Iaccarino, G., & Xiao, H. (2019). Turbulence modeling in the age of data. *Annual Review of Fluid Mechanics*, *51*(1), 357–377.

Durbin, P. A. (2018). Some recent developments in turbulence closure modeling. *Annual Review of Fluid Mechanics*, *50*(1), 77–103.

Flack, K. A., & Schultz, M. P. (2014). Roughness effects on wall-bounded turbulent flows. *Physics of Fluids*, *26*(10), 101305.

Flack, K. A., & Schultz, M. P. (2023). Hydraulic characterization of sandpaper roughness. *Experiments in Fluids*, *64*(1), 3.

Flack, K. A., Schultz, M. P., Barros, J. M., & Kim, Y. C. (2016). Skin-friction behavior in the transitionally-rough regime. *International Journal of Heat and Fluid Flow*, *61*, 21–30.

Forooghi, P., Stroh, A., Magagnato, F., Jakirlić, S., & Frohnapfel, B. (2017). Toward a universal roughness correlation. *Journal of Fluids Engineering*, *139*(12), 121201.

Fowler, M., Zaki, T. A., & Meneveau, C. (2022). A lagrangian relaxation towards equilibrium wall model for large eddy simulation. *Journal of Fluid Mechanics*, *934*, A44.

Gao, T., Jin, J., Ke, Z. T., & Moryoussef, G. (2025). A comparison of deepseek and other LLMs. arXiv preprint arXiv: 2502.03688.

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

Gao, R., & Lv, Y. (2025). A novel large-eddy simulation framework with consistently enforced wall models by a near-wall dynamic correction procedure. *Physics of Fluids*, *37*(3), 035103.

Goc, K. A., Lehmkuhl, O., Park, G. I., Bose, S. T., & Moin, P. (2021). Large eddy simulation of aircraft at affordable cost: A milestone in computational fluid dynamics. *Flow*, *1*, E14.

Goc, K. A., Moin, P., Bose, S. T., & Clark, A. M. (2024). Wind tunnel and grid resolution effects in large-eddy simulations of the high-lift common research model. *Journal of Aircraft*, *61*(1), 267–279.

Guo, D., Yang, D., Zhang, H., Song, J., Zhang, R., Xu, R., Zhu, Q., Ma, S., Wang, P., Bi, X., et al. (2025). Deepseek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning. arXiv preprint arXiv: 2501.12948.

Huang, X. L., & Yang, X. I. A. (2021). A Bayesian approach to the mean flow in a channel with small but arbitrarily directional system rotation. *Physics of Fluids*, *33*(1), 015103.

Jiang, Q., Gao, Z., & Karniadakis, G. E. (2025). Deepseek vs. ChatGPT vs. Claude: A comparative study for scientific computing and scientific machine learning tasks. *Theoretical and Applied Mechanics Letters*, *15*(3), 100583.

Johnston, J. P., Halleent, R. M., & Lezius, D. K. (1972). Effects of spanwise rotation on the structure of two-dimensional fully developed turbulent channel flow. *Journal of Fluid Mechanics*, *56*(3), 533–557.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., Hassabis, D. (2021). Highly accurate protein structure prediction with alphafold. *Nature*, *596*(7873), 583–589.

Kawai, S., & Larsson, J. (2012). Wall-modeling in large eddy simulation: Length scales, grid resolution, and accuracy. *Physics of Fluids*, *24*(1), 015105.

Larsson, J., Kawai, S., Bodart, J., & Bermejo-Moreno, I. (2016). Large eddy simulation with modeled wall-stress: Recent progress and future directions. *Mechanical Engineering Reviews*, *3*(1), 15-00418.

Ling, J., Kurzawski, A., & Templeton, J. (2016). Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, *807*, 155–166.

Lozano-Durán, A., & Bae, H. J. (2023). Machine learning building-block-flow wall model for large-eddy simulation. *Journal of Fluid Mechanics*, *963*, A35.

Lv, Y., Huang, X. L., Yang, X., & Yang, X. I. A. (2021). Wall-model integrated computational framework for large-eddy simulations of wall-bounded flows. *Physics of Fluids*, *33*(12), 125120.

Ma, R., & Lozano-Durán, A. (2025). Machine-learning wall-model large-eddy simulation accounting for isotropic roughness under local equilibrium. *Journal of Fluid Mechanics*, *1007*, A17.

Mani, M., & Dorgan, A. J. (2023). A perspective on the state of aerospace computational fluid dynamics technology. *Annual Review of Fluid Mechanics*, *55*(1), 431–457.

