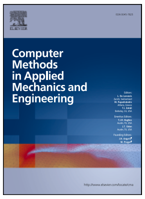
Computer Methods in Applied Mechanics and Engineering 449 (2026) 118506

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma



Reduced subgrid scale terms in three-dimensional turbulence

Rik Hoekstra *, Wouter Edeling

Centrum Wiskunde & Informatica, Scientific Computing group, Science Park 123, Amsterdam, 1098 XG, Netherlands

ARTICLE INFO

Keywords:
Reduced subgrid scale modelling
Large eddy simulation
Turbulence
Stochastic

ABSTRACT

Large eddy simulation (LES) has become a central technique for simulating turbulent flows in engineering and applied sciences, offering a compromise between accuracy and computational cost by resolving large-scale motions and modeling the effects of smaller, unresolved scales through a subgrid scale (SGS) model. The fidelity and robustness of LES depends critically on the SGS model, particularly in coarse simulations where much of the turbulence spectrum remains unresolved.

In this work, we extend the tau-orthogonal (TO) method, a data-driven SGS modeling framework, to three-dimensional turbulent flows. The method reformulates the high-dimensional SGS closure problem as a low-dimensional prediction task focused on spatially-integrated, scale-aware quantities of interest (QoIs). We extend the model to incorporate QoI-state dependence and temporal correlations by combining regularized least-squares regression with a multivariate Gaussian residual model. This yields a simple yet effective stochastic time-series prediction model, with orders-of-magnitude fewer parameters than typical deep learning approaches which try to directly learn the high-dimensional SGS closure.

We demonstrate the effectiveness of the new TO model in three-dimensional forced isotropic turbulence, turbulent channel flow and on a Taylor-Green vortex. The model achieves accurate long-term QoI distributions, robust performance across hyperparameter settings, and good reproduction of key flow features such as kinetic energy spectra and coherent structures, despite being trained solely on QoI trajectories. Comparisons against classical SGS models, including Smagorinsky and WALE formulations, highlight the new model's balance of accuracy and computational efficiency.

1. Introduction

Turbulent flows are everywhere. They play an important role in mixing of the Earth's atmosphere and oceans, as well as in aerodynamics. These flows span a vast range of spatial and temporal scales, which are hard to fully resolve in real-world applications. Capturing the smallest scales would require a number of grid points far beyond current computing capabilities. To circumvent this problem in turbulence simulations, large eddy simulation (LES) has become a popular method [1]. In LES, only the large-scale motions are explicitly resolved; whereas the smaller, unresolved scales are modeled using a subgrid scale (SGS) model. The performance of these large eddy simulations, depends heavily on the quality of the SGS model and the portion of the turbulent scales that needs to be modeled.

Classical SGS models, such as the Smagorinsky model [2] and its dynamic variants [3] have been widely employed due to their simplicity and empirical success, especially when the majority of the turbulent scales are resolved. They model the influence of the

---

* Corresponding author.
E-mail addresses: rik@cwi.nl (R. Hoekstra), edeling@cwi.nl (W. Edeling).

https://doi.org/10.1016/j.cma.2025.118506
Received 21 July 2025; Received in revised form 23 September 2025; Accepted 17 October 2025

subgrid scales as a function of the resolved field. While these models are useful, they have notable limitations when a larger portion of the scales needs to be modeled. In particular, they often struggle to adapt to different flow conditions, and they tend to oversimplify turbulent dynamics.

At the other end of the modeling spectrum lie Reynolds averaged Navier-Stokes (RANS) models, which model all turbulent fluctuations. The RANS equations essentially solve none of the turbulent scales, but are computationally cheap. A wide variety of RANS models have been developed to handle different flow scenarios [4]. However, they cannot easily leverage the growing availability of computational power, which now makes it feasible to resolve a substantial portion of the turbulence spectrum in many applications.

This evolving computational landscape opens the door for more flexible and expressive LES-SGS models, that can faithfully represent subgrid dynamics regardless of how much of the turbulence spectrum is resolved. This motivates our focus on test cases designed to emulate very coarse LES conditions.

Recent advances in data-driven modeling have offered promising tools to construct flexible SGS models for LES, particularly through deep learning techniques [5–10]. These models learn subgrid dynamics from training data sets, which can be obtained from high-resolution simulations. While these models provide greater adaptability to complex turbulent behaviors compared to classical SGS models, they also come with drawbacks. They are computationally more expensive and often act as “black boxes”, lacking interpretability and guarantees on generalization. Therefore, addressing these challenges is crucial for the wider adoption of deep learning in SGS modeling [11].

Over the recent years, several approaches have been explored to mitigate the drawbacks of deep learned SGS models. We mention some of these approaches to sketch the background for the methods introduced in this paper. Major efforts have been made on constraining the SGS models with known physical laws, using soft constraints or hard constraints [12,13] which do considerably improve the generalization of the models. Furthermore, it has turned out to be of importance that the training data, used to train the SGS model, is as consistent as possible with the data the model needs to process when employed in a solver. The authors of [14] and [15] have addressed the training-model data inconsistency in the filters of implicit LES, and explicit LES respectively.

Besides that, the community has gained insights into the advantages of *a-posteriori* learning compared to *a-priori learning*, highlighting a trade-off between training costs and model accuracy [16,17]. In a-priori learning, the SGS model is trained to reproduce subgrid terms given resolved flow snapshots, without feedback from the full simulation; whereas in a-posteriori learning, the model is trained in the context of a running solver, a computationally very costly procedure. A particularly promising middle ground between these strategies could involve implementing a nudging procedure during the learning process [18,19].

Additionally, slightly changing the framework to “ideal LES” [20] leads naturally to the concept of stochastic SGS models. This framework emphasizes that, given the resolved scales, the unresolved scales might still be in any configuration. Therefore, predicting the next state of the system should not be seen as a deterministic mapping. This perspective encourages the search for stochastic closure terms, which have been developed in simple form by [21] or can be obtained by training generative deep networks [22]. The “ideal” LES framework has also been unified with the field of data assimilation to obtain a simple stochastic SGS model [23].

Our research primarily addresses the high computational costs associated with training and evaluating the large deep learning models. Our approach, the tau-orthogonal method, is grounded in the observation that many practical applications of LES focus on a limited set of quantities of interest (QoIs)-such as average energy or enstrophy-rather than the full, high-dimensional flow field. By focusing on these QoIs, we reformulated the SGS modeling task as a low-dimensional (reduced) learning problem, significantly reducing the computational complexity while improving model interpretability [24].

The key to our approach lies in representing unresolved dynamics using a minimal set of scalar time series, one for each QoI. To obtain statistics of these time series, we use a nudging approach that forces LES simulations toward reference QoI trajectories. In earlier work [25], we demonstrated that this method can successfully reproduce the long-term distributions of four QoIs in a simple two-dimensional turbulence test case, using a random noise model for the time series, provided sufficient training data is available. In this formulation we enforce instantaneous matching of spatially-averaged QoIs. This choice mirrors common practice in deep learning based SGS models, which are typically trained on instantaneous subgrid stresses, and avoids the added complexity of temporal averaging. However, if fully converged long-term statistics are available, one could adopt approaches such as [26], where the SGS term enforces convergence to the correct invariant measure.

In this paper, we derive reduced SGS models for three-dimensional turbulent flows. Besides increasing the problem dimension, we move to a staggered-grid solver with a new coarse-graining procedure, demonstrating that Fourier-space-based, scale-aware QoIs can be coupled to solvers operating entirely in physical space. To better capture the temporal correlations of QoIs in three dimensions, we introduce a linear regression model, combining a small number of lagged QoI values with an additive noise term. This keeps the learning task low-dimensional and tractable, independent of spatial resolution. In contrast, applying conventional deep learning approaches to map directly from the resolved field to the SGS term becomes even less practical in three dimensions, as both input and output dimensions grow cubically with grid size, requiring vast datasets and significant compute resources.

The remainder of this paper is organized as follows. Section 2 introduces the coarse-grained Navier-Stokes equations and the QoIs. Section 3 presents our methodology in detail. Sections 4–6 validate the proposed model on three-dimensional isotropic turbulence, channel flow and a Taylor-Green vortex, comparing its performance to classical SGS models. After a discussion on generalization in Section 7, we conclude in Section 8.

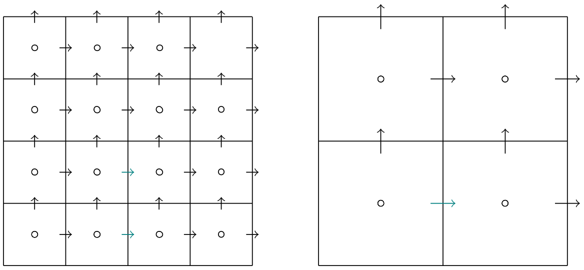


Fig. 1. Coarse-graining a two-dimensional staggered grid with face-averaging. The arrows indicate velocity components, and the dot denotes the location of the pressure.

2. Governing equations

The incompressible Navier-Stokes equations describe the time evolution of the velocity field $\mathbf{u} = (u_1, u_2, u_3)^T$ and the pressure field p in three dimensions. They are given by:

$$\nabla \cdot \mathbf{u} = 0, \tag{1}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u}\mathbf{u}^T) = -\nabla p + \nu \nabla^2 \mathbf{u} + \mathbf{f}, \tag{2}$$

where ν denotes the kinematic viscosity, and $\mathbf{f}$ is a forcing term used to sustain turbulence.

We employ an energy-conserving, incompressible Navier-Stokes solver [27] to solve these equations. This solver discretizes the equations on a staggered Cartesian grid. On this grid, velocity components are defined at cell faces, whereas pressure is defined at cell centers. A projection method eliminates the pressure term, resulting in the discretized pressure-free evolution equation (for details, see [15]):

$$\frac{d\mathbf{u}}{dt} = P\mathbf{F}(\mathbf{u}), \tag{3}$$

$$\mathbf{F}(\mathbf{u}) = -\mathbf{G}(\mathbf{u}\mathbf{u}^T) + \nu \mathbf{D}\mathbf{u} + \mathbf{f}, \tag{4}$$

where P is the pressure projection operator, $\mathbf{D}$ is the discretized divergence operator, $\mathbf{G}$ is the discretized gradient operator, and $\mathbf{F}(\mathbf{u})$ represents the combined effects of the convective, viscous, and forcing terms. The pressure projection operator ensures that the velocity field $\mathbf{u}$ remains divergence-free.

Direct numerical simulations (DNS) of simple turbulent flows are achievable using this solver. However, resolving all scales of motion requires a very fine grid resolution, rendering DNS computationally demanding. We distinguish DNS and high-fidelity (HF) reference simulations. Where the former resolves all turbulent scales, the latter does not fully do so, yet its resolution is sufficiently high to act as a trustworthy reference for our low-dimensional TO method.

2.1. *Coarse-graining and SGS term*

To alleviate the computational cost of simulations on a very fine grid, we aim to solve the equations on a coarser grid. In order to move from the solution at high-fidelity (or DNS) resolution to a solution on a coarser grid we used a filtering operation, denoted by

$$\bar{\mathbf{u}} = \Phi \mathbf{u}, \tag{5}$$

where $\mathbf{u}$ represents the HF velocity field, and $\bar{\mathbf{u}}$ is the coarse-grained velocity field. Following [15] we used a discrete face-averaging filter, which averages velocity components on the faces of coarser cells, as illustrated in Fig. 1. This filter maintains the divergence-free property of the coarse-grained velocity field with respect to the discretized divergence operator, in contrast to standard volume averaging-filters, which conserve divergence of the continuous divergence operator.

Applying this filtering operation to the pressure-free evolution equation yields the filtered equation:

$$\frac{d\bar{\mathbf{u}}}{dt} = \bar{P}\bar{\mathbf{F}}(\bar{\mathbf{u}}) + \mathbf{c}(\mathbf{u}, \bar{\mathbf{u}}), \tag{6}$$

where $\bar{P}$ and $\bar{F}$ are operators on the coarse grid, and $\mathbf{c}(\mathbf{u}, \bar{\mathbf{u}}) = \Phi P\mathbf{F}(\mathbf{u}) - \bar{P}\bar{\mathbf{F}}(\bar{\mathbf{u}})$ is the commutator error arising from the filtering operation. This term, commonly referred to as the subgrid scale (SGS) term, effectively represents the influence of unresolved small scales on the resolved scales. Two properties of this SGS term should be noted: it depends on the specific discretization of the solver, and it is divergence-free with respect to the discretized divergence operator on the coarse grid.

However, the SGS term depends on the HF field, which is not available in coarse simulations. Consequently, it must be approximated by a model $\mathbf{m}$ that depends solely on resolved scales. This leads to the low-fidelity model equation:

$$\frac{d\mathbf{v}}{dt} = \bar{P}(\bar{\mathbf{F}}(\mathbf{v}) + \mathbf{m}(\mathbf{v})), \tag{7}$$

where $\mathbf{v}$ denotes the solution of the low-fidelity simulation. We include the SGS model within the projection to ensure $\mathbf{v}$ remains (discretely) divergence-free.

The main challenge of LES lies in developing an SGS model and solver that satisfy the approximation:

$$\mathbf{v}(t) \approx \bar{\mathbf{u}}(t). \tag{8}$$

The interpretation of this approximation varies. While some models aim to replicate the filtered HF solution pointwise, a task rendered impractical over extended simulations due to the chaotic nature of turbulence, others focus on reproducing statistical properties. In this study, we focus on the time evolution of integrated quantities of the solution fields, which will be elaborated in the subsequent section.

2.2. Quantities of interest and scale-awareness

Instead of reproducing the full flow field, we focus on a limited set of physically relevant quantities of interest (QoIs) that encapsulate the key turbulent dynamics. We select QoIs, such as kinetic energy and enstrophy, for their ability to characterize turbulence across different scales. By targeting these integrated quantities, our approach reduces the dimensionality of the closure problem by several orders of magnitude, irrespective of the number of spatial dimensions in the problem, or the degrees of freedom in the filtered governing equations.

A QoI is defined as an integrated function of the coarse-grained solution:

$$Q = \int_{\Omega} q(\bar{\mathbf{u}})\mathrm{d}\boldsymbol{x}. \tag{9}$$

In this paper we examined two well-known quantities of such form; kinetic energy and enstrophy.
The total kinetic energy in the coarse-grained solution on computational domain Ω is given by

$$E = \frac{1}{2}\int_{\Omega} \|\bar{\mathbf{u}}\|^2 \,\mathrm{d}\boldsymbol{x}. \tag{10}$$

To capture dynamics across scales, we decomposed the energy into wavenumber bins in Fourier space. To this end, we introduce the total energy in Fourier space as

$$\hat{E} = \frac{1}{2}\frac{|\Omega|}{(N_x N_y N_z)^2}\sum_{\mathbf{k}} \hat{\mathbf{u}}_{\mathbf{k}} \cdot \mathrm{conj}(\hat{\mathbf{u}}_{\mathbf{k}}), \tag{11}$$

where $|\Omega|$ denotes the volume of the computational domain, $\hat{\mathbf{u}}$ is the Fourier-transformed coarse velocity field, $\mathbf{k}$ are the wavenumber vectors in Fourier space, and N_x denotes the number of grid points in the x-direction used for the Fourier expansion. This expression follows from applying quadrature in Fourier space, see Appendix A. We define the energy in wavenumber bin $[l, m]$ as follows

$$\hat{E}[l,m] = \frac{1}{2}\frac{|\Omega|}{(N_x N_y N_z)^2}\sum_{\mathbf{k}} \hat{R}_{[l,m]}\hat{\mathbf{u}}_{\mathbf{k}} \cdot \mathrm{conj}(\hat{R}_{[l,m]}\hat{\mathbf{u}}_{\mathbf{k}}). \tag{12}$$

Here we introduced the scale-aware sharp Fourier filter $R_{[l,m]}$

$$\begin{aligned} R_{[l,m]} &= \mathcal{F}^{-1}\hat{R}_{[l,m]}\mathcal{F}, \\ \hat{R}_{[l,m]}\,\hat{\mathbf{u}}_{\mathbf{k}} &= \begin{cases} \hat{\mathbf{u}}_{\mathbf{k}} & \text{if} \quad l - \frac{1}{2} \leq \|\mathbf{k}\|_2 < m + \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}. \end{aligned} \tag{13}$$

Similarly, the enstrophy is defined as the integral over the squared vorticity

$$Z = \int_{\Omega} \|\bar{\boldsymbol{\omega}}\|^2 \,\mathrm{d}\boldsymbol{x} = \int_{\Omega} \|\nabla \times \bar{\mathbf{u}}\|^2 \,\mathrm{d}\boldsymbol{x}, \tag{14}$$

which has a corresponding scale-aware form:

$$\hat{Z}[l,m] = \frac{|\Omega|}{(N_x N_y N_z)^2}\sum_{\mathbf{k}} \hat{R}_{[l,m]}\hat{\boldsymbol{\omega}}_{\mathbf{k}} \cdot \mathrm{conj}(\hat{R}_{[l,m]}\hat{\boldsymbol{\omega}}_{\mathbf{k}}). \tag{15}$$

where $\hat{\boldsymbol{\omega}}$ is the Fourier-transformed coarse vorticity field. This spectral decomposition of energy and enstrophy allows us to analyze turbulence characteristics across distinct spatial scales, improving the accuracy of the SGS model by focusing on key statistical properties of turbulent flow.

