Region-Aware Hybrid LSTM-GRU Modeling for Wake Flows Induced by Offshore Wind-Turbine Foundations

Results in Engineering 29 (2026) 109518

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Region-aware hybrid LSTM-GRU modeling for wake flows induced by offshore wind-turbine foundations

Yifan Ma [a], Zhaoyue Xu [b,*], Hua-Dong Yao [b]

[a] Department of Materials Science and Engineering, Ångström Laboratory, Uppsala University, Uppsala, SE-75103, Sweden
[b] Division of Marine Technology, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, SE-41296, Sweden

ARTICLE INFO

Keywords:
Wake flow modeling
Region-aware learning
Proper orthogonal decomposition
LSTM-GRU networks
Grad-CAM

ABSTRACT

The wake dynamics downstream of offshore wind turbine foundations exhibit strong nonlinearity, temporal variability, and three-dimensional disturbances. Accurately and efficiently modeling these unsteady flows has long been a challenge. Traditional recurrent neural network (RNN) methods often struggle to capture localized, nonstationary wake structures due to limited spatial awareness and sensitivity to variations in input dimensionality. This study presents a region-aware hybrid modeling framework to address these limitations. The approach integrates three key components: automatic wake-region detection, Proper Orthogonal Decomposition (POD) for dimensionality reduction, and a hybrid LSTM-GRU network for temporal sequence modeling. A CNN enhanced with Grad-CAM identifies dynamically active wake regions. POD is applied separately to both the global domain and the extracted wake subdomain to derive dominant modal coefficients. A dual-branch temporal prediction model then forecasts future low-dimensional flow representations for both domains. The training process employs a composite loss function combining time-domain mean squared error (MSE) with frequency-domain Fourier loss. To prioritize key unsteady features, higher weighting is applied to wake-region predictions. Results show that the proposed model effectively captures key wake dynamics, including vortex shedding patterns, streamwise wake decay, and localized disturbances. For full-field predictions, the mean absolute relative error (MARE) is reduced from 15.6% for the conventional global POD with a standard LSTM baseline to 12.2% for the final blended prediction of the proposed region-aware hybrid framework, indicating a substantial improvement in predictive accuracy. The applicability of the proposed approach to different Reynolds numbers, alternative foundation geometries, and unsteady tidal inflow conditions remains to be investigated.

1. Introduction

The rapid expansion of offshore wind turbines has raised increasing concerns regarding the environmental impacts of their foundation structures, which can significantly modify local hydrodynamic conditions [1,2]. In seasonally stratified waters, submerged foundations generate underwater wakes characterized by vortex shedding and localized shear layers. Such disturbances influence sediment transport, nutrient redistribution, and near-bottom mixing processes, potentially affecting benthic ecosystems and broader marine ecological dynamics [1,3]. Although extensive research has focused on atmospheric wakes induced by rotor-atmosphere interactions [4–6], hydrodynamic wakes generated by submerged wind turbine foundations remain comparatively underexplored. Addressing this gap is essential for the environmentally sustainable development of offshore wind energy systems.

High-fidelity computational fluid dynamics (CFD) provides detailed insight into wake dynamics but entails substantial computational cost, particularly for resolving unsteady three-dimensional flows around offshore foundations [4,7]. To alleviate this burden, reduced-order modeling (ROM) techniques have been developed to capture dominant flow structures at significantly lower cost [8,9]. Among these approaches, Proper Orthogonal Decomposition (POD) is widely used to extract energy-dominant spatial modes from high-fidelity simulations [10]. VerHulst et al. [11] demonstrated that a limited number of POD modes effectively capture vertical kinetic energy transport in wind farms. Bastin et al. [12] showed that low-order POD modes reconstruct thrust and power with reasonable accuracy, though energy-ranked modes are not always optimal for dynamic load prediction. Hamilton et al. [13] further developed a data-driven POD-based ROM in which modal coefficients were evolved using polynomial ODE systems with regularization and inflow reinitialization. Collectively, these studies establish POD-based

* Corresponding author.
E-mail addresses: yifanma_aero@outlook.com (Y. Ma), zhaoyue@chalmers.se (Z. Xu), huadong.yao@chalmers.se (H.-D. Yao).

https://doi.org/10.1016/j.rineng.2026.109518
Received 27 November 2025; Received in revised form 8 February 2026; Accepted 8 February 2026
Available online 13 February 2026

ROM as a versatile and computationally efficient framework for wake analysis and wind-energy applications.

Recent advances in machine learning have further enhanced reduced-order and surrogate modeling in fluid dynamics [14,15]. Long Short-Term Memory (LSTM) networks have shown particular promise in learning nonlinear temporal evolution from reduced-order representations [16]. Nazvanova et al. [17] employed a one-LSTM-per-mode strategy for vortex-induced vibration modeling, achieving high accuracy and significant computational savings. Raj et al. [18] demonstrated that LSTM-based ROMs outperform linear approaches such as DMD for chaotic flow prediction. Mashhadi et al. [19] introduced a mode-pairing LSTM training strategy to mitigate high-frequency interference in complex 3D wake fields. Yousif et al. [20] compared LSTM and BiLSTM architectures for POD-coefficient prediction, highlighting architectural sensitivity in wake modeling. Li et al. [21] and Hasegawa et al. [22] incorporated CNN encoders and attention mechanisms into LSTM-based ROMs, demonstrating improved spatiotemporal prediction accuracy and robustness across varying Reynolds numbers. These developments indicate a clear progression toward more expressive hybrid architectures while maintaining computational efficiency.

Despite these advances, most LSTM-based reduced-order models operate on global flow domains or predefined spatial supports, lacking explicit spatial adaptivity. Consequently, localized wake structures, particularly those generated near turbine foundations, may be underrepresented. Even when CNN encoders or attention mechanisms are employed, the absence of explicit region awareness allows dynamically inactive far-field regions to dominate the representation, limiting modeling efficiency and physical interpretability.

To enhance spatial adaptivity, gradient-based saliency techniques have recently been explored in CFD. Grad-CAM, originally developed for computer vision [23], generates gradient-weighted activation maps from convolutional layers to highlight regions most influential to model predictions without altering network architecture. Morimoto et al. [24] and Xu et al. [25] applied Grad-CAM to sensor placement and flow reconstruction in complex geometries, including periodic-hill and axisymmetric-body flows. Jagodinski et al. [26] and Ajaya et al. [27] further demonstrated that Grad-CAM can improve wake-region modeling through multi-layer saliency fusion and data rebalancing strategies. These studies confirm Grad-CAM's potential for physics-informed region identification in fluid flows.

Grad-CAM has also been coupled with recurrent networks; however, existing applications primarily focus on flow-regime classification rather than state prediction. For example, Wen et al. [28] integrated Grad-CAM with LSTM networks to localize salient buffet structures and classify subsonic buffet regimes over high-angle-of-attack airfoils. While demonstrating interpretability benefits, systematic integration of Grad-CAM into reduced-order state prediction frameworks remains limited.

Motivated by these limitations, this study proposes a region-aware hybrid LSTM framework that integrates wake-region identification, reduced-order modeling, and hybrid temporal learning to improve the accuracy and efficiency of unsteady wake prediction. By explicitly separating dynamically significant wake regions from the global flow field and modeling them in parallel, the framework enhances localized representation while preserving full-field coherence.

The overall framework, illustrated in Fig. 1, consists of four main stages:

Step 1: Wake Region Identification: A Grad-CAM-based mechanism extracts a statistically dominant wake subdomain from the full computational domain.

Step 2: Flow Decomposition: The global domain and wake subdomain are independently standardized and decomposed using POD to obtain compact reduced-order representations.

Step 3: Temporal Prediction: The evolution of the POD coefficients is modeled using a hybrid LSTM-GRU architecture to capture both long-term periodic behavior and short-term residual dynamics.

Step 4: Flow Reconstruction and Blending: Predicted coefficients are projected back to physical space via inverse POD, and the reconstructed global and wake fields are combined using a wake-global blending strategy to ensure smooth and physically consistent full-field predictions.

The research ultimately aims to support wind farm siting optimization, foundation design, and environmental impact evaluations.

2. Method

2.1. CFD simulation setup and data acquisition

The flow configuration consists of a circular cylinder representing a monopile foundation submerged in water. The cylinder diameter is $D = 8$ m. The incoming freestream velocity is $(u_\infty, v_\infty, w_\infty) = (0.5, 0, 0)$ m/s. The fluid is assumed incompressible with density $\rho = 997.561\ \mathrm{kg/m^3}$. The corresponding Reynolds number based on cylinder diameter is approximately $Re = 4 \times 10^6$, indicating a fully turbulent separated wake regime.

The computational domain is shown in Fig. 2. Its dimensions are $(63.5D \times 53D \times 6.25D)$ in the streamwise (x), spanwise (y), and vertical (z) directions. The dimensions of the computational domain have been validated in preliminary studies. The cylinder is positioned at the mid-span of the domain and centered at $(x, y) = (0, 0)$ m, with $7.25D$ upstream to the inlet and $56.25D$ downstream to the outlet.

The boundary condition at the inlet is the velocity inlet condition that has a uniform distribution of the freestream velocity. The outlet is set with the pressure-outlet condition. The lateral and top boundaries are modeled as symmetry planes, and the cylinder surface is imposed with a no-slip wall condition.

The flow is solved using STAR-CCM+ based on the finite volume method. The improved delayed detached eddy simulation (IDDES) model is employed for turbulence closure [29]. The IDDES formulation is based on the Spalart-Allmaras (SA) RANS model in the near-wall region, coupled with LES treatment in separated wake regions. Standard wall functions are used to model the near-wall behavior. This hybrid RANS-LES approach enables efficient resolution of large-scale turbulent structures in the separated wake while maintaining computational feasibility.

A segregated pressure-velocity coupling scheme using the SIMPLE algorithm is adopted. Spatial discretization is second-order accurate, and temporal integration employs a second-order implicit scheme. The time step is set to $\Delta t = 0.5$ s, ensuring that the average Courant number remains below one throughout the domain. The simulation is run for a total physical time of 4000 s, providing sufficiently long time histories to capture the development and statistical characteristics of the turbulent wake.

The unstructured mesh generation employs hybrid trimmed meshing and near-wall prism layers. The mesh refinement follows a multi-region strategy focused on capturing the near-field and wake flow. A mesh sensitivity study had been conducted in previous work. A cut plane of the mesh chosen for the simulation is shown in Fig. 3, where the plane is located in the middle of the domain height. The total number of nodes in the mesh is approximately 10.3×10^6. Further details regarding mesh validation, near-wall resolution, and data fidelity are provided in A.

