Data-driven RANS closures using a relative importance term analysis based classifier for 2D and 3D separated flows

Buchanan, Tyler; Lăcătuș, Monica; West, Alastair; Dwight, Richard P.

DOI
10.1016/j.compfluid.2025.106899

Publication date
2026

Document Version
Final published version

Published in
Computers and Fluids

Citation (APA)
Buchanan, T., Lăcătuș, M., West, A., & Dwight, R. P. (2026). Data-driven RANS closures using a relative importance term analysis based classifier for 2D and 3D separated flows. *Computers and Fluids*, *305*, Article 106899. https://doi.org/10.1016/j.compfluid.2025.106899

Computers and Fluids 305 (2026) 106899

Contents lists available at ScienceDirect

Computers and Fluids

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

Data-driven RANS closures using a relative importance term analysis based classifier for 2D and 3D separated flows

Tyler Buchanan [a,b,*], Monica Lăcătuş [c], Alastair West [b], Richard P. Dwight [a]

[a] Aerodynamics Group, Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, Delft, 2600GB, Zuid-Holland, The Netherlands
[b] Williams Racing Formula One Team, Grove, OX12 0DQ, Oxfordshire, United Kingdom
[c] Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, Delft, 2628 CD, Zuid-Holland, The Netherlands

ARTICLE INFO

Keywords:
Data-driven turbulence modelling
Machine learning
RANS
Zonally-augmented $k-\omega$SST

ABSTRACT

This study presents a novel approach for enhancing Reynolds-averaged Navier-Stokes (RANS) turbulence modeling through the application of a Relative Importance Term Analysis (RITA) methodology to develop a new zonally-augmented $k-\omega$ SST model. Traditional Linear Eddy Viscosity Models often struggle with separated flows. Our approach introduces a physics-based binary classifier that systematically identifies separated shear layers requiring correction by analyzing the relative magnitudes of terms in the turbulent kinetic energy equation. Using symbolic regression, we develop compact correction terms for Reynolds stress anisotropy and turbulent kinetic energy production. Trained on 2D configurations, our model demonstrates significant improvements in predicting separation dynamics while maintaining baseline performance in fully attached flows. Generalization tests on Ahmed body and Faith hill 3D configurations confirm robust transferability, establishing an effective methodology for targeted enhancement of RANS predictions in separated flows.

1. Introduction

Turbulence modeling stands at a crossroads in Computational Fluid Dynamics (CFD). Industry-standard Reynolds-averaged Navier-Stokes (RANS) simulations are widely adopted for their computational efficiency but often lack accuracy in predicting complex flow phenomena. Conversely, high-fidelity methods such as Large Eddy Simulation (LES) and Direct Numerical Simulation (DNS) offer superior accuracy in resolving turbulent scales but remain computationally prohibitive for routine engineering applications. This dichotomy between computational efficiency and predictive accuracy motivates the development of innovative turbulence modeling techniques that bridge the gap between RANS simulations and high-fidelity methods.

Despite growing computational power, RANS is expected to remain the workhorse of engineering CFD for decades [1], primarily due to its ability to cheaply perform design optimization workflows. In contrast, LES remains computationally prohibitive for such iterative design processes. However, RANS approaches face critical limitations when predicting fundamental flow phenomena such as separation and reattachment, particularly in the presence of strong adverse pressure gradients. These limitations also extend to favorable pressure gradient scenarios, where traditional RANS models exhibit systematic deficiencies in predicting boundary layer development, heat transfer rates, and skin friction under strong acceleration conditions [2,3]. Linear Eddy Viscosity Models (LEVMs), which assume a linear relationship between the Reynolds stress tensor and mean strain rate tensor, do not account for turbulence anisotropy [4,5], resulting in insufficient turbulent mixing in shear layers [6,7].

These limitations have spurred the development of data-driven turbulence modeling approaches in recent years. Early data-driven RANS investigations focused primarily on addressing model coefficient optimization and uncertainty quantification [8–11]. Subsequently, researchers tackled more fundamental structural uncertainties by developing corrective terms or new model structures from high-fidelity data [12–15].

While these data-driven methods improved accuracy, they often lacked physical interpretability. Weatheritt and Sandberg [16] addressed this limitation by developing a sparse symbolic regression framework using Gene Expression Programming (GEP). Their approach leveraged Pope's tensor basis theory to derive interpretable algebraic equations for Reynolds stress suitable for direct implementation in RANS solvers. Schmelzer et al. [17] further improved this approach by developing SpaRTA (Sparse Regression of Turbulence Anisotropy), which introduced the k-corrective-frozen RANS approach that maintained

* Corresponding author.
E-mail addresses: t.s.b.buchanan@tudelft.nl (T. Buchanan), m.i.l.lacatus@tudelft.nl (M. Lăcătuş), alastair.west@williamsf1.com (A. West), r.p.dwight@tudelft.nl (R.P. Dwight).

https://doi.org/10.1016/j.compfluid.2025.106899
Received 31 March 2025; Received in revised form 6 October 2025; Accepted 4 November 2025
Available online 13 November 2025

interpretability while enabling efficient derivation of correction terms. Recent developments include Symbolic Bayesian Learning for SpaRTA (SBL-SpaRTA), which combines symbolic regression with Bayesian inference for comprehensive uncertainty quantification [18].

Despite these methodological advances, which improve upon generalizability, interoperability, and compute time, recent studies [19–22] have highlighted the critical importance of selectively applying corrections in separated flows, particularly in shear layers. Rumsey [6] and Li et al. [7] demonstrated that errors in RANS models for separated flows predominantly occur outside the boundary layer, especially in shear layers where turbulent flow behavior exhibits strong anisotropy, non-equilibrium conditions, and coherent structure dynamics. Specifically, Rumsey's investigation of an ad hoc modification of the $k-\omega$ model utilized a multiplier that adjusted the ω-destruction term based on the ratio of turbulence production to destruction, to help correct underestimated turbulent shear stress in separation bubbles. While this approach showed promising improvements in specific cases, it lacked universal effectiveness across different separated flow scenarios.

Building on these insights, Srivastava et al. [19] developed refined FIML-based augmentation strategies for separated flows, while Wu et al. [20] introduced the Conditioned Field Inversion (FI-CND) method to maintain boundary layer accuracy while enhancing separated flow predictions. Nishi et al. [21] proposed a closed-form model correction using Gaussian radial basis functions for incorporating local flow features into RANS models. The implementation of such corrections presents inherent challenges, particularly with the $k-\omega$ SST model. As emphasized by Wu et al. [20] and Nishi et al. [21], boundary and shear layer corrections must be implemented cautiously to preserve the model's accuracy in the log-law region, while still improving predictions in separated flow regions.

These implementation challenges highlight a broader issue in the field: the significant difficulty in developing corrections using full-field data that generalize effectively across different flow configurations [22,23]. While various methods have demonstrated success in specific classes of flows, achieving broader applicability remains elusive due to the complexity of training with diverse flow conditions and the challenge of identifying consistent error patterns across different geometries.

With the aim of addressing these fundamental limitations, the present study introduces the Relative Importance Term Analysis (RITA) methodology inspired by Callaham et al. [24]. Rather than attempting to develop a universally generalizable correction model, we provide a systematic method for identifying and classifying the shear layer regions where RANS models require improvement in separated flows. RITA focuses on quantitatively analyzing the ratios of terms in the k-equation of the $k-\omega$ SST model, enabling the development of a physics-based binary classifier that isolates the shear layer using three criteria: (1) the RITA ratio of turbulence destruction to production, (2) the turbulent-to-total kinetic energy ratio, and (3) a vorticity-based Reynolds number.

Building on this classification, correction terms are derived using the SpaRTA framework introduced by Schmelzer et al. [17] and applied selectively within the identified shear layer regions, resulting in a zonally-augmented $k-\omega$ SST model. This approach preserves baseline RANS accuracy in regions where the $k-\omega$ SST model already performs well, such as attached boundary layers and the free-stream, while targeting improvements in shear layers where deficiencies arise. We validated the resulting approach on 2D separated flows-periodic hill, periodic hills of parametrized geometries, NASA wall-mounted hump (henceforth NASA-Hump), backward-facing step (henceforth BFS) and a curved backward-facing step (henceforth CBFS)-before extending to more complex 3D geometries such as the Faith hill and Ahmed body, where it also demonstrated good performance.

The remainder of this manuscript is organized as follows: Section 2 presents the formulation of the RITA methodology and its theoretical development, along with the construction of the zonally-augmented $k-\omega$ SST model using SpaRTA-derived corrections. Section 3 provides a comprehensive evaluation of the resulting model through training and generalization tests in 2D benchmark cases, along with an extension to 3D geometries. Finally, Section 4 summarizes the key findings and identifies promising directions for future investigation.

2. Methodology

This section presents our framework for enhancing RANS turbulence modeling via a new zonally-augmented $k-\omega$ SST model. We begin by quantifying model-form errors between baseline RANS solutions and high-fidelity data using the k-corrective-frozen RANS method (Section 2.1). Observing that the largest corrections predominantly localize in shear layers and wake regions, we develop the Relative Importance Term Analysis (RITA) methodology (Section 2.2) to systematically identify these regions through physics-based classification criteria. Following classifier validation, we employ symbolic regression through SpaRTA (Section 2.3) to derive compact correction terms applied exclusively in classified regions. Finally, we outline our comprehensive dataset selection and numerical setup (Section 2.4 and 2.5), featuring both 2D training cases and 3D generalization test cases to evaluate the robustness of our approach. Throughout this section, we emphasize on how each component contributes to our goal of improving predictions in separated flows while preserving baseline model performance in well-predicted regions.

2.1. Extracting model-form errors of RANS

This study starts from the SpaRTA framework developed by Schmelzer et al. [17]. The SpaRTA framework operates by (*i*) injecting high-fidelity data into RANS closure equations to obtain residual correction fields using the k-corrective-frozen RANS method and (*ii*) approximating these fields as a function of local flow quantities. The high-fidelity data used for this study consists of DNS and LES results, providing accurate representations of turbulent flows and capturing complex phenomena that RANS models often struggle with, particularly in separated flow regions.

Following Schmelzer et al. [17], we apply SpaRTA to our baseline $k-\omega$ SST model. This enables the identification and correction of model-form errors in both the Reynolds stress tensor and the RANS closure equations. Within the RANS closure equations, the Reynolds stress tensor τ_{ij}^B based on Boussinesq approximation is approximated as:

$$\tau_{ij} \simeq \tau_{ij}^B := 2k\left(\frac{1}{3}\delta_{ij} + b_{ij}^B\right), \tag{1}$$

where k is the turbulent kinetic energy (TKE), δ_{ij} is the Kronecker delta, and b_{ij}^B denotes the anisotropy tensor in the baseline $k-\omega$ SST model. This formulation thus divides the Reynolds stress tensor into an isotropic part, $\frac{2}{3}k\delta_{ij}$, and an anisotropic part, $a_{ij} = 2kb_{ij}^B$. The anisotropic component a_{ij} captures momentum transport in different flow directions. In contrast, the isotropic component is absorbed into a modified mean pressure, contributing to the pressure-like behavior of turbulence. The Boussinesq hypothesis approximates this anisotropy tensor in LEVMs as a linear function of the mean velocity gradient:

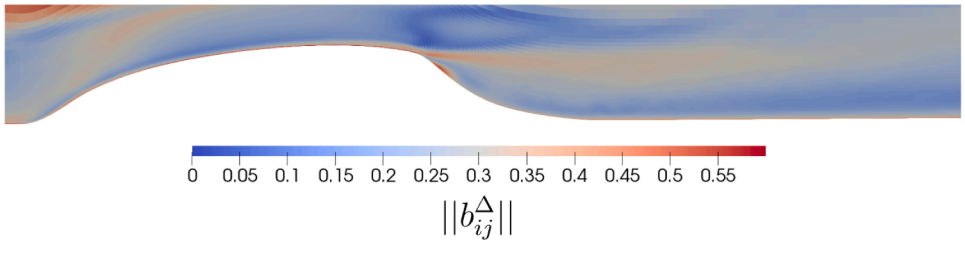
$$b_{ij}^B = -\frac{\nu_t}{k}S_{ij}, \tag{2}$$

where ν_t is the eddy viscosity, and $S_{ij} := \frac{1}{2}(\partial_j U_i + \partial_i U_j)$ is the mean-strain tensor. To capture the model-form error between the high-fidelity data's Reynolds stress tensor τ_{ij}^* and the baseline $k-\omega$ SST model, a residual term b_{ij}^Δ is introduced into (1) as

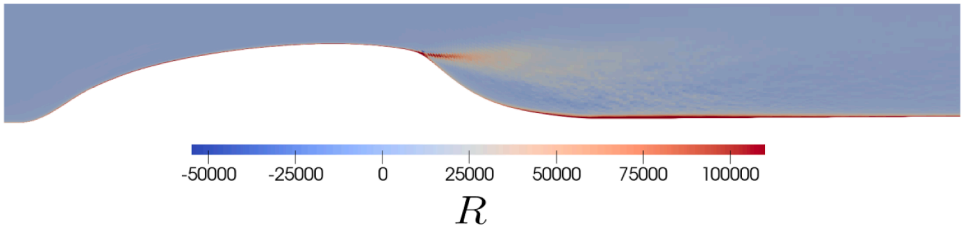
$$\tau_{ij}^* = 2k\left(\frac{1}{3}\delta_{ij} + b_{ij}^B + b_{ij}^\Delta\right), \tag{3}$$

where b_{ij}^Δ represents the additional anisotropy correction needed to account for the nonlinear viscosity relationship [25]. This correction requires specification of both ν_t and ω.

We employ the k-corrective-frozen RANS method [17] to extract model-form errors from high-fidelity DNS or LES data. In this approach,



(a) b_{ij}^{Δ} correction field.



(b) R correction field.

Fig. 1. Correction fields obtained using the k-corrective-frozen RANS approach on the NASA-Hump training case. The b_{ij}^{Δ} field (top) shows the Frobenius norm of anisotropic stress correction, while the R field (bottom) indicates additional turbulence defect in the k-budget.

the high-fidelity variables U^*, k^*, and b_{ij}^* are prescribed as inputs to the governing equations while the ω-equation is solved iteratively to determine the eddy viscosity ν_t. The eddy viscosity ν_t is computed using the frozen k^* and evolved specific dissipation rate ω, while the strain rate tensor is derived from the frozen velocity field U^*. This approach enables the systematic extraction of the correction field b_{ij}^{Δ} that quantifies anisotropy-related model-form errors in the baseline $k-\omega$ SST model. The mathematical formulation of this approach involves augmented k and ω transport equations:

$$\underbrace{\frac{\partial k^*}{\partial t}+U_i^*\frac{\partial k^*}{\partial x_i}}_{C_k-\text{Convection}}=\underbrace{P_k}_{P_k-\text{Production}}+R-\underbrace{\beta^*\omega k^*}_{D_k-\text{Destruction}}+\underbrace{\frac{\partial}{\partial x_i}\left[(\nu+\sigma_k^*\nu_t)\frac{\partial k^*}{\partial x_i}\right]}_{d_k-\text{Diffusion}}, \tag{4}$$

$$\frac{\partial \omega}{\partial t}+U_i^*\frac{\partial \omega}{\partial x_i}=\frac{\gamma}{\nu_t}(P_k+R)-\beta\omega^2+\frac{\partial}{\partial x_i}\left[(\nu+\sigma_\omega\nu_t)\frac{\partial \omega}{\partial x_i}\right]+CD_{kw}, \tag{5}$$

where β is a model coefficient, and σ_k and σ_ω are turbulent Prandtl numbers controlling the diffusive transport of k and ω respectively. The production term P_k incorporates the anisotropy correction b_{ij}^{Δ} and is defined as:

$$P_k=\min\left(-2k\left(b_{ij}^{B}+b_{ij}^{\Delta}\right)\frac{\partial U_i^*}{\partial x_j},10\beta^*\omega k^*\right). \tag{6}$$

However, modifying the baseline model's production term with b_{ij}^{Δ} alone cannot fully address the model-form error in the k-equation, as this modification does not ensure that solving the k-equation will reproduce the same k values present in the high-fidelity data [17]. Therefore, a residual correction term R is introduced in Eqs. (4) and (5) to address errors in both the production and dissipation terms of the k and ω equations. This correction can be represented either as a field $R(\mathbf{x})$ or as a function of local flow variables $R(\nabla U, k, \omega, \ldots)$.

This methodology is validated using the 2D NASA-Hump case [26], a benchmark configuration for evaluating turbulence models under flow separation and reattachment conditions with adverse pressure gradients. Fig. 1 presents the correction fields b_{ij}^{Δ} and R generated by the k-corrective-frozen-RANS method, highlighting regions where the baseline $k-\omega$ SST model deviates from high-fidelity DNS/LES data.

The magnitude and spatial distribution of the b_{ij}^{Δ} correction field (Fig. 1(a)) reveals two critical regions requiring significant adjustment to the baseline $k-\omega$ SST model's linear eddy-viscosity assumption: the shear layer emanating from the hump's apex and the near-wall recovery region downstream. The most substantial corrections occur in the separated shear layer, indicating significant anisotropy that the baseline model fails to capture. This observation is reinforced by the corresponding R correction field (Fig. 1(b)), which shows strong positive corrections concentrated in similar regions. The spatial correlation between these two correction fields suggests that the baseline model mispredicts the Reynolds stress anisotropy and, even taking that into account, still significantly underestimates TKE production in the separated region. Most notably, the corrections persist well into the recovery region downstream of the hump, highlighting the baseline model's limitations in predicting both the immediate separation and subsequent flow reattachment physics. These results demonstrate the necessity for non-linear turbulence modeling approaches that can better capture both the turbulent anisotropy and corrections to the k-budget in separated flows.

Despite requiring corrections in separated flows, the $k-\omega$ SST model has robust performance in predicting wall-bounded turbulence. However, implementing corrections within the boundary layer requires caution to maintain the model's accuracy in predicting the log-law of the wall [27]. Wu et al. [20] emphasize in their work on FI-CND that corrections, within the boundary layer should be minimized or avoided to maintain the log-law's accuracy. These recent studies, including the k-corrective-frozen RANS method, collectively highlight the importance of selectively applying corrections in separated flows, particularly in shear layers. However, accurately identifying these critical regions while preserving the model's performance in well-predicted areas remains challenging. To address this, we introduce RITA, aligning with the strategies proposed by Srivastava et al. [19] and Wu et al. [20], as highlighted next.

2.2. Relative importance term analysis (RITA)

RITA is a physics-based classification framework that isolates specific flow features-in this case, shear layers-by analyzing the relative importance of terms in the k-equation of the $k-\omega$ SST turbulence model.[1]

The method builds on the approach of Callaham et al. [28], who identified dominant flow physics by examining the balance of terms in the

[1] The underlying idea of RITA is not limited to this model. It can be extended naturally to other turbulence models such as $k-\epsilon$, $k-\omega$, or Spalart-Allmaras.