3. The tau-orthogonal method for reduced SGS modeling

The tau-orthogonal (TO) method offers a computationally efficient approach to SGS modeling focusing on a limited set of QoIs. Instead of modeling the SGS term pointwise, the TO method approximates the SGS term as a weighted sum of spatial patterns, each corresponding to a QoI:

$$\mathbf{m}(\mathbf{v},t) = \sum_{i=1}^{N_Q} \tau_i(t)\mathbf{O}_i(\mathbf{v}), \tag{16}$$

where N_Q is the number of QoIs, $\tau_i(t)$ are time-dependent scalar coefficients representing the unresolved contributions to the ith QoI, and $\mathbf{O}_i(\mathbf{v})$ are spatial patterns chosen such that each QoI is independently controlled.

We derive the TO method for a periodic computational domain. By projecting the governing equations onto the selected QoIs, we derive a reduced system that evolves their dynamics directly. We start from the ODEs for the selected QoIs:

$$\frac{\mathrm{d}Q_i}{\mathrm{d}t} = \int_\Omega \frac{\mathrm{d}q_i(\mathbf{v})}{\mathrm{d}t}\mathrm{d}\boldsymbol{x} = \int_\Omega \mathbf{V}_i \cdot \frac{\partial R_i \mathbf{v}}{\partial t}\mathrm{d}\boldsymbol{x}, \tag{17}$$

where R_i is the scale-aware filter used for the ith QoI, and $\mathbf{V}_i$ is the weak derivative of q_i w.r.t. $R_i\mathbf{v}$, i.e. $\int_\Omega q_i \frac{\partial \boldsymbol{\varphi}}{\partial R_i \mathbf{v}}\mathrm{d}\boldsymbol{x} = -\int_\Omega \mathbf{V}_i \boldsymbol{\varphi}\mathrm{d}\boldsymbol{x}$ for the set of test functions $\boldsymbol{\varphi}$. More detail on this can be found in [24,25]. We substitute the filtered Eq. (7) as expression for the time derivative:

$$\frac{\mathrm{d}Q_i}{\mathrm{d}t} = \int_\Omega \mathbf{V}_i \cdot R_i \bar{P}(\bar{\mathbf{F}}(\mathbf{v}) + \mathbf{m}(\mathbf{v}))\mathrm{d}\boldsymbol{x}. \tag{18}$$

We construct the spatial patterns in (16) such that they each affect only a single QoI. This is achieved by enforcing the following orthogonality conditions:

$$\int_\Omega \mathbf{V}_i \cdot \mathbf{O}_j \, d\boldsymbol{x} = 0 \quad \text{for } i \neq j. \tag{19}$$

Decomposing the evolution of each QoI into resolved and unresolved contributions leads to a system, where only the τ_i time series are unclosed:

$$\frac{\mathrm{d}Q_i}{\mathrm{d}t} = \frac{\mathrm{d}Q_i^r}{\mathrm{d}t} + \frac{\mathrm{d}Q_i^u}{\mathrm{d}t} = \frac{\mathrm{d}Q_i^r}{\mathrm{d}t} + \tau_i(t)\int_\Omega \bar{P}\mathbf{V}_i \cdot R_i \mathbf{O}_i \mathrm{d}\boldsymbol{x}. \tag{20}$$

This formulation allows us to model unresolved dynamics as a low-dimensional time-series problem, shifting the burden of subgrid modeling from high-dimensional spatial fields to a set of reduced-order coefficients.

3.1. Constructing spatial patterns

To fulfill the orthogonality constraints (19), we construct the basis functions $\mathbf{O}_i$ using weak derivatives $\mathbf{V}_i$ of the QoIs. For the scale-aware energy we derive, using (12):

$$\begin{aligned}\frac{\mathrm{d}E_{[l,m]}}{\mathrm{d}t} &= \frac{1}{2}C\int_\Omega \sum_{\alpha=1}^{3} \frac{\partial(R_{[l,m]}v_\alpha R_{[l,m]}v_\alpha)}{\partial R_{[l,m]}v_\alpha}\frac{\partial R_{[l,m]}v_\alpha}{\partial t}\mathrm{d}\boldsymbol{x} \\ &= C\int_\Omega R_{[l,m]}\mathbf{v} \cdot \frac{\partial R_{[l,m]}\mathbf{v}}{\partial t}\mathrm{d}\boldsymbol{x},\end{aligned} \tag{21}$$

where $C = |\Omega|/(N_x N_y N_z)^2$. Therefore $\mathbf{V}_i = R_{[l,m]}\mathbf{v}$.

Finding the weak derivative for the scale-aware enstrophy is less straightforward. Using the chain rule and integration by parts we find:

$$\begin{aligned}\frac{\mathrm{d}Z_{[l,m]}}{\mathrm{d}t} &= C\int_\Omega \sum_{\alpha=1}^{3} \frac{\partial(R_{[l,m]}\omega_\alpha R_{[l,m]}\omega_\alpha)}{\partial R_{[l,m]}\omega_\alpha}\frac{\partial R_{[l,m]}\omega_\alpha}{\partial t}\mathrm{d}\boldsymbol{x} \\ &= C\int_\Omega 2R_{[l,m]}\boldsymbol{\omega} \cdot \frac{\partial R_{[l,m]}(\nabla \times \mathbf{v})}{\partial t}\mathrm{d}\boldsymbol{x} \\ &= C\int_\Omega (2\nabla \times R_{[l,m]}\boldsymbol{\omega}) \cdot \frac{\partial R_{[l,m]}\mathbf{v}}{\partial t}\mathrm{d}\boldsymbol{x},\end{aligned} \tag{22}$$

where $C = |\Omega|/(N_x N_y N_z)^2$. So for the enstrophy $\mathbf{V}_i = 2\nabla \times R_{[l,m]}\boldsymbol{\omega}$.

Finally, the spatial patterns are constructed as:

$$\mathbf{O}_i = \sum_{j=1}^{N_Q} c_{ij} T_j(\mathbf{v}, \mathbf{x}), \quad \text{for} \quad 1 \leq i \leq N_Q \tag{23}$$

Here, the T_j are user-specified resolved basis functions, depending on the resolved solution $\mathbf{v}$. The c_{ij} are determined from the orthogonality constraints in (19), additionally requiring $c_{ii} = 1$. We choose $T_j = V_j$. This simple choice of basis functions has given good results in previous work on two-dimensional turbulence and allows to write the spatial patterns in closed form [24,25].

3.2. Predictor-corrector setup

The TO method is implemented in a predictor-corrector framework, where the low-fidelity solver first advances the system without the SGS term. The intermediate state, $\mathbf{v}^{n^*}$, is then corrected using the TO model:

$$\mathbf{v}^{n^*} = S(\mathbf{v}^{n-1}),$$

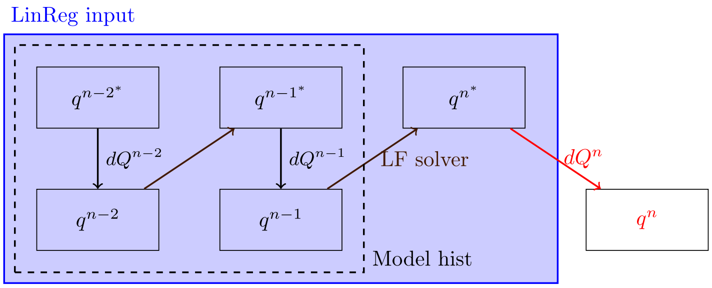


Fig. 2. Predictor-corrector subgrid scale term setup. During tracking, the corrections, dQ's, are computed using Eq. (25). We train models to predict either dQ^n or Q^n from historical inputs.

$$
\begin{aligned}
\mathbf{v}^n &= \mathbf{v}^{n^*} + \int_{t^{n-1}}^{t^n} \mathbf{m}(\mathbf{v}^{n^*})\mathrm{d}t, \\
&= \mathbf{v}^{n^*} + \mathbf{M}(\mathbf{v}^{n^*}),
\end{aligned}
\tag{24}
$$

where S denotes the low fidelity solver, and $\mathbf{M}$ represents a model for the correction term.

Tracking. The predictor-corrector split allows the low-fidelity simulation to track a reference trajectory of QoIs, $\mathbf{Q}^{\mathrm{ref}}(t^n)$, obtained from high-fidelity data. Given the predicted QoIs, $Q_i(\mathbf{v}^{n^*})$, we determine $\tau_i(t^n)$ from (20):

$$
\begin{aligned}
\tau_i(t^n) &= \frac{\mathrm{d}Q_i^u}{\mathrm{d}t} \Big/ \int_\Omega \bar{P}\mathbf{V}_i \cdot R_i \mathbf{O}_i \mathrm{d}\boldsymbol{x}, \\
\frac{\mathrm{d}Q_i^u}{\mathrm{d}t} &= \frac{1}{\Delta t}\left(Q_i^{\mathrm{ref}}(t^n) - Q_i(\mathbf{v}^{n^*})\right).
\end{aligned}
\tag{25}
$$

This nudging approach ensures that the low-fidelity solver remains dynamically consistent with the reference trajectory while allowing us to extract statistical properties of the SGS term.

Predicting corrections. An effective subgrid scale model should predict the SGS term without relying on reference trajectories. To this end, we train multivariate time-series models to predict the corrections based on historical data. These models learn the relationship between past QoI states and the unresolved-scale corrections.

The overall setup is summarized in Fig. 2, illustrating the correction process within the predictor-corrector framework. Here we introduce $dQ_i^n = \mathrm{d}Q_i^u/\mathrm{d}t|_{(t=t^n)} \cdot \Delta t$ as the SGS correction to the ith QoI at time t^n.

3.3. Time-series models for SGS correction

To model the unresolved dynamics, we adopt stochastic time-series models. Both the Mori-Zwanzig formalism [28] and the ideal LES framework [20] emphasize that, in a coarse-grained simulation, a single resolved state may correspond to multiple high-fidelity realizations. This inherent uncertainty suggests that the SGS correction should be treated probabilistically.

Thus, rather than modeling a deterministic mapping from $\mathbf{v}^{n^*}$ to dQ_i^n, we construct a distribution of possible corrections. Advancing the simulation then involves sampling from this learned distribution, effectively introducing a stochastic SGS term.

We explore two modeling approaches for predicting the SGS correction: a simple data-driven noise model and a linear regression model:

- **Data-driven noise model (DDM)**
  In previous work [25], we used a multivariate Gaussian model trained on the dQ corrections obtained from tracking. At each time step, the SGS correction was drawn from this learned distribution. While simple, this model does not account for temporal correlations or even a dependence on the current state of the system.
- **Linear regression with stochastic residuals (LRS)**
  To improve predictive accuracy, we extend the data-driven noise model with a linear regression term. This model predicts the corrected QoIs using a history of previous states:

$$
\begin{aligned}
\mathbf{LRS}(\underline{\mathbf{q}}_h) &= \underline{\mathbf{q}}_h C + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma), \\
\underline{\mathbf{q}}_h &= [\mathbf{q}^{n-1}, \dots, \mathbf{q}^{n-h}, \mathbf{q}^{n^*}, \mathbf{q}^{n-1^*}, \dots, \mathbf{q}^{n-h^*}, 1],
\end{aligned}
\tag{26}
$$

where $\underline{\mathbf{q}}_h$ is a vector with all model inputs for history length h, $C \in \mathbb{R}^{len(\underline{\mathbf{q}}_h)\times N_Q}$ is a learnable regression matrix, and $\mathcal{N}$ is a N_Q-dimensional multivariate Gaussian distribution. The "1" in the model inputs allows us to learn a bias term. We fit the model to

Table 1
High-fidelity parameters HIT (homogeneous isotropic turbulence).

N	Δt	$1/\nu$	T_L	Δt_f	e^*	k_f
512^3	$2.5 \cdot 10^{-4}$	2000	0.01	$2.5 \cdot 10^{-3}$	0.1	$\sqrt{2}$

predict the corrected QoIs $\mathbf{q}^n$. We solve a regularized least squares problem to find C. Where we regularize using the Frobenius-norm of C to promote long-term stability of the coupled LES-SGS system, since the Frobenius-norm penalizes large singular values:

$$C = \arg\min_{X \in \mathbb{R}^{len(\underline{\mathbf{q}}_h) \times N_Q}} \frac{1}{2} \|\mathbf{Q}_h X - \mathbf{Q}\|_F^2 + \lambda \|X\|_F, \tag{27}$$

where the nth rows of $\mathbf{Q}_h$ and $\mathbf{Q}$ are $\underline{\mathbf{q}}_h^n$ and $\mathbf{q}^n$ from the training data points, and λ is the regularization parameter. Before fitting the model, we standardize the inputs and outputs by dividing by the standard deviation of the reference trajectories.

3.4. *Final integration into LES*

By replacing explicit tracking with self-generated predictions for dQ_i, the final model achieves a fully independent SGS closure. The TO LRS model has only two hyper parameters: the history length h and the regularization strength λ. We explore their influence in the next section. The model has between 100 and 1000 learnable parameters, depending on the number of QoIs and the history length. This makes it computationally cheaper than traditional deep learning based SGS models, which can easily contain many millions of learnable parameters. Moreover, the time series model predicts corrections to the QoIs which are more interpretable.

4. Turbulence in a box

To evaluate the performance of our reduced SGS model and find suitable hyper parameter settings we first conduct long-term simulations of three-dimensional homogeneous isotropic turbulence (HIT) in a periodic box.

4.1. *Forcing*

Sustaining isotropic turbulence requires external forcing. The choice of forcing scheme influences large-scale flow structures and the resulting turbulence spectrum. A simple choice, forcing only one wavenumber in one direction, results in Kolmogorov flow [29]. However, this introduces anisotropy in the large scales.

To avoid this, we adopt a stochastic forcing scheme based on Ornstein-Uhlenbeck (OU) processes [30], following the efficient implementation by [31]. The forcing is applied in Fourier space, where only modes with wavenumber magnitude $\|\boldsymbol{k}\| \leq C_F$ are excited. Each forced Fourier coefficient $\hat{\boldsymbol{f}}_{\boldsymbol{k}}$ evolves independently as:

$$\hat{\boldsymbol{f}}_{\boldsymbol{k}}(t + \Delta t_f) = \hat{\boldsymbol{f}}_{\boldsymbol{k}}(t)\left(1 - \frac{\Delta t_f}{T_L}\right) + W(t)\left(2\sigma^2 \frac{\Delta t_f}{T_L}\right)^{1/2}, \tag{28}$$

where Δt_f is the forcing time step, T_L is the characteristic timescale, $W(t)$ denotes white noise, and σ^2 controls the variance. Instead of directly prescribing σ, we specify the energy injection rate as $e^* = \sigma^2 T_L$, see Table 1. The forcing field in physical space is obtained by an inverse Fourier transform and subsequently projected to ensure incompressibility.