The solver outputs were post-processed to extract the instantaneous three-dimensional velocity fields recorded at regular time intervals to capture the unsteady wake dynamics. A horizontal cross-section at $z = 4.7D$ was selected for further analysis. The simulation produces 8000 snapshots. Each snapshot was saved on a fixed spatial grid, enabling time-resolved modeling of wake evolution. To balance computational

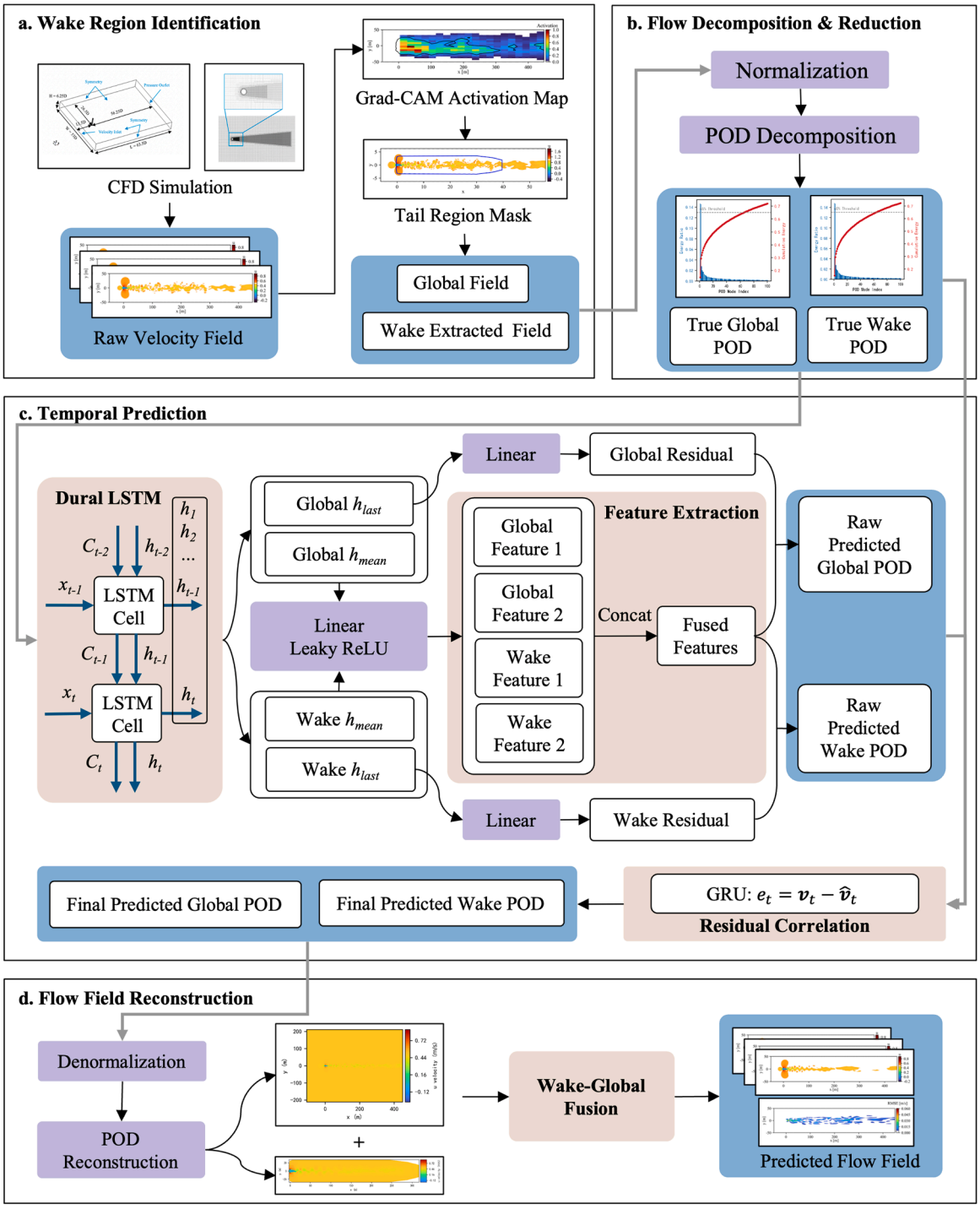

Fig. 1. Overall framework of the region-aware hybrid LSTM-GRU modeling for wake flow prediction.

efficiency with learning accuracy, the dataset was temporally subsampled by selecting every eighth snapshot, resulting in a temporal resolution of 4 s per frame. This procedure yields 1000 effective snapshots. To eliminate initial transient effects, the first 25 snapshots were discarded, leaving 975 samples for subsequent reduced-order modeling and training. Given that the estimated vortex shedding period is approximately 80 s, this sampling strategy provides roughly 20 data points per shedding cycle, which is sufficient to resolve the dominant temporal dynamics of the unsteady wake.

2.2. Wake region identification

To identify spatial regions that contribute most strongly to the overall unsteady wake dynamics, a data-driven wake localization strategy based on convolutional neural networks (CNNs) and gradient-weighted activation mapping (Grad-CAM) is employed. The objective is to extract a statistically dominant wake region that is most relevant to the global unsteadiness of the flow, thereby enabling region-focused ROM in subsequent stages.

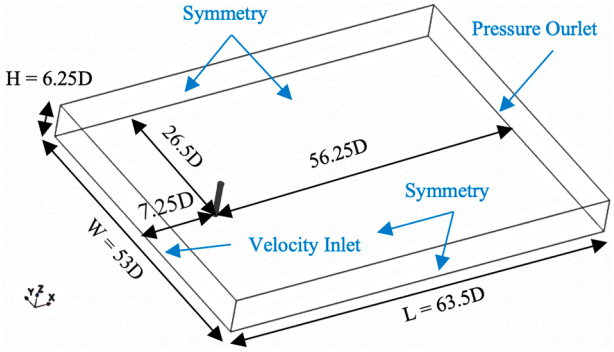

Fig. 2. Setup of the computational domain in the CFD simulation.

The raw CFD output consists of unstructured velocity vectors (u, v, w) defined on a spatial mesh. To enable CNN-based feature extraction, the three velocity components from a horizontal two-dimensional slice are interpolated onto a regular Cartesian grid. The grid resolution is automatically determined according to the physical aspect ratio of the computational domain to preserve spatial consistency. Each instantaneous snapshot is represented as a multi-channel image of size $(H, W, 3)$. Here, H and W denote the spatial resolution of the interpolated Cartesian grid in the two in-plane directions, and each channel corresponds to one velocity component, yielding a dataset of T time-resolved input images.

To characterize the overall level of flow unsteadiness, a fluctuation-intensity field is first computed at each spatial location as

$$\sigma(x, y) = \sqrt{\frac{1}{N_t} \sum_{t=1}^{N_t} \left(|\mathbf{u}(x, y, t)| - \overline{|\mathbf{u}(x, y)|} \right)^2}, \tag{1}$$

where $|\mathbf{u}(x, y, t)|$ denotes the instantaneous velocity magnitude and the overbar indicates temporal averaging. Rather than performing pixel-wise regression, this spatial field is aggregated into a single scalar quantity,

$$\sigma_{\text{global}} = \frac{\sum_{i=1}^{N} w_i \, \sigma(x_i, y_i)}{\sum_{i=1}^{N} w_i}, \tag{2}$$

which represents a global indicator of flow unsteadiness for each snapshot. Here, w_i denotes the area weight associated with spatial point i; for a uniform Cartesian grid, this expression reduces to a simple arithmetic mean. This scalar quantity, denoted as σ_{global}, serves as the regression target for training the CNN.

For numerical stability and improved training efficiency, the regression target is linearly rescaled by a constant factor (600 in the present study), such that its numerical magnitude is comparable to typical image-based learning targets. This rescaling improves optimization stability and does not alter the relative gradients that underpin the subsequent Grad-CAM analysis.

The CNN architecture is tailored to the anisotropic characteristics of wake-dominated flows. Specifically, the network consists of a sequence of convolutional layers with elongated kernels of size (3×1), designed to enhance sensitivity to cross-stream variations while preserving streamwise resolution. Unidirectional max-pooling layers with a pooling size of (1×2) and a corresponding stride of (1×2) are applied after selected convolutional blocks, reducing resolution only in the streamwise direction while fully preserving cross-stream spatial information. Three convolutional blocks with 32, 64, and 128 filters are employed, interleaved with unidirectional max-pooling layers applied only in the streamwise direction. The convolutional backbone is followed by a fully connected layer with 128 neurons and a linear output layer. The network is trained using a mean squared error loss function and the Adam optimizer.

After training, Grad-CAM is applied by computing the gradients of the scalar regression output σ_{global} with respect to the feature maps of the final convolutional layer. Although the Grad-CAM heatmap is computed for an individual instantaneous snapshot, the CNN itself is trained to infer a time-aggregated unsteadiness indicator from a large ensemble of snapshots. As a result, the Grad-CAM highlights spatial regions that are consistently informative for predicting the same global statistical measure of wake unsteadiness, rather than instantaneous flow features.

To define a compact wake subdomain, the Grad-CAM heatmaps are thresholded using a fixed value of 0.2 and post-processed to extract the largest connected region.

In the present study, the resulting wake mask is treated as static and is not updated at each time step. Accordingly, the term “automatic wake-region detection” refers to a data-driven identification of a statistically dominant wake region based on learned flow features, rather than instantaneous, frame-by-frame segmentation of the flow field.

2.3. Flow decomposition and subdomain extraction

To extract dominant coherent structures and enable efficient learning of temporal dynamics, POD is applied to the simulated velocity fields. Prior to POD, each velocity component (u, v, w) is standardized by removing the temporal mean per spatial location. Then, a snapshot matrix $\widetilde{\mathbf{U}} \in \mathbb{R}^{N_t \times 3N}$ is assembled by stacking the three components from a horizontal slice with N spatial points over several time steps. A truncated singular value decomposition (SVD) yields:

$$\widetilde{\mathbf{U}} \approx \mathbf{V}\mathbf{\Phi}. \tag{3}$$

Here $\mathbf{V} \in \mathbb{R}^{N_t \times r}$ contains the temporal POD coefficients, and $\mathbf{\Phi} \in \mathbb{R}^{r \times 3N}$ holds the leading r spatial POD modes chosen based on an energy criterion. The POD is performed in two spatial contexts: the global full domain and the local subdomain of the wake identified with Grad-CAM.

2.4. Hybrid LSTM-GRU for temporal prediction

The temporal evolution of flow structures, which are captured by the reduced-order POD representation, is modeled with a hybrid recurrent network built from LSTM-GRU blocks. The model learns the temporal dynamics of POD coefficients from both global domain and wake subdomain. A fusion stage integrates these streams to improve accuracy and interpretability.

2.4.1. Input construction

The temporal predictor ingests two synchronized streams of POD coefficients: one from the global domain and one from the wake subdomain. To capture temporal dependencies, a sliding window approach is utilized, with a lookback horizon of 30 frames (120 s), covering approximately one vortex-shedding period. After subsampling and discarding the initial transient segment, a total of 975 frames are retained for modeling. The dataset is then split chronologically into training and validation sets at an 80%–20% ratio, corresponding to 780 frames for training and 195 frames for validation. By applying the same lookback across both streams, the model is trained to exploit temporal dependencies at two spatial scales. This dual-resolution input strategy, followed by a dedicated fusion stage, improves the prediction of multi-scale dynamics in wake evolution.

2.4.2. Network architecture

As illustrated in Fig. 1(c), the proposed fusion network processes two sequences of POD coefficients corresponding to the global domain and the wake subdomain using two independent three-layer LSTM encoders. Both encoders share the same architecture and hidden dimension $H = 64$. Given an input sequence of length L, each encoder produces a hidden-state tensor of size (B, L, H), where the batch size is fixed to $B = 256$, L denotes the sequence length, and H the LSTM hidden dimension.