RANS momentum equations across different regimes, such as the transition of a laminar boundary layer to turbulence. RITA instead targets the k-equation, which governs the evolution of TKE. Focusing on this equation makes it possible to directly assess how turbulence is generated, transported, and dissipated. Since these processes vary systematically across flow regions, the analysis provides a natural way to distinguish between them.

Concretely, RITA evaluates the relative ratio of the terms in the k-equation: convection (C_k), production (P_k), destruction (D_k), and diffusion (d_k). The central idea is that while the absolute magnitudes of these terms vary strongly with flow conditions, the ratios between them remain characteristic of particular flow regions. By exploiting this property, RITA systematically uncovers coherent flow structures across different flow configurations. The ratios we found most effective for characterizing different flow regions are:

$$\phi_{D_k/P_k} = \frac{|D_k|}{|P_k| + |D_k|}, \tag{7}$$

$$\phi_{D_k/C_k} = \frac{|D_k|}{|C_k| + |D_k|}, \tag{8}$$

$$\phi_{D_k/d_k} = \frac{|D_k|}{|d_k| + |D_k|}. \tag{9}$$

Among these, the destruction-to-production ratio ϕ_{D_k/P_k}, proved especially effective. In shear layers, ϕ_{D_k/P_k} consistently falls below 0.55, reflecting the dominance of turbulence production driven by the large velocity gradients and Reynolds shear stress characteristic of these regions. In attached boundary layers, on the other hand, destruction dominates, driving ϕ_{D_k/P_k} above 0.55. Finally, in free-stream regions, where turbulence production is negligible, the ratio approaches unity.

As detailed in Lăcătuş [29], the ϕ_{D_k/P_k} ratio alone is insufficient for robust shear layer identification. In our analysis of 2D separated flows (the periodic hill, NASA-Hump, and CBFS), we found that additional criteria related to the content of TKE and rotational effects were needed to reliably distinguish shear layers from the outer regions of attached boundary layers. This motivated the introduction of two additional ratios: ϕ_k, the turbulent-to-total kinetic energy ratio (TTKER), and Re_Ω, the vorticity Reynolds number. They are defined as:

$$\phi_k = \frac{k}{k + 0.5|U|^2}, \quad Re_\Omega = \frac{d_w^2|\Omega|}{\nu}, \tag{10}$$

where d_w is the wall distance, $|U|$ is the mean velocity magnitude, and $|\Omega|$ is the vorticity magnitude.

TTKER measures the relative contribution of turbulence fluctuations to the overall kinetic energy. This provides a clear way to distinguish the shear layer from the outer portion of an attached boundary layer, since both regions can exhibit production rates exceeding destruction. In the boundary layer, the mean kinetic energy remains much larger than the TKE, so ϕ_k stays relatively small even if $\phi_{D_k/P_k} \leq 0.55$. In contrast, in the shear layer, strong turbulence fluctuations develop while the mean velocity is reduced, so the turbulent fraction of the total energy becomes large. This leads to higher values of ϕ_k in shear layers, allowing them to be clearly separated from the outer boundary layer in a way that the ϕ_{D_k/P_k} ratio alone cannot. We note here that since ϕ_k is computed from the mean velocity magnitude, it is not Galilean invariant and may change in simulations performed in different reference frames. However, this does not limit its usefulness for distinguishing between flow regions, provided that all comparisons are made within the same reference frame.

The vorticity Reynolds number identifies regions of strong rotational flow and was introduced primarily to improve the stability of the corrective model in 3D flows by preventing corrections from being applied too close to the wall. Its formulation includes the wall distance d_w, which shifts the shear layer classification away from the wall by a few computational cells.

These three ratios, ϕ_{D_k/P_k} , ϕ_k and Re_Ω define a binary classifier:

$$\sigma_{SL} = \begin{cases} 1, & \phi_{D_k/P_k} \leq 0.55 \wedge \phi_k \geq 0.12 \wedge \phi_{Re_\Omega} \geq 0.02 \\ 0, & \text{otherwise.} \end{cases}$$

where ϕ_{Re_Ω} is the min-max normalization of Re_Ω,

$$\phi_{Re_\Omega} = \frac{Re_\Omega - Re_{\Omega,\min}}{Re_{\Omega,\max} - Re_{\Omega,\min}}. \tag{11}$$

Because this normalization is performed using the case-wide extrema of Re_Ω, ϕ_{Re_Ω} is a global quantity within each case. The advantage of this scaling is that it consistently maps ϕ_{Re_Ω} to the interval $[0, 1]$, enabling a single threshold to be applied for the classifier across different cases.

The classifier thresholds were first determined through physical reasoning and then adjusted empirically. Each threshold was motivated by the expected balance of terms in different flow regions: for instance, the cutoff $\phi_{D_k/P_k} \leq 0.55$ reflects the dominance of production in shear layers, higher ϕ_k values indicate the buildup of turbulence intensity compared to attached boundary layers, and higher ϕ_{Re_Ω} values mark regions of strong rotational flow. Starting from these physics-based expectations, the precise values were tuned across our three 2D separated flows (periodic hill, NASA-Hump, and CBFS). The tuning was performed iteratively to ensure that only shear layer regions were identified, while attached boundary layers and free-stream regions were excluded.

Together, these criteria form a physics-based classifier that identifies shear layers across cases with varying geometries and Reynolds numbers. Cells meeting these criteria are classified as part of the shear layer ($\sigma_{SL} = 1$), receiving full b_{ij}^Δ and R corrections, while cells outside the shear layer ($\sigma_{SL} = 0$) receive no corrections. Notably, this classification approach ensures that corrections are not applied to fundamental flow cases such as flat plate boundary layer and channel flows, where the baseline $k-\omega$ SST model already provides accurate predictions (see Appendix F). This selective correction strategy maintains the established performance of the baseline model in these canonical cases. Fig. 2 demonstrates the accuracy of this classification approach across our training cases, clearly identifying the shear layer regions in the periodic hill (Fig. 2(a)), NASA-Hump (Fig. 2(b)), and CBFS (Fig. 2(c)) configurations.

To evaluate the effect of the classifier in practice, we compare two approaches. In the full-field Propagation approach, the correction terms from the k-corrective-frozen RANS method are applied throughout the entire domain. In contrast, the shear layer Propagation (SL-Propagation) approach restricts the corrections to cells classified as shear layer regions ($\sigma_{SL} = 1$). Cells outside the shear layer ($\sigma_{SL} = 0$) remain governed by the baseline $k-\omega$ SST model. The resulting modified model equations are shown below, with σ_{SL} applied to the R and b_{ij}^Δ terms:

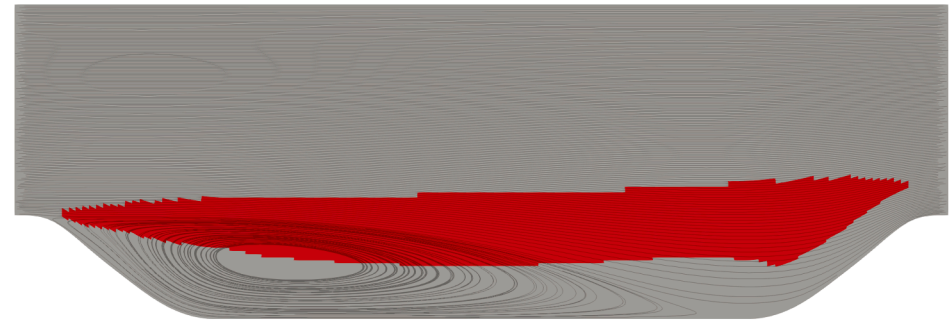
$$\frac{\partial k}{\partial t} + U_j \frac{\partial k}{\partial x_j} = P_k + \sigma_{SL} R - \beta^* w k + \frac{\partial}{\partial x_j}\left[(\nu + \sigma_k \nu_t)\frac{\partial k}{\partial x_j}\right], \tag{12}$$

$$\frac{\partial \omega}{\partial t} + U_j \frac{\partial \omega}{\partial x_i} = \frac{\gamma}{\nu_t}(P_k + \sigma_{SL} R) - \beta \omega^2 + \frac{\partial}{\partial x_i}\left[(\nu + \sigma_\omega \nu_t)\frac{\partial \omega}{\partial x_i}\right] + C D_{k\omega}, \tag{13}$$

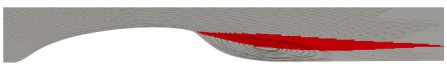
$$P_k = \min\left(-2k(b_{ij}^B + \sigma_{SL} b_{ij}^\Delta)\frac{\partial U_i}{\partial x_j}, 10 P_\omega k\right). \tag{14}$$

An important implementation detail is that the σ_{SL} classifier is updated dynamically during SL-Propagation simulations. At every timestep, the classifier is recomputed to account for changes in turbulence production and intensity caused by the corrective terms R and b_{ij}^Δ. The R term can locally increase or decrease the production of TKE: when $R > 0$, its contribution is added to the P_k term in the ϕ_{D_k/P_k} ratio, and when $R < 0$, it is added to the D_k term.

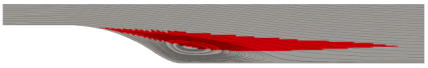
As a note, the binary classifier σ_{SL} is never differentiated in our formulation, neither in the transport equations nor in the momentum



(a) Periodic hill case.

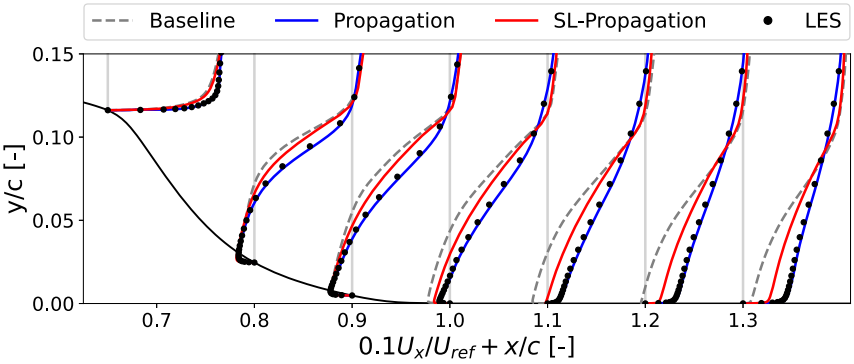


(b) NASA-Hump case.

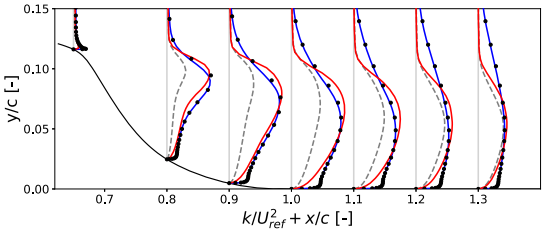


(c) CBFS case.

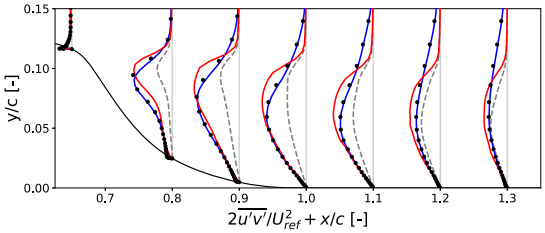
Fig. 2. Classification outcome of the σ_{SL} classifier on the 2D separated flow cases (red region: $\sigma_{SL} = 1$, gray region: $\sigma_{SL} = 0$).



(a) Axial velocity profiles.



(b) TKE profiles.



(c) Reynolds shear stress profiles.

Fig. 3. Performance comparison between full-field Propagation and SL-Propagation on the NASA-hump case.

equation. Consequently, its discrete nature (0 or 1) does not introduce immediate numerical instabilities at the interfaces between corrected and uncorrected regions. We compared SL-Propagation with full-field Propagation for the NASA-Hump case, where the former applies corrections exclusively to regions identified by $\sigma_{SL} = 1$. Fig. 3 presents axial velocity, TKE, and Reynolds shear stress profiles at multiple streamwise locations. While SL-Propagation shows reduced accuracy compared to full-field Propagation, primarily due to the absence of corrections in the region above the hill, it still substantially improves upon the baseline model in capturing the shear layer development and reattachment characteristics.

Further validation of the method's performance is provided by the skin friction coefficient comparison shown in Fig. 4. Although SL-Propagation shows some deviation in the recovery region ($0.7 \leq x/c \leq 1.0$), it captures the key separation and reattachment features well, and notably improves prediction in the wake region ($x/c > 1.0$) compared to the baseline model. This trade-off in accuracy is offset by RITA's ability to preserve the baseline model's behavior in well-predicted regions, such as the upstream boundary layer and freestream. The targeted strategy improves computational efficiency while maintaining the model's established capabilities where corrections are unnecessary.

Having established a zonal classifier for identifying regions requiring correction, the next challenge is developing generalizable models for these corrections. This motivates our application of symbolic regression to discover compact, physics-based expressions for the b_{ij}^{Δ} and R correction terms within $\sigma_{SL} = 1$ identified regions.

2.3. Model discovery

The symbolic regression method SpaRTA [17] is employed to formulate algebraic models for the b_{ij}^{Δ} and R correction fields within the identified $\sigma_{SL} = 1$ regions. Following Pope [30], the anisotropic Reynolds stress correction adopts a nonlinear generalization of the eddy viscosity concept, where b_{ij} depends only on the strain rate tensor $S_{ij} =$

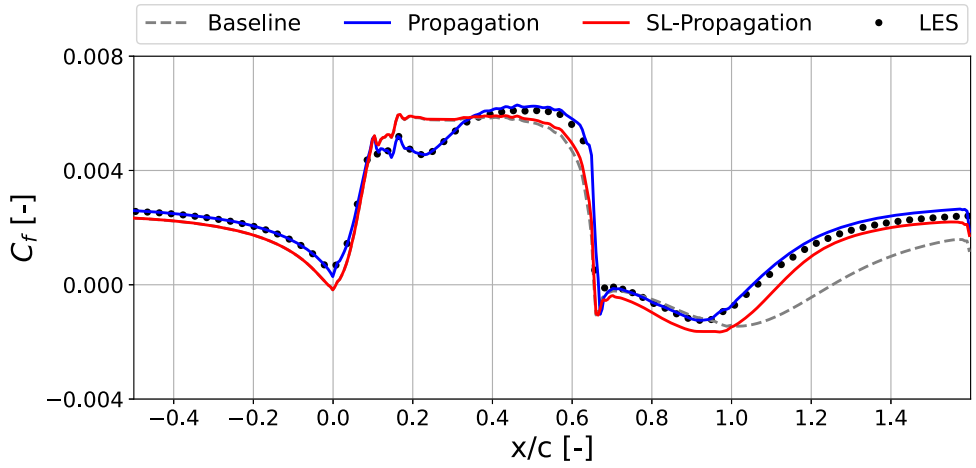


Fig. 4. Skin friction comparison plot between full-field Propagation and SL-Propagation on the NASA-hump case.

Table 1
Additional scalar features (q_m) and RITA terms added to b_{ij}^{Δ} and R feature libraries.

Description	Symbol	Equation
Time Scale Ratio	ϕ_{TS}	$\frac{k\|S\|}{\epsilon+k\|S\|}$
Q-criterion	$Q_{\text{criterion}}$	$\frac{\|\Omega\|^2-\|S\|^2}{\|\Omega\|^2+\|S\|^2}$
Ratio of Total to Normal Reynolds Stresses	$\tau_{ij,ratio}$	$\frac{\|\tau_{ij}\|}{10k+\|\tau_{ij}\|}$
Shear Parameter	S_k	$\frac{k\|\nabla U\|}{\|}$
RITA: D_k/P_k Ratio	ϕ_{D_k/P_k}	$\frac{\|D_k\|}{(\|D_k\|+\|P_k\|)}$
RITA: D_k/C_k Ratio	ϕ_{D_k/C_k}	$\frac{\|D_k\|}{(\|D_k\|+\|C_k\|)}$
RITA: D_k/d_k Ratio	ϕ_{D_k/d_k}	$\frac{\|D_k\|}{(\|D_k\|+\|d_k\|)}$
RITA: C_k/d_k Ratio	ϕ_{C_k/d_k}	$\frac{\|C_k\|}{(\|C_k\|+\|d_k\|)}$
RITA: P_k/d_k Ratio	ϕ_{P_k/d_k}	$\frac{\|P_k\|}{(\|P_k\|+\|d_k\|)}$
RITA: P_k/C_k Ratio	ϕ_{P_k/C_k}	$\frac{\|P_k\|}{(\|P_k\|+\|C_k\|)}$

$\tau\frac{1}{2}(\partial_j U_i + \partial_i U_j)$ and rotational rate tensor $\Omega_{ij} = \tau\frac{1}{2}(\partial_j U_i - \partial_i U_j)$ with timescale $\tau = 1/\omega$. Using the Cayley-Hamilton theorem, the anisotropic part of the Reynolds stress takes the general form:

$$b_{ij}(S_{ij}, \Omega_{ij}) = \sum_{n=1}^{10} \alpha_n(I_1, I_2, \dots, I_m) T_{ij}^{(n)}, \tag{15}$$

where α_n is a set of coefficients for a given value of n invariants. The terms $T_{ij}^{(n)}$ and I_m correspond to the ten non-linear basis tensors and five invariants, as defined in detail in Appendix A. This formulation, with its basis tensors (A.10) and invariants (A.11), provides the foundation for expressing the correction terms b_{ij}^{Δ} and R. The library extends beyond Pope's original basis tensors and invariants by incorporating additional scalar invariants (q_m) introduced by Wang et al. [31] and RITA terms as defined in Table 1.