4.2. *High-fidelity simulation*

To generate reference data, we performed a high-fidelity simulation at $N = 512^3$, with a time step of $\Delta t = 2.5 \cdot 10^{-4}$. The forcing term was updated every 10th time step, to enable consistent forcing on low-fidelity simulations. Table 1 summarizes the key parameters.

The velocity field was initialized at rest. Turbulence developed over a spin-up period of four time units. Fig. 3 shows the evolution of kinetic energy, illustrating how energy accumulates before reaching a statistically steady state. The flow statistics at $t = 4$ are listed in Table 2, and the corresponding energy spectrum is shown in Fig. 4. In the figure, we included the Kolmogorov length scale, the Taylor length scale and the grid size. We note that the Kolmogorov scale is close to the grid spacing, so this setup should be interpreted as a high-fidelity reference rather than a strict DNS. However, as shown in Appendix B, comparison with a fully resolved spin-up simulation on a 800^3 grid gives nearly unchanged turbulence statistics and energy spectra. Since the focus of this work is on evaluating whether reduced SGS models can learn corrections relative to a HF-LF gap, the 512^3 grid provides an appropriate and computationally tractable long-term reference simulation.

4.3. *Low-fidelity simulation*

For the low-fidelity (LF) simulations, we used a solver with resolution $N = 64^3$ and a time step ten times larger than that of the HF simulation, $\Delta t_{\mathrm{LF}} = 2.5 \cdot 10^{-3}$. Fig. 5 displays the energy spectrum of the coarse-grained initial velocity field, taken from the HF simulation at $t = 4$.

Table 2
Turbulence statistics at the end of HF HIT spin-up: Mean velocity, dissipation rate, Kolmogorov length scale, Taylor length scale, integral length scale, Reynolds number, Taylor-scale Reynolds number, large-eddy turnover time.

u_{avg}	ϵ	η	λ	L	Re_{int}	Re_{λ}	t_{int}
3.72	3.78	0.0024	0.096	0.15	1100	411	0.04

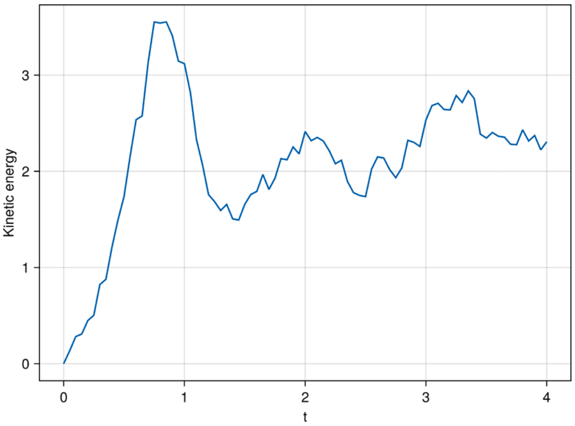


Fig. 3. Kinetic energy during spin-up of high-fidelity HIT solver.

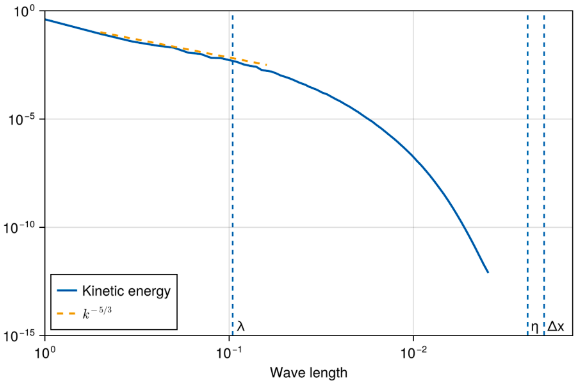


Fig. 4. Energy spectrum at $t = 4$ in high-fidelity HIT solver.

To assess the impact of different SGS models, we analyze six QoIs: scale-aware energy and enstrophy in three wavenumber bands: $[0, 6]$, $[7, 15]$, and $[16, 32]$. Fig. 7 shows the long-term distributions of these QoIs over 100 time units for a baseline LF simulation without an SGS model, compared against the HF reference data. The largest discrepancies are in the high wavenumber energy $E_{[16,32]}$ and enstrophy $Z_{[16,32]}$, where the long-term distributions shifted to the right, indicating accumulation of energy near the cutoff scale of the LF simulation.

To quantify such discrepancies throughout the paper, we use the Kolmogorov-Smirnov (KS) distance, which compares two distributions by measuring the maximum difference between their cumulative density functions $F(x)$ and $G(x)$:

$$KS\Big(F(x), G(x)\Big) = \max_{x} \Big|F(x) - G(x)\Big|. \tag{29}$$

The KS-distance is especially well-suited for this context because it does not require integration over the distributions' support, making it robust when comparing QoIs with different magnitudes or units. This allows us to summarize discrepancies across all QoI distributions in one number; the summed KS-distance. Table 3 reports the KS-distances between the long-term distributions for the QoIs in the low-fidelity and high-fidelity solver.

Table 3
KS-distance between long-term distributions in HF HIT simulation and LF HIT simulation with or without Smagorisky model.

	$Z_{[0,6]}$	$E_{[0,6]}$	$Z_{[7,15]}$	$E_{[7,15]}$	$Z_{[16,32]}$	$E_{[16,32]}$	sum
No model	0.16	0.047	0.17	0.29	0.87	0.80	2.34
Smag 0.071	0.13	0.078	0.22	0.19	0.067	0.021	0.705

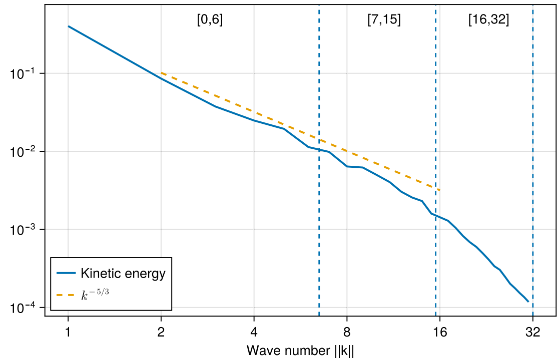


Fig. 5. HIT energy spectrum coarse-grained initial field.

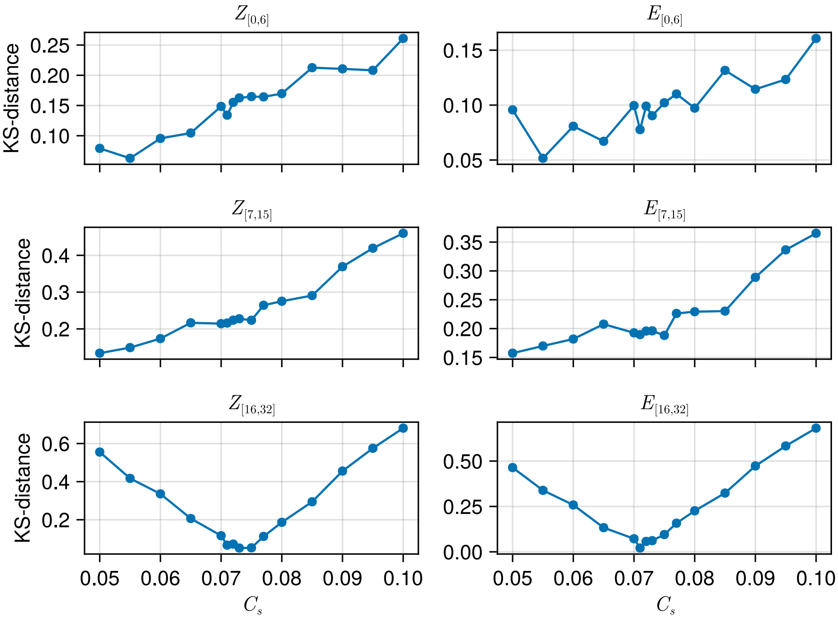


Fig. 6. Distance between 100 time unit QoI distributions in LF HIT simulations with Smagorinsky model and reference distributions, for different values of C_s.

4.4. Smagorinsky model

The Smagorinsky model [2] addresses the energy accumulation at high wavenumbers of the LF simulation by introducing an eddy-viscosity term based on the local strain rate:

$$\mathbf{m}(\mathbf{v}, C_s) = \nabla \cdot (2\nu_t S_{ij}), \tag{30}$$

$$\nu_t = C_s^2 \Delta^2 \sqrt{2 S_{ij} S_{ij}}. \tag{31}$$

Here, $S_{ij} = 1/2(\partial v_i/\partial x_j + \partial v_j/\partial x_i)$ is the strain rate tensor, Δ is the filter width (set to the LF grid size), and C_s is the Smagorinsky constant. We tuned C_s for best agreement with the reference distributions, as shown in Fig. 6. The best agreement, in terms of the

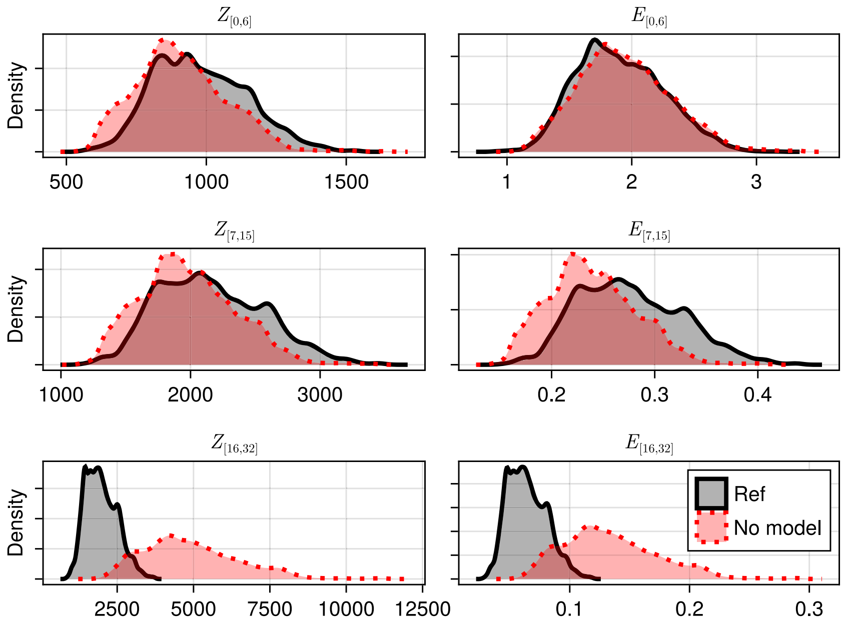


Fig. 7. Long-term HIT QoI distributions LF solver without SGS term.

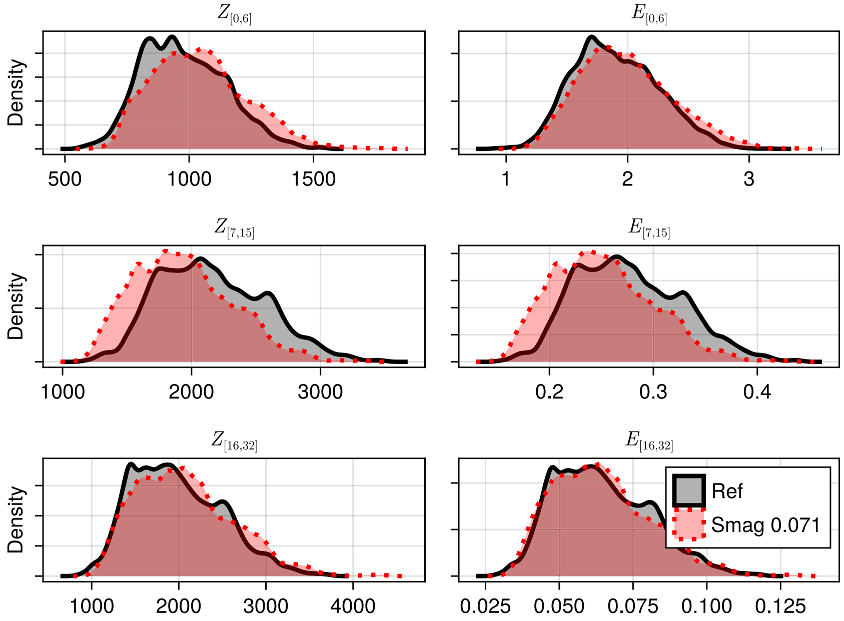


Fig. 8. Long-term HIT QoI distributions with Smagorinsky model.

summed KS-distance across all QoIs, was obtained with $C_s = 0.071$, yielding a summed KS-distance of 0.705. The corresponding long-term distributions are plotted in Fig. 8.

4.5. TO model

We trained the TO method using time series of the six QoIs from the HF simulation. Specifically, we used the first 10 time units of data to train the models and evaluated their performance over 100 time units. Training data for dQ was obtained by running a 10 time unit LF simulation that tracked the reference trajectories via Eq. (25). The effect of reducing the amount of training data is explored in Section 4.7.

Subsequently, we fitted linear regression models with stochastic residuals (LRS) using different history lengths. We also trained a data-driven noise model (DDN), which only fits a multivariate Gaussian distribution to the dQ data. These models were evaluated as

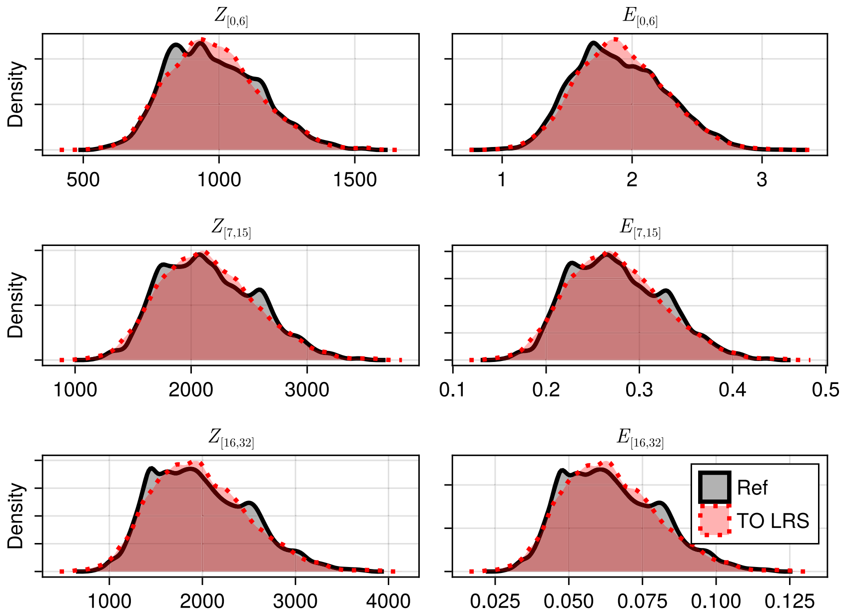


Fig. 9. Long-term HIT QoI distributions of 5-member ensemble with TO SGS term, using LRS with history length 5.

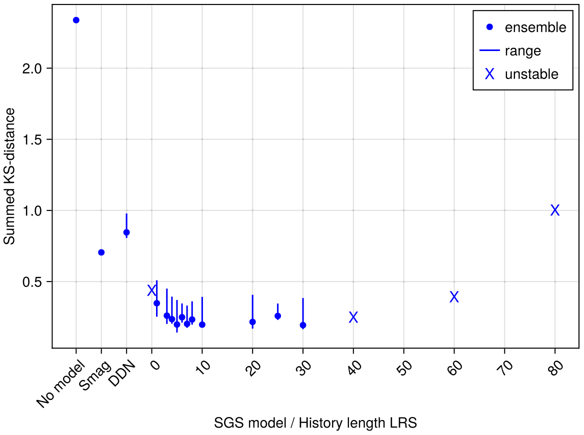


Fig. 10. Predictive quality of SGS models on HIT, extrapolating from a 10 time unit training domain to 100 time units. DDN refers to the data-driven noise model, LRS refers to the linear regression with stochastic residuals, and smag is the Smagorinsky model with $C_s = 0.071$. The plotted ranges are based on 5 replica simulations. The ensemble KS-distance results from aggregating the QoI data of all ensemble members. "X" markers indicate at least one unstable replica, their location is based on the ensemble KS-distance of the stable part of the trajectories.

subgrid scale term in LF simulations over 100 time units. Since these models are stochastic, we ran five replica simulations of each model with a different random seed.