From each branch, two complementary temporal descriptors are extracted. The final hidden state $\mathbf{h}_{\text{last}} \in \mathbb{R}^{256 \times 64}$ captures short-term temporal information, while the mean hidden state over the sequence $\mathbf{h}_{\text{mean}} \in$

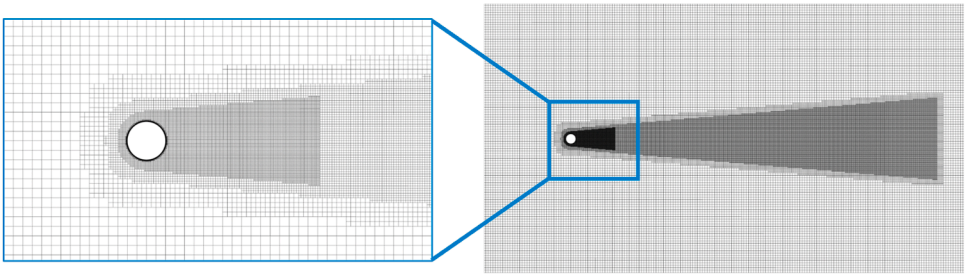

Fig. 3. The mesh for the CFD simulation.

$\mathbb{R}^{256\times64}$ captures long-term temporal trends. Each descriptor is passed through a fully connected layer followed by a LeakyReLU activation (negative slope 0.1), yielding four feature vectors in total: two from the global branch and two from the wake branch. All four feature vectors have dimension $(256, 64)$.

The four feature vectors are concatenated along the feature dimension to form a combined latent representation of size $(256, 256)$. This representation is processed by a two-layer fusion block. The first fusion layer maps $(256, 256)$ to $(256, 128)$, followed by a LeakyReLU activation and a dropout layer with rate 0.3. The second fusion layer reduces the dimensionality from $(256, 128)$ to $(256, 64)$, producing a compact fused latent representation that integrates information from both spatial domains and temporal scales.

To enhance stability and preserve branch-specific temporal information, residual connections are introduced from the final hidden state of each LSTM branch. Specifically, the last hidden state from the global branch and the wake branch is projected through separate linear layers of size $(256, 64) \rightarrow (256, 64)$ and added element-wise to the fused latent representation. The resulting residual-enhanced features are then passed through two separate output heads. The global output head maps $(256, 64)$ to $(256, 40)$ to predict the next-step POD coefficients for the global domain, while the wake output head maps $(256, 64)$ to $(256, 40)$ to predict the corresponding coefficients for the wake subdomain, where $r_g = r_w = 40$ denote the retained numbers of POD modes in both domains.

Finally, the raw predictions are refined using a GRU-based residual correction module with hidden dimension 64. This module learns from the temporal discrepancy between the raw predictions and the reference POD coefficients and outputs corrected POD coefficient sequences for both branches. This modular design enables the model to exploit both global dynamics and localized wake features while providing additional temporal refinement and error correction.

2.4.3. Learning strategy

To guide training, a composite loss function is formulated to jointly capture the temporal accuracy and spectral fidelity of the predicted POD coefficients. The total loss L_{total} consists of two components: time-domain loss, measured by mean squared error (MSE), and frequency-domain loss, computed on the discrete Fourier transform to preserve temporal dynamics of magnitude and phase. These two components are linearly combined to form the final objective:

$$L_{\text{total}} = \alpha L_{\text{MSE}} + (1-\alpha) L_{\text{Fourier}}. \tag{4}$$

The weighting factor α is set to 0.9, to prioritize time-domain accuracy while still preserving the dominant spectral content.

To account for spatially distinct learning objectives, the total loss of the region-aware hybrid model is computed by combining the losses from the global global domain and local wake subdomain. Specifically, the final training objective is defined as:

$$L_{\text{final}} = (1-\beta)\cdot L_{\text{global}} + \beta \cdot L_{\text{wake}}, \tag{5}$$

where L_{global} and L_{wake} are the total losses (each computed using the hybrid formulation of MSE and Fourier loss) for the global and wake-region POD coefficients, respectively. The weighting parameter $\beta = 0.6$ emphasizes the wake-region loss, placing greater focus on learning the dynamically rich and structurally complex wake features, while still retaining contributions from the global flow field.

All models are implemented in PyTorch 2.4.1 with CUDA 12.4 (cu124). The network is trained for a maximum of 300 epochs using the Adam optimizer, with an initial learning rate of 1×10^{-3} for the fusion network ($lr_{\text{fusion}} = 1\times10^{-3}$) and a smaller learning rate for the residual-correction module ($lr_{\text{residual}} = 5\times10^{-4}$). To reduce training stochasticity and improve reproducibility, the random seed is fixed to 42. Gradient clipping is applied with a maximum norm of 1 to stabilize optimization and prevent exploding gradients. Early stopping is employed based on the validation loss with a minimum improvement threshold of 1×10^{-5}, and training is terminated when no further reduction in validation loss is observed. Throughout training, the validation loss is monitored to assess generalization performance. The final model checkpoint retains the best-performing LSTM and GRU weights corresponding to the lowest validation loss, which are subsequently used for flow-field prediction and reconstruction.

2.5. Flow-field reconstruction and blending

2.5.1. Velocity-field recovery from POD coefficients

After forecasting the final temporal POD coefficients $\widetilde{\boldsymbol{v}}_t \in \mathbb{R}^r$ for both the global domain and wake subdomain representations, the original high-dimensional velocity field can be approximately reconstructed using the previously obtained spatial POD modes. Let $\boldsymbol{\Phi} \in \mathbb{R}^{r\times3N}$ denote the truncated POD basis, consisting of the top r energy-ranked spatial modes. The reconstructed (normalized) velocity field $\widetilde{\boldsymbol{U}}_t \in \mathbb{R}^{3N}$ at time t is given by:

$$\widetilde{\boldsymbol{U}}_t = \sum_{i=1}^{r} \widetilde{v}_{t,i}\boldsymbol{\phi}_i = \widetilde{\boldsymbol{v}}_t\boldsymbol{\Phi}. \tag{6}$$

Here, $\widetilde{\boldsymbol{v}}_t = [\widetilde{v}_{t,1}, \widetilde{v}_{t,2}, \ldots, \widetilde{v}_{t,r}]$ is the POD coefficient vector at time t, and $\boldsymbol{\phi}_i \in \mathbb{R}^{3N}$ is the ith spatial mode vector. The resulting vector $\widetilde{\mathbf{U}}_t$ contains the concatenated, normalized values of the three velocity components u, v, and w across all N spatial locations.

The actual physical velocity values are recovered by inverting the standardization using the StandardScaler parameters (means and standard deviations) stored during the training phase:

$$\mathbf{U}_t = \boldsymbol{\sigma} \odot \widetilde{\mathbf{U}}_t + \boldsymbol{\mu}. \tag{7}$$

Here, $\boldsymbol{\mu}, \boldsymbol{\sigma} \in \mathbb{R}^{3N}$ denote the concatenated temporal means and standard deviations of all three velocity components (u, v, w) at each spatial point, respectively. The operator $\odot$ denotes the Hadamard (element-wise) product.

This reconstruction process involves first projecting the predicted coefficients onto the POD basis. It is followed by component-wise inverse normalization, resulting in the recovered velocity field in its original physical scale and units.

2.5.2. Region-aware flow-field blending

To balance the subdomain prediction accuracy and the full-domain coverage, a spatial fusion procedure is applied. This method selectively blends the wake-region reconstruction U^w with the global reconstruction U^g, based on geometric proximity and spatial context.

Let $U_t^g(i) \in \mathbb{R}^3$ be the velocity vector at global point i at time t, and $U_t^w(j) \in \mathbb{R}^3$ be the velocity at wake-region point j. Denote $p_i = (x_i, y_i, z_i)$ as the spatial coordinate of point i in the global domain and q_j as its nearest neighbor in the wake region, determined via KD-tree matching. The spatial fusion procedure is as follows.

1. Spatial Matching: For each global point p_i located within the bounding box of the identified wake subdomain, find its nearest neighbor q_j.

$$j = \arg\min_k \left\| \mathbf{p}_i - \mathbf{q}_k \right\|. \tag{8}$$

2. Compute Bledding Weights: A blending factor $\lambda_i \in [0, 1]$ is calculated based on the spatial distance between p_i and q_j and proximity to the boundaries of the wake subdomain.

$$\lambda_i = \max\left(0,\ 1 - \frac{d_{ij}}{d_{\max}}\right) \max\left(0,\ 1 - \frac{d_{\text{border}}(i)}{w_{\text{blend}}}\right), \tag{9}$$

where $d_{ij} = \|\mathbf{p}_i - \mathbf{q}_j\|$ denotes the Euclidean distance between the global point $\mathbf{p}_i$ and its nearest wake-region point $\mathbf{q}_j$, $d_{\text{border}}(i)$ is the minimum distance from $\mathbf{p}_i$ to the extracted wake boundary, $d_{\max}$ is the maximum influence radius for blending (e.g., 2.0 m), and w_{blend} specifies the width of the transition zone near the boundary.

3. Apply Blended Reconstruction: The final reconstructed velocity at point i is defined as:

$$\mathbf{U}_t^{\text{mre}}(i) = \lambda_i \cdot \mathbf{U}_t^w(j) + (1 - \lambda_i) \cdot \mathbf{U}_t^g(i). \tag{10}$$

This region-aware fusion enables the model to retain high-resolution predictions in critical subdomains but ensuring full-domain coverage smooth transitions across the boundaries between subdomains.

3. Results

3.1. Wake region detection performance

Flow passing through a circular cylinder is at a Reynolds number of approximately 4×10^6. This is deduced from a representative scenario where the monopile foundation of an offshore wind turbine is submerged in water. At the Reynolds number, the wake is fully turbulent, characterized by unsteady large-scale structures, broadband fluctuations, and shear layer instabilities. These features highlight the strongly nonlinear and unstable behavior of the wake, making it particularly challenging for ROM due to the nonlinear and time-varying nature.

The analysis is focused on the flow data in the cut plane across the middle span of the cylinder, which was obtained from three-dimensional CFD simulations. Fig. 4(a) shows a snapshot of the transient streamwise velocity u in the middle cut plane. All velocity fields are normalized by the freestream velocity U_∞. A long wake extending more than $50D$ meters presents downstream of the cylinder. This observation confirms that in real-world offshore wind farms, foundation-induced underwater turbulent wakes persist over long downstream distances and consequently affect marine environments, although relevant research quantifying this effect remains limited in the literature. On the other hand, turbine-induced wakes in the atmosphere are known to propagate long distances [30]. The wake exhibits a clear velocity deficit along the centerline, which gradually weakens as the flow recovers downstream. The POD of the CFD data provides modal coefficients for training and assessing the predictive model.

To extract transient wake regions for ROM, Grad-CAM heatmaps are utilized. A snapshot of a transient heatmap is shown in Fig. 4(b). A threshold value of 0.2 is set to discern boundaries for the extraction. The regions that contribute most to the network's prediction of velocity fluctuations are highlighted. Notably, the regions encircled by the boundaries closely match the distribution of vortices in both shape and location, particularly around the near-wake where vortex shedding is strongest. The activation intensity progressively diminishes downstream, consistent with the physical decay of coherent wake structures over distance. This spatial alignment validates the effectiveness of Grad-CAM in identifying regions of dynamic importance.

To enhance interpretability and avoid jagged outlines, the mask is constructed by extracting the largest connected region from the Grad-CAM heatmap and simplifying it using a smoothed polygonal contour. Fig. 4(c) overlays the final extracted wake region mask on a representative u-velocity field from the simulation, with the extracted boundary shown in blue. The resulting region extends downstream from the bluff body, maintaining symmetry and alignment with the core wake path. Spatially, the mask covers approximately 12% of the total mesh area, providing a compact yet dynamically representative subset of the flow field. This demonstrates a substantial reduction in computational scope while preserving essential flow features. The smooth and cohesive boundary avoids fragmented or noisy edges, indicating robust and physically meaningful region segmentation.

