

A Spatio-Temporal Prediction Model for Wall Turbulence Based on Hybrid Neural Networks

Yushan Yao*, Yuhang Guan#

School of Mathematics and Information, China West Normal University, Nanchong, China

Email: yaoy2023@126.com, #2533180563gyh@gmail.com

How to cite this paper: Yao, Y.S. and Guan, Y.H. (2026) A Spatio-Temporal Prediction Model for Wall Turbulence Based on Hybrid Neural Networks. *Applied Mathematics*, 17, 200-218.
<https://doi.org/10.4236/am.2026.173012>

Received: February 6, 2026

Accepted: March 15, 2026

Published: March 18, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The spatio-temporal evolution of wall-bounded turbulence is characterized by high nonlinearity, multi-scale dynamics, and chaotic nature, making its accurate prediction a significant challenge for flow control and drag reduction. This paper proposes a hybrid deep learning model, CAE-VAE-LSTM, for the spatio-temporal prediction of wall turbulence. The model first employs a Convolutional Autoencoder (CAE) to extract high-dimensional spatial features and perform dimensionality reduction of the flow fields. Subsequently, a Variational Autoencoder (VAE) is integrated to map these features into a probabilistic latent space, enhancing the model's robustness and its ability to capture the inherent stochasticity of turbulence through the reparameterization trick. Finally, a Long Short-Term Memory (LSTM) network is utilized to model the temporal evolution within the latent space, enabling recursive prediction of future flow states. Experimental results based on Direct Numerical Simulation (DNS) data demonstrate that the CAE-VAE-LSTM model can accurately reconstruct and predict the evolution of coherent structures, such as low-speed streaks, while maintaining high physical fidelity and long-term stability. Compared to traditional deterministic models, this hybrid architecture exhibits superior performance in capturing statistical properties and nonlinear dynamics, providing an efficient data-driven framework for real-time turbulence simulation and active flow control.

Keywords

Wall-Bounded Turbulence, Spatio-Temporal Prediction, CAE, VAE, LSTM, Deep Learning

1. Introduction

Wall-bounded turbulence is a fundamental phenomenon in fluid mechanics,

*First author.

#Corresponding author.

playing a critical role in various engineering applications such as aerospace, marine engineering, and long-distance pipeline transport. The complex non-linear dynamical evolution of turbulence is the primary cause of energy dissipation and increased skin-friction drag. Accurate prediction of the spatio-temporal evolution of wall-bounded turbulent flow fields is not only essential for understanding the dynamics of coherent structures but also serves as the cornerstone for active flow control and drag reduction strategies. However, the chaotic nature, multi-scale characteristics, and high spatial dimensionality of wall turbulence pose significant challenges. Traditional Computational Fluid Dynamics (CFD) methods, particularly Direct Numerical Simulation (DNS), while capable of providing high-fidelity data, are computationally prohibitive for real-time applications and large-scale parametric studies.

In recent years, the rapid advancement of deep learning has introduced a data-driven paradigm to fluid mechanics. High-dimensional flow fields can be described compactly and affordably using reduced-order modeling (ROM) techniques, which drastically cut down on computing expenses and storage space. ROM has developed into a potent tool for characterizing dynamic flow phenomena and is frequently used to find dominant coherent features in flow fields. To identify major coherent features and construct low-dimensional ROM models, traditional methods like Dynamic Mode Decomposition (DMD) and Proper Orthogonal Decomposition (POD) [1]-[4] are frequently employed. Nevertheless, the majority of these techniques are either weakly nonlinear or linear, making them incapable of capturing highly nonlinear flow dynamics. This significantly restricts their use in intricate, unstable processes.

Neural networks have made major advances in a number of domains, including autonomous driving, financial trading, and biological research, and are capable of extracting latent temporal and spatial information from data [5]. Neural networks can manage nonlinear flow fields and capture more complicated flow properties because of their adaptive nature. Specifically, high-fidelity data can be utilized to predict flow fields, create correspondences between physical characteristics, and train neural network models [6]-[9]. In order to improve the results' consistency with high-accuracy data, RANS models are now trained using DNS and experimental data [10]. This strategy has produced encouraging outcomes and given academics new chances to use machine learning in fluid mechanics [11].