Fig. 10 reports the KS-distances between the resulting long-term QoI distributions and the HF reference. We show the range between the KS-distance of the worst and best replica as well as the ensemble KS-distance, which results from aggregating the QoI data of all ensemble members in one EDF (empirical density function). Models leading to unstable trajectories in at least one replica are marked with an "X". The DDN model performs slightly worse than the optimized Smagorinsky model. In contrast, the LRS models exhibit good performance, clearly outperforming the Smagorinsky model, and achieving near-optimal accuracy across a wide range of history lengths. This indicates that the method is not highly sensitive to this hyperparameter. Instabilities are observed only for models with zero history (i.e., relying only on $\mathbf{q}^{n^*}$) or very long history lengths.

Fig. 9 contains the long-term distributions of one of the best performing models: LRS with history length 5. Here we plotted the histogram which combines the QoI trajectories of all five ensemble members.

All LRS models in this section were trained without regularization. Appendix C shows that adding L2 regularization improves stability for long history lengths, but at the cost of degraded long-term distribution accuracy.

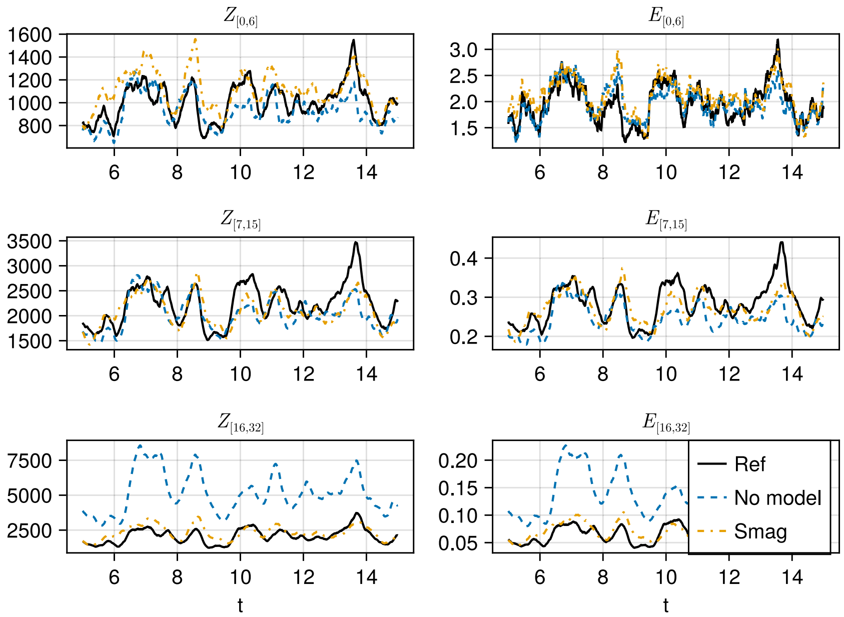


Fig. 11. HIT QoI trajectories without SGS model and with optimized Smagorinsky model.

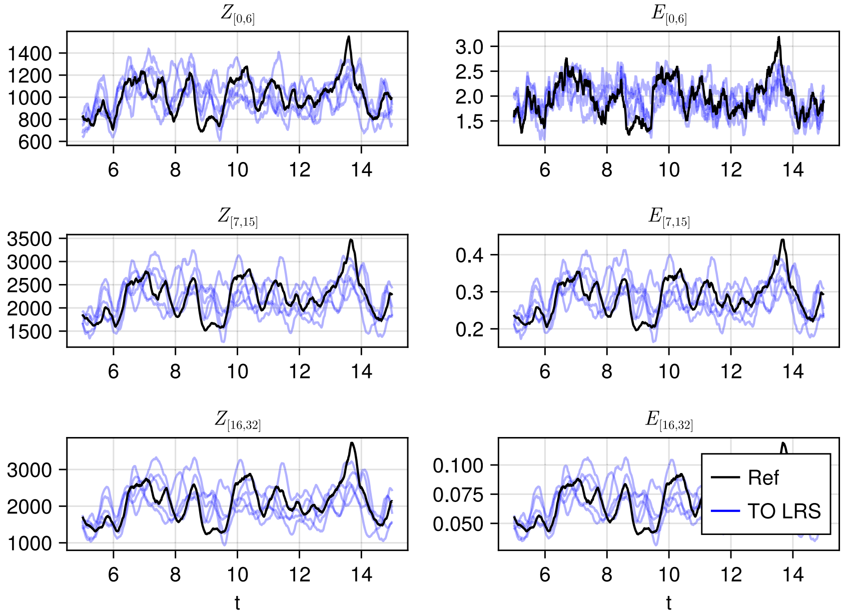


Fig. 12. HIT QoI trajectories with TO LRS model with history length 5.

4.6. Beyond QoI distributions

Thus far, we have shown that the TO method, combined with a linear regression model using a history length of five, accurately reproduces the long-term statistics of the QoIs it was trained on. In this section, we extend the analysis by examining segments of the resulting short-term QoI trajectories and evaluating additional turbulence characteristics; the energy spectrum and coherent flow structures.

Figs. 11 and 12 show short segments of the QoI trajectories for different SGS models. In Fig. 11, the no model case clearly deviates from the reference across most QoIs. The Smagorinsky model reduces this discrepancy, although some deviation in amplitude remains. Fig. 12 displays the TO model with history length five, plotted as five ensemble members. While the TO model generally follows the reference trajectory more closely and preserves variability in most bands, occasional underestimation of amplitudes shows remaining shortcomings in its predictive accuracy.

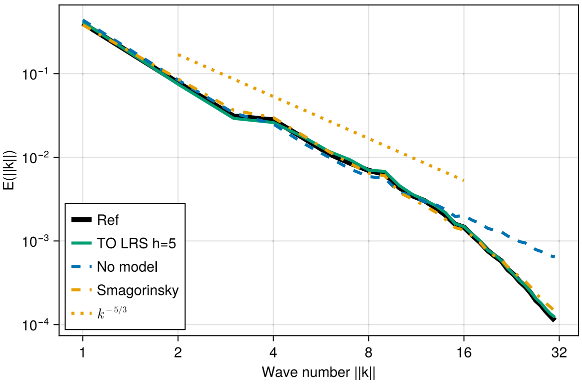
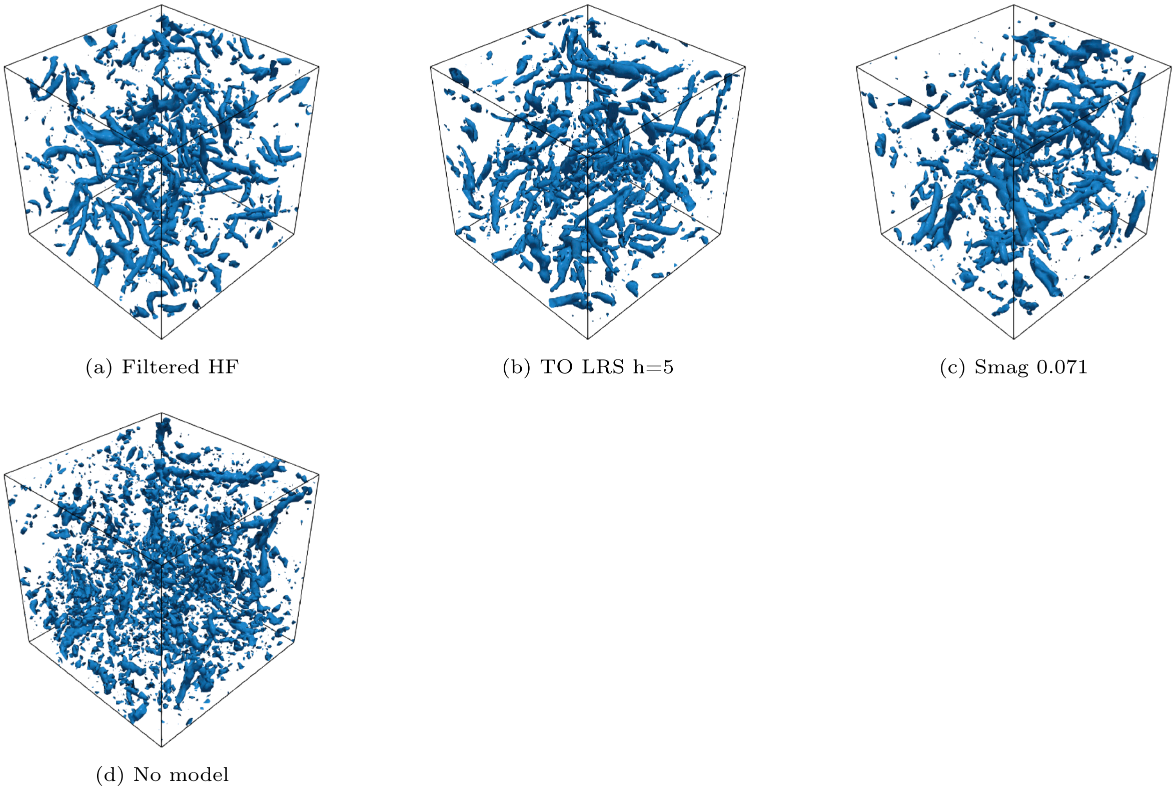


Fig. 13. HIT energy spectrum for various SGS models.


(a) Filtered HF

(b) TO LRS h=5

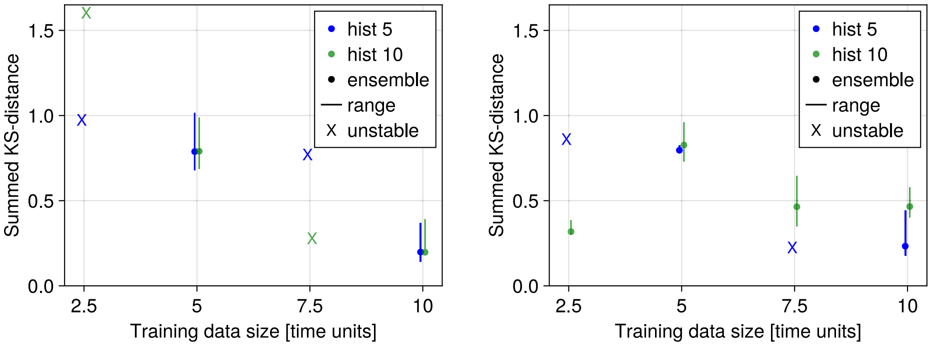
(c) Smag 0.071

(d) No model

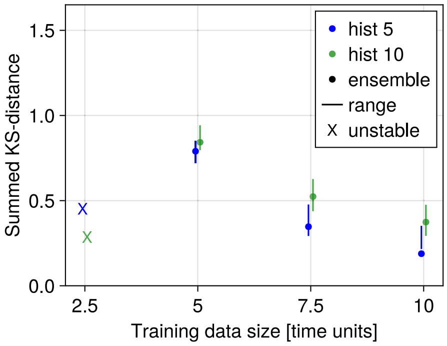
Fig. 14. Turbulent vortices at $T = 100$ in long-term HIT simulations, visualized via isocountours $Q = 2000$. Here $Q = \frac{1}{2}(\|\Omega\|^2 - \|S\|^2)$, and can be seen as the local balance between vorticity magnitude and shear strain rate.

Fig. 13 presents the time-averaged energy spectrum, computed over ten snapshots taken between $t = 75$ and $t = 100$ in the long-term simulations. For the TO model, we show results from the first ensemble member. The LF simulation without an SGS model shows clear energy accumulation at high wavenumbers. The TO model achieves a slightly improved match to the HF reference spectrum compared to the Smagorinsky model.

Fig. 14 visualizes the turbulent structures in the final fields of the long-term simulations using isocontours of the second invariant of the velocity gradient tensor, a standard method for identifying coherent vortices [1]. Both the TO and Smagorinsky models preserve these structures, while the simulation without an SGS model exhibits smaller, more numerous vortices.



(a) $\lambda = 0$ (b) $\lambda = 0.01$



(c) $\lambda = 0.1$

Fig. 15. Predictive quality of the TO LRS model on HIT with less training data, for various regularization strengths. Ranges are based on 5 replica simulations. "X" markers indicate at least one unstable replica, their location is based on the ensemble KS-distance of the stable part of the trajectories.

4.7. *Performance of TO LRS on limited training data*

To evaluate the data efficiency of the TO LRS method, we examined how model performance varies with the amount of available training data. Specifically, we trained models using history lengths of 5 and 10, across a range of training data sizes. We also explored the effect of different regularization strengths on performance.

The results of these experiments are presented in Fig. 15. We observe a decline in performance as the amount of training data decreases. Without regularization, models trained on 7.5 time units of data already exhibit instability in simulations. Introducing regularization mitigates this issue and results in a smoother dependence of performance on training data size.

A methodological note: we consistently excluded the initial unit of simulation time from the training set to avoid bias due to transient dynamics at the start of the low-fidelity simulations. As such, a model reported as trained on 2.5 time units effectively used data from $1 \leq t \leq 2.5$.

5. Channel flow

To demonstrate the versatility of the TO LRS approach, we apply it to a turbulent channel flow. Channel flow is a classical benchmark in turbulence research due to its relatively simple geometry and the presence of well-known turbulent structures. However, it presents a significant modeling challenge due to the different turbulence layers. Recall that the TO LRS model relies solely on the history of QoIs and is largely agnostic to geometry. This makes channel flow a stringent test of its generalizability.

5.1. *Setup*

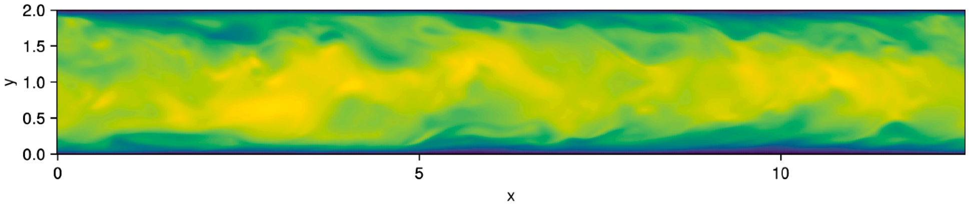
We ran simulations following the general setup described in [32], who performed direct numerical simulations of turbulent channel flow at a friction Reynolds number of $Re_\tau = 180$. They compared these simulations to existing data bases, among which [33]. We will use the publicly available DNS data from [32] to verify our SGS models for this test case.

The channel has dimensions $L_x \times L_y \times L_z = 4\pi \times 2 \times 4/3\pi$, where x, y, and z denote the streamwise, wall-normal, and spanwise directions, respectively. The domain is periodic in the streamwise and spanwise directions, with no-slip boundary conditions applied

Table 4
HF parameters channel flow.

$L_x \times L_y \times L_z$	$N_x \times N_y \times N_z$	Δt	Re_τ	ν	y^+	$\mathbf{f}$
$4\pi \times 2 \times 4/3\pi$	$512 \times 512 \times 256$	$5 \cdot 10^{-4}$	180	1/180	$180y$	$(1,0,0)^T$

(a) High resolution.



(b) Low resolution (coarse-grained).

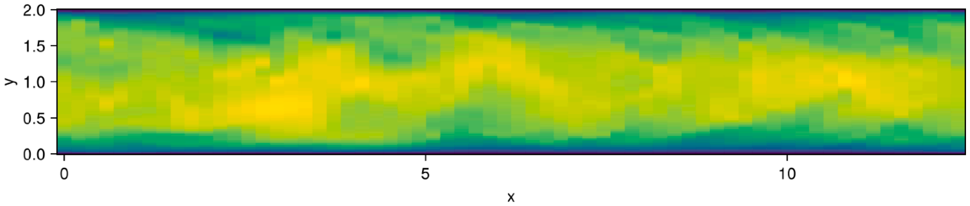


Fig. 16. Initial turbulent field for channel flow experiments, cross section of x-velocity $\mathbf{u}_1$ at $z = 0$.

at the walls ($y = 0$ and $y = 2$). A constant mean pressure gradient in the streamwise direction is imposed to drive the flow. We express y in wall units $y^+ = 180y$.