3.2. Field decomposition and reconstruction

The decomposition of the flow field using POD yields a ranked set of orthogonal modes, each associated with a corresponding energy level. Fig. 5 compares the POD energy spectra and cumulative energy distributions obtained from the global domain and from the wake-restricted subdomain identified via Grad-CAM.

In the global domain, the cumulative energy increases rapidly, and the first 39 modes already preserve more than 65% of the total fluctuation energy. In contrast, when POD is applied only to the extracted wake region, the energy decay is noticeably slower, and approximately 65 modes are required to reach the same 65% energy threshold. This behavior is expected, as the wake region is dominated by spatially localized and dynamically rich structures, such as vortex shedding and shear-layer instabilities, which exhibit higher effective dimensionality and therefore require more modes to be represented accurately.

When a fixed number of 40 POD modes is retained, an inherent trade-off arises between representing globally smooth background regions and resolving localized unsteady dynamics. Under this constraint, the wake-restricted POD remains effective because the decomposition is performed over a spatial domain that is explicitly focused on the dynamically active wake. As a result, approximately 55% of the total fluctuation energy within the wake region is preserved, which is sufficient to resolve the dominant coherent structures while maintaining computational efficiency. By contrast, the global POD basis spans the entire computational domain and allocates a substantial fraction of its modal energy to background regions with weak unsteadiness. When the modal budget is limited, this leads to a dilution of modal efficiency in the wake. Although the wake-restricted POD captures a smaller fraction of the total global energy under the same number of modes, it provides superior local representativeness within the region of interest, which is critical for accurate downstream modeling of dominant wake dynamics.

The first two POD modes of the streamwise (u) and cross-stream (v) velocity components are shown in Fig. 6. In both components, modes contribute to the total kinetic energy and reflect coherent wake dynamics. In the u component, the leading modes exhibit an anti-symmetric distribution along the spanwise direction, while the v component shows a symmetric pattern. Moreover, the first and second modes are highly similar, differing mainly by a streamwise phase shift. This pairing structure is characteristic of periodic vortex shedding and suggests strong modal coupling in the near-wake region. These two mode pairs dominate the coherent dynamics of the near-wake, collectively accounting for over 27% of the total flow energy. The remaining higher modes contribute to finer-scale structures and asymmetric perturbations further downstream.

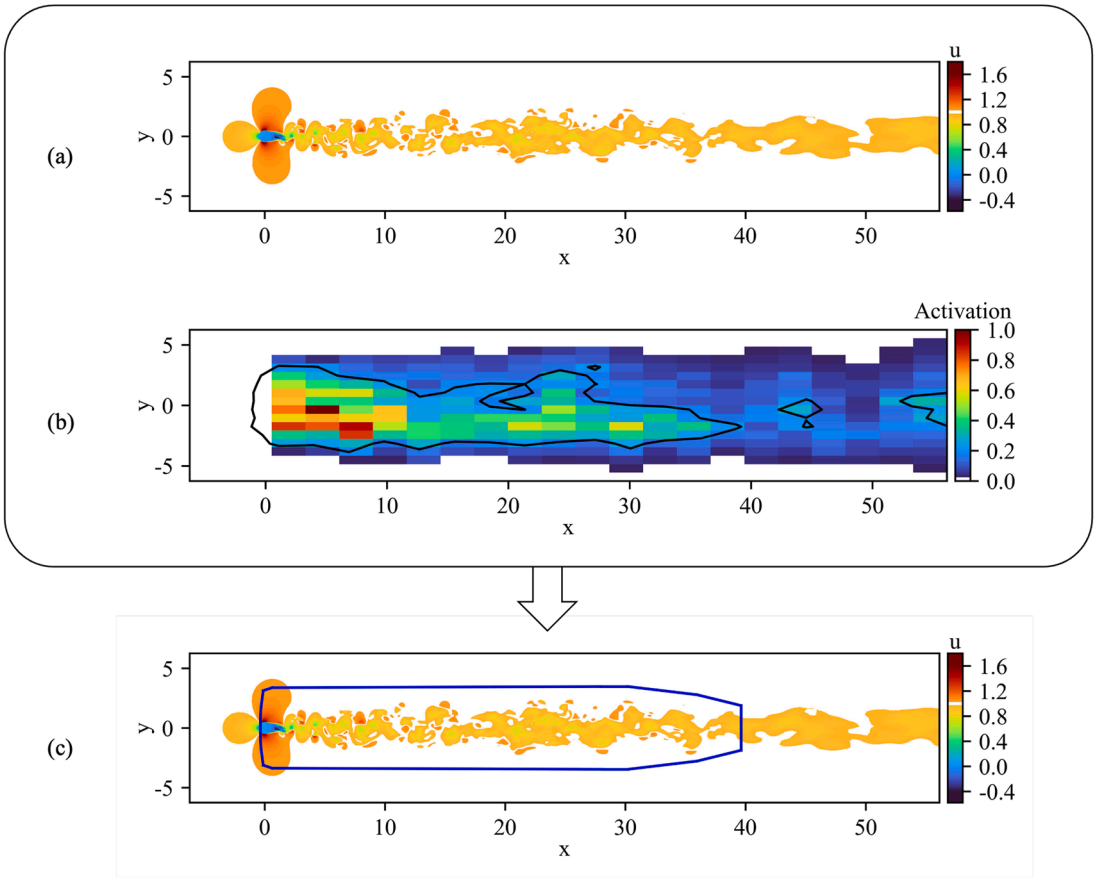

Fig. 4. Identification of the wake region using Grad-CAM.

Fig. 7 illustrates contours of the streamwise velocity component that are reconstructed from the modes of the global and local wake-restricted domains. Both reconstructions successfully reproduce the primary wake features, including the alternating vortex shedding and shear layers downstream of the cylindrical foundation. These dominant structures are well preserved due to their strong energy content being captured in the leading POD modes. However, when compared with the original flow field shown in Fig. 4, it becomes evident that some fine-scale details such as local eddy deformation and subtle shear gradients are smoothed out during reconstruction. This is a known limitation of POD, as high-frequency modes associated with localized turbulence are often truncated for efficiency.

The accuracy of the POD reconstructions is quantified using three error metrics: the root-mean-square error (RMSE), the relative reconstruction error (RRE), and the mean absolute relative error (MARE). All metrics are evaluated by comparing the reconstructed velocity fields against reference snapshots reconstructed using the same number of retained POD modes, with 40 modes used consistently in all cases.

The RMSE is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(u_i^{\mathrm{rec}} - u_i^{\mathrm{ref}}\right)^2}, \tag{11}$$

where u_i^{rec} and u_i^{ref} denote the reconstructed and reference velocities at spatial location i, respectively.

The relative reconstruction error (RRE) is defined in an energy-norm sense as

$$\mathrm{RRE} = \frac{\left\|\mathbf{u}^{\mathrm{rec}} - \mathbf{u}^{\mathrm{ref}}\right\|_2}{\left\|\mathbf{u}^{\mathrm{ref}}\right\|_2}, \tag{12}$$

where $\mathbf{u}$ denotes the velocity vector assembled over all spatial locations.

Table 1
POD reconstruction errors relative to the high-fidelity CFD flow field using global and wake-restricted bases.

Criteria	Unit	Global domain	Wake subdomain
RMSE	[m/s]	0.015	0.022
RRE	[%]	3.13	4.52
MARE	[%]	15.56	32.88

To better assess prediction accuracy in low-velocity wake regions, the mean absolute relative error (MARE) is additionally introduced and defined as

$$\mathrm{MARE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|u_i^{\mathrm{rec}} - u_i^{\mathrm{ref}}\right|}{\left|u_i^{\mathrm{ref}}\right| + \varepsilon}, \tag{13}$$

where ε is a small constant introduced to avoid numerical singularities.

Table 1 summarizes the POD reconstruction errors relative to the high-fidelity CFD reference, comparing global-domain and wake-subdomain bases in terms of RMSE, RRE, and MARE.

3.3. Prediction performance

3.3.1. Overall prediction quality assessment

Contours of the streamwise velocity predicted by the proposed hybrid LSTM–GRU model, together with the corresponding absolute prediction errors relative to a reference field reconstructed from the first 40 POD modes, are shown in Fig. 8. The predicted field recovers the key unsteady structures such as the location of the velocity deficit zone, the alternating vortex street pattern, and the gradual streamwise decay of the flow. It indicates that the model has learned the dominant temporal evolution of the POD modes, with high spatiotemporal coherence and

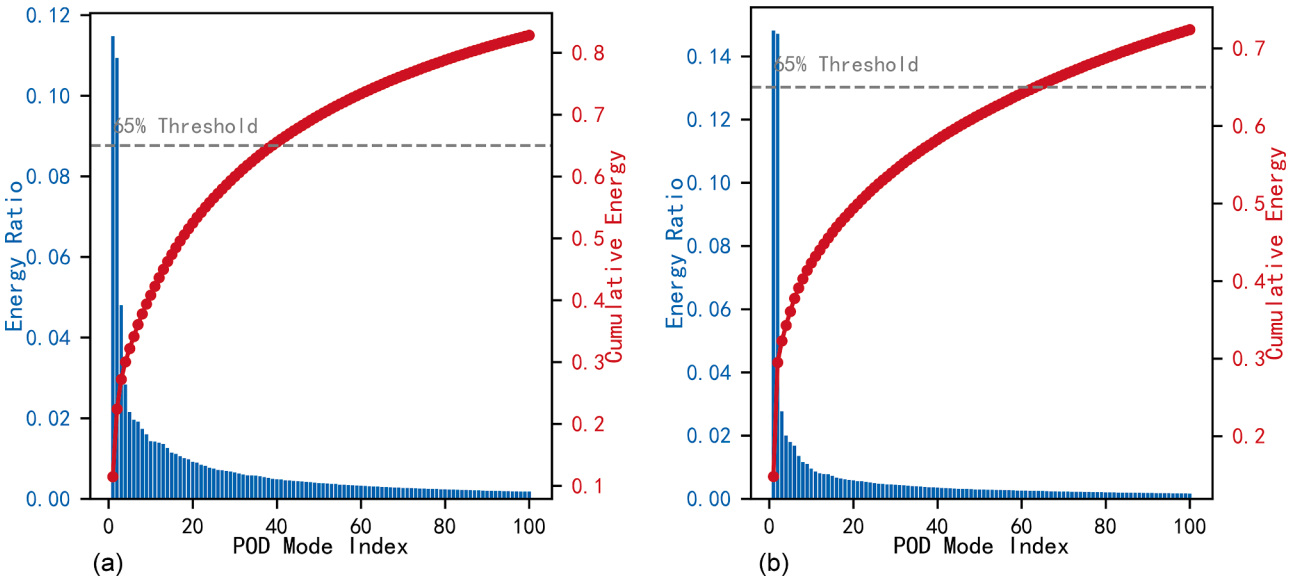

Fig. 5. POD modal energy distributions for the global domain and the Grad-CAM-identified wake subdomain. Blue bars denote the normalized energy contribution of individual modes, while the red curves indicate the cumulative energy. The dashed line marks the 65% cumulative energy threshold. (a) Global-domain POD energy spectrum and cumulative energy. (b) Wake-restricted POD energy spectrum and cumulative energy. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