Martinez-Tossas, L. A., Churchfield, M. J., Yilmaz, A. E., Sarlak, H., Johnson, P. L., Sørensen, J. N., Meyers, J., & Meneveau, C. (2018). Comparison of four large-eddy simulation research codes and effects of model coefficient and inflow turbulence in actuator-line-based wind turbine modeling. *Journal of Renewable and Sustainable Energy*, *10*(3), 033301.

Marusic, I., & Monty, J. P. (2019). Attached eddy model of wall turbulence. *Annual Review of Fluid Mechanics*, *51*(1), 49–74.

Maulik, R., San, O., Jacob, J. D., & Crick, C. (2019). Sub-grid scale model classification and blending through deep learning. *Journal of Fluid Mechanics*, *870*, 784–812.

Medjnoun, T., Rodriguez-Lopez, E., Ferreira, M., Griffiths, T., Meyers, J., & Ganapathisubramani, B. (2021). Turbulent boundary-layer flow over regular multiscale roughness. *Journal of Fluid Mechanics*, *917*, A1.

Meneveau, C., & Katz, J. (2000). Scale-invariance and turbulence models for large-eddy simulation. *Annual Review of Fluid Mechanics*, *32*(1), 1–32.

Moody, L. F. (1944). Friction factors for pipe flow. *Journal of Fluids Engineering*, *66*(8), 671–678.

Nair, S. S., Wadhai, V. A., Kunz, R. F., & Yang, X. I. A. (2024). Rough surfaces in underexplored surface morphology space and their implications on roughness modelling. *Journal of Fluid Mechanics*, *999*, A78.

Ni, B., & Buehler, M. J. (2024). MechAgents: Large language model multi-agent collaborations can solve mechanics problems, generate new data, and integrate knowledge. *Extreme Mechanics Letters*, *67*, 102131.

Nikuradse, J. (1950). *Laws of flow in rough pipes*. Technical Report Technical Memorandum 1292, National Advisory Committee for Aeronautics.

Pandey, S., Schumacher, J., & Sreenivasan, K. R. (2020). A perspective on machine learning in turbulent flows. *Journal of Turbulence*, *21*(9-10), 567–584.

Pandey, S., Xu, R., Wang, W., & Chu, X. (2025). OpenFOAMGPT: A retrieval-augmented large language model (LLM) agent for OpenFOAM-based computational fluid dynamics. *Physics of Fluids*, *37*(3), 035120.

Parish, E. J., & Duraisamy, K. (2016). A paradigm for data-driven predictive modeling using field inversion and machine learning. *Journal of Computational Physics*, *305*, 758–774.

Park, G. I., & Moin, P. (2014). An improved dynamic non-equilibrium wall-model for large eddy simulation. *Physics of Fluids*, *26*(1), 015108.

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

Perry, A., Marusic, I., & Jones, M. (2002). On the streamwise evolution of turbulent boundary layers in arbitrary pressure gradients. *Journal of Fluid Mechanics*, *461*, 61–91.

Piomelli, U., & Balaras, E. (2002). Wall-layer models for large-eddy simulations. *Annual Review of Fluid Mechanics*, *34*(1), 349–374.

Piomelli, U., Ferziger, J., Moin, P., & Kim, J. (1989). New approximate boundary conditions for large eddy simulations of wall-bounded flows. *Physics of Fluids*, *1*(6), 1061–1068.

Pope, S. B. (2001). *Turbulent flows*. Cambridge University Press.

Pozuelo, R., Li, Q., Schlatter, P., & Vinuesa, R. (2022). An adverse-pressure-gradient turbulent boundary layer with nearly constant. *Journal of Fluid Mechanics*, *939*, A34.

Pröhl, T., Putzier, E., & Zarnekow, R. (2024). Benchmarking of LLM detection: Comparing two competing approaches, arXiv preprint arXiv: 2406.11670.

Rezwana, J., & Maher, M. L. (2022). Identifying ethical issues in ai partners in human-AI co-creation. arXiv preprint arXiv: 2204.07644.

Schultz, M. P. (2007). Effects of coating roughness and biofouling on ship resistance and powering. *Biofouling*, *23*(5), 331–341.

Schumann, U. (1975). Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli. *Journal of Computational Physics*, *18*(4), 376–404.

Shan, X., Liu, Y., Cao, W., Sun, X., & Zhang, W. (2023). Turbulence modeling via data assimilation and machine learning for separated flows over airfoils. *AIAA Journal*, *61*(9), 3883–3899.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., Hassabis, D. (2017). Mastering the game of go without human knowledge. *Nature*, *550*(7676), 354–359.