We adapted the initial conditions from [34] and used these in a 15 time unit spin-up simulation to create a turbulent initial field for our experiments, see Fig. 16a.

5.2. *High-fidelity simulation*

To generate training data (QoI trajectories), we ran our own high-fidelity simulation, using a $512 \times 512 \times 256$ equidistance grid. The setup is summarized in Table 4. We choose to use an equidistance grid to keep the coarse-graining operation and the expressions for the scale-aware QoIs simple. This choice yields a wall-normal cell size of approximately $0.7y^+$, which is comparable to the smallest cells used in the coarsest DNS of [32], who employed a $256 \times 128 \times 128$ grid with a tangent hyperbolic stretching in the wall-normal direction.

5.3. *Low-fidelity simulation*

The low-fidelity simulations were performed on a $64 \times 64 \times 32$ grid with a timestep of $\Delta t_{LF} = 5 \cdot 10^{-3}$. This setup results in a wall-normal cell size of approximately $5.6y^+$. We used face-averaging to coarsen the HF initial velocity field to the LF initial field. The coarse-grained initial field is plotted in Fig. 16b.

5.3.1. *Scale-aware QoIs*

As in the previous test case, we define six scale-aware QoIs based on energy and enstrophy content in three Fourier bands: $[0,3]$, $[4,10]$, and $[11,17]$. To ensure periodicity in all directions (a requirement for Fourier analysis), we symmetrically extend the solution field in the wall-normal direction. Specifically, a vector field $\mathbf{f}$ on the domain $4\pi \times 2 \times 4/3\pi$ is extended to a periodic vector field $\tilde{\mathbf{f}}$ on $4\pi \times 4 \times 4/3\pi$ by mirroring in the wall at $y = 2$:

$$\tilde{\mathbf{f}}(x,y,z) = \begin{cases} \mathbf{f}(x,y,z), & \text{if } y \leq 2, \\ -\mathbf{f}(x,4-y,z), & \text{if } y > 2. \end{cases} \tag{32}$$

Due to the domain size, the resulting Fourier wavenumbers are fractional. For instance, the longest non-constant wave in the streamwise direction corresponds to a wavenumber of $1/(4\pi)$. To illustrate the spatial features captured by the different Fourier bands, we visualized the velocity and vorticity fields which result from filtering the initial field with the different scale-aware filters in Appendix D.

Table 5
Calibrated eddy-viscosity model constants for LF channel flow.

	HF	Smag $C_s = 0.13$	WALE $C_w = 0.53$
mean $v_x(y=1)$	18.267	18.300	18.265

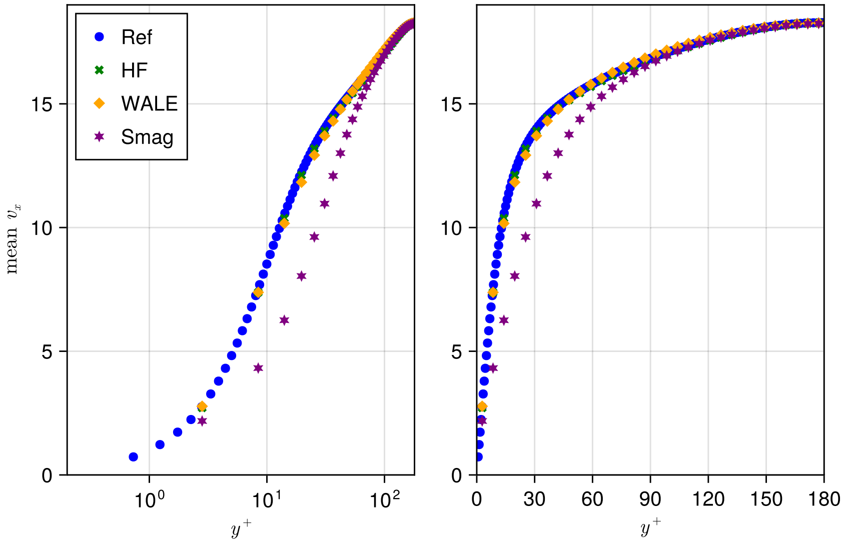


Fig. 17. Mean x-velocity profiles over 10 time units in LF channel flow solver with tuned eddy-viscosity models compared to coarse-grained HF solution and reference.

5.4. WALE model

As baseline method we use the WALE model. The WALE model is an eddy-viscosity model based on the square of the velocity gradient tensor [35]. It has two advantages over the Smagorinsky model in a channel flow simulation. Firstly, it takes into account both strain and rotation rates. And more importantly, its computed eddy-viscosity goes to zero near the channel walls. For incompressible flow, the model is given by

$$\mathbf{m}(\mathbf{v}, C_w) = \nabla \cdot (2\nu_t S_{ij}), \tag{33}$$

$$\nu_t = C_w^2 \Delta^2 \frac{(\mathfrak{S}_{ij}^d \mathfrak{S}_{ij}^d)^{3/2}}{(S_{ij}S_{ij})^{5/2} + (\mathfrak{S}_{ij}^d \mathfrak{S}_{ij}^d)^{5/4}}, \tag{34}$$

where the symmetric part of the square of the velocity gradient tensor is given by

$$\mathfrak{S}_{ij}^d = \frac{1}{2}(g_{ik}g_{kj} + g_{jk}g_{ki}). \tag{35}$$

Here, $g_{ij} = \partial v_i/\partial x_j$ is the velocity gradient tensor, $S_{ij} = 1/2(\partial v_i/\partial x_j + \partial v_j/\partial x_i)$ is the strain rate tensor, Δ is the filter width (again set to the LF grid size), and C_w is the WALE constant.

5.5. Fitting eddy-viscosity model constants

To calibrate the WALE and Smagorinsky models in the low-fidelity simulation, we adjust their constants so that the mean streamwise velocity at the channel center matches that of the high-fidelity simulation over the first 10 time units. Table 5 summarizes the tuned values. Fig. 17 compares the resulting mean x-velocity profiles, plotted against wall distance in both linear and logarithmic scales. Here we show the profile obtained from DNS validation data in [32], the mean profile from our 10 time unit HF simulation and the profiles of the calibrated Smagorinsky and WALE simulations.

5.6. TO model

The TO LRS method was trained on the same 10 time units of the high-fidelity simulation. First, we ran a low-fidelity simulation which tracked the reference trajectories of the six QoIs, obtaining a series of subgrid scale corrections. Then we trained an LRS model

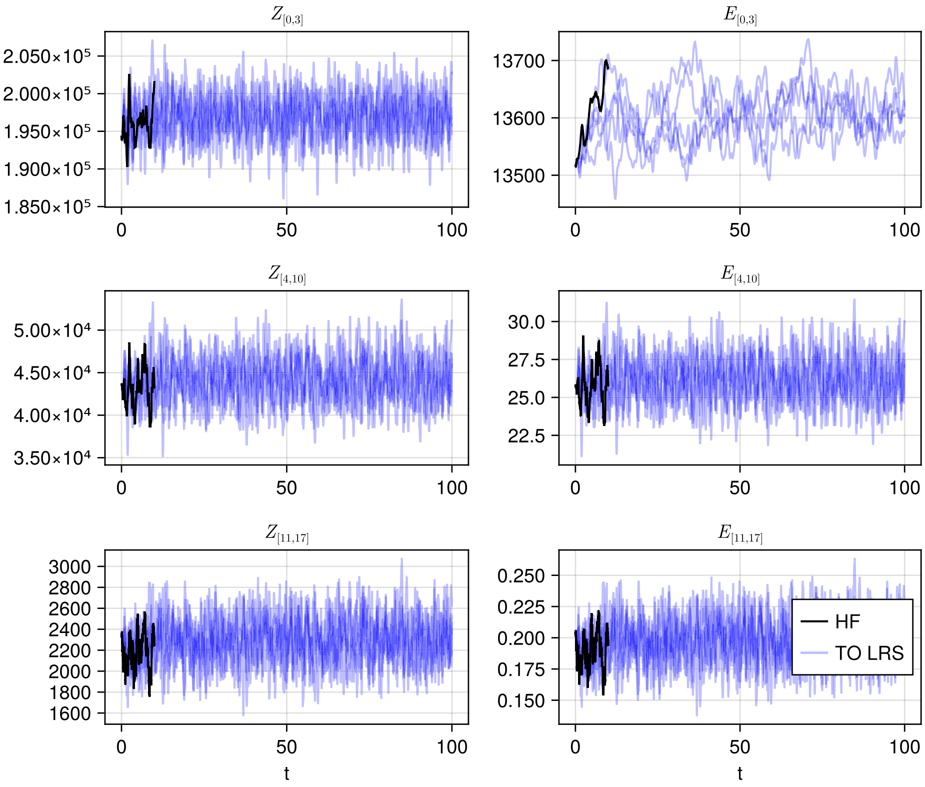


Fig. 18. Channel flow QoI trajectories for TO LRS model with history length 5 and for the HF training data.

to predict these subgrid scale corrections, see Section 3.3. We evaluated models with history lengths of 5 and 10, which previously performed well on the HIT test case.

On this test case these models lead to instable simulations if there is no regularization. However, introducing a small regularization ($\lambda = 10^{-4}$) was sufficient to get stable simulations.

5.7. Results

All subgrid scale models were evaluated in low-fidelity simulations over 100 time units. For the TO LRS models we again ran five replica simulations.

Figs. 18 and 19 show the trajectories of the six scale-aware QoIs for TO LRS with history lengths of 5 and 10, respectively. These plots include the corresponding trajectories from the HF simulation, which served as the training data. Across all five replicas, the TO LRS models produced stable simulations, and the QoI trajectories remained close to the range of values seen during training. This indicates that the TO LRS models generalize consistently over longer time horizons, without changing the distribution of the QoIs much.

In contrast, Fig. 20 presents the QoI trajectories for simulations using the eddy-viscosity models (Smagorinsky and WALE) as well as a simulation without SGS model. The Smagorinsky model and the no model simulations both exhibit significant loss of energy in the largest scales. additionally, the build-up of energy in the smaller scales is apparent for the simulation without SGS model. Interestingly, a similar small-scale energy accumulation is observed for the WALE model. This leads to trajectories with long-term distributions that clearly diverge from the HF data. However, the WALE model performs well when looking at the large-scale energy.

To further assess model performance, Fig. 21 shows the mean streamwise velocity profile across the channel for various SGS models. The profile from the first TO LRS replica with history length 5 is shown, though other replicas yielded nearly indistinguishable profiles. Among all models, WALE achieves the best match to the reference profile, particularly near the wall. The TO LRS model also performs well, accurately capturing the velocity profile away from the wall, even though it was not explicitly trained to do so. However, it leads to a to low streamwise velocity in the 4 grid points closest to the wall. In contrast, the Smagorinsky model fails to reproduce the correct profile, reflecting its known difficulty in handling wall-bounded turbulence.

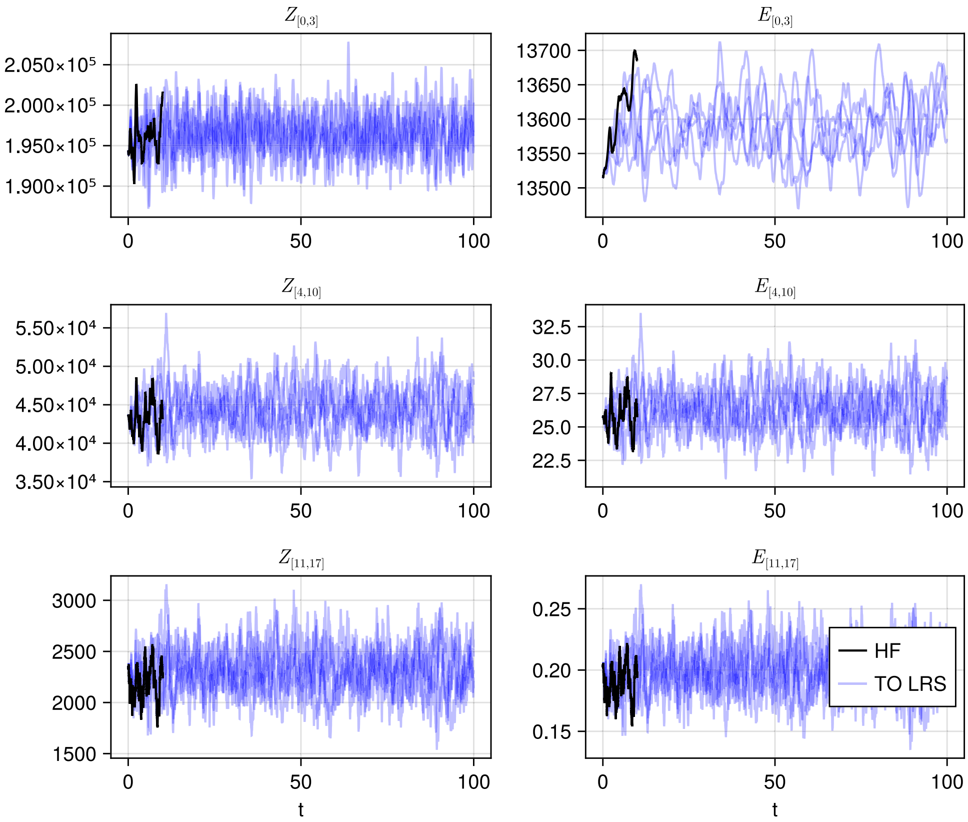


Fig. 19. Channel flow QoI trajectories for TO LRS model with history length 10 and for the HF training data.

Finally, Fig. 22 visualizes the turbulent structures at the end of the LF simulations using isocontours of the second invariant of the velocity gradient tensor. These are compared against the filtered HF solution at $T = 10$. The difference between the LF simulations is less pronounced than in the HIT test case. The LF simulation without SGS term again shows more numerous vortices. But in this test case the filtered HF solution also has more vortices than the LF simulations with SGS terms. Still, a notable result is that the stochastic corrections introduced by the TO LRS method do not break the coherent vortex structures.

6. Taylor-Green vortex

The Taylor-Green vortex is a demanding test case for SGS models due to its laminar-turbulent transition [36]. Classical SGS models are known to have difficulty predicting the onset of turbulence [37]. We consider this test case to show that the TO LRS model can handle a flow which is not in a statistically steady state. This test case shows that the TO LRS model can reproduce transition dynamics if they are part of the training data, yet it offers no possibility for extrapolation in time.

In its standard three-dimensional configuration, the flow is initialized on a triple-periodic cube of side-length 2π with:

$$u_x(x,y,z) = \sin(x)\cos(y)\sin(z), \quad u_y(x,y,z) = -\cos(x)\sin(y)\sin(z), \quad u_z(x,y,z) = 0. \tag{36}$$

We set $\nu = 1/800$. The initially laminar flow transitions into quasi-isotropic turbulence after 9 time units.

6.1. High-fidelity simulation

To generate reference data, we performed a high-fidelity simulation at $N = 512^3$, with a time step of $\Delta t = 5 \cdot 10^{-3}$. Fig. 23 shows the second invariant of the velocity gradient tensor at $t = 0$, $t = 10$ and $t = 20$, illustrating the onset and development of turbulence. Corresponding energy spectra are provided in Appendix E.

6.2. Low-fidelity simulation

The low-fidelity simulations were performed on a 64^3 grid with timestep $\Delta t_{\mathrm{LF}} = 5 \cdot 10^{-2}$. We defined again six scale-aware QoIs: energy and enstrophy in the wavenumber bands [0,1], [2,3] and [4,5]. Here the wavenumbers are expressed in unit length of the simulation, so the smallest nonzero wavenumber is $1/(2\pi)$ and the largest resolved wavenumber is $32/(2\pi)$.