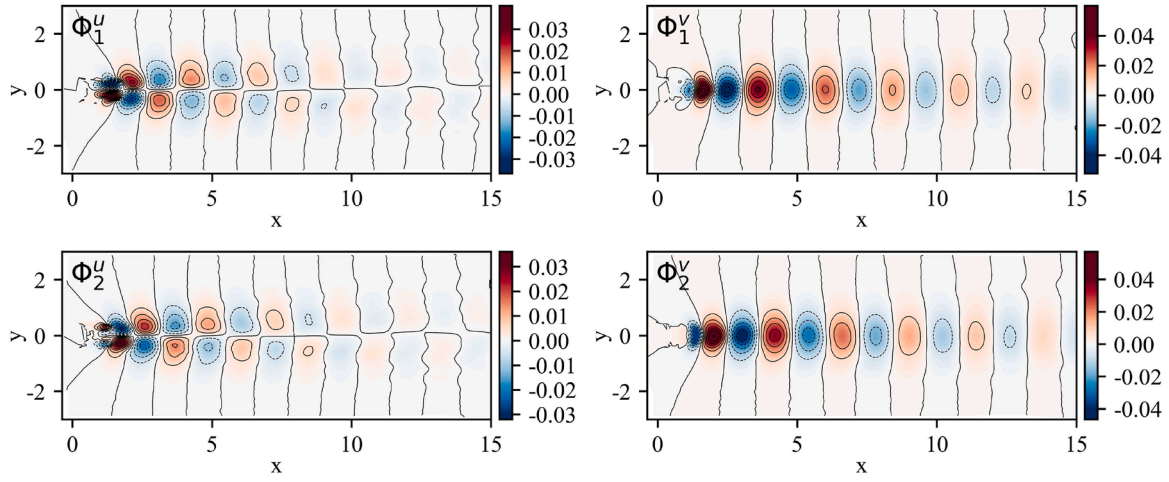

Fig. 6. First and second POD modes of the streamwise (Φ_1^u, Φ_2^u) and cross-stream (Φ_1^v, Φ_2^v) velocity fields.

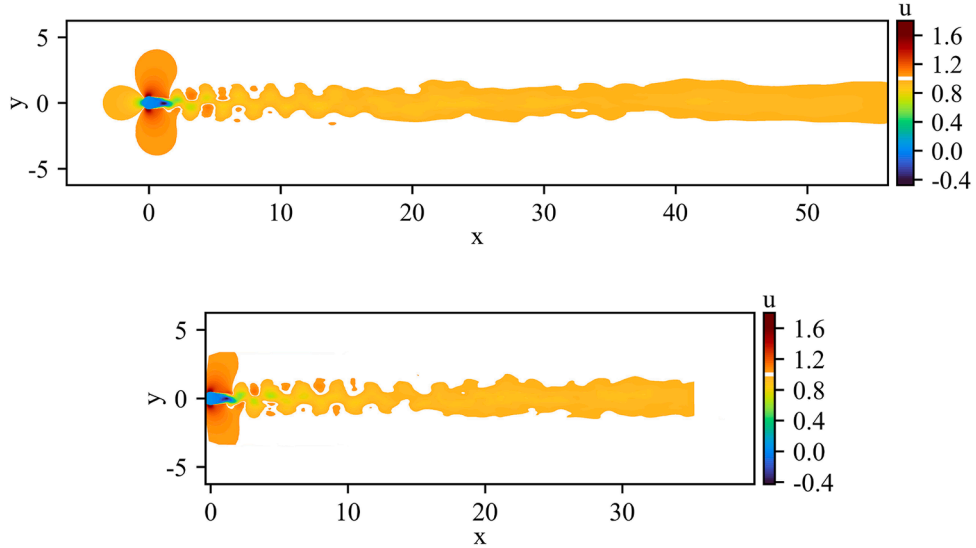

Fig. 7. Snapshot of the reconstructed streamwise velocity, u, from the POD modes of (top) the global and (bottom) the local wake regions.

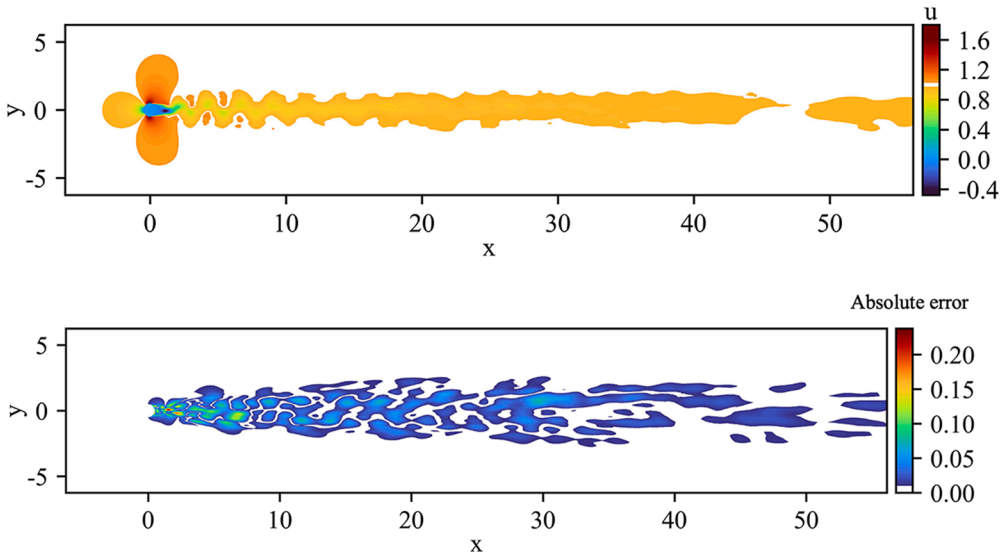

Fig. 8. Hybrid LSTM–GRU prediction of the streamwise velocity field. Top: predicted velocity contour. Bottom: absolute error relative to the reference field reconstructed from the first 40 POD modes.

physical fidelity. Errors primarily localize the vortex street downstream of the cylinder and the adjacent shear layers, where high-frequency unsteadiness and steep velocity gradients challenge the prediction. In contrast, the background and far-wake regions exhibit significantly lower errors due to their quasi-steady behavior. Thus, the major prediction difficulty lies in the dynamically active wake, where unsteady features are more pronounced and physically important.

Fig. 9 illustrates the phase portrait of the first two POD coefficients $(a_1(t), a_2(t))$ extracted from the tail wake region. The black dots represent the true POD coefficients, while the blue dots indicate those predicted by the surrogate model. Two concentric circles visualize the average amplitude (Euclidean norm) of the POD coefficient vectors: the solid red circle (radius ≈ 15) corresponds to the true trajectory, and the dashed red circle (radius ≈ 12) corresponds to the predicted one. The predicted trajectory captures the general periodic behavior of the true dynamics, forming a closed-loop structure that reflects vortex-shedding-induced cycling between the modes. However, the predicted average amplitude is underestimated by approximately 20%, as indicated by the smaller radius of the dashed circle. This discrepancy suggests a mild underestimation of energy in the predicted near-wake region.

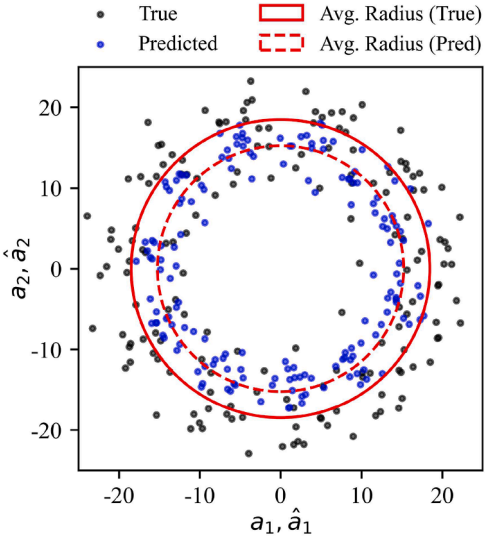

Fig. 9. Phase-space comparison of the first two POD coefficients from the true and predicted wake fields.

3.3.2. *Ablation study on model components*

To quantify the contribution of different model components, an ablation study is conducted considering the following five model variants:

1. a baseline model combining global POD with a standard LSTM (BL);
2. region-aware POD coupled with a standard LSTM (RA-POD + LSTM);
3. region-aware POD with a bidirectional LSTM (RA-POD + BiLSTM);
4. region-aware POD with a gated recurrent unit (RA-POD + GRU);
5. the proposed model integrating region-aware POD with a hybrid LSTM–GRU architecture (RA-POD + HLG).

All model variants are trained and evaluated under identical settings to ensure a fair comparison. Specifically, the same dataset, POD truncation levels, input sequence length, hidden dimension, optimizer, learning rate, batch size, and training protocol are used for all cases. The only differences among the variants lie in the spatial decomposition strategy (global vs. region-aware) and the recurrent network architecture, so that the impact of each modeling component can be isolated. All implementation and training details follow the configuration described in Section 2.4.

Table 2 reports the error metrics RMSE, RRE, and MARE evaluated over the global domain, the identified wake subdomain, and the blended field obtained by merging the global and wake-restricted predictions. All error metrics are computed by comparing the predicted and reference POD coefficient vectors, with 40 retained POD modes used consistently in all cases.

As expected, all models exhibit larger errors in the wake subdomain, where small-scale turbulent structures and low-velocity magnitudes pose significant challenges for data-driven prediction. Compared with the global baseline, all region-aware models consistently achieve lower errors in the wake region, demonstrating the effectiveness of explicitly incorporating wake-focused representations. Furthermore, the region-aware formulation produces a merged field by blending the global and wake-restricted predictions, allowing the wake-focused improvements to be reflected in the final full-field reconstruction while maintaining smooth transitions across the wake boundary. Among all model variants, the proposed model attains the lowest area-averaged errors across all metrics, confirming that the hybrid architecture most effectively exploits the complementary strengths of the global and wake-restricted predictions, leading to improved accuracy both within the wake region

Table 2
Ablation study of area-averaged prediction errors evaluated over different spatial domains, computed by comparing predicted and reference velocity fields reconstructed from the same set of 40 retained POD modes.

Model	Domain	RMSE [m/s]	RRE [%]	MARE [%]
BL (G-POD + LSTM)	Global	1.75e-2	3.59	15.6
	Wake	2.69e-2	6.28	47.0
RA-POD + LSTM	Global	1.75e-2	3.59	15.6
	Wake	2.53e-2	5.28	36.0
	Blended	1.75e-2	3.60	14.2
RA-POD + BiLSTM	Global	1.73e-2	3.57	17.4
	Wake	2.51e-2	5.24	41.3
	Blended	1.74e-2	3.57	16.5
RA-POD + GRU	Global	1.75e-2	3.59	19.8
	Wake	2.52e-2	5.27	46.4
	Blended	1.74e-2	3.59	18.0
RA-POD + HLG (Proposed)	Global	1.80e-2	3.69	14.6
	Wake	2.43e-2	5.07	27.7
	Blended	1.68e-2	3.46	12.2

Table 3
Dominant frequencies and band-averaged spectral coherence between predicted and reference POD coefficients.

Model	Mode	f_{peak}	Coherence
BL (G-POD + LSTM)	1	0.99	0.83
	2	0.98	0.77
	3	0.98	0.26
	4	0.98	0.37
RA-POD + LSTM	1	0.99	0.89
	2	0.98	0.50
	3	0.98	0.32
	4	0.98	0.44
RA-POD + BiLSTM	1	0.99	0.94
	2	0.98	0.97
	3	0.98	0.31
	4	0.98	0.55
RA-POD + GRU	1	0.99	0.36
	2	0.98	0.98
	3	0.98	0.30
	4	0.98	0.55
RA-POD + HLG (Proposed)	1	0.99	0.74
	2	0.98	0.75
	3	0.98	0.25
	4	0.98	0.31

and over the entire computational domain.