For the scalar correction R, the tensor basis functions are converted to scalar features through double dot products with the mean velocity gradient tensor. In addition, the turbulence dissipation rate (ϵ) is added as a basis function in the R correction. The dissipation rate ϵ is computed differently depending on the methodology stage: during k-corrective-frozen-RANS, $\epsilon = \beta^*\omega k^*$ where k^* is frozen high-fidelity data; during SpaRTA regression, $\epsilon = \omega k$ from the propagated solutions without the β^* factor to allow SpaRTA to learn the optimal coefficient; and in final implementation, $\epsilon = \beta^*\omega k$ using evolved transport variables. The final form for b_{ij}^{Δ} and R is expressed

$$b_{ij}^{\Delta}(S_{ij}, \Omega_{ij}, q_m) = \sum_{n=1}^{10} \beta_n(I_1, \dots, I_5, q_m) T_{ij}^{(n)}, \tag{16}$$

$$R(S_{ij}, \Omega_{ij}, q_m) = 2k \sum_{n=1}^{10} \alpha_n(I_1, \dots, I_5, q_m) T_{ij}^{(n)} \partial_j U_i + a_n(I_n, q_m)\epsilon, \tag{17}$$

where β_n and α_n are the coefficient functions for the respective corrections, $\beta_n : \mathbb{R}^5 \to \mathbb{R}$ and $\alpha_n : \mathbb{R}^5 \to \mathbb{R}$. Given the distinct nature of our two correction terms, separate candidate libraries are constructed. Following Schmelzer et al. [17], the libraries forming the columns of $C_{b_{ij}^{\Delta}}$ and C_R are generated using the FFX algorithm [32]. In this approach, the invariants and basis functions are combined using unary functions to generate the candidate functions that are regressed to the target data. The library degree for both corrections is set to 1, resulting in monovariate candidate functions (single invariant transformed by a unary function and multiplied by a basis function). This design choice achieves an effective balance for robust generalization, prioritizing numerical stability and model transferability over maximum functional expressiveness. While bivariate and trivariate functions offer greater expressiveness during training, our preliminary investigations revealed that such complex functional forms tend to exhibit numerical instabilities when applied to flow configurations significantly more complex than the training cases, particularly when generalizing from 2D training geometries to complex 3D test cases. The monovariate approach maintains sufficient functional diversity while ensuring the numerical robustness necessary for practical RANS applications across varied flow conditions. The resulting libraries contain 2510 candidate functions for b_{ij}^{Δ} and 399 candidate functions for R, utilizing the unary functions (f_i) detailed in B. All features are normalized by their standard-deviations to ensure numerical stability during regression. The regression problem is subsequently formulated in matrix form:

$$b_{ij}^{\Delta} = C_{b_{ij}^{\Delta}} \Theta_{b_{ij}^{\Delta}} = \left[f_1(I_1)T_{ij}^{(1)}, f_2(I_1)T_{ij}^{(1)}, \dots, f_6(q_m)T_{ij}^{(10)} \right] \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{2510} \end{bmatrix}, \tag{18}$$

$$R = C_R \Theta_R = \left[f_1(I_1)T_{ij}^{(1)}\partial_j U_i, \dots, f_6(q_m)T_{ij}^{(10)}\partial_j U_i, \dots, f_6(q_m)\epsilon \right] \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_{399} \end{bmatrix}, \tag{19}$$

where $C_{b_{ij}^{\Delta}}$ or C_R are the feature matrices containing the candidate functions from each library and $\Theta_{b_{ij}^{\Delta}}$ and Θ_R are the coefficient matrices for each correction term. These coefficients are determined through elastic net optimization:

$$\Theta_{\Delta} = \underset{\hat{\Theta}_{\Delta}}{argmin} \left\| C_{\Delta}\hat{\Theta}_{\Delta} - \Delta^* \right\|_2^2 + \lambda\rho \left\| \hat{\Theta}_{\Delta} \right\|_1 + 0.5\lambda(1-\rho) \left\| \hat{\Theta}_{\Delta} \right\|_2^2 \tag{20}$$

Table 2
Training and test cases with their numerical configurations: dimensionality *Dim*, Reynolds number Re_H, hill width-scaling factor α, cell count N, and source of the reference data.

Case	*Dim*	Re_H	α	N	Reference data
Training Cases					
NASA-Hump	2D	9.3×10^5	–	5.1×10^4	Uzun and Malik [39]
Periodic hill	2D	1.0×10^4	1.0	1.5×10^4	Breuer et al. [40]
CBFS	2D	1.3×10^4	–	2.1×10^4	Bentaleb et al. [41]
Test Cases					
BFS	2D	3.6×10^4	–	1.10×10^5	Driver and Seegmiller [42]
Periodic hill	2D	5.6×10^3	0.5	1.5×10^4	Xiao et al. [43]
Periodic hill	2D	5.6×10^3	0.8	1.5×10^4	Xiao et al. [43]
Periodic hill	2D	5.6×10^3	1.0	1.5×10^4	Xiao et al. [43]
Periodic hill	2D	5.6×10^3	1.2	1.5×10^4	Xiao et al. [43]
Periodic hill	2D	5.6×10^3	1.5	1.5×10^4	Xiao et al. [43]
Faith hill	3D	5.0×10^5	–	1.5×10^6	Bell et al. [44]
Ahmed body	3D	7.6×10^5	–	1.1×10^7	Lienhart et al. [45]

where the subscript $\Delta = b_{ij}^{\Delta}$ or R specifies the corrective field, and the regression target is denoted as $\Delta^* = b_{ij}^{\Delta}$ or R, depending on which optimization is being performed. The parameter $\lambda > 0$ controls the regularization weight, and $0 < \rho < 1$ sets the balance between Lasso and Ridge penalties. This formulation promotes sparsity while ensuring numerical stability [33].

To ensure physical interpretability and numerical stability, the regression process employs the two-step approach of Schmelzer et al. [17]. First, candidates undergo standardization before elastic-net regression to assess their relative significance independent of magnitude. Second, Ridge regression is applied using the original, unstandardized candidate functions to maintain appropriate units and small refit coefficients in the OpenFOAM solver [17]. Through systematic evaluation of various basis function and feature combinations, we identified optimal formulations for both R and b_{ij}^{Δ} corrections.

While *a priori* regression metrics (R^2 values) are commonly used for initial model validation, they often fail to capture critical aspects of model performance in practice. This limitation stems from fundamental challenges mentioned in Mandler and Weigand [22]: the disconnect between model development and application environments [34–36], and the compounding effects of numerical uncertainties during simulations [25,37,38]. Therefore, we focus our analysis on comprehensive *a posteriori* testing, beginning with the 2D training cases themselves to establish baseline performance before examining generalization capabilities.

2.4. Dataset selection and description

The development and validation of our correction terms requires careful selection of training cases that span diverse separation mechanisms. As summarized in Table 2, the model is trained using the same three 2D separated flow configurations (the periodic hill, NASA-Hump, and CBFS) previously employed to validate the σ_{SL} classifier. Each case presents distinct flow physics - the periodic hill ($Re_H = 1.0\times10^4$) features sustained separation and reattachment cycles, the NASA-Hump ($Re_c = 9.3\times10^5$) introduces high Reynolds number effects such as thin turbulent boundary layers and smooth-surface separation under adverse pressure gradients, while the CBFS ($Re_H = 1.3\times10^4$) combines geometric and curvature-induced separation mechanisms.

To assess model generalization, we constructed a test suite that systematically increases complexity beyond the training cases. The BFS case provides a particularly challenging test of the model's ability to predict flow separation over sharp geometric discontinuities, representing a fundamentally different separation mechanism from the smooth and curved geometries used in training. The parametric periodic hill series varies the hill width-scaling factor ($\alpha = 0.5 - 1.5$), testing robustness to geometric modifications while maintaining similar flow physics. This generalization capability of SpaRTA models has already been demonstrated in a non-zonal model framework by [46].

We include two 3D flow cases: the Faith hill (Fig. 5(a)) and the Ahmed body (Fig. 5(b)). These configurations significantly differ from the 2D training cases in both geometric complexity and flow physics, providing a rigorous evaluation of the model's ability to extend learned corrections beyond 2D.

The Faith hill case ($Re_H = 5.0\times10^5$) presents a 3D separation scenario driven by an adverse pressure gradient over a smooth surface, similar in principle to the NASA-Hump training case but with the added complexity of a necklace vortex system, spiral nodes in the wake region, and significant spanwise flow variations. The higher Reynolds number produces stronger shear layers with varying skin friction intensities, complex saddle points, and 3D breakdown of coherent structures at reattachment.

The Ahmed body ($Re_H = 7.6\times10^5$) represents a bluff-body wake flow, characterized by strong counter-rotating C-pillar vortices from the slant edges, large-scale recirculation regions behind the vertical base, and unsteady shear layer development. Unlike the training cases, which predominantly feature wall-bounded separation, the Ahmed body introduces a distinct 3D wake structure where the interaction between longitudinal vortices and separated flow creates asymmetric pressure distributions and a complex dynamic wake topology.

2.5. Numerical setup

All simulations were performed using OpenFOAM with second-order accurate discretization schemes. Grid independence was ensured through systematic refinement studies, with final mesh resolutions detailed in Table 2. Detailed numerical setup parameters, grid descriptions and boundary conditions for each configuration are provided in C.

3. Results and discussion

3.1. SpaRTA-identified resulting model

The application of SpaRTA within the shear-layer regions identified by the σ_{SL} classifier, for the training cases listed in Table 2, yielded compact, physics-based correction models for both b_{ij}^{Δ} and R. The model for R takes the following form:

$$R = 0.426\left(\frac{\phi_{D_k/C_k}/0.1248}{1+(\phi_{D_k/C_k}/0.1248)^2}\right)\epsilon. \tag{21}$$

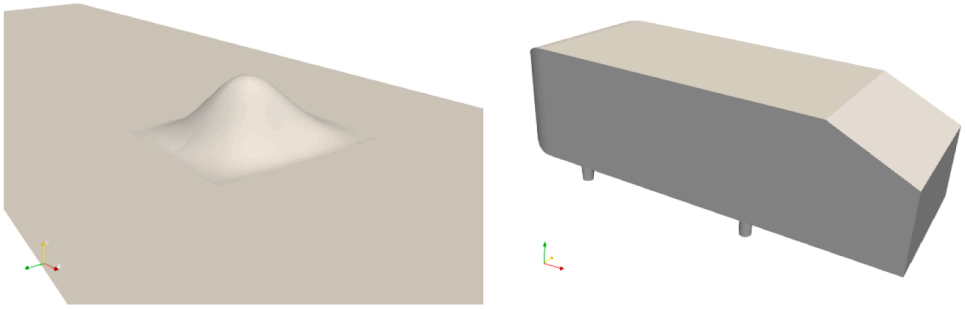
The structure of this scalar correction directly addresses the imbalance between TKE production and dissipation in separated shear flows. The correction scales with the turbulence dissipation rate ϵ and effectively reduces dissipation in the k-equation, similar to how Rumsey [6] adjusts the ω-destruction term in separation regions. It is modulated by the RITA ratio of destruction-to-convection ϕ_{D_k/C_k}, capturing the local balance between turbulent transport and dissipation. The nonlinear limiting term prevents excessive corrections in regions of extreme destruction to convection imbalance.

The anisotropic Reynolds stress correction is expressed as:

$$b_{ij}^{\Delta} = 3.69\phi_{D_k/P_k}T_{ij}^{(2)} - 5.092\left(\frac{I_2/0.01247}{1+(I_2/0.01247)^2}\right)T_{ij}^{(2)}. \tag{22}$$

This formulation combines two physically distinct mechanisms: the first term captures shear layer anisotropy through the RITA ratio of destruction-to-production ϕ_{D_k/P_k}, while the second term, scaled by the rotation invariant I_2, accounts for the effects of strong rotation on turbulent stress alignment. Both terms utilize Pope's second basis tensor $T_{ij}^{(2)}$, which represents the interaction between mean strain and rotation rate.

The complete proposed, zonally-augmented $k-\omega$ SST model (henceforth SL-Model), incorporating these corrections through the shear layer



(a) Faith hill geometry. (b) Ahmed body geometry.

Fig. 5. Geometries of 3D test cases.

classifier σ_{SL}, is presented in Box 1. This framework shows how the correction terms are integrated with the baseline model equations while maintaining the original model structure through appropriate limiters and blending functions. The standard $k-\omega$ SST model coefficients, blending functions and auxiliary functions are provided in D.

Having developed these physically-motivated corrections that leverage RITA ratios for local flow characterization, we now turn to their validation through comprehensive *a posteriori* testing. Our analysis focuses particularly on the model's ability to predict separation and reattachment dynamics in complex geometries, first examining performance on training configurations before assessing generalization capabilities on independent test cases.

3.2. Performance on training cases

The performance of the SL-Model was evaluated through systematic *a posteriori* testing on the three training cases: periodic hill, NASA-Hump and CBFS. Results were compared against two references: the baseline $k-\omega$ SST model, and the SL-Propagation model. The latter applies the R and b_{ij}^{Δ} fields obtained from the k-corrective-frozen RANS approach with dynamic σ_{SL} classification, and represents the upper limit of performance achievable with zonal corrections. In contrast, the SL-Model uses the SpaRTA discovered R and b_{ij}^{Δ} corrections together with dynamic σ_{SL} classification. Predictions from all models were validated against high-fidelity DNS/LES data to assess the accuracy of the proposed SL-Model.

Fig. 6 presents the performance of the SL-Model for the NASA-Hump training case. The axial velocity profiles (Fig. 6(a)) demonstrate that the SL-Model reduces the separation bubble length overprediction, characteristic of the baseline $k-\omega$ SST model, showing improved agreement with LES data particularly in the recovery region ($x/c \geq 1.2$), where the baseline model typically overpredicts the separation bubble extent. The consistency between SL-Model and SL-Propagation results indicates that the correction terms effectively capture the key separation flow physics.

The TKE profiles (Fig. 6(b)) reveal an overprediction compared to both the baseline model and LES data, particularly in the wake region ($x/c \geq 1.0$). While the SL-Model achieves improved mean flow predictions, this suggests that the correction mechanisms may compensate for other model deficiencies through enhanced turbulent mixing. The Reynolds shear stress distributions (Fig. 6(c)) show better alignment with LES data in terms of peak locations, though their magnitudes are influenced by these elevated turbulence levels. The SL-Model reproduces much of the performance of the SL-Propagation model, while also revealing a limitation in the current approach, where improved mean flow predictions come at the cost of TKE accuracy.

The skin friction distribution (Fig. 7) quantitatively validates separation and reattachment prediction. The SL-Model shows improved prediction of skin friction at the point of separation ($x/c \approx 0.65$) and in the far wake region ($x/c \geq 1.0$), reducing the reattachment length

Box 1. Zonally-Augmented $k-\omega$ SST Model (SL-Model)

σ_{SL} Classifier Criteria:

$$\sigma_{SL} = \begin{cases} 1, & \text{if } \phi_{D_k/P_k} \leq 0.55 \ \wedge \ \phi_k \geq 0.12 \ \wedge \ \phi_{Re_\Omega} \geq 0.02, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

where ϕ_{Re_Ω} is the min-max normalization of Re_Ω,

$$\phi_{Re_\Omega} = \frac{Re_\Omega - Re_{\Omega,\min}}{Re_{\Omega,\max} - Re_{\Omega,\min}} \tag{24}$$

such that $\phi_{Re_\Omega} \in [0,1]$ across all cases.

Correction Term Equations:

$$R = 0.426\left(\frac{\phi_{D_k/C_k}/0.1248}{1+(\phi_{D_k/C_k}/0.1248)^2}\right)\epsilon. \tag{25}$$

$$b_{ij}^{\Delta} = 3.69\phi_{D_k/P_k}T_{ij}^{(2)} - 5.092\left(\frac{I_2/0.01247}{1+(I_2/0.01247)^2}\right)T_{ij}^{(2)}. \tag{26}$$

Continuity and Momentum Equation:

$$\frac{\partial U_i}{\partial x_i} = 0 \tag{27}$$

$$\frac{\partial U_i}{\partial t} + U_j\frac{\partial U_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \frac{\partial}{\partial x_j}\left[\nu\left(\frac{\partial U_i}{\partial x_j}+\frac{\partial U_j}{\partial x_i}\right) - 2k(b_{ij}^{B})\right] + \sigma_{SL}\frac{\partial}{\partial x_j}(2kb_{ij}^{\Delta}) \tag{28}$$

Turbulent Kinetic Energy Equation:

$$\frac{\partial k}{\partial t} + U_j\frac{\partial k}{\partial x_j} = P_k + \sigma_{SL}R - \beta^* wk + \frac{\partial}{\partial x_j}\left[(\nu+\sigma_k\nu_t)\frac{\partial k}{\partial x_j}\right] \tag{29}$$

Specific Dissipation Rate Equation:

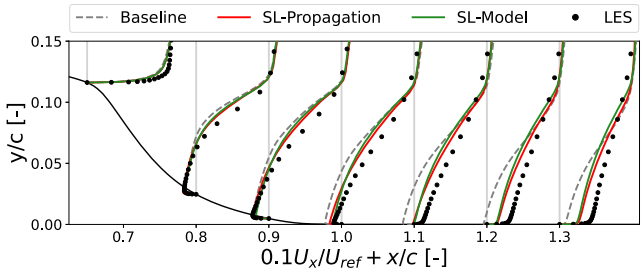
$$\frac{\partial \omega}{\partial t} + U_j\frac{\partial \omega}{\partial x_i} = \frac{\gamma}{\nu_t}(P_k+\sigma_{SL}R) - \beta\omega^2 + \frac{\partial}{\partial x_i}\left[(\nu+\sigma_\omega\nu_t)\frac{\partial \omega}{\partial x_i}\right] + CD_{k\omega} \tag{30}$$

Eddy Viscosity Definition:

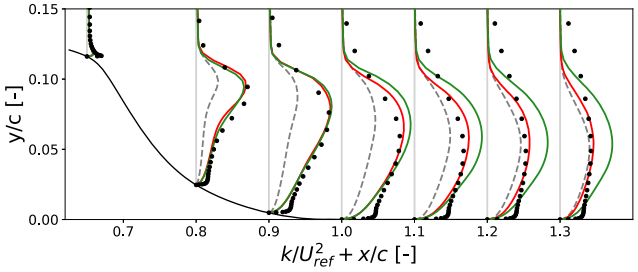
$$\nu_t = \frac{a_1 k}{\max(a_1\omega, SF_2)} \tag{31}$$

Turbulent Production Term:

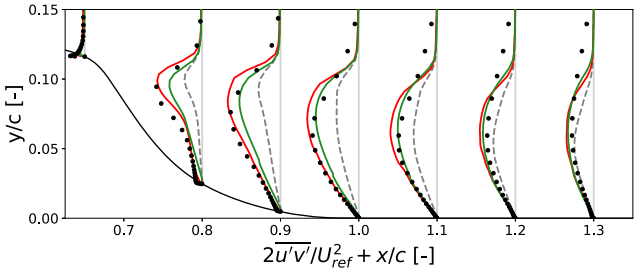
$$P_k = \min\left(-2k(b_{ij}^{B}+\sigma_{SL}b_{ij}^{\Delta})\frac{\partial U_i}{\partial x_j}, 10P_\omega k\right) \tag{32}$$



(a) Axial velocity profiles.



(b) TKE profiles.



(c) Reynolds shear stress profiles.

Fig. 6. Performance of the SL-Model on the NASA-hump training case.

overprediction characteristic of the baseline model. A slight degradation in prediction is observed in the immediate post-separation region ($0.7 \leq x/c \leq 0.9$), which can be attributed to the vorticity Reynolds number criterion used in the σ_{SL} classifier limiting corrections near the wall. Despite this local trade-off, the SL-Model successfully preserves the baseline model's accurate prediction of the attached boundary layer upstream of separation ($x/c \leq 0.5$) while achieving better overall wake recovery predictions.

Similar validation was performed for the periodic hill and CBFS cases, with detailed results presented in E. Across all three training cases-NASA-Hump (smooth-surface separation), periodic hill (periodic separation-reattachment), and CBFS (combined geometric and curvature effects)-the SL-Model demonstrates consistent performance in capturing separation locations and recovery characteristics. While improved mean flow predictions come at the cost of systematically elevated TKE levels (due to enhanced turbulent mixing from the R correction), this trade-off indicates that the discovered corrections effectively capture essential shear layer physics through a simplified algebraic formulation.