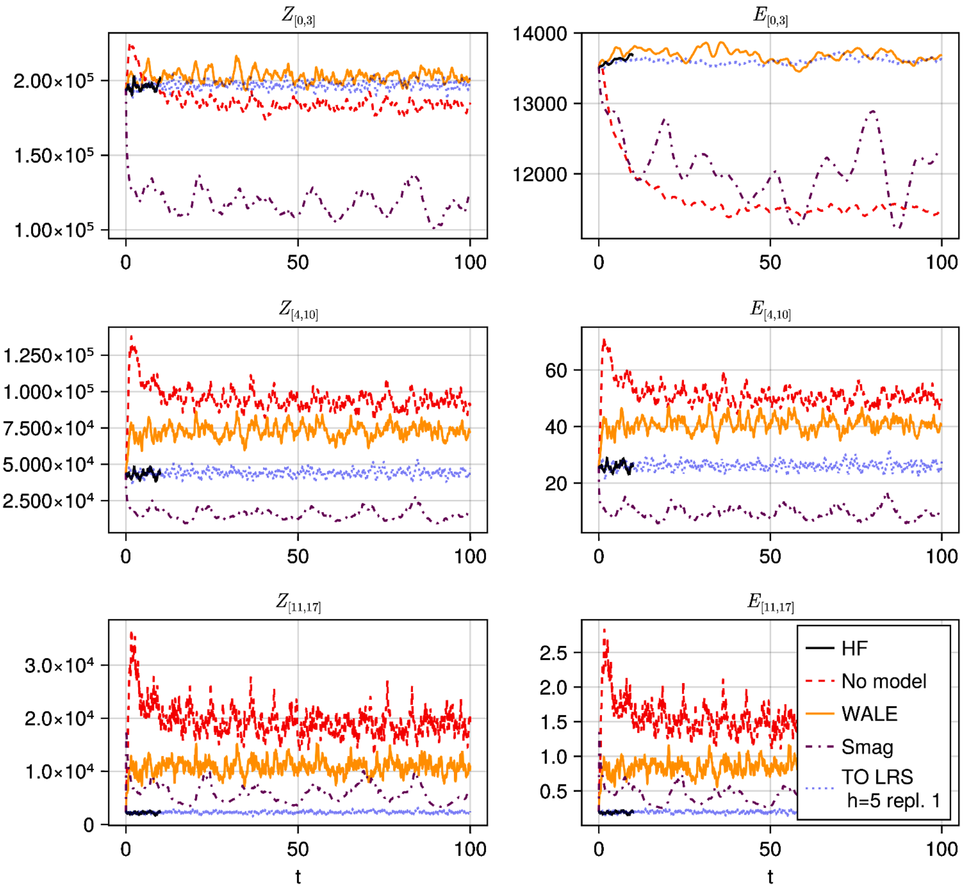


Fig. 20. Channel flow QoI trajectories for eddy-viscosity models, for the HF training data, and the first replica of the TO LRS model with history length 5.

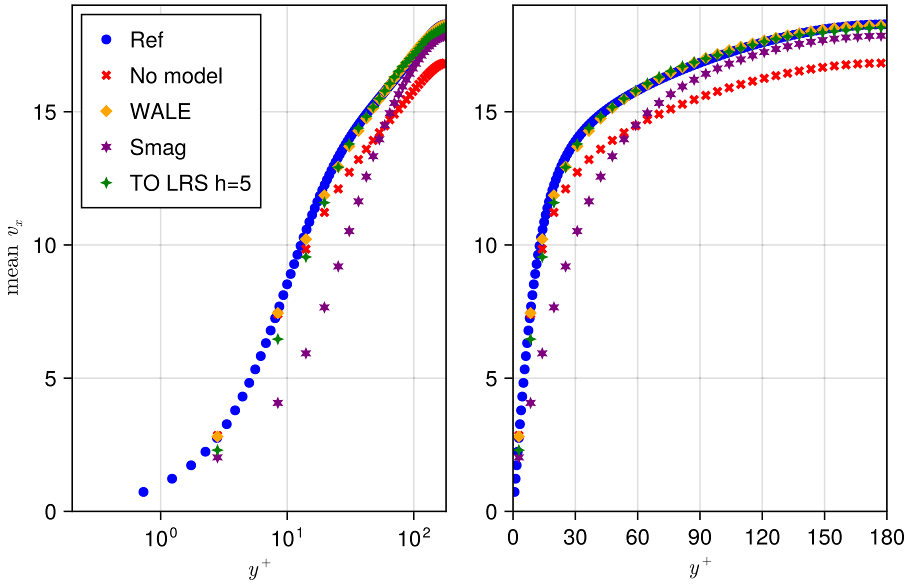
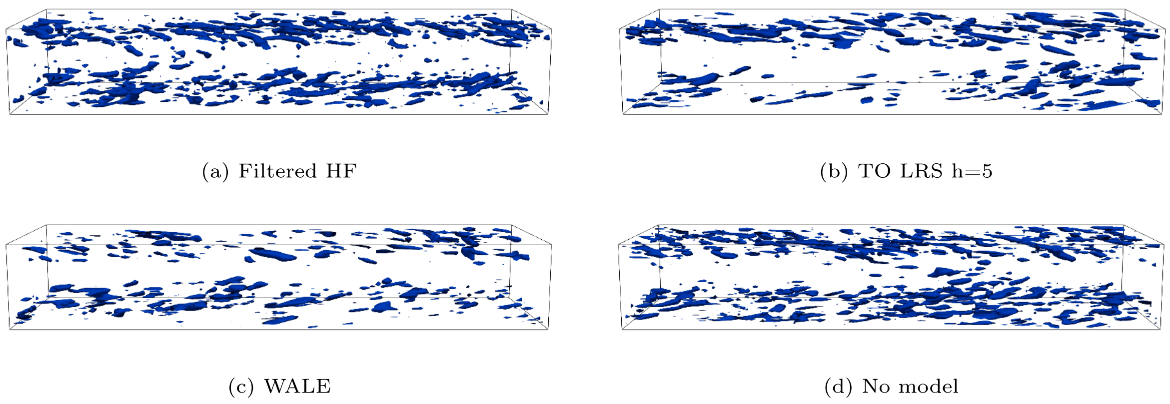


Fig. 21. Mean x-velocity profiles over 100 time units in LF channel flow simulations with various subgrid scale terms compared to reference.


(a) Filtered HF (b) TO LRS h=5

(c) WALE (d) No model

Fig. 22. Turbulent vortices in long term channel flow simulations at $T = 100$ compared to the filtered HF solution at $T = 10$, visualized via isocontours $Q = 100$.

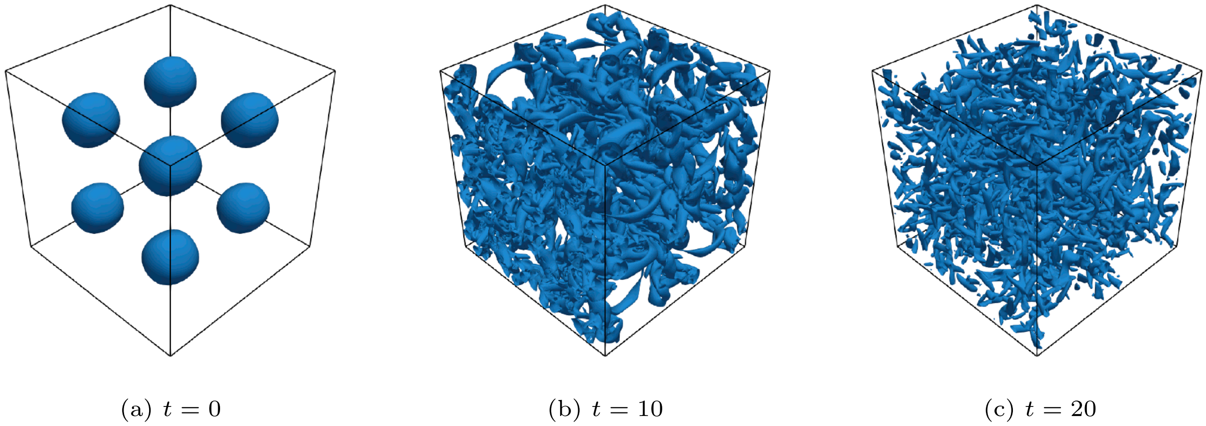

(a) $t = 0$ (b) $t = 10$ (c) $t = 20$

Fig. 23. Flow structures in HF Taylor-Green simulation, visualized via isocountours $Q = 0.5$.

6.3. Smagorinsky and WALE models

We ran low-fidelity simulations with the Smagorinsky and WALE models across a range of model coefficients. Figs. 24 and 25 compare their QoI trajectories with those of the HF simulation and with an LF simulation without SGS model.

Without SGS model, the LF simulation shows excess energy and enstrophy at the smallest scales. Around $t = 7$, energy and enstrophy in the mid scales grow too steeply, which suppresses the peak in large-scale enstrophy and leaves the simulation under predicting large-scale enstrophy later on.

For suitably chosen model constants the Smagorinsky and WALE model can dissipate the energy and enstrophy surplus in the smallest scales. However, neither the Smagorinsky model, nor the WALE model obtain satisfactory behavior on the mid- and large-scale QoIs.

6.4. TO model

The Taylor-Green vortex begins in a laminar state, with no energy or enstrophy at high wavenumbers. To avoid division by zero in the orthogonalization procedure, we set $dQ = 0$ in the TO method whenever any scale-aware QoI is zero. This choice is physically consistent: in the absence of small-scale eddies, no subgrid scale effects are expected.

We trained the TO LRS models on the full reference trajectories. First, we tracked the reference trajectories, then we fitted the LRS model to predict the corrections dQ on the data points where all the QoI values after the prediction step of the solver are non-zero, see Section 3.2.

We trained and evaluated LRS models with history lengths 5 and 10, combined with regularization strengths of 0, 0.005, and 0.02. Each configuration was tested in ensembles of five replica simulations. Models without regularization consistently lead to unstable simulations, that blew up within 6 time units. The QoI trajectories of the regularized models with history length 5 are shown in Figs. 26 and 27. In Appendix F (Figs. F.1 and F.2) we show the trajectories of the models with history length 10.

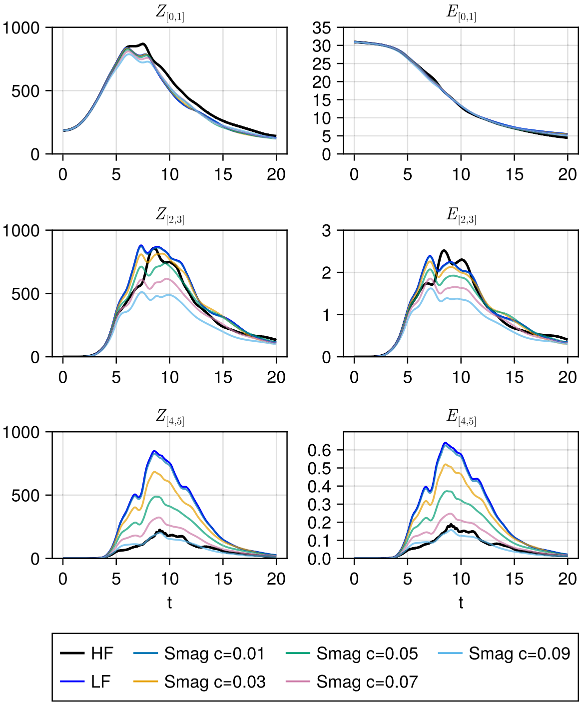


Fig. 24. Taylor-Green QoI trajectories Smagorinsky model compared to HF simulation and LF simulation without SGS model.

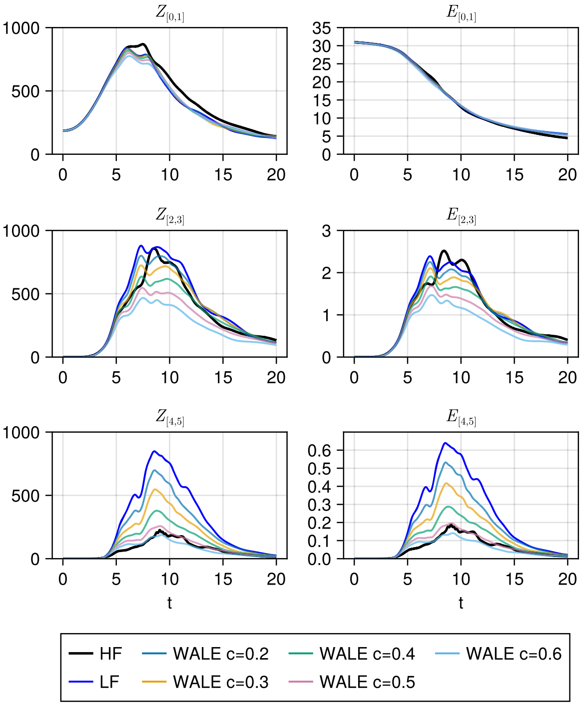


Fig. 25. Taylor-Green QoI trajectories WALE model compared to HF simulation and LF simulation without SGS model.

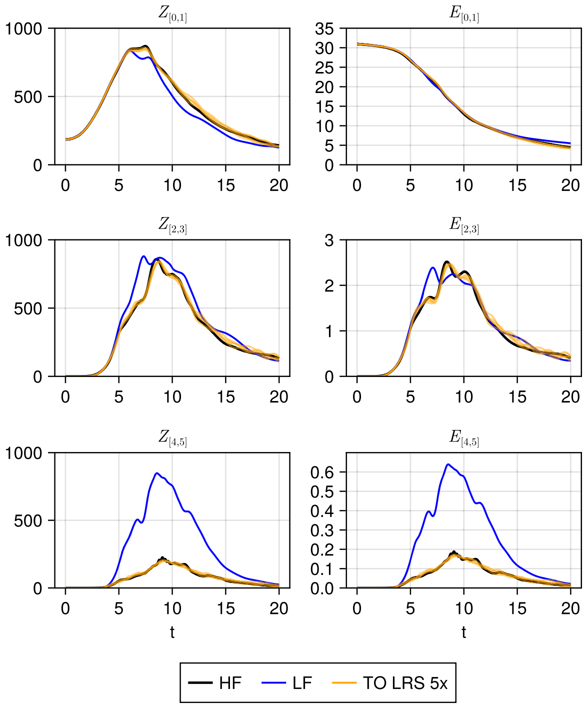


Fig. 26. Taylor-Green QoI trajectories TO LRS model with history length 5 and $\lambda = 0.005$ compared to HF simulation and LF simulation without SGS model.

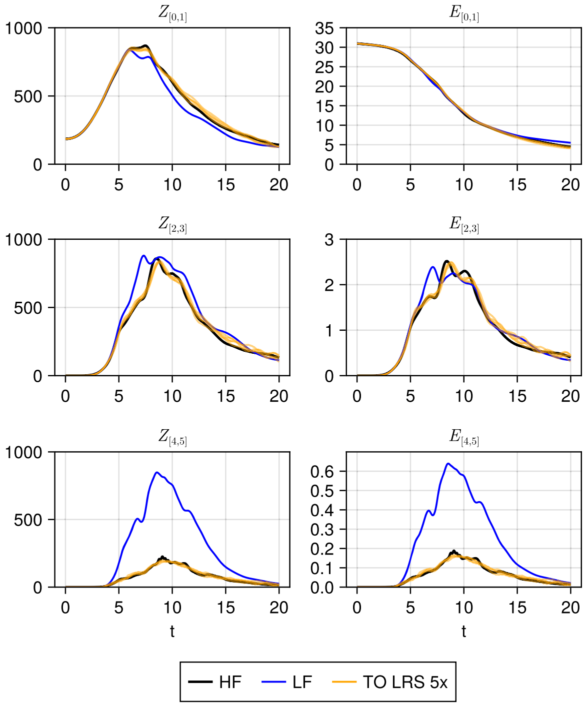


Fig. 27. Taylor-Green QoI trajectories TO LRS model with history length 5 and $\lambda = 0.02$ compared to HF simulation and LF simulation without SGS model.