To further assess the temporal consistency of the predicted dynamics, spectral coherence between predicted and reference POD coefficients is evaluated for the dominant wake modes. Table 3 summarizes the non-dimensional dominant frequencies and the corresponding band-averaged coherence values.

All models exhibit non-dimensional dominant frequencies close to unity, indicating consistent alignment with the reference temporal scale defined by the ground-truth POD coefficients. Differences between model variants are therefore primarily reflected in the band-averaged coherence, which measures phase consistency and spectral concentration rather than frequency location. BiLSTM-based models achieve the highest coherence for the dominant oscillatory modes due to their enhanced phase alignment, whereas GRU-based models tend to concentrate spectral energy on a single phase mode. In contrast, the proposed hybrid model yields slightly lower coherence for the leading modes but maintains a more balanced representation across higher-order and low-frequency modes, indicating improved robustness to multi-scale wake dynamics.

Fig. 10 presents the frequency-domain comparison of the first two POD coefficients extracted from the wake region. The black lines represent the true values of a_1 and a_2, while the blue lines denote the predicted values $\hat{a}_1$ and $\hat{a}_2$. In both modes, the model clearly captures the dominant frequency components. As shown in panel (a), the true coefficient a_1 exhibits a sharp peak at the normalized frequency $f = 0.0015$, which is accurately reproduced by $\hat{a}_1$ with a frequency shift of less than 1%. Similarly, panel (b) shows good agreement for the second mode, with both a_2 and $\hat{a}_2$ peaking at approximately the same $f = 0.0015$. The vertical dashed lines mark this dominant frequency for the true and predicted signals. These results validate the model's ability to preserve key periodic structures in the wake flow, particularly the vortex shedding frequency, for the temporal dynamics of both spatial modes.

3.3.3. Pointwise velocity analysis

To further evaluate the spatiotemporal accuracy of the hybrid model, the pointwise comparison reveals that it effectively captures the dominant temporal dynamics across both velocity components and varying wake regions, while highlighting limitations in reproducing high-frequency content and exact amplitudes (Fig. 11). Here, locations are: Point 1 ($x = 5D$, $y = 0$; centerline), Point 2 ($x = 5D$, $y = 0.75D$; shear layer), and Point 3 ($x = 15D$, $y = 0$; downstream recovery), with D as the reference cylinder diameter.

The model demonstrates strong capabilities in reproducing primary periodic structures. For the v-velocity component, which exhibits prominent sinusoidal oscillations due to vortex shedding, both the POD reconstruction and hybrid LSTM-GRU prediction closely match the phase and frequency of the ground truth at all three points. At Point 1, the model maintains good tracking over multiple cycles. At Point 2, it resolves the main oscillatory behavior, though fine details like spikes are attenuated. Even at Point 3, where oscillation energy has decayed to low-amplitude waves, the dominant frequency and phase are effectively replicated. Similarly, for the u-velocity component, clear periodicity at Point 2 is successfully captured by both reconstructions, aligning well with CFD. At Point 3 with non-periodic fluctuations, all three curves show close agreement in overall trend.

Despite this accurate periodicity, a consistent observation across locations and components is the model's tendency to smooth high-frequency fluctuations and underestimate amplitudes. For the u-component at Point 1, where CFD shows low-amplitude, irregular fluctuations lacking dominant periodicity, the POD captures the general trend but underestimates local peaks due to mode truncation. For the u-velocity component at Point 2, high-frequency details are attenuated in both reconstructions. For the v-component at Point 2, both reconstructions significantly smooth sharp spikes, contributing to amplitude underestimation.

This analysis confirms the hybrid LSTM-GRU model's robustness in preserving key periodic structures in the wake flow, particularly the vortex shedding frequency, across both velocity components. However, the amplitude discrepancies and smoothing of sharp features align with POD's low-pass filtering via mode truncation and the LSTM-GRU's preference for dominant, smooth trends.

3.4. Computational efficiency

The computational efficiency of the proposed framework is evaluated by comparing the wall-clock cost of data-driven prediction with that of high-fidelity CFD simulations over the same physical time interval, and by benchmarking against a simpler global-only baseline model.

The CFD simulations advance the flow using a physical time step of $\Delta t_{CFD} = 0.5$ s. Based on the measured runtime, extrapolation to a physical duration of 660 s yields a total CFD computational cost of t_{CFD} on the order of 3.2×10^5 s.

In contrast, the trained region-aware hybrid LSTM–GRU model predicts 165 snapshots at a sampling interval of $\Delta t_{ML} = 4$ s, corresponding to the same physical duration of 660 s, in approximately $t_{ML} = 10$ s of wall-clock time. This inference cost is incurred after a one-time offline training stage, which requires approximately 65 s in total.

For comparison, the global-only POD–LSTM baseline requires approximately 34 s for training and 7 s for predicting the same 660 s physical duration. This indicates that the additional computational cost in-

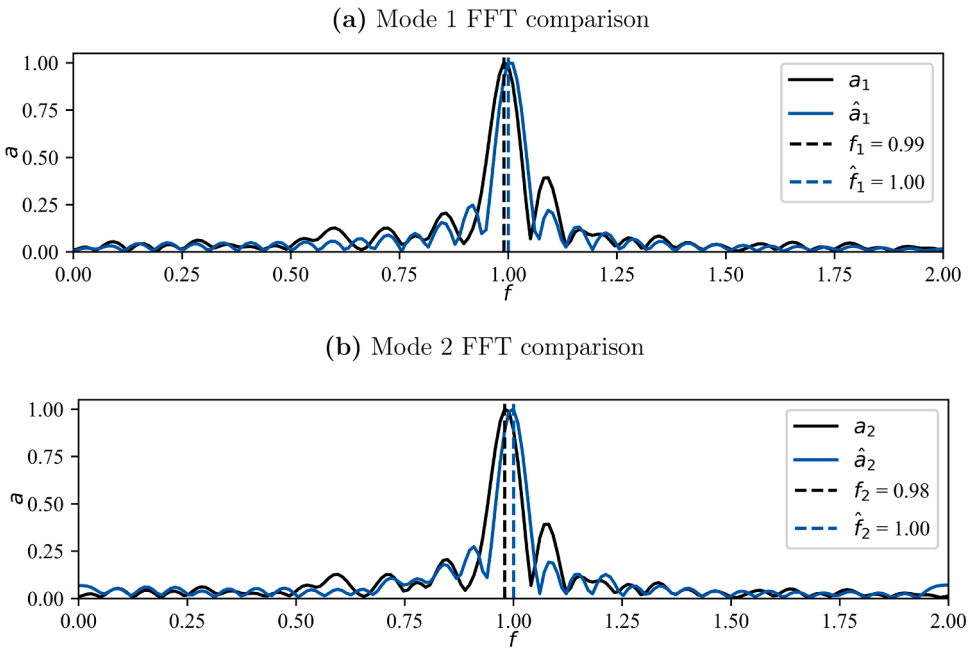

Fig. 10. Frequency-domain comparison of the first two POD coefficients in the wake region. (a) shows the FFT spectrum of mode 1 (a_1) and its prediction ($\hat{a}_1$); (b) shows the same for mode 2 (a_2 and $\hat{a}_2$). Vertical dashed lines indicate the dominant frequencies $f_1, \hat{f}_1$ in (a) and $f_2, \hat{f}_2$ in (b). **(a)** Mode 1 FFT comparison **(b)** Mode 2 FFT comparison.

troduced by region-aware modeling and hybrid temporal prediction remains modest, while providing substantial gains in predictive accuracy.

Based on these measurements, the ratio between CFD and data-driven prediction time for the same physical duration is approximately

$$\frac{t_{\mathrm{CFD}}}{t_{\mathrm{ML}}} \approx 3.2 \times 10^4,$$

demonstrating several orders-of-magnitude reduction in computational cost during the prediction phase. All machine-learning models are executed with GPU acceleration. While high-fidelity CFD remains indispensable for generating training data, the proposed framework enables rapid wake prediction over long time horizons at negligible computational expense once trained.

4. Discussion

4.1. Mechanistic interpretation of region-aware modeling

This study presents a region-aware reduced-order modeling framework that integrates Grad-CAM-based wake localization with POD dimensionality reduction and a hybrid LSTM–GRU temporal predictor to forecast unsteady hydrodynamic wakes behind an offshore wind turbine foundation. The results demonstrate clear gains in predictive fidelity, with the proposed hybrid fusion strategy achieving area-averaged RMSE values as low as 0.0168 m/s and a MARE of 12.2% over the merged field. Compared with all other model variants considered in the ablation study, including global POD with a standard LSTM and region-aware POD combined with LSTM, BiLSTM, and GRU architectures, the proposed framework consistently yields lower prediction errors in the merged-field evaluation, underscoring the advantage of combining region-aware flow representations with a hybrid recurrent architecture that captures both long-term and short-term temporal dynamics.

A central contribution of the proposed framework lies in its use of Grad-CAM as an integral component of the reduced-order modeling pipeline, rather than as a post hoc interpretability tool. Unlike conventional ROM approaches that operate on the entire computational domain and dilute modeling capacity in dynamically inactive regions, the Grad-CAM activation maps consistently isolate physically meaningful wake regions associated with vortex shedding and shear-layer dynamics. By identifying statistically dominant regions that govern the global unsteadiness, the proposed approach enables region-focused modeling by concentrating predictive capacity on approximately 12% of the mesh that contains the majority of the unsteady energy, thereby reducing redundancy in far-field regions and facilitating subsequent regionalized modeling. While prior CFD studies have employed Grad-CAM primarily for interpretability, sensor placement, or regime classification [24–28], its systematic integration into reduced-order state prediction frameworks remains limited. The present work advances beyond these efforts by embedding Grad-CAM directly into the spatial definition of the reduced-order state space used for temporal forecasting.

The POD analysis further highlights the benefits of this spatial separation. In the wake-restricted domain, the slower convergence of modal energy reflects the increased complexity of local instabilities, while enabling superior reconstruction of coherent structures, including paired anti-symmetric modes associated with vortex shedding. In contrast, global POD exhibits faster energy convergence but suffers from modal dilution, leading to reduced fidelity in the near-wake. The hybrid LSTM-GRU architecture effectively exploits these complementary characteristics: the LSTM component captures long-term temporal dependencies linked to periodic shedding, while the GRU-based residual correction refines short-term discrepancies. As a result, the model preserves phase and dominant frequency content across key monitoring locations, while improving local reconstruction quality in dynamically active regions. Similar modal energy distributions in turbulent wake flows have been reported in [20].

Through distance-weighted blending of global and wake predictions, the proposed hybrid fusion strategy achieves spatially coherent reconstructions with minimal boundary artifacts. This approach yields the lowest overall errors among the tested configurations, reducing the relative reconstruction error from 15.6% in the global-only baseline to 12.2% in the fused solution. In the wake subdomain, MARE is further reduced from 47.0% to 27.7%, indicating a substantial improvement in wake-region prediction accuracy.