Stevens, R. J., & Meneveau, C. (2017). Flow structure and turbulence in wind farms. *Annual Review of Fluid Mechanics*, *49*(1), 311–339.

Vadrot, A., Yang, X. I. A., & Abkar, M. (2023a). Survey of machine-learning wall models for large-eddy simulation. *Physical Review Fluids*, *8*(6), 064603.

Vadrot, A., Yang, X. I. A., Bae, H. J., & Abkar, M. (2023b). Log-law recovery through reinforcement-learning wall model for large eddy simulation. *Physics of Fluids*, *35*(5), 055122.

Volino, R. J. (2020). Reynolds number dependence of zero pressure gradient turbulent boundary layers including third-order moments and spatial correlations. *Journal of Fluids Engineering*, *142*(5), 051303.

Vreman, A. (2004). An eddy-viscosity subgrid-scale model for turbulent shear flow: Algebraic theory and applications. *Physics of Fluids*, *16*(10), 3670–3681.

Wang, J.-X., Wu, J.-L., & Xiao, H. (2017). Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data. *Physical Review Fluids*, *2*(3), 034603.

Womack, K. M., Volino, R. J., Meneveau, C., & Schultz, M. P. (2022). Turbulent boundary layer flow over regularly and irregularly arranged truncated cone surfaces. *Journal of Fluid Mechanics*, *933*, A38.

Wu, C., Zhang, S., & Zhang, Y. (2025). Development of a generalizable data-driven turbulence model: Conditioned field inversion and symbolic regression. *AIAA Journal*, *63*(2), 687–706.

Xia, Z., Shi, Y., & Chen, S. (2016). Direct numerical simulation of turbulent channel flow with spanwise rotation. *Journal of Fluid Mechanics*, *788*, 42–56.

Xie, C., Wang, J., & W., E. (2020). Modeling subgrid-scale forces by spatial artificial neural networks in large eddy simulation of turbulence. *Physical Review Fluids*, *5*(5), 054606.

Yang, X. I. A., Abkar, M., & Park, G. (2024a). Grid convergence properties of wall-modeled large eddy simulations in the asymptotic regime. *Journal of Fluids Engineering*, *146*(8), 081501.

Yang, X. I. A., Chen, P. E., Zhang, W., & Kunz, R. (2024b). Predictive near-wall modelling for turbulent boundary layers with arbitrary pressure gradients. *Journal of Fluid Mechanics*, *993*, A1.

Yang, X. I. A., & Griffin, K. P. (2021). Grid-point and time-step requirements for direct numerical simulation and large-eddy simulation. *Physics of Fluids*, *33*(1), 015108.

Yang, X. I. A., Park, G. I., & Moin, P. (2017). Log-layer mismatch and modeling of the fluctuating wall stress in wall-modeled large-eddy simulations. *Physical Review Fluids*, *2*(10), 104601.

Yang, X. I. A., Pirozzoli, S., & Abkar, M. (2020a). Scaling of velocity fluctuations in statistically unstable boundary-layer flows. *Journal of Fluid Mechanics*, *886*, A3.

Yang, X. I. A., Sadique, J., Mittal, R., & Meneveau, C. (2015). Integral wall model for large eddy simulations of wall-bounded turbulent flows. *Physics of Fluids*, *27*(2), 025112.

Yang, X. I. A., Xia, Z.-H., Lee, J., Lv, Y., & Yuan, J. (2020b). Mean flow scaling in a spanwise rotating channel. *Physical Review Fluids*, *5*(7), 074603.

Yang, X. I. A., Zafar, S., Wang, J.-X., & Xiao, H. (2019). Predictive large-eddy-simulation wall modeling via physics-informed neural networks. *Physical Review Fluids*, *4*(3), 034602.

Yang, X. I. A., Zhang, W., Yuan, J., & Kunz, R. F. (2023). In search of a universal rough wall model. *Journal of Fluids Engineering*, *145*(10), 101302.

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press

Yurtsever, E., Lambert, J., Carballo, A., & Takeda, K. (2020). A survey of autonomous driving: Common practices and emerging technologies. *IEEE Access*, *8*, 58443–58469.

Zhou, Z., He, G., & Yang, X. (2021). Wall model based on neural networks for les of turbulent flows over periodic hills. *Physical Review Fluids*, *6*(5), 054610.

Cite this article: Yang Z., Bin Y., Shi Y. and Yang X. I. A. (2025). Large language model driven development of turbulence models. *Flow*, *5*, E40. https://doi.org/10.1017/flo.2025.10032

https://doi.org/10.1017/flo.2025.10032 Published online by Cambridge University Press