All these simulations are stable and follow the reference trajectories closely. The spread between replica simulations is small and only noticeable near the end of the simulations. The TO LRS models with history length 10 tracked the reference trajectories slightly closer than those with history length 5. This illustrates the TO LRS method's ability to accurately capture transition dynamics.

7. Extrapolation and generalization

The main strength of the tau-orthogonal method is its ability to extrapolate in time on statistically stationary turbulence. From a short HF trajectory, the model can produce long LF runs that recover the correct long-term distributions of selected QoIs. As shown in Appendix G, this provides nearly an order-of-magnitude cost reduction compared to running the HF simulation for the same duration.

It is important to distinguish between temporal extrapolation and generalization across flow configurations. The TO method, in its current form, does not generalize across flows: each new configuration requires fresh HF reference trajectories for training. For example, altering the geometry, boundary conditions, or Reynolds number necessitates new HF data. This limitation contrasts with classical SGS models, which can often be applied more broadly without retraining, albeit at the expense of reduced accuracy.

Another aspect concerns solver dependence. The SGS term in Eq. (6) depends explicitly on the discretization of the solver. As a result, the TO closure is solver-specific. Switching to a different LF solver requires retraining the TO model. However, assuming the HF reference data as ground truth, no additional HF simulations are needed. This makes solver adaptation comparatively inexpensive.

For non-uniform grids, defining scale-aware QoIs is less straightforward than in the uniform case. The Fourier-based filtering used here is no longer directly applicable. One alternative is to employ localized kernels with varying widths to extract scale-dependent quantities in physical space. Developing such approaches would broaden the applicability of the TO framework to more complex discretizations and geometries, and represents an important direction for future work.

8. Conclusion

We have shown how the tau-orthogonal method can be extended to three-dimensional flows, a challenge for many data-driven approaches due to cubic scaling of the number of (unclosed) degrees of freedom that need to be modeled. As in the original two-dimensional method, we therefore reformulated the problem of modeling the high-dimensional SGS term as a closure problem for a small set of scale-aware QoIs: energy and enstrophy in large, mid, and small scales. Moreover, we extended the method to be QoI-state dependent, including history. This was achieved by combining a regularized least squares fit with a multivariate Gaussian distribution fitted to the residual (the LRS model). This new time-series prediction model can easily be fitted to training data obtained from tracking reference QoI trajectories, since it involves only two hyperparameters and a modest number of trainable parameters (in the order of 100 - 1000).

We conducted a hyperparameter study on three-dimensional homogeneous isotropic turbulence, from which we concluded that the LRS model exhibits robustness to the choice of history length, with a broad range of near-optimal values. We also saw that the regularization improves stability for long history lengths, but at the cost of degraded long-term distribution accuracy. The LRS model outperformed the Smagorinsky model in terms of obtaining correct long-term distributions for the QoIs. Furthermore, it improved on our previously introduced data-driven noise model, which had no dependence on QoI state.

Remarkably, despite being trained only on reproducing QoI trajectories, the TO LRS model appears to capture other salient features of the turbulent flow. Specifically, it reproduced the averaged kinetic energy spectrum slightly better than the Smagorinsky model. Moreover, it preserved coherent turbulent structures in the simulated fields.

To test the applicability of this approach to heterogeneous flows, we applied the TO LRS model to a turbulent channel flow. Despite being agnostic to geometry and wall-normal distance, the TO LRS model produced stable simulations with QoI trajectories that remained consistent with the training data. In contrast, the WALE model led to a slight energy build-up in the small-scale QoIs. The TO LRS model captured the mean velocity profile in the channel well, except near the wall, where its performance degraded slightly.

Finally, we tested the TO LRS model on a Taylor-Green vortex, a flow which is not in statistically steady state. We showed that the model can accurately capture transition dynamics, provided these are represented in the training data. In contrast, the Smagorinsky and WALE models failed to reproduce the correct evolution of mid- and large-scale QoIs.

In summary, the TO LRS method provides a simple, data-driven approach to modeling SGS effects that generalizes in time beyond the training domain for statistically steady state turbulence. Moreover, it can capture transient dynamics provided these are present in the training data. This makes it a promising alternative to classical SGS models and traditional deep learning approaches for a broad range of flow configurations. Main avenues for future work will involve improving the basis functions in the TO method for better near-wall behavior and extending the method to non-structured grids.

CRediT authorship contribution statement

Rik Hoekstra: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Wouter Edeling:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Data availability

Code to reproduce all results is publicly available at Zenodo.org and cited in the manuscript.

Software & data availability

The implementation of the methods which are presented in this paper is publicly available in the repository “TO_Incompressible-NavierStokes.jl” [38]. This repository contains scripts and data to reproduce all plots in this paper.

Funding

This research is funded by the Netherlands Organization for Scientific Research (NWO) through the ENW-M1 project “Learning small closure models for large multiscale problems” (OCENW.M.21.053). This work used the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-9277.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to rewrite sections of the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Quadrature in the Fourier domain

Our scale-aware QoIs are based on local energy/enstrophy density in Fourier space. In this section we show that these densities are related to the total energy in physical space via quadrature. We rely on the Discrete Fourier transform of the velocity field, which gives Fourier coefficients $\hat{\mathbf{u}}_{\mathbf{k}}$ such that

$$\mathbf{u}(x, y, z) \approx \frac{1}{N_x N_y N_z} \sum_{k_1=0}^{N_x-1} \sum_{k_2=0}^{N_y-1} \sum_{k_3=0}^{N_z-1} \exp\left[2\pi i\left(\frac{x k_1}{L_x N_x} + \frac{y k_2}{L_y N_y} + \frac{z k_3}{L_z N_z}\right)\right]\hat{\mathbf{u}}_{\mathbf{k}}, \tag{A.1}$$

where N_x denotes the number of grid points in the x-direction and L_x denotes the length of the computational domain in this direction.

Let us now consider the total kinetic energy of a real-valued velocity field $\mathbf{u}$

$$\begin{aligned} E &= \frac{1}{2}\int_\Omega \mathbf{u}\cdot\mathbf{u}\,\mathrm{d}\boldsymbol{x}, \\ &\approx \frac{1}{2}\int_\Omega \frac{1}{(N_x N_y N_z)^2} \sum_{\mathbf{k}} \sum_{\mathbf{l}} \exp\left[2\pi i\left(\frac{x_1(k_1+l_1)}{L_x N_x} + \frac{x_2(k_2+l_2)}{L_y N_y} + \frac{x_3(k_3+l_3)}{L_z N_z}\right)\right]\hat{\mathbf{u}}_{\mathbf{k}}\cdot\hat{\mathbf{u}}_{\mathbf{l}}\,\mathrm{d}\boldsymbol{x}, \\ &= \frac{1}{2}\frac{|\Omega|}{(N_x N_y N_z)^2}\sum_{\mathbf{k}} \hat{\mathbf{u}}_{\mathbf{k}}\cdot\hat{\mathbf{u}}_{-\mathbf{k}}, \\ &= \frac{1}{2}\frac{|\Omega|}{(N_x N_y N_z)^2}\sum_{\mathbf{k}} \hat{\mathbf{u}}_{\mathbf{k}}\cdot\mathrm{conj}(\hat{\mathbf{u}}_{\mathbf{k}}), \end{aligned} \tag{A.2}$$

where $|\Omega|$ denotes the volume of the computational domain. Here, we used the orthogonality of Fourier modes to get from the second to the third expression. The last step only holds for fields that are real-valued in physical space.

Appendix B. HIT on a finer grid

We compare the final field of the high-fidelity spin-up simulation with the final field of a DNS with higher resolution. The DNS is performed on an 800^3 grid and covers 4 time units, see Table B.1. Table B.2 contains the turbulence statistics of the final fields of both simulations. Fig. B.1 shows the energy spectra of the final fields. The difference between the energy spectra and turbulence statistics of the simulations is small. Indicating that our high-fidelity simulation will contain most turbulence characteristics found in the full DNS.

Table B.1
DNS parameters HIT (homogeneous isotropic turbulence).

N	Δt	$1/\nu$	T_L	Δt_f	e^*	k_f
800^3	$2.5 \cdot 10^{-4}$	2000	0.01	$2.5 \cdot 10^{-3}$	0.1	$\sqrt{2}$

Table B.2
HF and DNS turbulence statistics at the end of HIT spin-up: Mean velocity, dissipation rate, Kolmogorov length scale, Taylor length scale, integral length scale, Reynolds number, Taylor-scale Reynolds number, large-eddy turnover time.

N	u_{avg}	ϵ	η	λ	L	Re_{int}	Re_λ	t_{int}
512^3	3.72	3.78	0.0024	0.096	0.15	1100	411	0.04
800^3	3.59	3.28	0.0025	0.099	0.15	1056	411	0.04

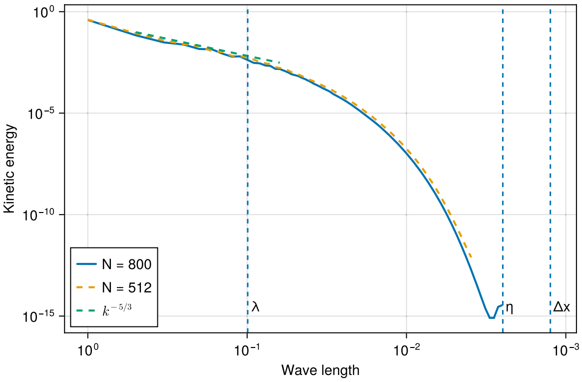


Fig. B.1. HIT energy spectra at $t = 4$, for DNS and HF simulation. The Taylor length scale, Kolmogorov length scale and gird spacing of the DNS ($N = 800^3$) are plotted.

Appendix C. Regularization in the LRS model

We examine the effect of L2-regularization on the TO LRS models in the HIT test case. Fig. C.1 shows the outcome of applying a mild regularization term ($\lambda = 0.01$). The results indicate that regularization stabilizes models with longer history lengths. Conversely, it causes instability in models with shorter history.

Overall, the addition of regularization tends to degrade model performance. This is further illustrated in Fig. C.2, which compares the long-term QoI distributions for the regularized and non-regularized TO LRS model with a history length of 10. Regularization in this case results in a to low variance.

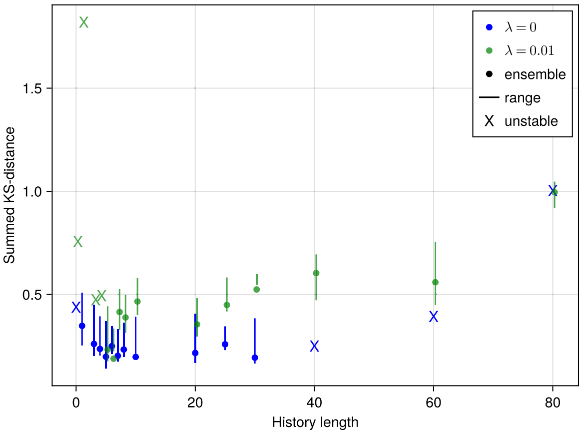
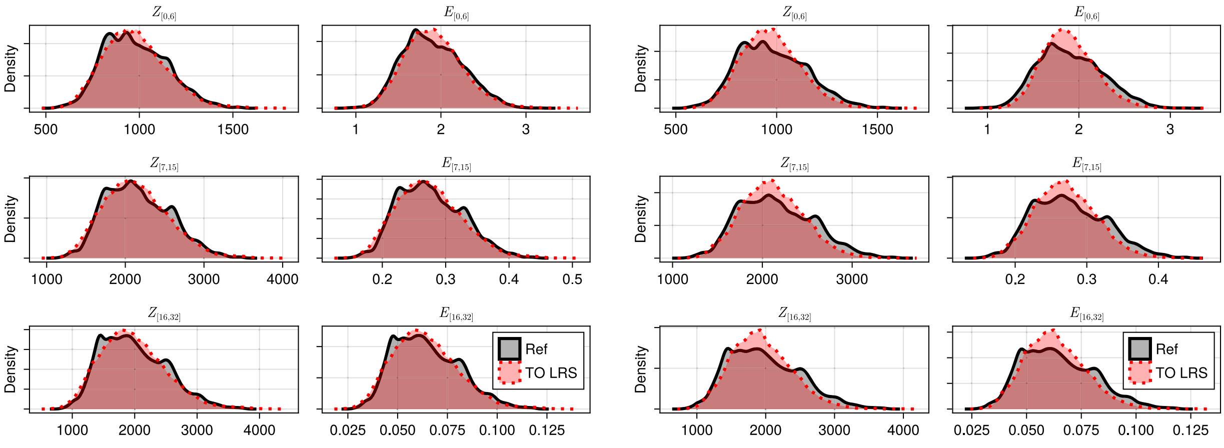


Fig. C.1. Influence of regularization on predictive quality LRS model in the HIT test case, extrapolating from a 10 time unit training domain to 100 time units. The plotted ranges are based on 5 replica simulations. The ensemble KS-distance results from aggregating the QoI data of all ensemble members. “X” markers indicate at least one unstable replica, their location is based on the ensemble KS-distance of the stable part of the trajectories.



(a) $\lambda = 0$ (b) $\lambda = 0.01$

Fig. C.2. Long-term HIT QoI distributions of 5-member ensemble TO LRS with history length 10, with and without regularization.

Appendix D. Scale-aware filters applied to initial field channel flow

For our channel flow test case we defined six scale-aware QoIs: energy and enstrophy in three wavenumber bands: $[0,3]$, $[4,10]$, and $[11,17]$. To get a feeling for these QoIs, we plotted the velocity and vorticity fields which result from filtering. Fig. D.1 shows the velocity and vorticity magnitudes of the scale-aware filtered initial field for: $R_{[l,m]}\mathbf{u}$ and $R_{[l,m]}\boldsymbol{\omega}$. We see that the large-scale energy is located in the bulk of the channel. The other QoIs are highest near the walls, as expected.

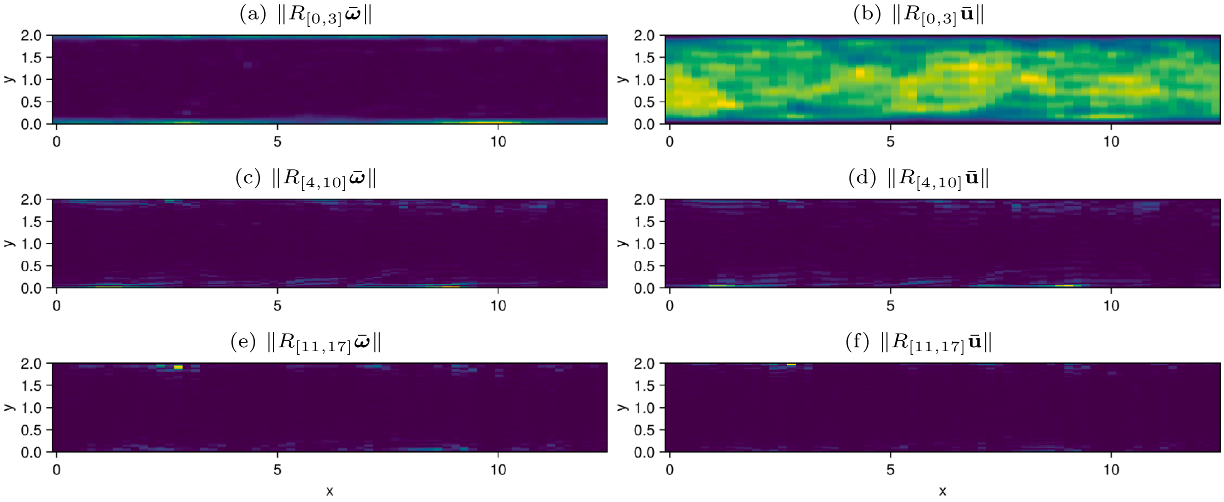


Fig. D.1. Scale-aware filters applied to initial turbulent field of channel flow test case.