At the same time, the reduced-order framework achieves an approximate 90% reduction in computational cost relative to full CFD simulations, while maintaining accurate representation of dominant unsteady

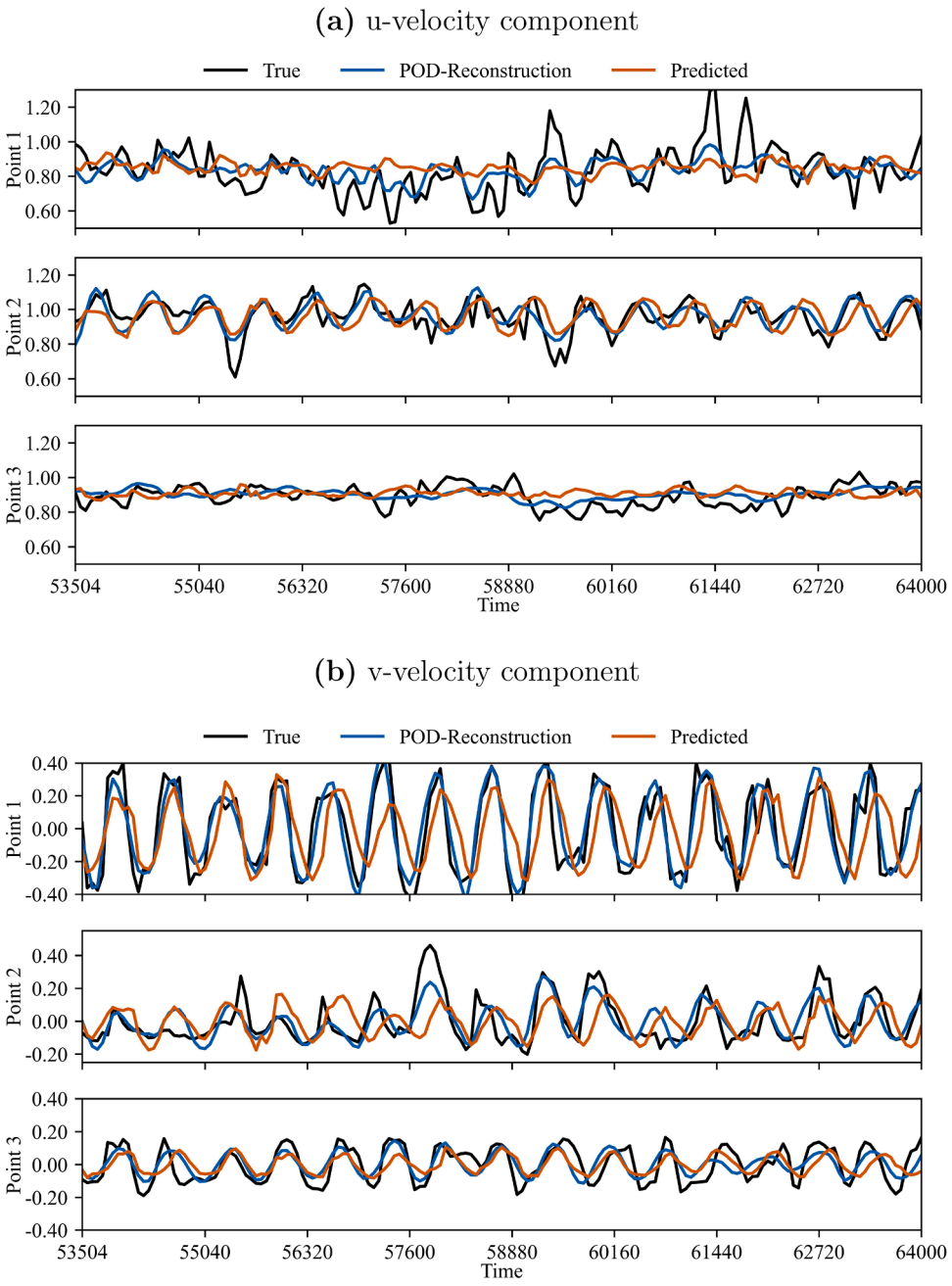

Fig. 11. Time-series comparison of u- and v-velocity components at three representative wake locations: Point 1 ($x = 5D, y = 0$), Point 2 ($x = 5D, y = 0.75D$), and Point 3 ($x = 15D, y = 0$). Each subplot shows the ground-truth velocity (black), POD reconstruction (blue), and model prediction (orange). **(a)** u-velocity component **(b)** v-velocity component. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

dynamics and outperforming linear DMD and baseline LSTM models in the present nonlinear turbulent regime.

4.2. Error characteristics and amplitude attenuation

In addition to these accuracy improvements, the results reveal a systematic underestimation of fluctuation amplitudes, most evident in phase portraits and pointwise velocity analyses. This behavior is a known limitation of reduced-order recurrent frameworks and arises from a combination of three factors.

First, POD modal truncation constrains the representational bandwidth of the reduced-order space. Although the temporal prediction is performed on truncated POD coefficients, retaining a finite number of modes limits the expressiveness of intermittent, high-energy fluctuations, which can amplify the smoothing tendency of recurrent regression models operating in the reduced space. This smoothing effect due to modal truncation has been widely discussed in POD-based reduced-order modeling literature [31].

Second, the use of a mean-squared-error (MSE) loss emphasizes energy-dominant dynamics in an averaged sense, inherently penalizing large deviations and thereby attenuating peak amplitudes. Similar amplitude attenuation under MSE-based regression has been observed in literature [18,20,32].

Third, the gated recurrent architecture itself (LSTM/GRU) introduces an inductive bias toward smooth temporal evolution, which favors stable, low-frequency trajectories and further suppresses sharp transient features. Recurrent neural networks are known to favor smooth temporal trajectories in chaotic systems, as reported in [14].

Together, these effects lead to amplitude smoothing while largely preserving phase and dominant frequency content. Future work may mitigate this limitation through energy-weighted or mode-adaptive loss formulations, as well as architectural modifications designed to better retain intermittent high-amplitude dynamics, without compromising numerical stability.

4.3. Limitations and practical implications

From a broader application perspective, several modeling assumptions constrain the generalizability of the present framework. First, the analysis is based on two-dimensional horizontal slices extracted from an inherently three-dimensional flow field. While this simplification is motivated by computational efficiency and the dominance of horizontally coherent wake structures in the current configuration, it neglects vertical flow features such as three-dimensional vortex stretching and cross-plane momentum exchange. These effects may become significant in stratified flows or in the presence of surface-wave interactions, potentially altering wake dynamics. Second, the model is validated for a single high-Reynolds-number case ($Re \approx 4 \times 10^6$) and a single monopile geometry. Its performance for other Reynolds numbers, alternative foundation designs (e.g., jacket or tripod structures), or tidally varying inflow conditions remains untested and cannot be assumed.

Despite these limitations, the proposed framework shows strong potential for offshore wind engineering applications under conditions consistent with the present assumptions. Accurate wake prediction can support foundation siting analyses, improve array-level layout design by better capturing hydrodynamic interactions, and contribute to environmental impact assessments aligned with regulatory frameworks such as the EU Marine Strategy Framework Directive. Extending the methodology to fully three-dimensional formulations and validating it across a broader range of Reynolds numbers and foundation geometries constitute important directions for future research. Incorporating attention-based mechanisms or physics-informed constraints may further enhance robustness and pave the way toward operational reduced-order modeling tools for marine renewable energy applications.

5. Conclusion

This study proposed a region-aware reduced-order modeling framework for predicting unsteady wakes behind offshore wind turbine foundations. By integrating Grad-CAM-based wake localization, spatially separated POD decomposition, and a hybrid LSTM-GRU temporal predictor, the framework enables targeted modeling of a fixed wake subdomain that captures the dominant unsteady dynamics.

Quantitative results demonstrate clear accuracy improvements. Compared with a conventional global POD-LSTM baseline, the fused region-aware hybrid model reduces the full-field mean absolute relative error (MARE) from 15.6% to 12.2%. Within the wake subdomain, MARE decreases from 47.0% to 27.7%, highlighting the effectiveness of explicitly concentrating modeling capacity on dynamically active regions. In addition, the proposed approach achieves orders-of-magnitude computational savings relative to high-fidelity CFD, while maintaining accurate representation of dominant unsteady dynamics.

A key contribution of this work lies in embedding Grad-CAM directly into the reduced-order modeling pipeline. Rather than using activation maps solely for interpretability or classification, the extracted wake region is incorporated into the predictive workflow, enabling region-focused decomposition and temporal modeling. The hybrid LSTM–GRU architecture further improves temporal consistency by combining long-term dependency capture with residual refinement.

The present study is restricted to two-dimensional slices of a single high-Reynolds-number monopile configuration. Extension to other Reynolds numbers, alternative foundation geometries (e.g., jacket or tripod structures), and tidally varying inflow conditions remains to be investigated. Future work will focus on fully three-dimensional formulations and broader validation to enhance generalizability.

Overall, the proposed framework provides an efficient and physically interpretable strategy for wake prediction, with potential applications in offshore wind foundation design, wake interaction analysis, and reduced-order modeling of marine renewable energy systems.

CRediT authorship contribution statement

Yifan Ma: Writing - original draft, Visualization, Software, Resources, Methodology, Investigation, Formal analysis, Data curation; **Zhaoyue Xu:** Writing - review & editing, Writing - original draft, Validation, Supervision, Software, Methodology, Conceptualization; **Hua-Dong Yao:** Writing - review & editing, Supervision, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization.

Data availability

Data will be made available on request.

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.

Acknowledgements

The computations and data handling were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725. The CFD simulation setup was supported by Andhika Pratama based on his thesis work supervised by Yanlin Shao, Ida-Maja Hassellöv, and Hua-Dong Yao.

This research was supported by the Clean Energy Transition Partnership, an initiative co-funded by the European Union, through the project "Enhancing Shared mOoring systeM design for flOating Offshore wind faRms (ESOMOOR)" (grant agreement no. 2024-206305).

Appendix A. Numerical setup and data fidelity

This appendix summarizes the numerical setup of the CFD simulations and the temporal sampling strategy adopted for the reduced-order and data-driven modeling in the present study. The objective is to clarify the fidelity level of the training data and to ensure consistency between the underlying CFD resolution and the subsequent machine-learning framework.

A.1. Computational mesh and turbulence modeling

The flow database employed in this study is generated using the improved delayed detached eddy simulation (IDDES) approach. The computational mesh is inherited from a previously validated large-eddy simulation (LES) study conducted within the same research group, in which the mesh resolution and near-wall treatment were systematically examined through a grid-independence analysis.

In the original study, three progressively refined meshes were considered, namely a base mesh, a 1.25-times finer mesh, and a 1.5-times finer mesh. These meshes share identical domain extents and refinement layouts, while differing primarily in the cell sizes within the near-wall region and the wake refinement zones. The detailed grid parameters and total cell counts are reported in [33] and are summarized in Table 3.1 of the present manuscript for reference.

The grid-independence assessment demonstrated that the base mesh provides a satisfactory balance between numerical accuracy and computational cost for capturing the dominant wake dynamics. Since the

primary objective of the present work is reduced-order modeling and data-driven prediction rather than further turbulence-model or mesh-development studies, this validated base mesh is directly adopted in combination with the IDDES formulation.

The adopted mesh consists of approximately 10.3×10^6 control volumes and covers a computational domain of $63D \times 53D \times 6D$, where D denotes the cylinder diameter. The mesh is partitioned into several regions, including a near-wall region, two wake refinement regions, and an anisotropic far-field region. The near-wall resolution is designed to ensure that the dimensionless wall distance remains below unity over the entire cylinder surface, enabling the IDDES model to operate in a wall-resolved mode in the vicinity of the solid boundary. A statistical distribution of the resulting y^+ values is provided in Fig. A.12 to further document the near-wall mesh quality.