3.3. Generalization to 2D and 3D test cases

To test the generalization performance of the SL-Model, we evaluated it on several 2D and 3D test cases: BFS, parametric periodic hills, the Ahmed body, and Faith hill. We first of all examine the classification of the σ_{SL} classifier to determine whether it can generalize beyond the training cases on which it was tuned. We then compare the SL-Model predictions with those of the baseline $k-\omega$ SST model and with high-fidelity DNS/LES or experimental data, in order to evaluate the overall generalization accuracy of the SL-Model.

3.3.1. Parametric periodic hills (2D generalization test)

The classification results of the σ_{SL} classifier on the parametric periodic hill cases is shown in Fig. 8. The classifier effectively identifies shear layer regions requiring corrections, as indicated by the red regions corresponding to $\sigma_{SL} = 1$. The spatial extent of these regions systematically evolves with α: for larger hill spacing ($\alpha = 1.5$), the shear layer region extends further downstream due to delayed reattachment, while for smaller spacing ($\alpha = 0.5$), the classifier captures the more compact separation region and stronger interaction between successive hills. This systematic adaptation demonstrates the classifier's ability to identify physically relevant correction regions across weakly varying geometries.

The performance of the SL-Model on the parametric periodic hills test cases is summarized in Fig. 9, comparing axial velocity and TKE profiles for various α values. The SL-Model improves mean flow velocity predictions compared to the baseline $k-\omega$ SST model for most geometric configurations, with the magnitude of improvement varying systematically with geometry. However, this improvement typically comes at the cost of elevated TKE levels, and at the smallest spacing ($\alpha = 0.5$), the model struggles with both velocity and TKE predictions due to complex hill-to-hill interactions.

For larger hill spacing ($\alpha = 1.5$), where the flow has more room to develop, the velocity profiles show significant improvement in the extended separation region ($4 \leq x/H \leq 6$). The SL-Model reduces the baseline's underprediction of velocity within the recirculation zone, better capturing the delayed reattachment characteristic of widely-spaced hills. The TKE overprediction is most pronounced in the extended shear layer region, where the correction term R enhances mixing to achieve improved mean flow prediction.

For intermediate cases ($\alpha = 1.2$ and $\alpha = 1.0$), where the separation bubble size gradually decreases, the SL-Model shows particular improvement in capturing the separation point and initial shear layer development. At $x/H \approx 2.5$, the model better captures the velocity inflection point, reducing the overestimation of recirculation intensity observed in the baseline model. The TKE predictions show better spatial distribution compared to larger α cases, though still elevated compared to the DNS data.

For the smallest hill spacings ($\alpha = 0.8$ and $\alpha = 0.5$), where hill-to-hill interaction becomes stronger, the SL-Model shows more variation in performance. While it improves the prediction of the compact separation region, there is a tendency to overpredict velocity ($x/H > 3$) due to enhanced mixing. The TKE profiles reflect the intensified interaction between successive separations, with elevated levels particularly evident in the regions between hills. Despite these challenges, the model maintains better prediction of shear layer development compared to the baseline.

The particularly poor performance at $\alpha = 0.5$ can be attributed to the formation of secondary recirculation bubbles that are fundamentally different from the single, well-defined separation events used in model training. At the smallest spacing, the flow does not have sufficient distance to fully reattach between hills, leading to complex dual-bubble dynamics where a secondary recirculation zone forms at the foot of the second hill. The SL-Model's correction mechanism applies enhanced mixing uniformly across the identified shear layer region, but this uniform enhancement disrupts the delicate balance required to capture these multi-bubble interactions. The enhanced turbulent transport effectively suppresses the secondary bubble structure by over-energizing the flow field, preventing the formation of the weak recirculation zone at the second hill's base that is present in the DNS data. This limitation demonstrates that the current correction approach is most effective for single, well-defined separation events rather than complex multi-bubble configurations.

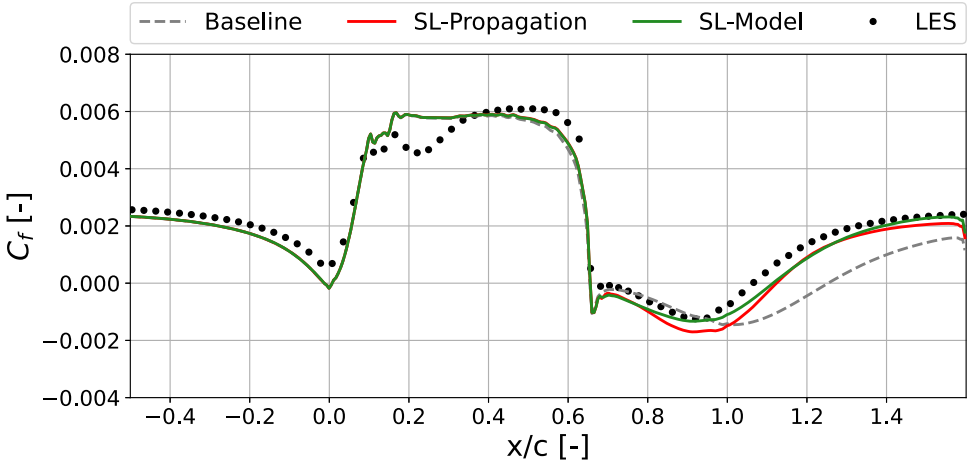


Fig. 7. Skin friction comparison plot for the SL-Model on the NASA-hump training case.

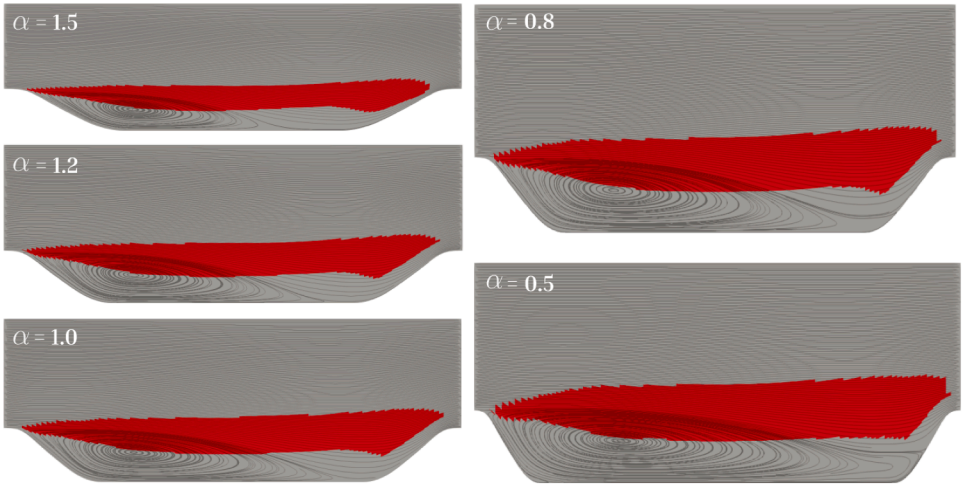


Fig. 8. Classification outcome of the σ_{SL} classifier on the parametric periodic hill cases (red region: $\sigma_{SL} = 1$, gray region: $\sigma_{SL} = 0$).

Having examined the detailed flow features across individual cases, we now assess whether these local observations translate to systematic trends in global prediction accuracy. Figs. 10(a) and (b) provide this aggregate perspective for the spacing parametrization studied here, where α varies hill separation and slope.

The root mean square error (RMSE) trend in Fig. 10(a) quantitatively confirms the geometry-dependence observed in the profiles. Across this spacing variation, the SL-Model exhibits minimum error at $\alpha = 1.2$, with elevated RMSE at both extremes: $\alpha = 0.5$ where complex multi-bubble interactions challenge the correction mechanism, and a slight increase at $\alpha = 1.5$. This contrasts sharply with the baseline's monotonic degradation, indicating that the correction mechanism responds differently to changes in hill spacing and slope than the underlying SST model.

Fig. 10(b) reveals the underlying mechanism: as hills become more widely spaced in this constant-height sequence, the gap between SL-Model TKE predictions and DNS data systematically narrows. This convergence suggests that for these geometries, the enhanced mixing strategy becomes increasingly compatible with the flow physics as separation regions extend and hill slopes become gentler. At tight spacings, the persistently elevated TKE levels quantify the overmixing that manifests as velocity overprediction in the profiles.

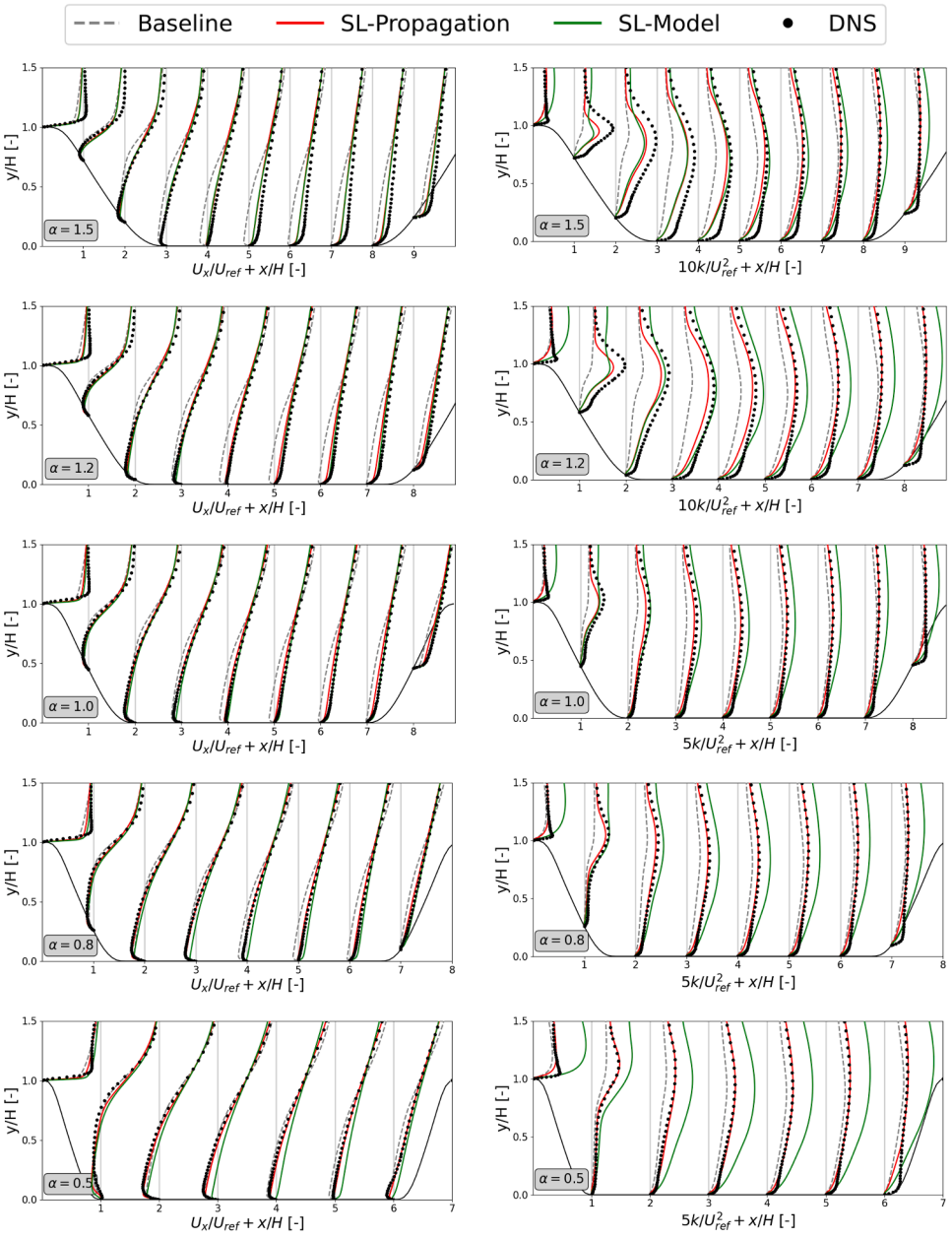
These complementary metrics reveal an important trade-off in the SL-Model's behavior: improved mean flow prediction comes through enhanced turbulent mixing that sometimes overestimates TKE magnitude, particularly at smaller α values where complex hill-to-hill interactions present greater modeling challenges. The model achieves minimum velocity error at intermediate spacings ($\alpha \approx 1.0$–1.2) while showing degraded performance at very tight spacings where multi-bubble dynamics dominate.

3.3.2. *Backwards facing step (2D generalization test)*

The parametric periodic hills demonstrated that model performance degrades as separation becomes steeper and hill-to-hill interactions intensify (particularly at $\alpha = 0.5$). The BFS represents the limiting case of this progression, with instantaneous separation at a sharp geometric discontinuity. This provides a critical test of whether corrections trained on smooth and curved separations can transfer to abrupt, edge-induced separation.

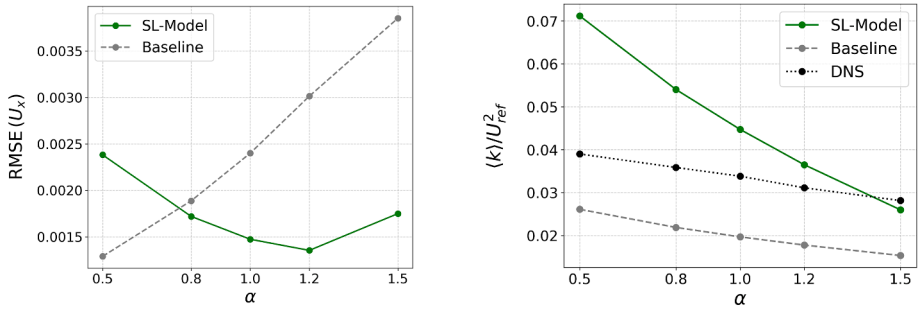
The BFS geometry provides a critical assessment of the SL-Model's ability to generalize from smooth and curved separation mechanisms to sharp-edge induced separation. Unlike the training cases, which feature either smooth surface separations (NASA-Hump) or curved geometries (periodic hill, CBFS), the sharp BFS presents an instantaneous separation at a well-defined geometric discontinuity. This fundamental difference in separation physics tests whether correction terms learned from pressure gradient and curvature-driven separations can successfully transfer to geometrically triggered separation.

Fig. 11 presents the classification results of the σ_{SL} classifier on the sharp BFS case. The classifier successfully identifies the shear layer region extending from the step edge ($x/H = 0$) downstream to the reattachment point, as indicated by the red region corresponding to $\sigma_{SL} = 1$. Notably, the classifier captures the characteristic thin shear layer ema-



(a) Axial velocity profiles. (b) TKE profiles.

Fig. 9. Performance of the SL-Model across parametric periodic hills cases with varying width-scaling factors ($\alpha = 0.5 - 1.5$). DNS data (black dots), baseline $k - \omega$ SST (dashed), SL-Propagation (solid red), and SL-Model (solid green).



(a) Root-mean-square error (RMSE) of axial velocity for baseline $k - \omega$ SST and SL-Model. (b) Average TKE in the shear layer region for baseline $k - \omega$ SST, SL-Model, and DNS data.

Fig. 10. Comparison of RANS model performance metrics across parametric periodic hills cases.

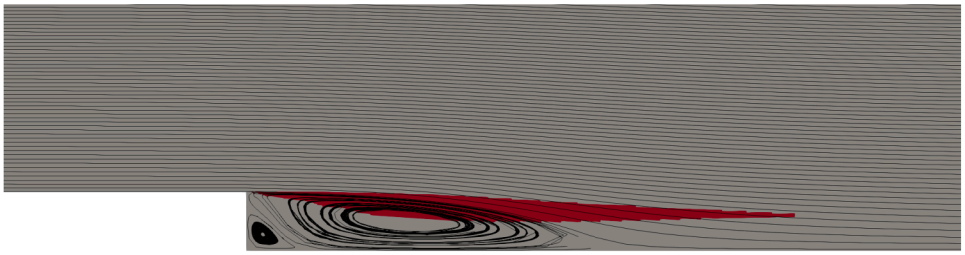
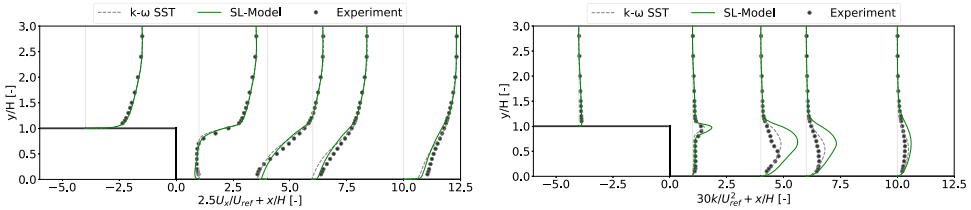


Fig. 11. Classification outcome of the σ_{SL} classifier on the BFS case (red region: $\sigma_{SL} = 1$, gray region: $\sigma_{SL} = 0$).



(a) Axial velocity profiles. (b) TKE profiles.

Fig. 12. Performance of the SL-Model on the BFS case. Experimental data (black dots), baseline $k - \omega$ SST (dashed), and SL-Model (solid).

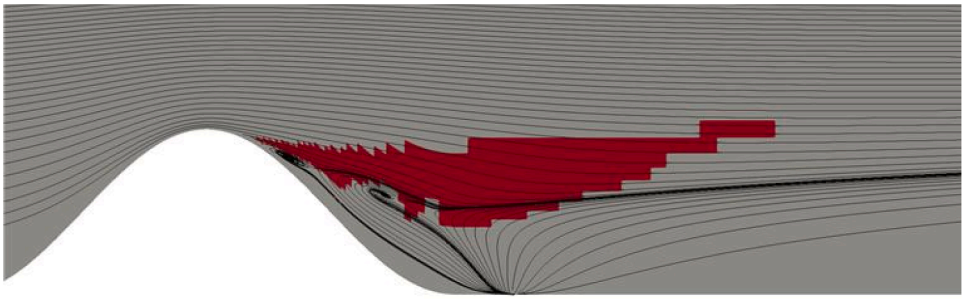
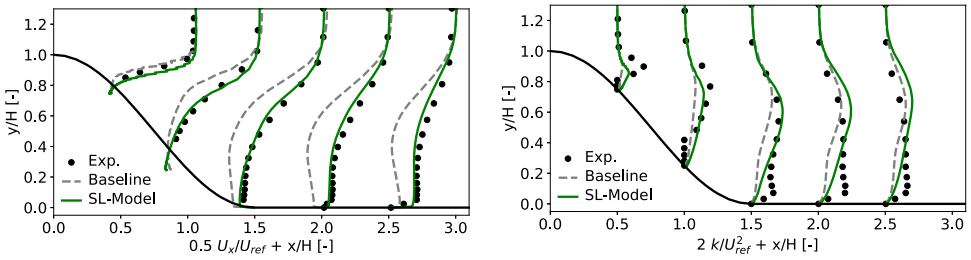


Fig. 13. Classification outcome of the σ_{SL} classifier on the Faith hill case (red region: $\sigma_{SL} = 1$, gray region: $\sigma_{SL} = 0$).