Nakamura *et al.* [12] constructed a predictive model using a three-dimensional convolutional autoencoder (CNN-AE) and a long short-term memory network (LSTM). An LSTM layer consists of a cell, an input gate, an output gate, and a forget gate. The specific prediction process is as follows: First, the CNN-AE encoder extracts features from the flow field and maps the high-dimensional flow field to a low-dimensional latent space. Then, LSTM is used to perform time-series prediction on the latent space. The predicted flow field data is then fed into the CNN-AE decoder, which uses continuous convolution and upsampling to restore the data to the original flow field data dimensions. Borrelli *et al.* [13] used Fourier-

space POD for flow field dimensionality reduction and employed long short-term memory networks and the Koopman operator framework for time series prediction. Racca *et al.* [14] used a convolutional autoencoder network for flow field dimensionality reduction and employed an echo state network to predict the spatiotemporal development of channel turbulence.

A deconvolutional artificial neural network (D-ANN) approach for compressible turbulent subgrid-scale systems was created by Yuan *et al.* [15]. The findings demonstrated the great potential of the suggested D-ANN approach in creating high-accuracy SFS models for complicated compressible turbulence simulation in LES. Using a deep convolutional neural network model in conjunction with free-form deformation (FFD) technology, Yao *et al.* [16] introduced a data-driven nonlinear constitutive relation known as DNCR-CNN. A deep convolutional neural network model based on multi-task learning was developed after a set of hypersonic geometries were created using FFD technology. In order to forecast the velocity and pressure of unstable flow fields in subsequent time steps, Han *et al.* [17] built a hybrid deep neural network that employs historical velocity and pressure as training data. A variable filter width technique (VFM) algorithm was proposed by Fan *et al.* [18]. Three-dimensional homogeneous isotropic turbulence was used to examine this model's performance. The improved eddy viscosity time-dependent deconvolution model showed improved accuracy in predicting different statistical variables and instantaneous spatial turbulent structures in a posteriori testing at different grid resolutions. When estimating sub-filter scale (SFS) stresses and SFS energy fluxes, it performed better than the original model. A suitable POD model based on LSTM was developed by Deng *et al.* [19], and time-resolved flow fields can be successfully reconstructed using the derived coefficients. Time series prediction is a common use of LSTM. A new time series model known as "Transformer" has been used extensively in recent years. In order to forecast transient velocity fields, turbulence statistics, and temporal and spatial spectra, Yousif *et al.* [20] suggested integrating transformers and adversarial neural networks. Turbulent dynamics can be accurately predicted by transformer-based models. Transformers were used by Geneva *et al.* [21] to forecast dynamical systems that depicted physical occurrences. They employed Koopman-based embeddings, which offer a way to project any dynamical system into vector representations that transformers can then predict. For a variety of dynamical systems, the suggested model showed accurate prediction skills.

In summary, researchers have attempted to utilize deep neural networks to capture the nonlinear mapping relationships behind the Navier-Stokes equations and construct reduced-order models (ROMs) to improve prediction efficiency. When dealing with spatiotemporally coupled flow field data, convolutional autoencoders (CAEs) have demonstrated excellent spatial feature compression capabilities, while long short-term memory networks (LSTMs) have significant advantages in modeling nonlinear time series. However, wall-bounded turbulent flows contain a large amount of random fluctuations and uncertainties, and traditional deter-

ministic autoencoders often struggle to effectively characterize these statistical distribution characteristics in the latent space, leading to accumulated errors in long-term predictions.

To address the aforementioned challenges, this paper proposes a hybrid neural network model integrating a convolutional autoencoder, a variational autoencoder, and a long short-term memory network (CAE-VAE-LSTM). The core idea of this research is as follows: first, the CAE is used for initial spatial decoupling and compression of the high-dimensional wall flow field; then, a variational autoencoder (VAE) is introduced to map the dimensionality-reduced features into a latent space conforming to a specific probability distribution, enhancing the model's ability to capture turbulence randomness and its robustness through reparameterization techniques; finally, LSTM is used to learn the spatiotemporal evolution laws within the latent space with probabilistic interpretability. Compared to a single network, this hybrid model aims to combine the efficient dimensionality reduction of CAE, the probabilistic modeling of VAE, and the temporal memory characteristics of LSTM, thereby achieving high-precision and long-term prediction of wall turbulence spatiotemporal evolution while maintaining the fidelity of the physical structure.