Appendix E. High-fidelity energy spectra Taylor-Green vortex

Fig. E.1 shows the energy spectra of the high-fidelity Taylor-Green simulation at $t = 10$ and $t = 20$. At $t = 10$, just after the onset of turbulence, the solution field contains the most small-scale eddies. We see that our grid is just fine enough to resolve them. Later in the simulation the turbulent kinetic energy decreases.

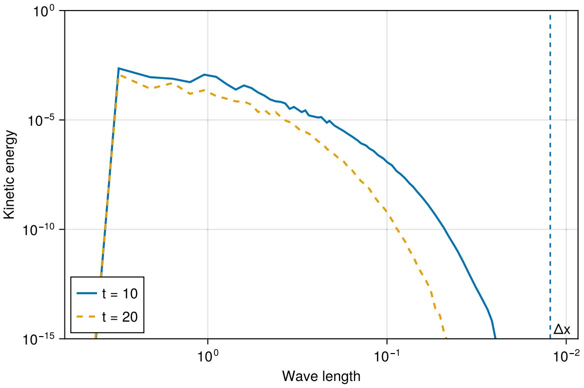


Fig. E.1. Energy spectra in HF Taylor-Green simulation.

Appendix F. Additional TO LRS results on Taylor-Green vortex

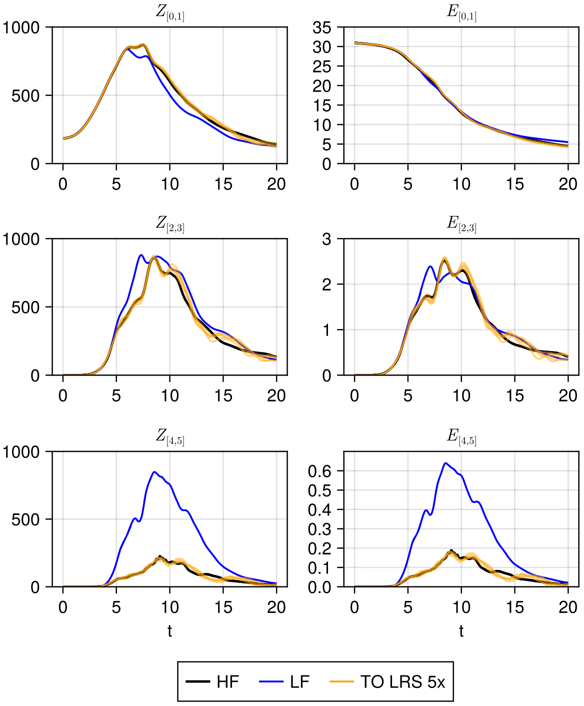


Fig. F.1. Taylor-Green QoI trajectories TO LRS model with history length 10 and $\lambda = 0.005$ compared to HF simulation and LF simulation without SGS model.

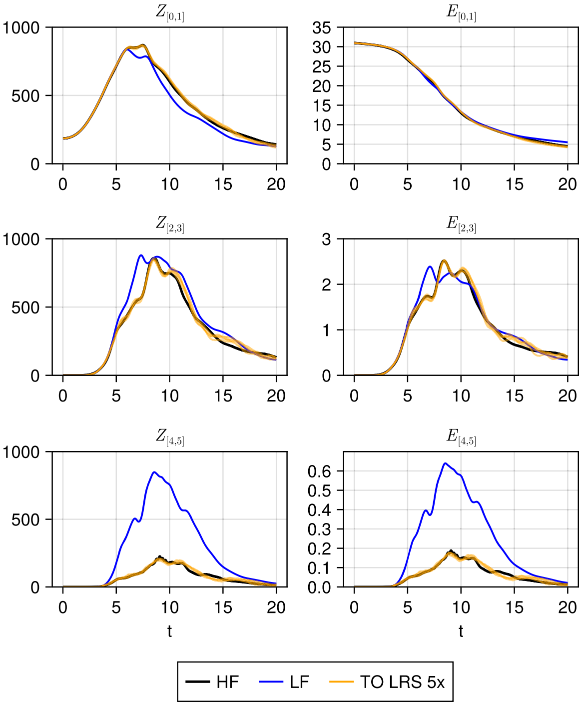


Fig. F.2. Taylor-Green QoI trajectories TO LRS model with history length 10 and $\lambda = 0.02$ compared to HF simulation and LF simulation without SGS model.

Appendix G. Computational cost

In this section we report on the computational cost of the simulations presented in this paper. All simulations were performed on GPU-accelerated nodes of the Dutch national supercomputer *Snellius*. The high-fidelity simulations were performed using a single NVIDIA H100-SXM5-80GB GPU. The low-fidelity simulations, being computationally much lighter, were run on A100-SXM4-40GB GPUs which were split into two using multi-Instance GPU technology. These “half” A100 GPUs are the smallest GPUs available on the cluster.

Table G.1 lists the wall time and cost per simulated time unit for the HIT and the channel flow simulations. The reported wall times are based on the average of three simulations, where we timed 100 time steps of the solvers, excluding initialization and compilation. “TO track” denotes TO simulations used for generating training data by tracking reference QoIs, while “TO LRS” refers to runs with the trained linear regression and noise model. On the HIT test case, the LF simulation is 320 times faster than the HF simulation. The subgrid scale models make the LF solver up to two times slower. The Smagorinsky model adds the smallest overhead. For channel flow, the LF solver achieves about 190× speedup, again reduced by about a factor of two when an SGS model is applied. The Taylor-Green vortex case uses the same grid as HIT and differs only in forcing; we therefore omit separate timing results.

The computational costs in System Billing Units (SBUs) are reported to take into account that running on a smaller GPU for the same time is cheaper. On *Snellius*, 1 h on a H100 GPU costs 192 SBUs and 1 h on an “half” A100 GPU costs 64 SBUs.

As an illustration, consider the HIT test case where accurate long-term QoI distributions (over 100 time units) were obtained using an ensemble of five TO LRS runs trained on 10 time units of HF data. The required steps are:

- One 10 time unit HF simulation: 479 SBUs,
- TO tracking of QoIs in LF simulation over 10 time units: 0.9 SBUs,
- Fitting the linear regression model: negligible (a few seconds),
- Five TO LRS LF simulations (100 time units each): 50 SBUs.

The total cost is approximately 530 SBUs, compared to 4790 SBUs for a 100 time unit HF simulation. This demonstrates nearly an order-of-magnitude cost reduction while retaining statistical accuracy.

Table G.1
Time spent / cost per simulated time unit.

Simulation	Time [s]	Cost [SBU]
HIT HF	897	47.9
HIT LF:		
No model	2.8	0.05
Smag	4.6	0.08
TO track	5.0	0.09
TO-LRS	5.8	0.10
Channel HF	440	23.5
Channel LF:		
No model	2.3	0.04
Smag	5.2	0.09
WALE	5.2	0.09
TO track	3.7	0.07
TO LRS	3.9	0.07

References

[1] M. Lesieur, O. Métais, P. Comte, Large-Eddy Simulations of Turbulence, Cambridge university press, 2005. https://doi.org/10.1017/CBO9780511755507
[2] J. Smagorinsky, General circulation experiments with the primitive equations: I. the basic experiment, Mon. Weather Rev. 91 (3) (1963) 99–164.
[3] M. Germano, U. Piomelli, P. Moin, W.H. Cabot, A dynamic subgrid-scale eddy viscosity model, Phys. Fluids A Fluid Dyn. 3 (7) (1991) 1760–1765. https://doi.org/10.1063/1.857955
[4] S.S. Girimaji, Turbulence closure modeling with machine learning: a foundational physics perspective, New J. Phys. 26 (7) (2024) 071201. https://doi.org/10.1088/1367-2630/ad6689
[5] J.M. Gimenez, F.M. Sívori, A.E. Larreteguy, S.I. Montaño, H.J. Aguerre, N.M. Nigro, S.R. Idelsohn, A multiscale pseudo-DNS approach for solving turbulent boundary-layer problems, Comput. Methods Appl. Mech. Eng. 437 (2025) 117804. https://doi.org/10.1016/j.cma.2025.117804
[6] Z. Li, W. Peng, Z. Yuan, J. Wang, Fourier neural operator approach to large eddy simulation of three-dimensional turbulence, Theor. Appl. Mech. Lett. 12 (6) (2022) 100389. https://doi.org/10.1016/j.taml.2022.100389
[7] N. Saura, T. Gomez, Subgrid stress tensor prediction in homogeneous isotropic turbulence using 3D-convolutional neural networks, Prepr. SSRN 4184202 (2022). https://doi.org/10.2139/ssrn.4184202
[8] M. Kurz, A. Beck, B. Sanderse, Harnessing equivariance: modeling turbulence with graph neural networks, (2025). arXiv preprint arXiv:2504.07741
[9] B. Liu, H. Yu, H. Huang, N. Liu, X. Lu, Investigation of nonlocal data-driven methods for subgrid-scale stress modeling in large eddy simulation, AIP Adv. 12 (6) (2022). https://doi.org/10.1063/5.0094316
[10] J. Park, H. Choi, Toward neural-network-based large eddy simulation: application to turbulent channel flow, J. Fluid Mech. 914 (2021) A16. https://doi.org/10.1017/jfm.2020.931
[11] B. Sanderse, P. Stinis, R. Maulik, S.E. Ahmed, Scientific machine learning for closure models in multiscale problems: a review, Found. Data Sci. 7 (1) (2025) 298–337. https://doi.org/10.3934/fods.2024043
[12] Y. Guan, A. Subel, A. Chattopadhyay, P. Hassanzadeh, Learning physics-constrained subgrid-scale closures in the small-data regime for stable and accurate LES, Physica D 443 (2023) 133568. https://doi.org/10.1016/j.physd.2022.133568
[13] T. Van Gastelen, W.N. Edeling, B. Sanderse, Energy-conserving neural network for turbulence closure modeling, J. Comput. Phys. 508 (2024) 113003. https://doi.org/10.1016/j.jcp.2024.113003
[14] M. Kurz, P. Offenhäuser, A. Beck, Deep reinforcement learning for turbulence modeling in large eddy simulations, Int. J. Heat Fluid Flow 99 (2023) 109094. https://doi.org/10.1016/j.ijheatfluidflow.2022.109094
[15] S.D. Agdestein, B. Sanderse, Discretize first, filter next: learning divergence-consistent closure models for large-eddy simulation, J. Comput. Phys. 522 (2025) 113577. https://doi.org/10.1016/j.jcp.2024.113577
[16] B. List, L. Chen, N. Thuerey, Learned turbulence modelling with differentiable fluid solvers: physics-based loss functions and optimisation horizons, J. Fluid Mech. 949 (2022) A25. https://doi.org/10.1017/jfm.2022.738
[17] B. List, L. Chen, K. Bali, N. Thuerey, Differentiability in unrolled training of neural physics simulators on transient dynamics, Comput. Methods Appl. Mech. Eng. 433 (2025) 117441. https://doi.org/10.1016/j.cma.2024.117441
[18] S. Rasp, Coupled online learning as a way to tackle instabilities and biases in neural network parameterizations: general algorithms and Lorenz 96 case study (v1. 0), Geosci. Model Dev. 13 (5) (2020) 2185–2196. https://doi.org/10.5194/gmd-13-2185-2020
[19] Y. Ling, A. Lozano-Duran, Numerically consistent data-driven subgrid-scale model via data assimilation and machine learning, in: AIAA SCITECH 2025 Forum, 2025, p. 1280. https://doi.org/10.2514/6.2025-1280
[20] J.A. Langford, R.D. Moser, Optimal LES formulations for isotropic turbulence, J. Fluid Mech. 398 (1999) 321–346. https://doi.org/10.1017/S0022112099006369
[21] A.P. Guillaumin, L. Zanna, Stochastic-deep learning parameterization of ocean momentum forcing, J. Adv. Model. Earth Syst. 13 (9) (2021) e2021MS002534. https://doi.org/10.1029/2021MS002534
[22] P. Perezhogin, L. Zanna, C. Fernandez-Granda, Generative data-driven approaches for stochastic subgrid parameterizations in an idealized ocean model, J. Adv. Model. Earth Syst. 15 (10) (2023) e2023MS003681. https://doi.org/10.1029/2023MS003681
[23] S.R. Ephrati, Probabilistic data-driven turbulence closure modeling by assimilating statistics, J. Comput. Phys. (2025) 114234. https://doi.org/10.1016/j.jcp.2025.114234
[24] W.N. Edeling, D.T. Crommelin, Reducing data-driven dynamical subgrid scale models by physical constraints, Comput. Fluids 201 (2020) 104470. https://doi.org/10.1016/j.compfluid.2020.104470
[25] R. Hoekstra, D.T. Crommelin, W.N. Edeling, Reduced data-driven turbulence closure for capturing long-term statistics, Comput. Fluids 285 (2024) 106469. https://doi.org/10.1016/j.compfluid.2024.106469
[26] S.R. Ephrati, P. Cifani, M. Viviani, B.J. Geurts, Data-driven stochastic spectral modeling for coarsening of the two-dimensional Euler equations on the sphere, Phys. Fluids 35 (9) (2023). https://doi.org/10.1063/5.0156942
[27] S.D. Agdestein, S. Ciarella, B. Sanderse, R. Hoekstra, IncompressibleNavierStokes.jl (v3.0.0) [Software], Zenodo, 2025. https://doi.org/10.5281/zenodo.15294602
[28] G.A. Gottwald, D.T. Crommelin, C.L.E. Franzke, Stochastic climate theory, in: C.L.E. Franzke, T.J. O'Kane (Eds.), Nonlinear and Stochastic Climate Dynamics, Cambridge University Press, 2017, p. 209-240. https://doi.org/10.1017/9781316339251.009
[29] J.V. Shebalin, S.L. Woodruff, Kolmogorov flow in three dimensions, Phys. Fluids 9 (1) (1997) 164–170. https://doi.org/10.1063/1.869159
[30] V. Eswaran, S.B. Pope, An examination of forcing in direct numerical simulations of turbulence, Comput. Fluids 16 (3) (1988) 257–278. https://doi.org/10.1016/0045-7930(88)90013-8

[31] A. Chouippe, M. Uhlmann, Forcing homogeneous turbulence in direct numerical simulation of particulate flow with interface resolution and gravity, Phys. Fluids 27 (12) (2015). https://doi.org/10.1063/1.4936274

[32] A.W. Vreman, J.G.M. Kuerten, Comparison of direct numerical simulation databases of turbulent channel flow at $Re_\tau = 180$, Phys. Fluids 26 (1) (2014). https://doi.org/10.1063/1.4861064

[33] R.D. Moser, J. Kim, N.N. Mansour, Direct numerical simulation of turbulent channel flow up to $Re_\tau = 590$, Phys. fluids 11 (4) (1999) 943–945. https://doi.org/10.1063/1.869966

[34] P. Moin, J. Kim, On the numerical solution of time-dependent viscous incompressible fluid flows involving solid boundaries, J. Comput. Phys. 35 (3) (1980) 381–392. https://doi.org/10.1016/0021-9991(80)90076-5

[35] F. Nicoud, F. Ducros, Subgrid-scale stress modelling based on the square of the velocity gradient tensor, Flow Turbul. Combust. 62 (3) (1999) 183–200. https://doi.org/10.1023/A:1009995426001

[36] M. Reissmann, J. Hasslberger, R.D. Sandberg, M. Klein, Application of gene expression programming to a-posteriori LES modeling of a Taylor Green vortex, J. Comput. Phys. 424 (2021) 109859. https://doi.org/10.1016/j.jcp.2020.109859

[37] I. Yilmaz, L. Davidson, Comparison of SGS models in large-eddy simulation for transition to turbulence in Taylor–Green flow, in: The 16th International Conference on Fluid Flow Technologies CMFF, 2015.

[38] R. Hoekstra, S.D. Agdestein, S. Ciarella, B. Sanderse, TO_IncompressibleNavierStokes.jl (v0.1) [Software], Zenodo, 2025. https://doi.org/10.5281/zenodo.15861632