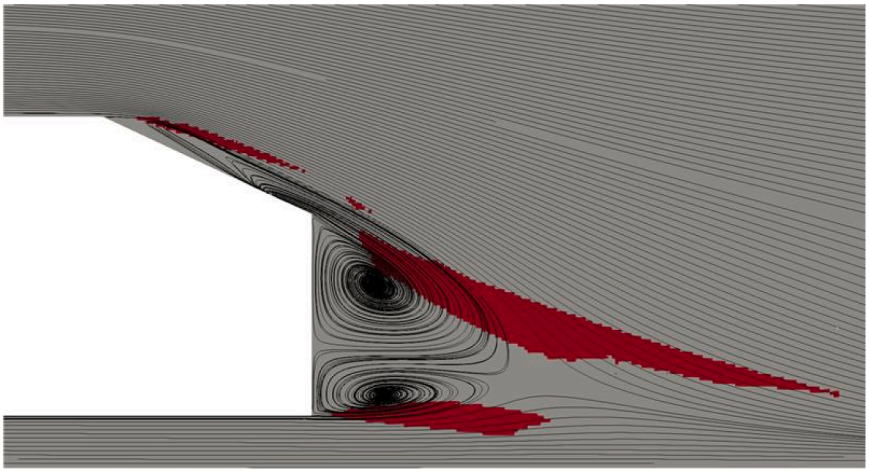
(a) Axial velocity profiles. (b) TKE profiles.

Fig. 14. Performance of the SL-Model on the Faith hill case at the symmetry plane: experimental data (black dots), baseline $k - \omega$ SST (dashed), and SL-Model (solid).

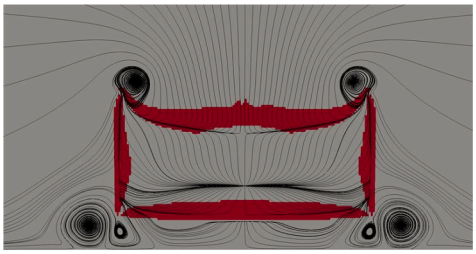
nating from the sharp edge, demonstrating that the classification criteria remain physically meaningful despite the geometric differences from the training configurations. The spatial extent of the correction region aligns well with the expected separated flow domain, providing confidence in the model's ability to identify relevant correction zones even for untrained separation mechanisms.

The performance comparison between the baseline $k - \omega$ SST model and the SL-Model on the sharp BFS case is presented in Fig. 12. The SL-Model demonstrates more modest improvements compared to the smooth geometry cases, with mixed performance in velocity prediction within the separation region. Notably, the baseline model actually performs better in the near-wall region at certain downstream locations ($x/H = 1$ and $x/H = 4$), while the SL-Model shows improvements primarily in the outer shear layer regions. This performance pattern suggests that the correction methodology is most effective for smooth geometry separations rather than sharp-edge induced separations.

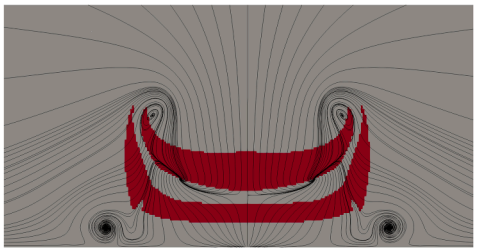
Consistent with observations from other cases, the SL-Model exhibits overproduction of TKE, particularly in the shear layer development region. The results indicate that while the SpaRTA framework can generalize across different separation mechanisms, the magnitude of improvement is geometry-dependent. The correction terms trained on smooth and curved separations provide significant benefits for similar flow physics but show diminished effectiveness for sharp geometric discontinuities. This suggests that the underlying separation mechanisms-whether driven by pressure gradients and curvature or by abrupt geometric changes influence the transferability of learned correction strategies within the SpaRTA framework.



(a) σ_{SL} classifier at symmetry plane.



(b) σ_{SL} classifier at $x/H = 0.27$



(c) σ_{SL} classifier at $x/H = 0.69$

Fig. 15. Classification outcome of the σ_{SL} classifier on the Ahmed body: visualization of shear layer regions ($\sigma_{SL} = 1$, red) at (a) symmetry plane showing separation from slant angle, (b,c) streamwise planes ($x/H = 0.27, 0.69$) showing how the classifier remains inactive in regions where 3D vortical structures develop.

3.3.3. Faith hill (3D generalization test)

Having examined 2D generalization across both smooth parametric geometries and sharp-edge separation, we now assess the model's performance on the Faith hill configuration-a fully 3D separated flow. This case presents a more challenging test of generalization, combining smooth-surface separation with specific 3D effects such as horseshoe vortex formation and spanwise flow variation that were not present in the training cases.

Fig. 13 shows the classification outcome of the σ_{SL} classifier at the symmetry plane. The classifier successfully identifies the separated shear layer region downstream of the hill's crest where the baseline $k-\omega$ SST model requires enhancement. The spatial extent of the identified region (shown in red) aligns with the expected separation zone, indicating that the classifier criteria remain physically relevant in 3D configurations.

The performance of the SL-Model at the symmetry plane is presented in Fig. 14. The axial velocity profiles show significant improvement over the baseline $k-\omega$ SST predictions, particularly in the separated region ($1.5 < x/H < 2.5$). The SL-Model better captures both the extent of the separation bubble and the recovery of the velocity profiles downstream. This improvement demonstrates that the tensor-based corrections, formulated using Pope's invariant basis, maintain their effectiveness when applied to 3D flows.

The TKE profiles in Fig. 14(b) exhibit behavior consistent with previous test cases. While the SL-Model shows improved prediction of TKE distribution compared to the baseline, particularly in capturing peak locations, it maintains the characteristic overprediction in the shear layer region. This consistency between 2D and 3D predictions suggests the correction mechanisms remain physically relevant while trained on 2D data due to their invariant formulation.

The SL-Model's ability to improve predictions for this 3D configuration stems from two key aspects: the use of frame-invariant quantities in both the classifier and correction terms, ensuring their physical meaning translates across dimensionality; and the robust formulation of the correction terms using Pope's tensor basis, which maintains consistent physical behavior in both 2D and 3D flows. This successful application to the Faith hill case demonstrates the model's capability to enhance predictions beyond its training scope.

3.3.4. Ahmed body (3D generalization test)

We examine the Ahmed body at slant angle of 25 ° [45], which features complex separation patterns including streamwise vortices shed from the sharp upper corners of the rear slant. At a 25 ° rear slant angle, these corner vortices have sufficient strength to reattach the flow half way down the slant. The Ahmed body, with its distinct slant angle and sharp edges, represents a significant departure from the training configurations, testing the model's ability to handle fully 3D separation mechanisms.

Fig. 15 presents the σ_{SL} classifier results across multiple planes. At the symmetry plane (Fig. 15(a)), it captures the separation from the slant angle and subsequent shear layer development. The streamwise planes (Figs. 15(b) and (c)) reveal how the classifier appropriately remains inactive in regions where streamwise vortices develop from the slant edges. This selective activation demonstrates the classifier's ability to distinguish between shear layer separation and vortical regions.

Fig. 16 illustrates the wake structure development at three streamwise locations behind the Ahmed body. At $x/H = 0$ (immediately behind the body), all three cases show the initial formation of the wake, with the baseline $k-\omega$ SST model predicting a larger recirculation region (blue contour enclosed by the black zero-velocity line) that is not present in the experimental data. The SL-Model shows improved prediction of the initial wake formation by correctly capturing the absence of this immediate recirculation.

At $x/H = 0.27$, a pronounced recirculation zone is visible in all three cases, clearly defined by the black zero-velocity contour line. The exper-

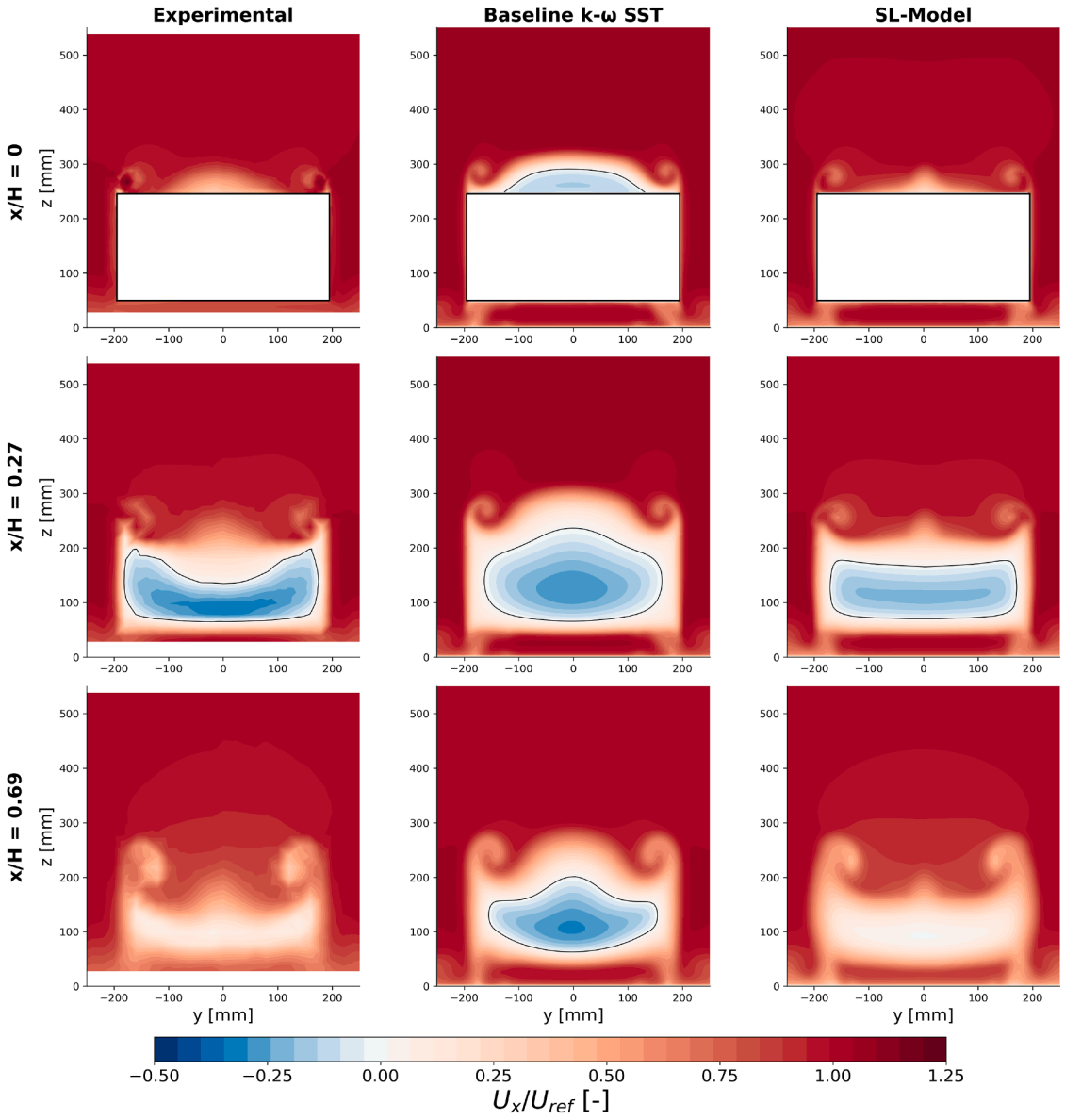


Fig. 16. Comparison of wake structure predictions at streamwise planes ($x/H = 0, 0.27, 0.69$). Columns show experimental data (left), baseline $k-\omega$ SST model (middle), and SL-Model (right). Contours show streamwise velocity normalized by freestream velocity. Black contour lines indicate where streamwise velocity is zero, delineating the separation region.

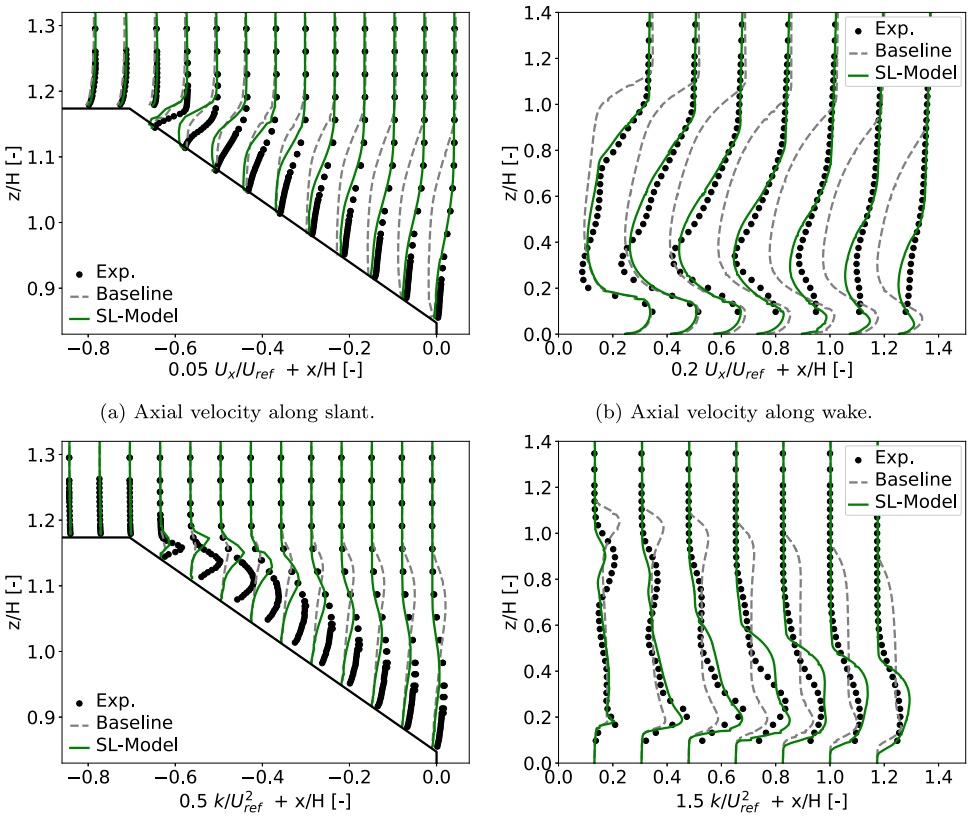
imental data reveals an asymmetric recirculation bubble with varying depths across the span. The baseline model overpredicts both the size and intensity of this recirculation region, showing a more symmetric, deeper, and more concentrated negative velocity region. In contrast, the SL-Model produces a shallower recirculation zone with dimensions and intensity that more closely match the experimental measurements, particularly in capturing the characteristic flattened shape of the upper boundary of the recirculation region.

Further downstream at $x/H = 0.69$, where wake recovery begins, significant differences persist. The baseline model continues to predict an overly strong and concentrated recirculation region, as indicated by the extent of its zero-velocity contour, while the experimental data shows a weakening and more diffuse negative velocity region. The SL-Model maintains better agreement with experimental data by correctly predicting a more moderate recirculation intensity and better capturing the spatial extent of the wake. This demonstrates the SL-Model's superior ability to predict both the initial formation and subsequent evolution of the wake structure, while appropriately preserving baseline model behavior in areas dominated by vortical structures outside the primary recirculation region.

Quantitative validation through symmetry plane profiles in Fig. 17 demonstrates these improvements. The velocity profiles along the slant angle ($-0.7 < x/H < 0$) show better prediction of shear layer development, particularly evident near $x/H = -0.65$ where the baseline model underpredicts the velocity gradient. In the wake region ($x/H > 0$), the SL-Model achieves closer agreement with experimental data for both velocity deficit magnitude and recovery rate.

The TKE profiles at the symmetry plane maintain trends observed in previous cases. While showing characteristic elevated levels compared to experimental data, the SL-Model better captures the spatial distribution of turbulence in the shear layer region. This indicates the correction mechanisms appropriately enhance mixing where intended, though with some overprediction of TKE.

The model's successful prediction of separated shear layer features, combined with its appropriate preservation of baseline behavior in vortical regions, demonstrates the effectiveness of the zonal correction ap-



(a) Axial velocity along slant.

(b) Axial velocity along wake.

(c) TKE along slant.

(d) TKE along wake.

Fig. 17. Performance of the SL-Model on the ahmed body case at the symmetry plane: experimental data (black dots), baseline $k-\omega$ SST (dashed), and SL-Model (solid).

proach. This selective enhancement suggests potential for developing targeted corrections for specific flow features, such as streamwise vortices, in future work.

4. Conclusion

This study introduced the RITA framework, a physics-based methodology for selectively improving RANS predictions in separated flows. By coupling a frame-invariant classifier with compact SpaRTA-derived correction models, we developed a systematic and interpretable approach of applying data-driven augmentation specifically within shear layer regions where the baseline model demonstrably fails. The resulting zonally-augmented $k-\omega$ SST model formulation preserved the model's accuracy in canonical flows while enhancing prediction of separation and reattachment dynamics across both 2D and 3D configurations.

The methodology integrates physically interpretable classification with targeted corrections: the σ_{SL} classifier isolates shear layer regions based on the balance of terms in the k-equation, and the derived anisotropic and scalar correction models for b^{Δ}_{ij} and R, address deficiencies in Reynolds stress anisotropy and turbulent production. Together, these mechanisms improved mean-flow prediction across geometries ranging from simple 2D cases such as the NASA-Hump and periodic hills to complex 3D configurations like the Ahmed body, confirming the robustness and transferability of the approach.

Nonetheless, some limitations remain. The tendency toward TKE overprediction highlights the need for refined balancing between enhanced mixing and production scaling. Moreover, the classifier's reliance on the turbulent-to-total kinetic energy ratio, which is not Galilean invariant, and the normalization of the vorticity Reynolds number as a global quantity may limit strict case-to-case consistency. Addressing these aspects through locally normalized and fully invariant criteria could further strengthen the generality and physical rigor of the σ_{SL} classifier.

For future work, several promising directions emerge. First, extending the training to include 3D cases could improve the model's performance in predicting complex 3D features. Second, development of specialized classifiers for vortical structures could complement the current shear layer focus, allowing comprehensive treatment of both separation and vortical regions. Third, further refinement of the correction terms might address the TKE overprediction while maintaining the improved mean flow predictions. Fourth, implementation of a smooth blending function to replace the binary classifier would eliminate potential discontinuities at region interfaces, improving robustness for implicit solver implementations.

In conclusion, the RITA methodology and resulting zonally-augmented $k-\omega$ SST model provide a viable approach for enhancing predictions of separated flows. The framework's success in handling various separation mechanisms while appropriately distinguishing between shear layers and vortical regions demonstrates that selective application of physically-motivated corrections offers an effective pathway toward improved RANS modeling for complex aerodynamic configurations.