A.2. Time integration and data sampling

The CFD simulations are advanced using a physical solver time step of $\Delta t_{\mathrm{CFD}} = 0.5$ s, which is selected to maintain an average Courant–Friedrichs–Lewy (CFL) number below unity throughout the computational domain. This choice is consistent with the spatial resolution of the base mesh and the stability requirements of the IDDES formulation, ensuring stable time integration while resolving the dominant unsteady flow features associated with wake development and vortex shedding.

For reduced-order analysis and data-driven modeling, the velocity fields are stored at a larger sampling interval of $\Delta t_{\mathrm{ML}} = 4$ s. This interval is determined based on the characteristic vortex shedding period of the flow, $T_s \approx 80$ s, corresponding to a shedding frequency of $f_s \approx 0.0125$ Hz. The resulting sampling frequency $f_{\mathrm{ML}} = 0.25$ Hz yields a Nyquist frequency of $f_N = 0.125$ Hz, providing nearly an order-of-magnitude margin relative to the dominant shedding frequency.

Consequently, each shedding cycle is represented by approximately twenty snapshots, which is sufficient to resolve the low-frequency wake dynamics targeted by the POD-based reduced-order and machine-learning models.

It is emphasized that this reduced sampling interval applies only to the stored flow snapshots used for POD decomposition and machine-learning training. The CFD solver itself resolves the flow evolution at a significantly finer temporal scale, and therefore no temporal under-resolution or aliasing is introduced at the level of the governing-equation integration. The adopted sampling strategy is thus consistent with the spectral bandwidth of the retained POD modes and the objective of predicting large-scale, energetically dominant wake structures rather than small-scale turbulent fluctuations.

A.3. Implications for reduced-order and data-driven modeling

The data-driven framework developed in this study operates on truncated POD coefficients, which predominantly represent large-scale, energetic, and low-frequency flow structures in the wake. As a result, the adopted temporal sampling strategy is consistent with the spectral bandwidth of the retained POD modes and the modeling objective of predicting dominant wake dynamics rather than small-scale turbulent fluctuations.

While grid coarsening in the far wake and POD truncation inevitably attenuate high-frequency content, these effects primarily impact small-scale turbulence that lies outside the intended scope of the reduced-order model. The retained dataset preserves the essential coherent structures governing wake evolution, rendering the numerical fidelity of the training data sufficient for the proposed region-aware reduced-order and machine-learning framework.

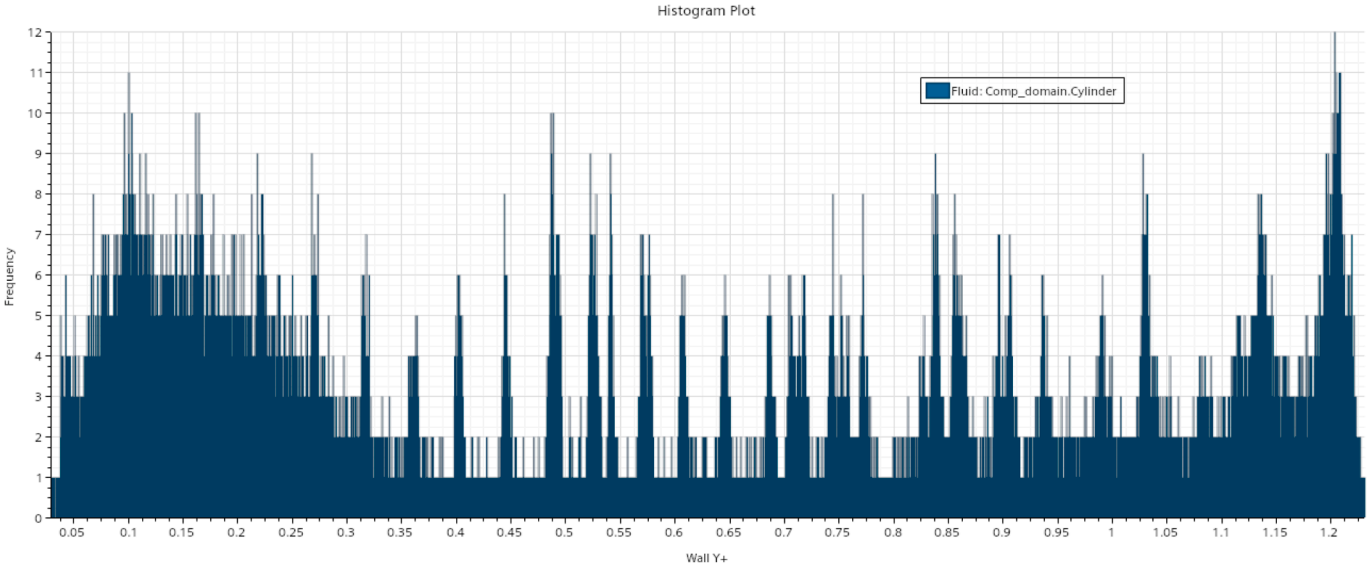

Fig. A.12. Histogram of the wall-normal resolution parameter y^+ over the cylinder surface for the base mesh. The distribution confirms that the majority of wall-adjacent cells satisfy $y^+ < 1$, ensuring adequate near-wall resolution for IDDES simulations.

References

[1] U. Daewel, et al, Offshore wind farms are projected to impact primary production and bottom water deoxygenation in the North Sea, Commun. Earth Environ. 3 (1) (2022) 292.

[2] E. Hendriks, et al. The impact of offshore wind turbine foundations on local hydrodynamics and stratification in the Southern North Sea, Front. Mar. Sci. 12 (2025) 1619577 .

[3] Q. Vanhellemont, K. Ruddick, Turbid wakes associated with offshore wind turbines observed with Landsat 8, Remote Sens. Environ. 145 (2014) 105–115.

[4] Z. Li, G. Dong, X. Yang, Onset of wake meandering for a floating offshore wind turbine under side-to-side motion, J. Fluid Mech. 934 (2022) A29.

[5] J. Bossuyt, et al, Floating wind farm experiments through scaling for wake characterization, power extraction, and turbine dynamics, Phys. Rev. Fluids 8 (12) (2023) 120501.

[6] X. Shao, et al, Near-wall modeling of forests for atmosphere boundary layers using lattice Boltzmann method on GPU, Eng. Appl. Computat. Fluid Mech. 16 (1) (2022) 2142–2155.

[7] R.J. A.M. Stevens, C. Meneveau, Flow structure and turbulence in wind farms, Annu. Rev. Fluid Mech. 49 (2017) 311–339.

[8] C.W. Rowley, S.T.M. Dawson, Model reduction for flow analysis and control, Annu. Rev. Fluid Mech. 49 (2017) 387–417.

[9] Z. Li, X. Yang, Resolvent-based motion-to-wake modelling of wind turbine wakes under dynamic rotor motion, J. Fluid Mech. 980 (2024) 1097.

[10] G. Berkooz, P. Holmes, J.L. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, Annu. Rev. Fluid Mech. 25 (1993) 539–575.

[11] C. Verhulst, C. Meneveau, Large Eddy Simulation Study of the Kinetic Energy Entrainment by Energetic Turbulent Flow Structures in Large Wind Farms, Vol. 26, AIP Publishing, 2014.

[12] D. Bastine, B. Witha, M. Wächter, J. Peinke, Towards a simplified dynamic wake model using POD analysis, Energies 8 (2) (2015) 895–920.

[13] N. Hamilton, B. Viggiano, M. Calaf, Tutkun, C. Murat, R. Bayoán, A generalized framework for reduced-order modeling of a wind turbine wake, Wind Energy 21 (6) (2018) 373–390.

[14] S.L. Brunton, B.R. Noack, P. Koumoutsakos, Machine learning for fluid mechanics, Annu. Rev. Fluid Mech. 52 (2020) 477–508.

[15] Z. Xu, X. Zhang, S. Wang, G. He, Artificial neural network based response surface for data-driven dimensional analysis, J. Comput. Phys. 459 (2022) 111145.

[16] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.

[17] A. Nazvanova, M. Ong, Chen, G. Yin, A data-driven reduced-order model based on long short-term memory neural network for vortex-induced vibrations of a circular cylinder, Phys. Fluids 35 (6) (2023).

[18] N. Raj, Ashwin, D. Tafti, N. Muralidhar, Comparison of reduced order models based on dynamic mode decomposition and deep learning for predicting chaotic flow in a random arrangement of cylinders, Phys. Fluids 35 (7) 073330.

[19] A. Mashhadi, Sohankar, M.M. Moradmand, Three-dimensional wake transition of rectangular cylinders and temporal prediction of flow patterns based on a machine learning algorithm, Phys. Fluids 36 (9) (2024) 094138.

[20] M.Z. Yousif, H.-C. Lim, Reduced-order modeling for turbulent wake of a finite wall-mounted square cylinder based on artificial neural network, Phys. Fluids 34 (1) (2022) 015116.

[21] S. Li, J. Yang, P. Teng, Data-driven prediction of cylinder-induced unsteady wake flow, Appl. Ocean Res. 150 (2024) 104114.

[22] K. Hasegawa, K. Fukami, T. Murata, K. Fukagata, CNN-LSTM based reduced order modeling of two-dimensional unsteady flows around a circular cylinder at different Reynolds numbers, Fluid Dyn. Res. 52 (6) (2020) 65501.

[23] R.R. Selvaraju, et al, Grad-CAM: visual explanations from deep networks via gradient-based localization, in: Proceedings of the IEEE International Conference on Computer Vision, the IEEE International Conference on Computer Vision, 2017, pp. 618–626.

[24] M. Morimoto, K. Fukami, K. Zhang, K. Fukagata, Generalization techniques of neural networks for fluid flow estimation, Neural Comput. Appl. 34 (5) (2022) 3647–3669.

[25] Z. Xu, S. Wang, Xin-Lei, Zhang, G. He, Optimal sensor placement for ensemble-based data assimilation using gradient-weighted class activation mapping, J. Comput. Phys. 514 (2024) 113224.

[26] E. Jagodinski, X. Zhu, S. Verma, Inverse identification of dynamically important regions in turbulent flows using three-dimensional convolutional neural networks, Phys. Rev. Fluids 8 (9) (2023) 94605.

[27] A. Kumar, A. Assam, Ashwani, Addressing performance improvement of a neural network model for Reynolds-averaged Navier-Stokes solutions with high wake formation, Eng. Computat. 41 (7) (2024) 1740–1765.

[28] S. Wen, et al, Feature identification in complex fluid flows by convolutional neural networks, Theor. Appl. Mech. Lett. 13 (6) (2023) 100482.

[29] J. Lin, H.-D. Yao, Modified Magnus effect and vortex modes of rotating cylinder due to interaction with free surface in two-phase flow, Phys. Fluids 35 (2023) 123614.

[30] B.C. nadillas, et al, Offshore wind farm wake recovery: airborne measurements and its representation in engineering models, Wind Energy 23 (2020) 1249–1265.

[31] K. Taira, M.S. Hemati, S.L. Brunton, Y. Sun, K. Duraisamy, S. Bagheri, S.T.M. Dawson, C.-A. Yeh, Modal analysis of fluid flows: applications and outlook, AIAA J. 58 (3) (2020) 998–1022.

[32] H. Kim, J. Kim, S. Won, C. Lee, Unsupervised deep learning for super-resolution reconstruction of turbulence, J. Fluid Mech. 910 (2021) A29.

[33] A. Pratama, Satria, Prediction of Wake Flows Induced by Offshore Wind-Turbine Foundations, Technical Report, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, Sweden, 2024.