CRediT authorship contribution statement

Tyler Buchanan: Writing - original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Monica Lăcătuş:** Writing - review & editing, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization; **Alastair West:** Writing – review & editing, Visualization, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis; **Richard P. Dwight:** Writing – review & editing, Supervision, Software, Resources, Project administra-

tion, Methodology, Investigation, Funding acquisition, Conceptualization.

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.

Acknowledgement

The authors gratefully acknowledge Williams Grand Prix Engineering Limited for financial support.

Appendix A. Non-Linear Basis Tensors and Invariants

The ten non-linear basis tensors $T_{ij}^{(n)}$ and five invariants I_m are defined following the work of Pope [30]:

$$T_{ij}^{(1)} = S \tag{A.1}$$

$$T_{ij}^{(2)} = S\Omega - \Omega S \tag{A.2}$$

$$T_{ij}^{(3)} = S^2 - \frac{1}{3} I\{S^2\} \tag{A.3}$$

$$T_{ij}^{(4)} = \Omega^2 - \frac{1}{3} I\{\Omega^2\} \tag{A.4}$$

$$T_{ij}^{(5)} = \Omega S^2 - S^2\Omega \tag{A.5}$$

$$T_{ij}^{(6)} = \Omega^2 S + S\Omega^2 - \frac{2}{3} I\{S\Omega^2\} \tag{A.6}$$

$$T_{ij}^{(7)} = \Omega S\Omega^2 - \Omega^2 S\Omega \tag{A.7}$$

$$T_{ij}^{(8)} = S\Omega S^2 - S^2\Omega S \tag{A.8}$$

$$T_{ij}^{(9)} = \Omega^2 S^2 + S^2\Omega^2 - \frac{2}{3} I\{S^2\Omega^2\} \tag{A.9}$$

$$T_{ij}^{(10)} = \Omega S^2\Omega^2 - \Omega^2 S^2\Omega \tag{A.10}$$

where the corresponding five invariants are:

$$I_1 = \{S^2\}, \quad I_2 = \{\Omega^2\}, \quad I_3 = \{S^3\}, \quad I_4 = \{\Omega^2 S\}, \quad I_5 = \{\Omega^2 S^2\}. \tag{A.11}$$

Note: S represents the symmetric part of the velocity gradient tensor, Ω represents the antisymmetric part, and $I\{\cdot\}$ denotes the isotropic part of the tensor.

Appendix B. Unary Functions for Feature Library

The following unary transformations were applied to the invariants in constructing the feature libraries for b_{ij}^{Δ} and R:

$$f_1(x) = x \quad \text{(Identity)} \tag{B.1}$$

$$f_2(x) = x^2 \quad \text{(Square)} \tag{B.2}$$

$$f_3(x) = |x| \quad \text{(Absolute value)} \tag{B.3}$$

$$f_4(x) = \sqrt{|x|} \quad \text{(Square root of absolute value)} \tag{B.4}$$

$$f_5(x) = \tanh(x) \quad \text{(Hyperbolic tangent)} \tag{B.5}$$

$$f_6(x) = \frac{x}{1 + x^2} \quad \text{(Regularized division)} \tag{B.6}$$

$$f_7(x) = \log(|x| + 1) \quad \text{(Regularized logarithm)} \tag{B.7}$$

The regularized functions provide numerical stability by avoiding singularities. An example of a candidate function in the library would be $\sqrt{|q_k|}T_{ij}^1$, where q_k is an invariant and T_{ij}^1 is a basis tensor.

Appendix C. Simulation Setup of Flow Cases

This section provides an overview of the OpenFOAM simulation setups used for the different cases in this study. Unless stated otherwise, the boundary conditions described here apply to all simulation types: baseline k-ω SST, k-corrective-frozen RANS, SL-Propagation, and SL-Model simulations. The initial conditions depend on the simulation type. For the baseline k-ω SST simulations, the standard initial conditions are applied. In the k-corrective-frozen RANS simulations, the high-fidelity fields k^* and U^* are used to initialize the solution, while ν_t and ω are initialized in the same way as in the baseline case. For the SL-Propagation and SL-Model simulations, the initial conditions are taken from the converged baseline k-ω SST solution.

Table C.1
Overview of the flow parameters specified for the NASA-Hump case.

Transport Property	Parameter	Value
Reynolds number based on chord length	Re_c	936,000
Kinetic viscosity	ν	1.55 x 10^{-5} m^2 s^{-1}
Free-stream reference velocity	U_{ref}	34.6 m s^{-1}
Reference kinetic energy	k_{ref}	0.00107 m^2 s^{-2}
Reference specific dissipation rate	ω_{ref}	0.118 s^{-1}
Reference pressure	p_{ref}	0 m^2 s^{-2}

Table C.2
Boundary and initial conditions for the NASA-Hump case. Wall Function is abbreviated to W.F.

Boundary Conditions					
Location	U [m/s]	p [m^2/s^2]	k [m^2/s^2]	ω [s^{-1}]	ν_t [m^2/s]
Inlet	fixed Value	zero Gradient	fixed Value	fixed Value	calculated
Outlet	zero Gradient	fixed Value	zero Gradient	zero Gradient	calculated
Top	symmetry				
Bottom	no Slip	zero Gradient	kqR W.F.	omega W.F.	nutUSpalding W.F.
Initial Conditions					
Standard	[U_{ref} 0 0]	0	k_{ref}	ω_{ref}	0.009

C.1. NASA-Hump

The NASA-Hump case is part of NASA's 2D-separated flow validation cases [47]. Based on the Glauert-Goldschmied type body, it follows the experimental setup in Greenblatt et al. [48], with the OpenFOAM configuration from Hoefnagel [49]. At a reference Mach number of 0.1, the flow is treated as incompressible. The case features a turbulent boundary layer that accelerates over the hump under a favorable pressure gradient, separates at the hump's crest due to an adverse pressure gradient, then reattaches downstream.

C.1.1. Flow parameters

An overview of the main flow parameters specified for the NASA-Hump case is given in Table C.1. This case has the highest Reynolds number equal to 936,000, computed based on the chord length of the hump c and the free-stream reference velocity U_{ref}.

C.1.2. Initial and boundary conditions

The boundary conditions and initial conditions for the NASA-Hump case in OpenFOAM are outlined in Table C.2. The k-corrective-frozen RANS simulation uses high-fidelity LES data fields from Uzun and Malik [39].

C.2. Periodic hill

The periodic hill case setup and the high-fidelity LES data have been obtained from the study of Breuer et al. [40], who extensively studied this geometry across various Reynolds numbers. This configuration consists of a series of repeating hills separated by a flat surface region. The

Table C.3
Overview of the flow parameters specified for the periodic hill case.

Transport Property	Parameter	Value
Reynolds number based on hill height	Re_H	10,595
Kinetic viscosity	ν	9.44 x 10^{-5} m^2 s^{-1}
Free-stream reference velocity	U_{ref}	1 m s^{-1}
Reference kinetic energy	k_{ref}	0.00375 m^2 s^{-2}
Reference specific dissipation rate	ω_{ref}	0.110 s^{-1}
Reference pressure	p_{ref}	0 m^2 s^{-2}

Table C.4
Boundary and initial conditions for the periodic hill case. Wall Function is abbreviated to W.F.

Boundary Conditions					
Location	U [m/s]	p [m^2/s^2]	k [m^2/s^2]	ω [s^{-1}]	ν_t [m^2/s]
Inlet/Outlet	cyclic	cyclic	cyclic	cyclic	cyclic
Top/Bottom	no Slip	zero Gradient	kqR W.F.	omega W.F.	nutUSpalding W.F.
Initial Conditions					
Standard	[U_{ref} 0 0]	0	1×10^{-15}	ω_{ref}	0

Table C.5
Overview of the flow parameters specified for the CBFS case.

Transport Property	Parameter	Value
Reynolds number based on step height	Re_H	13,700
Kinetic viscosity	ν	7.23 x 10^{-5} m^2 s^{-1}
Free-stream reference velocity	U_{ref}	1 m s^{-1}
Reference kinetic energy	k_{ref}	0.00668 m^2 s^{-2}
Reference specific dissipation rate	ω_{ref}	0.110 s^{-1}
Reference pressure	p_{ref}	0 m^2 s^{-2}

spacing between the hills is calculated to enable the flow to reattach to the flat surface post-separation before encountering the next hill.

C.2.1. Flow parameters

An overview of the main flow parameters specified for the periodic hill case is given in Table C.3. This case has a Reynolds number equal to 10,595, computed based on the hill height H and reference bulk velocity U_{ref}.

C.2.2. Initial and boundary conditions

The OpenFOAM boundary and initial conditions for this case are outlined in Table C.4. To simulate the series of hills, periodic boundary conditions are applied at the inflow and outflow of the domain. The boundary conditions for k and ν_t differ from those specified for the NASA-Hump case. Here, k is specified to be zero at the walls. The `nutLowReWallFunction` also sets ν_t to zero.

C.3. CBFS

The CBFS closely resembles the NASA-Hump and periodic hill cases. It involves a turbulent boundary layer separating from a curved step under an adverse pressure gradient. The key distinction lies in the contour of the step, which has a much gentler curvature compared to the other two cases, thereby promoting the flow to remain attached for longer. The case setup, the high-fidelity LES data and computational grid have been obtained from the study of Bentaleb et al. [41].

C.3.1. Flow parameters

An overview of the main flow parameters specified for the CBFS case is given in Table C.5. This case has a Reynolds number equal to 13,700, computed based on the step height H and the free-stream inlet reference velocity U_{ref}.

Table C.6
Boundary and initial conditions for the CBFS case. Wall Function is abbreviated to W.F.

Boundary conditions					
Location	U [m/s]	p [m^2/s^2]	k [m^2/s^2]	ω [s^{-1}]	ν_t [m^2/s]
Inlet	fixed Value	zero Gradient	fixed Value	fixed Value	calculated
Outlet	zero Gradient	fixed Value	zero Gradient	zero Gradient	calculated
Top Wall	no Slip	zero Gradient	1 x 10^{-15}	omega W.F.	nutLowRe W.F.
Bottom Wall	no Slip	zero Gradient	1 x 10^{-15}	omega W.F.	nutLowRe W.F.
Initial Conditions					
Standard	[U_{ref} 0 0]	0	k_{ref}	ω_{ref}	0

Table C.7
Overview of the flow parameters specified for the BFS case.

Transport Property	Parameter	Value
Reynolds number based on step height	Re_H	3.6×10^4
Step height	H	1.27 cm
Kinetic viscosity	ν	1.56×10^{-5} m^2 s^{-1}
Free-stream reference velocity	U_{ref}	44.2 m s^{-1}
Reference kinetic energy	k_{ref}	0.00109 m^2 s^{-2}
Reference specific dissipation rate	ω_{ref}	181,728 s^{-1}
Reference pressure	p_{ref}	0 m^2 s^{-2}

C.3.2. Initial and boundary conditions

The OpenFOAM boundary and initial conditions for this case are detailed in Table C.6. These are similar to the ones prescribed for the other 2D-separated flow cases.

C.4. Parametric periodic hills

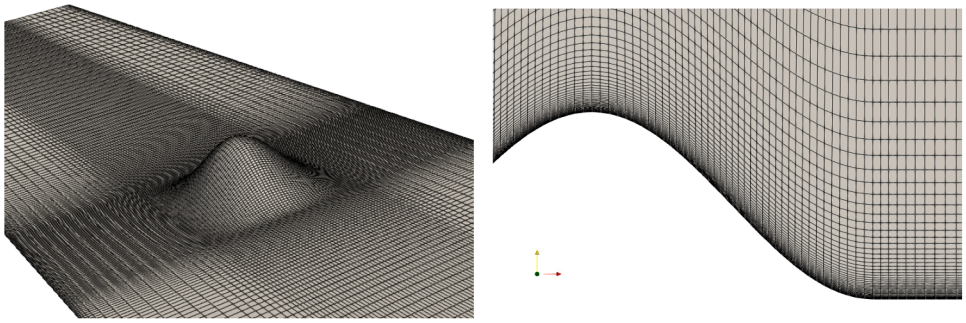
The setups of the parametric periodic hill cases and the high-fidelity DNS data have been obtained from the study of Xiao et al. [43], who systematically varied the geometry using a hill width-scaling factor ($\alpha = 0.5 - 1.5$). Detail on the case flow parameters and OpenFOAM boundary and initial conditions can be found in database from https://github.com/xiaoh/para-database-for-PIML.git.

C.5. BFS

The BFS case represents a fundamentally different separation mechanism from the smooth and curved geometries used in the training cases. It involves a turbulent boundary layer that undergoes instantaneous separation at a well-defined geometric discontinuity - the sharp step edge. This sharp-edge induced separation provides a critical assessment of the SL-Model's ability to generalize from smooth surface separations (NASA-Hump) and curved geometries (periodic hill, CBFS) to geometrically-triggered separation mechanisms. The case follows the NASA Langley Research Center validation case based on the experimental study of Driver and Seegmiller [42] , and represents a widely-tested benchmark configuration in the turbulence modeling community.

C.5.1. Flow parameters

An overview of the main flow parameters specified for the BFS case is given in Table C.7. This case has a Reynolds number equal to 3.6×10^4 based on step height H and reference velocity U_{ref}, corresponding to a Reynolds number of 5000 based on boundary layer momentum thickness prior to the step. The boundary layer thickness upstream of the step is approximately $1.5H$.



(a) Surface mesh. (b) Volume mesh at symmetry plane.

Fig. C.18. Surface and volume mesh of Faith hill case.

Table C.8
Boundary and initial conditions for the BFS case. wall function is abbreviated to W.F.

Boundary Conditions					
Location	U [m/s]	p [m^2/s^2]	k [m^2/s^2]	ω [s^{-1}]	ν_t [m^2/s]
Inlet	fixed Value	zero Gradient	fixed Value	fixed Value	calculated
Outlet	zero Gradient	fixed Value	zero Gradient	zero Gradient	calculated
Startup Walls	symmetry Plane	symmetry Plane	symmetry Plane	symmetry Plane	symmetry Plane
Upper/Lower Walls	no Slip	zero Gradient	kqR W.F.	omega W.F.	nutUSpalding W.F.
Front/Back	empty	empty	empty	empty	empty
Initial Conditions					
Standard	[U_{ref} 0 0]	0	k_{ref}	ω_{ref}	0

C.5.2. Initial and boundary conditions

The OpenFOAM boundary and initial conditions for this case are detailed in Table C.8 . The configuration follows the NASA validation case setup, with both bottom and top walls treated as viscous walls except for a short symmetry region near the inlet to avoid incompatibilities between freestream inflow and wall boundary conditions.

C.6. Faith hill

The *Fundamental Aero Investigates The Hill* experiment conducted at NASA Ames [44] is a complex 3D smooth-body separated aerodynamic flow. The hill has an axisymmetric cosine cross-section, and the oncoming boundary layer thickness was 1/3 the hill height H. The computational domain extents were the same as the wind tunnel test section, specifically $L_x/H = 20$, $L_y/H = 5.22$, $L_z/H = 8$ where the hill is placed at the center of the section [50]. The velocity at the inlet was prescribed as uniform and the boundary layer develops naturally. This matches well with the experimental measurements of [44]. The turbulence intensity matched the experiment at $TI = 0.13\,\%$ and the turbulent length scale used was that of the hill height. A mesh was used with 1.56×10^6 cells which was sufficient to obtain grid independence. The surface mesh resolution and volume mesh resolution at the symmetry plane can be seen in Fig. C.18.

C.6.1. Flow parameters

An overview of the main flow parameters specified for the Faith hill case is given in Table C.9. This case has a Reynolds number equal to 500,000, computed based on the hill height H and the free-stream inlet reference velocity U_{ref}.

Table C.9
Overview of the flow parameters specified for the Faith hill case.

Transport Property	Parameter	Value
Reynolds number based on hill height	Re_H	500,000
Kinetic viscosity	ν	2.0×10^{-6} m^2 s^{-1}
Free-stream reference velocity	U_{ref}	1 m s^{-1}
Reference kinetic energy	k_{ref}	2.535×10^{-5} m^2 s^{-2}
Reference specific dissipation rate	ω_{ref}	0.029 s^{-1}
Reference pressure	p_{ref}	0 m^2 s^{-2}

Table C.10
Boundary and initial conditions for the Faith hill case. Wall Function is abbreviated to W.F.

Boundary Conditions					
Location	U [m/s]	p [m^2/s^2]	k [m^2/s^2]	ω [s^{-1}]	ν_t [m^2/s]
Inlet	fixed Value	zero Gradient	fixed Value	fixed Value	calculated
Outlet	zero Gradient	fixed Value	zero Gradient	zero Gradient	calculated
Hill/Walls	no Slip	zero Gradient	kqR W.F.	omega W.F.	nutUSpalding W.F.
Symmetry	symmetry Plane	symmetry Plane	symmetry Plane	symmetry Plane	symmetry Plane
Initial Conditions					
Standard	[U_{ref} 0 0]	0	k_{ref}	ω_{ref}	0

C.6.2. Initial and boundary conditions

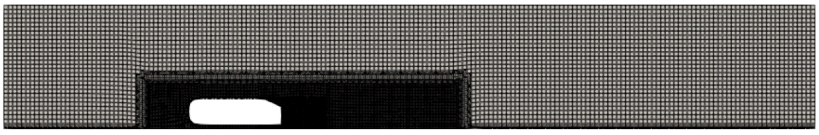
The OpenFOAM boundary and initial conditions for this case are detailed in Table C.10.

C.7. Ahmed body

The Ahmed body (Lienhart et al. [45]) comprises a flat front with rounded corners and a sharp slanted rear upper surface. The 25 ° rear slant angle was simulated including the stilts on which the model was mounted. An inlet condition is imposed 3m upstream of the body and an outlet condition is imposed 6m downstream. A no-slip wall condition is imposed on the ground plane and car body, with slip walls applied to the wind tunnel walls. An unstructured mesh of 11M cells was generated using snappyHexMesh (Fig. C.19). A high-y + approach was taken with 3 prism layers, resulting in a y^+ range of approximately 50–100 over the top surface of the body.

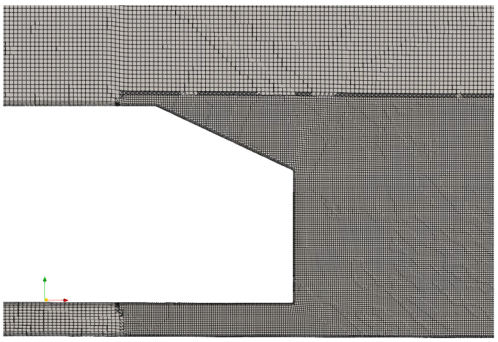
C.7.1. Flow parameters

An overview of the main flow parameters specified for the Ahmed body case is given in Table C.11. This case has a Reynolds number equal to 760,000 based on the body height $H = 0.288$m or 278,000 based on body length $L = 1.044$m with the free-stream inlet reference velocity U_{ref}.



(a) Domain volume refinement at symmetry plane.

(b) Body surface mesh.



(c) Close-up of rear refinement.

Fig. C.19. Surface and volume mesh of Ahmed body case.

Table C.11
Overview of the flow parameters specified for the Ahmed body case.

Transport Property	Parameter	Value
Reynolds number based on body height	Re_H	760,000
Reynolds number based on body length	Re_L	278,000
Kinetic viscosity	ν	15 x 10^{-6} m^2 s^{-1}
Free-stream reference velocity	U_{ref}	40 m s^{-1}
Reference kinetic energy	k_{ref}	0.00108 m^2 s^{-2}
Reference specific dissipation rate	ω_{ref}	0.110 s^{-1}
Reference pressure	p_{ref}	0 m^2 s^{-2}

Table C.12
Boundary and initial conditions for the ahmed body case. Wall function is abbreviated to W.F.

Boundary Conditions					
Location	U [m/s]	p [m^2/s^2]	k [m^2/s^2]	ω [s^{-1}]	ν_t [m^2/s]
Inlet	fixed Value	zero Gradient	fixed Value	fixed Value	calculated
Outlet	zero Gradient	fixed Value	zero Gradient	zero Gradient	calculated
Body/Ground Plane	no Slip	zero Gradient	kqR W.F.	omega W.F.	nutUSpalding W.F.
External Walls	Slip	zero Gradient	zero Gradient	omega W.F.	zero Gradient
Initial Conditions					
Standard	[U_{ref} 0 0]	0	k_{ref}	ω_{ref}	0

C.7.2. Initial and boundary conditions

The OpenFOAM boundary and initial conditions for this case are detailed in Table C.12.

Appendix D. Model Coefficients and Auxiliary Equations

Box 2. Zonally-Augmented $k-\omega$ SST Model Coefficients.

Model Coefficients: Let Φ_1 represent the coefficients in the $k-\omega$ model and Φ_2 those in the transformed $k-\epsilon$ model. The coefficients Φ in the $k-\omega$ SST model are found from : $\Phi = F_1\Phi_1 + (1-F_1)\Phi_2$. The Φ_1 set of coefficients:

$$\sigma_{k1} = 0.85, \quad \sigma_{\omega 1} = 0.5, \quad \beta_1 = 0.0750, \quad \beta^* = 0.09, \quad \gamma_1 = 5/9 \tag{D.1}$$

The Φ_2 set of coefficients:

$$\sigma_{k2} = 1.0, \quad \sigma_{\omega 2} = 0.856, \quad \beta_2 = 0.0828, \quad \beta^* = 0.09, \quad \gamma_2 = 0.44 \tag{D.2}$$

Auxiliary Equations:

$$F_1 = \tanh\left(arg_1^4\right) \text{ where } arg_1 = \min\left[\max\left(\frac{\sqrt{k}}{\beta^*\omega y}, \frac{500\nu}{y^2\omega}\right), \frac{4\rho\sigma_{\omega_2}k}{CD_{k\omega}y^2}\right] \tag{D.3}$$

$$CD_{k\omega} = \max\left(2\rho\sigma_{\omega_2}\frac{1}{\omega}\frac{\partial k}{\partial x_i}\frac{\partial \omega}{\partial x_i}, 10^{-20}\right) \tag{D.4}$$

$$F_2 = \tanh\left(arg_2^2\right) \quad \text{where} \quad arg_2 = \max\left(\frac{2\sqrt{k}}{\beta^*\omega y}, \frac{500\nu}{y^2\omega}\right) \tag{D.5}$$

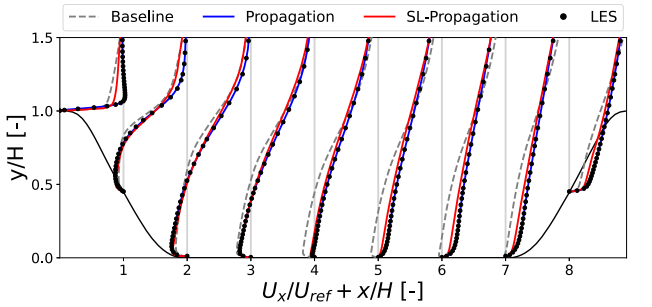
Appendix E. Training Case Performance for Periodic Hill and CBFS

E.1. Periodic hill

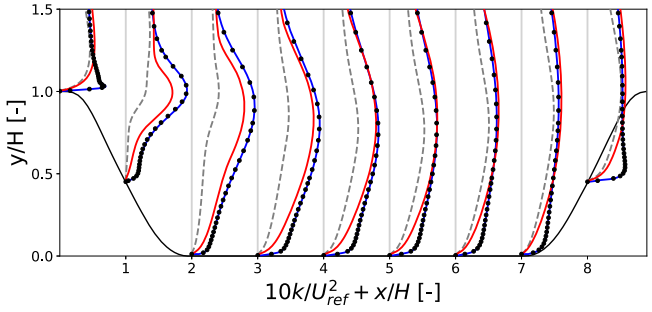
In this section, we examine the SL-Model's performance on the periodic hill configuration where the flow undergoes sustained cycles of separation and reattachment. Figs. E.20 and E.21 compare the performance of the full-field Propagation and SL-Propagation for this configuration.

The SL-Propagation case demonstrates comparable performance to full-field Propagation, with both approaches improving upon baseline predictions. The velocity profiles show better agreement with LES data throughout the periodic domain, particularly in capturing the recirculation zone extent between hills ($1 < x/H < 3$). The skin friction evolution confirms that the zonal framework maintains accuracy in predicting separation and reattachment behavior. Building on these results, Fig. E.22 presents the performance of the complete SL-Model implementation.

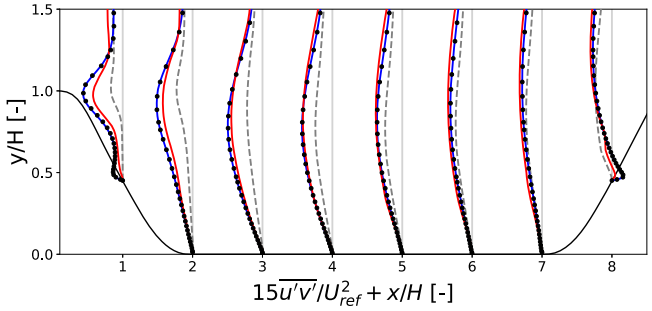
The SL-Model maintains the improved prediction of separation and reattachment behavior seen in the SL-Propagation case. Analysis of the turbulence quantities reveals characteristics similar to those observed in the NASA-Hump case. The TKE profiles (Fig. E.22(b)) show elevated levels compared to LES data, especially in the shear layer region above



(a) Axial velocity profiles.



(b) TKE profiles.



(c) Reynolds shear stress profiles.

Fig. E.20. Performance comparison between full-field Propagation and SL-Propagation on the periodic hill training case.

the hills, while the Reynolds shear stress distributions (Fig. E.22(c)) capture peak stress locations but with magnitudes influenced by the enhanced TKE prediction. This consistent behavior across both geometries points to an underlying mechanism in the correction approach where mean flow improvements are achieved through intensified turbulent transport.

Examination of the skin friction distribution (Fig. E.23) further validates the model's effectiveness across multiple separation-reattachment cycles. The SL-Model accurately reproduces peak skin friction at the hill crest ($x/H \approx 8.5$) while improving predictions in the recovery region between hills. It successfully captures both negative skin friction values in the separation zone and the gradual recovery trend upon reattachment. These results demonstrate that the zonal framework effectively handles the complex physics of periodic separation while maintaining appropriate behavior in attached flow regions.

E.2. CBFS

Following validation on smooth and periodic separation, we examine the model's performance on the CBFS, where the combination of geometric expansion and surface curvature creates a more complex separation mechanism. Figs. E.24 and E.25 compare the performance of full-field Propagation and SL-Propagation for this configuration.

The velocity profiles in Fig. E.24(a) demonstrate that SL-Propagation achieves comparable flow prediction to full-field Propagation, with notable improvements over the baseline model in the post-separation region ($x/H > 2.5$). The TKE profiles in Fig. E.24(b) show enhanced prediction of mixing layer development downstream of separation, while the Reynolds stress distributions in Fig. E.24(c) better capture the spatial evolution of turbulent transport through the curved section and subsequent expansion. This indicates that the zonal application of corrections effectively maintains the key physical mechanisms captured by full-field Propagation.

The skin friction distribution in Fig. E.25 validates these improvements quantitatively. Both correction approaches improve the prediction of separation onset and show significantly better agreement with LES data in the recovery region compared to the baseline model. While both correction approaches show similar performance, some differences emerge: SL-Propagation follows LES data more closely near $x/H = 5$, while full-field Propagation shows marginally better agreement in the early wake region ($7.5 \leq x/H \leq 10$). The σ_{SL} classifier demonstrates effective performance through two key behaviors: maintaining the baseline model's accurate prediction upstream of separation ($x/H < 0$) where corrections are not needed, and allowing a natural transition back to baseline behavior in the far wake region ($x/H > 10$). The close agreement between SL-Propagation and full-field Propagation confirms the σ_{SL} classifier's effectiveness in identifying regions requiring correction, particularly through the geometric transition.

Building on these results, the velocity profiles in Fig. E.26(a) show that the SL-Model successfully maintains the improved prediction of separation bubble size and recovery region. Analysis of the turbulence quantities reveals behavior consistent with previous test cases, but with important differences due to the geometric complexity. The TKE profiles in Fig. E.26(b) exhibit elevated levels in the shear layer region, particularly pronounced where the flow navigates the curved surface transition. The Reynolds shear stress distributions in Fig. E.26(c) better capture the mixing layer development and its spatial evolution, though their magnitudes reflect the enhanced TKE predictions seen in all training cases.

Analysis of the skin friction evolution in Fig. E.27 confirms the model's robust performance through the geometrically-induced separation. While the baseline model better captures the double-peak structure near $x/H = 2.5$, the SL-Model achieves similar improvements to SL-Propagation in overall separation prediction, particularly in capturing the minimum skin friction value ($x/H \approx 3.5$) and recovery rate ($5 < x/H < 7.5$). The model effectively preserves upstream prediction ($x/H < 0$) where the baseline $k-\omega$ SST performs well, and shows marginally better agreement with LES data than SL-Propagation in the far wake region ($x/H > 10$). This demonstrates that the discovered correction terms successfully reproduce the benefits of propagated corrections while maintaining appropriate baseline behavior through geometric transitions.

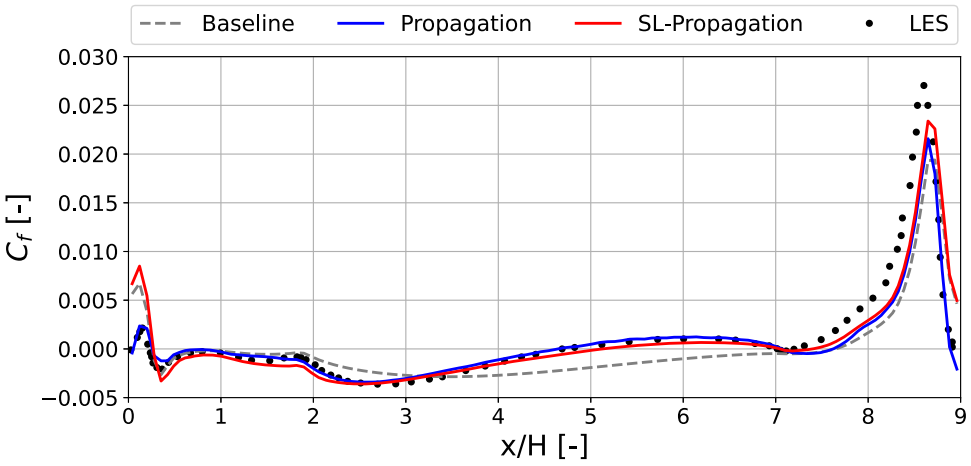
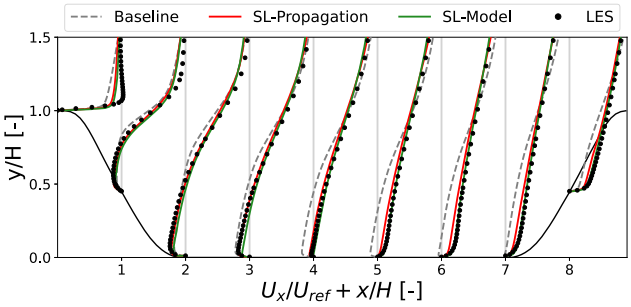
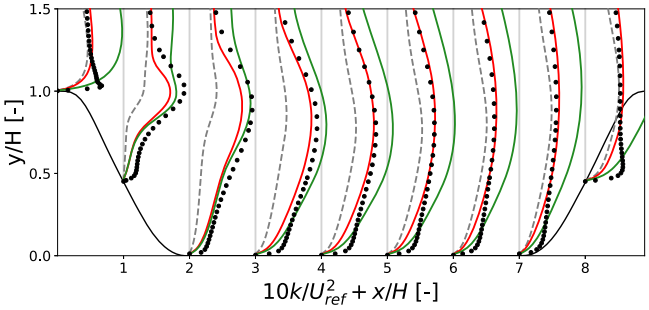


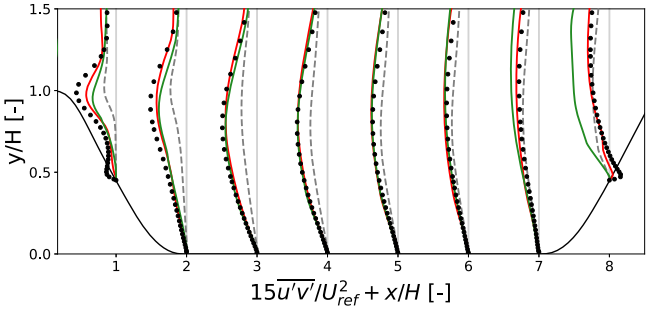
Fig. E.21. Skin friction comparison plot between full-field Propagation and SL-Propagation on the periodic hill training case.



(a) Axial velocity profiles.



(b) TKE profiles.



(c) Reynolds shear stress profiles.

Fig. E.22. Performance of the SL-Model on the periodic hill training case.

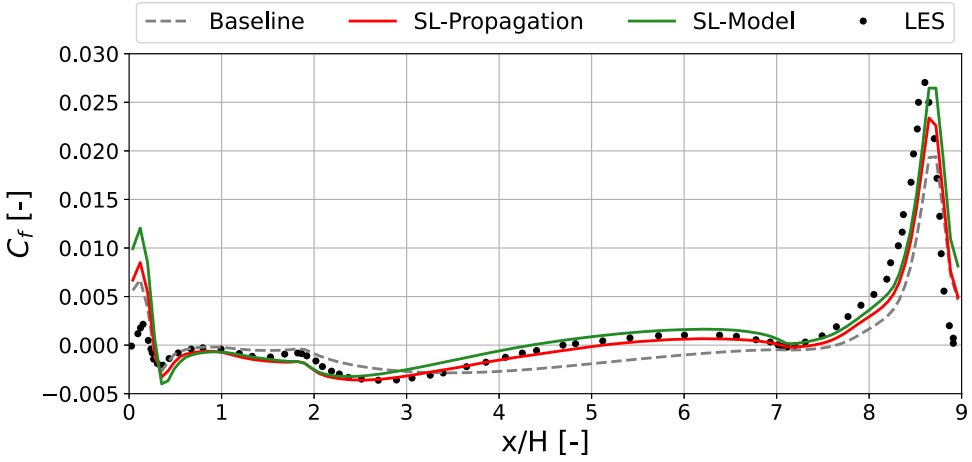
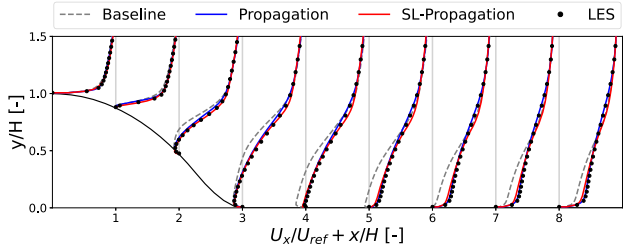
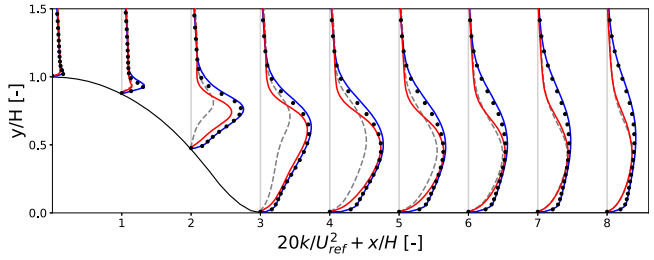


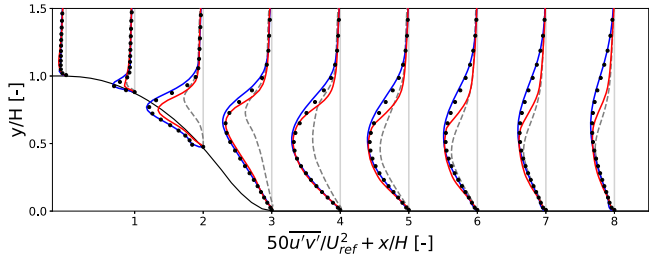
Fig. E.23. Skin friction comparison plot for the SL-Model on the periodic hill training case.



(a) Axial velocity profiles.



(b) TKE profiles.



(c) Reynolds shear stress profiles.

Fig. E.24. Performance comparison between full-field Propagation and SL-Propagation on the CBFS training case.

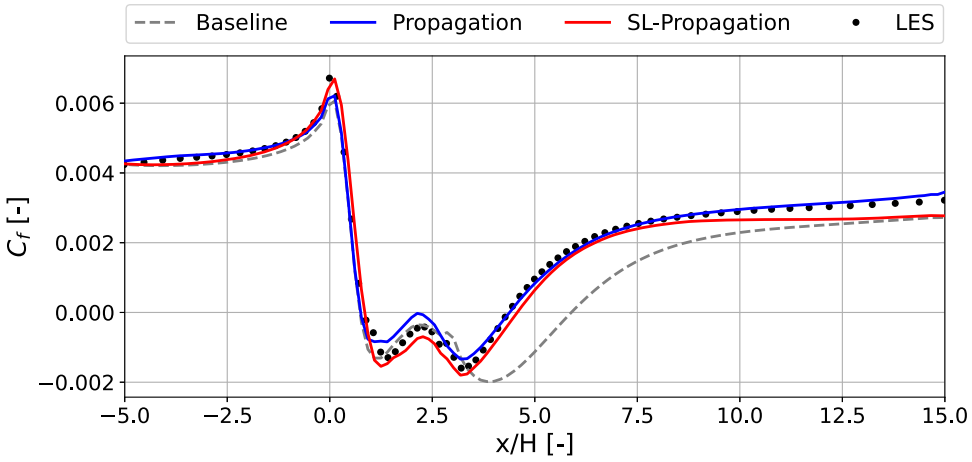
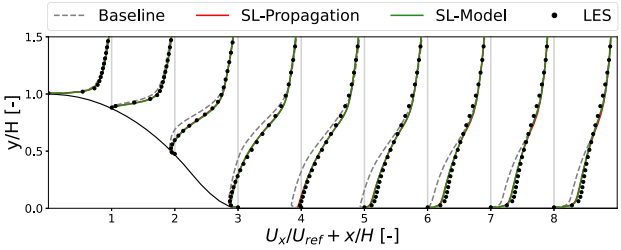
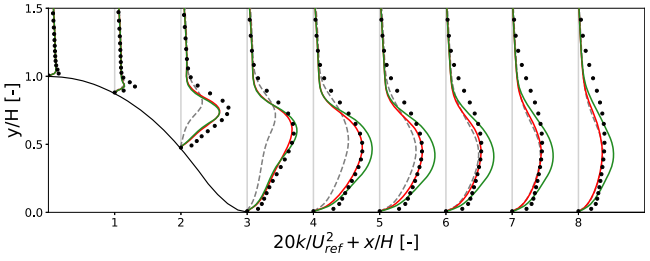


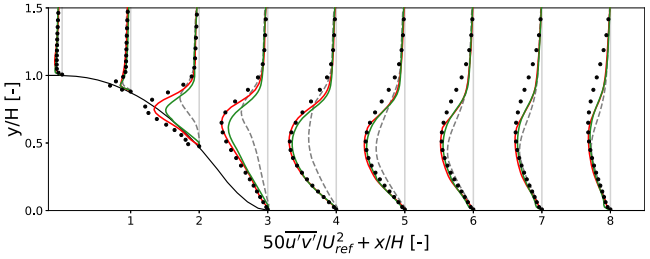
Fig. E.25. Skin friction comparison plot between full-field Propagation and SL-Propagation on the CBFS training case.



(a) Axial velocity profiles.



(b) TKE profiles.



(c) Reynolds shear stress profiles.

Fig. E.26. Performance of the SL-Model on the CBFS training case.

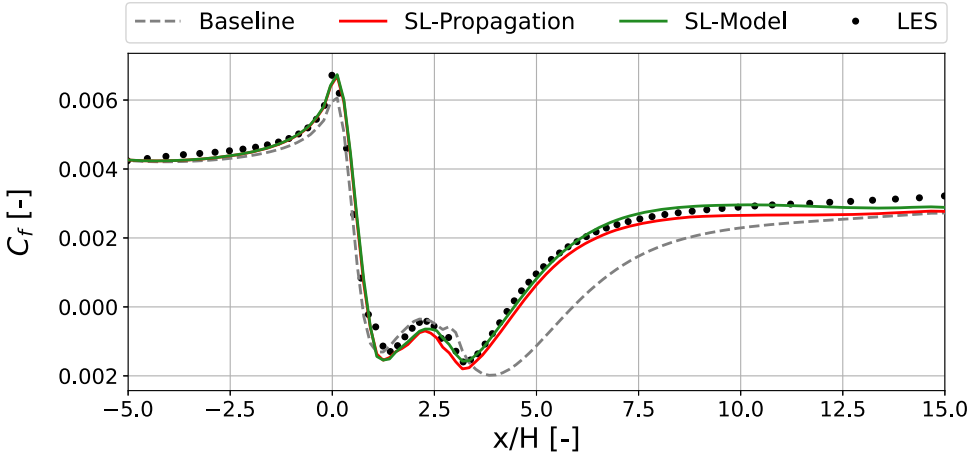


Fig. E.27. Skin friction comparison plot for the SL-Model on the CBFS training case.

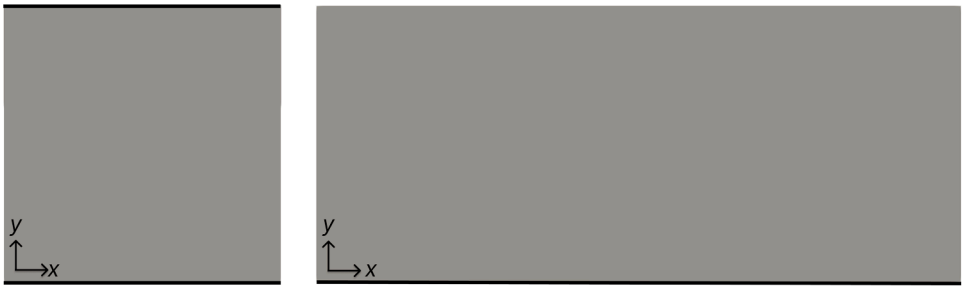
Appendix F. σ_{SL} Classifier Applied to Flat Plate and Channel Flow Cases

To verify that the σ_{SL} classifier does not activate in flow regions such as boundary layers and frees-stream, and thus preserves the accuracy of the baseline k-ω SST model in these regions, we tested the classification on a 2D zero-pressure-gradient flat plate and a fully-developed 2D channel flow. Both cases were run until statistical convergence was reached, and grid independence of the results was confirmed.

The flat plate simulation, with a length of $L = 2$ m, was run at a reference velocity of $U_{ref} = 69.4$ m/s, corresponding to a Reynolds number of $Re_L = 8.80 \times 10^6$ based on the plate length. The initial turbulence intensity was set to 0.039 %. A no-slip boundary condition was applied along the bottom plate.

The channel flow simulation used a reference bulk velocity of $U_{ref} = 65.8$ m/s. The Reynolds number in this case is $Re_H = 7.54 \times 10^7$, based on the channel height of $H = 17.3$ m. No-slip boundary conditions were applied at the top and bottom channel walls, while periodic boundary conditions were imposed at the inlet and outlet.

As shown in Fig. F.28, the σ_{SL} classifier does not activate in either case. This means that the SL-Model corrections remain fully inactive throughout the flow domain, ensuring that the baseline k-ω SST model is applied consistently across the entire computational domain.



(a) Channel flow. (b) Flat plate.

Fig. F.28. Classification outcome of the σ_{SL} classifier on the flat plate and channel flow cases: gray region: $\sigma_{SL} = 0$.

References

[1] Duraisamy K, Iaccarino G, Xiao H. Turbulence modeling in the age of data. Annu Rev Fluid Mech 2019;51. https://doi.org/10.1146/annurev-fluid-010518-040547

[2] Coleman GN, Garbaruk A, Spalart PR. Direct numerical simulation, theories and modelling of wall turbulence with a range of pressure gradients. Flow Turbul Combust 2015;95(2):261–76. https://doi.org/10.1007/s10494-015-9621-5

[3] Spalart PR, Watmuff JH. Experimental and numerical study of a turbulent boundary layer with pressure gradients. J Fluid Mech 1993;249:337–71. https://doi.org/10.1017/S002211209300120X

[4] Wilcox DC. Turbulence modeling for CFD. La Cañada, Calif: DCW Industries; 3rd ed ed.; 2006. ISBN 9781928729082.

[5] Schmitt FG. About Boussinesq's turbulent viscosity hypothesis: historical remarks and a direct evaluation of its validity. C R Méc 2007;335(9–10):617–27. https://doi.org/10.1016/j.crme.2007.08.004

[6] Rumsey CL. Exploring a method for improving turbulent separated-flow predictions with kappa-omega models. NASA Technical Memorandum NASA/TM-2009-215952; NASA Langley Research Center; Hampton, VA, United States; 2009. Work of the U.S. Government. Public Use Permitted.; https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20090042511.pdf.

[7] Li H, Zhang Y, Chen H. Aerodynamic prediction of iced airfoils based on modified three-equation turbulence model. AIAA J 2020;58(9):3863–76. Publisher: American Institute of Aeronautics and Astronautics; https://doi.org/10.2514/1.J059206

[8] Platteeuw PDA, Loeven GJA, Bijl H. Uncertainty quantification applied to the k-ϵ model of turbulence using the probabilistic collocation method. In: 49th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference. Schaumburg, IL: American Institute of Aeronautics and Astronautics. ISBN 978-1-60086-993-8; 2008, https://doi.org/10.2514/6.2008-2150

[9] Cheung SH, Oliver TA, Prudencio EE, Prudhomme S, Moser RD. Bayesian uncertainty analysis with applications to turbulence modeling. Quantif Margins Uncertainties 2011;96(9):1137–49. https://doi.org/10.1016/j.ress.2010.09.013

[10] Edeling WN, Cinnella P, Dwight RP, Bijl H. Bayesian estimates of parameter variability in the k-ϵ turbulence model. J Comput Phys 2014;258:73–94. https://doi.org/10.1016/j.jcp.2013.10.027

[11] Ray J, Lefantzi S, Arunajatesan S, Dechant L. Bayesian parameter estimation of a k-ϵ model for accurate jet-in-crossflow simulations. AIAA J 2016;54(8):2432–48. https://doi.org/10.2514/1.J054758

[12] Tracey BD, Duraisamy K, Alonso JJ. A machine learning strategy to assist turbulence model development. In: 53rd AIAA aerospace sciences meeting. Kissimmee, Florida: American Institute of Aeronautics and Astronautics. ISBN 9781624103438; 2015. https://doi.org/10.2514/6.2015-1287

[13] Gamahara M, Hattori Y. Searching for turbulence models by artificial neural network. Phys Rev Fluids 2017;2(5):054604. https://doi.org/10.1103/PhysRevFluids.2.054604

[14] Maulik R, San O. A neural network approach for the blind deconvolution of turbulent flows. J Fluid Mech 2017;831:151–81. Edition: 2017/10/13 Publisher: Cambridge University Press; https://doi.org/10.1017/jfm.2017.637

[15] Ling J, Kurzawski A, Templeton J. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. J Fluid Mech 2016;807:155–66. Edition: 2016/10/18 Publisher: Cambridge University Press. https://doi.org/10.1017/jfm.2016.615

[16] Weatheritt J, Sandberg RD. The development of algebraic stress models using a novel evolutionary algorithm. Int J Heat Fluid Flow 2017;68:298–318. https://doi.org/10.1016/j.ijheatfluidflow.2017.09.017

[17] Schmelzer M, Dwight RP, Cinnella P. Discovery of algebraic Reynolds-stress models using sparse symbolic regression. Flow Turbul Combust 2019;104(2–3):579–603. https://doi.org/10.1007/s10494-019-00089-x

[18] Cherroud S, Merle X, Cinnella P, Gloerfelt X. Sparse Bayesian learning of explicit algebraic Reynolds-stress models for turbulent separated flows. Int J Heat Fluid Flow 2022;98:109047. https://doi.org/10.1016/j.ijheatfluidflow.2022.109047

[19] Srivastava V, Rumsey CL, Coleman GN, Wang L. On generalizably improving RANS predictions of flow separation and reattachment. In: AIAA SCITECH 2024 Forum. Orlando, FL: American Institute of Aeronautics and Astronautics. ISBN 978-1-62410-711-5; 2024, https://doi.org/10.2514/6.2024-2520

[20] Wu C, Zhang S, Zhang Y. Development of a generalizable data-driven turbulence model: conditioned field inversion and symbolic regression. AIAA J 2024;:1–20 https://doi.org/10.2514/1.J064416

[21] Nishi Y, Krumbein A, Knopp T, Probst A, Grabe C. On the generalization capability of a data-driven turbulence model by field inversion and machine learning. Aerospace 2024;11(7):592. https://doi.org/10.3390/aerospace11070592

[22] Mandler H, Weigand B. Generalization limits of data-driven turbulence models. Flow Turbul Combust 2024; https://doi.org/10.1007/s10494-024-00595-7

[23] Ho J, Pepper N, Dodwell T. Probabilistic machine learning to improve generalisation of data-driven turbulence modelling. Comput Fluids 2024;284:106443. https://doi.org/10.1016/j.compfluid.2024.106443

[24] Callaham JL, Koch JV, Brunton BW, Kutz JN, Brunton SL. Learning dominant physical processes with data-driven balance models. Nat Commun 2021;12(1):1016. https://doi.org/10.1038/s41467-021-21331-z

[25] Wu J, Xiao H, Sun R, Wang Q. Reynolds-averaged Navier-Stokes equations with explicit data-driven Reynolds stress closure can be ill-conditioned. J Fluid Mech 2019;869:553–86. https://doi.org/10.1017/jfm.2019.205

[26] Seifert A, Pack LG. Active flow separation control on wall-mounted hump at high Reynolds numbers. AIAA J 2002;40(7):1363–72. https://doi.org/10.2514/2.1796

[27] Pope SB. Turbulent. Cambridge University Press; 1 ed.; 2000. https://doi.org/10.1017/CBO9780511840531

[28] Callaham JL, Koch JV, Brunton BW, Kutz JN, Brunton SL. Learning dominant physical processes with data-driven balance models. Nat Commun 2021;12:1016. https://doi.org/10.1038/s41467-021-21331-z

[29] Lăcătuș M. Improving data-driven RANS turbulence modelling for separated flow scenarios. Master's thesis; Delft University of Technology; 2024. https://resolver.tudelft.nl/uuid:755d3b8c-78d4-4926-9b1a-7516ef8b00e0.

[30] Pope SB. A more general effective-viscosity hypothesis. J Fluid Mech 1975;72(02):331. https://doi.org/10.1017/S0022112075003382

[31] Wang J-X, Wu J-L, Xiao H. Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data. Phys Rev Fluids 2017;2(3):034603. https://doi.org/10.1103/PhysRevFluids.2.034603

[32] McConaghy T. FFX: fast, scalable, deterministic symbolic regression technology. In: Riolo R, Vladislavleva E, Moore JH, editors. Genetic programming theory and practice IX. New York, NY: Springer New York. 2011, p. 235–60. https://doi.org/10.1007/978-1-4614-1770-5_13

[33] Zou H, Hastie T. Regularization and variable selection via the elastic Net. J R Stat Soc Ser B Stat Methodol 2005;67(2):301–20. https://doi.org/10.1111/j.1467-9868.2005.00503.x

[34] Sotgiu C, Weigand B, Semmler K. A turbulent heat flux prediction framework based on tensor representation theory and machine learning. Int Commun Heat Mass Transfer 2018;95:74–9. https://doi.org/10.1016/j.icheatmasstransfer.2018.04.005

[35] Duraisamy K. Perspectives on machine learning-augmented Reynolds-averaged and large eddy simulation models of turbulence. Phys Rev Fluids 2021;6(5):050504. https://doi.org/10.1103/PhysRevFluids.6.050504

[36] Kurz M, Beck A. Investigating model-data inconsistency in data-informed turbulence closure terms. In: 14th WCCM-ECCOMAS congress. CIMNE; 2021 https://doi.org/10.23967/wccm-eccomas.2020.115

[37] Thompson RL, Sampaio L EB, De Bragança Alves F AV, Thais L, Mompean G. A methodology to evaluate statistical errors in DNS data of plane channel flows. Comput Fluids 2016;130:1–7. https://doi.org/10.1016/j.compfluid.2016.01.014

[38] Melchers H, Crommelin D, Koren B, Menkovski V, Sanderse B. Comparison of neural closure models for discretised PDEs. Comput Math Appl 2023;143:94–107. https://doi.org/10.1016/j.camwa.2023.04.030

[39] Uzun A, Malik MR. Wall-resolved large-Eddy simulation of flow separation over NASA wall-mounted hump. In: 55th AIAA aerospace sciences meeting. Grapevine, Texas: American Institute of Aeronautics and Astronautics. ISBN 978-1-62410-447-3; 2017. https://doi.org/10.2514/6.2017-0538

[40] Breuer M, Peller N, Rapp C, Manhart M. Flow over periodic hills - numerical and experimental study in a wide range of Reynolds numbers. Comput Fluids 2009;38(2):433–57. https://doi.org/10.1016/j.compfluid.2008.05.002

[41] Bentaleb Y, Lardeau S, Leschziner MA. Large-eddy simulation of turbulent boundary layer separation from a rounded step. J Turbul 2012;13:N4. https://doi.org/10.1080/14685248.2011.637923

[42] Driver DM, Seegmiller HL. Features of a reattaching turbulent shear layer in divergent channelflow. AIAA J 1985;23(2):163–71. https://doi.org/10.2514/3.8890

[43] Xiao H, Wu J-L, Laizet S, Duan L. Flows over periodic hills of parameterized geometries: a dataset for data-driven turbulence modeling from direct simulations. Comput Fluids 2020;200:104431. https://doi.org/10.1016/j.compfluid.2020.104431

[44] Bell J, Heineck J, Zilliac G, Mehta R, Long K. Surface and flow field measurements on the FAITH hill model. In: 50th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition. Nashville, Tennessee: American Institute of Aeronautics and Astronautics. ISBN 978-1-60086-936-5; 2012, https://doi.org/10.2514/6.2012-704

[45] Lienhart H, Stoots C, Becker S. Flow and turbulence structures in the wake of a simplified car model (ahmed modell). In: Wagner S, Rist U, Heinemann H-J, Hilbig R, editors. New results in numerical and experimental fluid mechanics III. Berlin, Heidelberg: Springer Berlin Heidelberg. 2002, p. 323–30. https://doi.org/10.1007/978-3-540-45466-3_39

[46] Zhang Y, Dwight RP, Schmelzer M, Gómez JF, Han Z-h, Hickel S. Customized data-driven RANS closures for Bi-Fidelity LES-RANS optimization. J Comput Phys 2021;432:110153. https://doi.org/10.1016/j.jcp.2021.110153

[47] Rumsey CL, Coleman GN. Nasa symposium on turbulence modeling: roadblocks, and the potential for machine learning. NASA Technical Memorandum NASA/TM-20220015595; NASA Langley Research Center; Hampton, VA, United States; 2022. Work of the U.S. Government. Public use permitted.; https://ntrs.nasa.gov/api/citations/20220015595/downloads/NASA-TM-20220015595final.pdf.

[48] Greenblatt D, Paschal KB, Yao C-S, Harris J, Schaeffler NW, Washburn AE. Experimental investigation of separation control part 1: baseline and steady suction. AIAA J 2006;44(12):2820–30. https://doi.org/10.2514/1.13817

[49] Hoefnagel K. Multi-flow generalization in data-driven turbulence modeling: an exploratory study. Master's thesis; Delft University of Technology; Delft, Netherlands; 2023. https://resolver.tudelft.nl/uuid:324d4b2d-bf58-40a0-b60d-4e2e0b992797.

[50] Ho J, West A. Field inversion and machine learning for turbulence modelling applied to three-dimensional separated flows. In: AIAA AVIATION 2021 FORUM. VIRTUAL EVENT: American Institute of Aeronautics and Astronautics. ISBN 978-1-62410-610-1; 2021 https://doi.org/10.2514/6.2021-2903