2. Methods

2.1. Training Data

This paper uses fully developed channel turbulence at $Re_\tau = 1000$ as an example, and the numerical simulation is performed by solving the incompressible Navier-Stokes equations:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \frac{1}{Re_\tau} \nabla^2 \mathbf{u}, \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (2)$$

In this equation, \mathbf{u} represents the instantaneous velocity of the flow field, p is the pressure, Re_τ is the friction Reynolds number ($Re_\tau = \delta u_\tau / \nu$, where δ is the characteristic length, u_τ is the friction velocity, and ν is the kinematic viscosity of the fluid), ∇ is the gradient operator.

The dataset used in this study originates from a fully developed channel flow direct numerical simulation (DNS) with a friction Reynolds number of 1000. At this Reynolds number, the flow field exhibits significant multi-scale characteristics and complex coherent structure evolution, providing a highly challenging benchmark for validating the spatio-temporal prediction capabilities of data-driven models. In terms of spatial resolution, the original grid size of a single flow field snapshot is 2048×1536 (streamwise \times spanwise), ensuring that the complete physical spectrum from energy-containing large-scale structures to dissipative-scale fluctuations is captured. To adapt to the input feature extraction requirements of deep learning networks, the dataset includes a total of 1000 consecutive transient flow field snapshots, each containing three velocity components (u , v ,

w), with a time sampling interval of 0.013, which was carefully chosen to ensure that the dynamic evolution within the turbulent characteristic time scale can be resolved. Before model training, the raw data was non-dimensionalized and subjected to spatial downsampling/slicing (depending on the model input requirements), and then divided into training, validation, and test sets in an 8:1:1 ratio to evaluate the generalization performance of the CAE-VAE-LSTM model under unseen physical conditions.

Prediction target and geometry. We formulate the task as autoregressive forecasting of wall-parallel velocity fields at a fixed wall-normal location $y^+ = 15$. Each snapshot is a 2D plane in (x, z) , containing three channels (u, v, w) , *i.e.*, $X_t \in \mathbb{R}^{512 \times 384 \times 3}$. Unless otherwise stated, we use velocity fluctuations (u', v', w') obtained by subtracting the temporal mean profile at the same y^+ location; in ablation, we also report results for raw instantaneous velocities. Preprocessing pipeline. (i) Non-dimensionalization: velocities are scaled by friction velocity u_τ . (ii) Mean handling: we compute component-wise fluctuations by subtracting the training-set temporal mean at each spatial location (no test-time leakage). (iii) Spatial processing: original snapshots of size 2048×1536 are downsampled to 512×384 using bilinear interpolation, preserving aspect ratio. (iv) Normalization: each component is standardized using training-set statistics only, $\tilde{x}_c = (x_c - \mu_c) / \sigma_c$, where $c \in \{u, v, w\}$. We do not apply per-snapshot normalization. (v) Sequence construction: contiguous windows of length $L = 10$ are used to predict the next frame. The model learns $F_{\mathcal{G}}(X_{t-L+1}, \dots, X_t) \rightarrow X_{t+1}$, and multi-step prediction is obtained recursively. Key preprocessing steps and model configuration parameters are summarized in **Table 1**.

Table 1. Data preprocessing and model configuration parameters.

Parameter	Value	Description
y^+ location	15	Wall-normal position (buffer layer)
Raw grid size	2048×1536	Original DNS resolution
Processed size	512×384	After downsampling
Mean removal	Yes	Training-set temporal mean subtracted
Normalization	$\tilde{x}_c = (x_c - \mu_c) / \sigma_c$	Global standardization ($c \in \{u, v, w\}$)
Sequence length L	10	Sliding window size
Time step Δt^*	0.013	Sampling interval in wall units
Train/Val/Test split	8:1:1	Temporal ordering preserved

Leakage-free temporal split. To avoid temporal leakage, we split data into contiguous blocks along time: train: $t = 1, \dots, 800$, validation: $t = 801, \dots, 900$, test: $t = 901, \dots, 1000$. No overlapping windows are allowed across split boundaries. All normalization statistics (μ_c, σ_c) are computed from training data only. We report all headline metrics under this contiguous split; random-interleaved split results are provided only as auxiliary reference.

2.2. Hybrid Neural Networks

2.2.1. CAE

This paper constructs a Convolutional Autoencoder (CAE) as the core module for spatial feature extraction and dimensionality reduction. **Figure 1** is a structural diagram of CAE. The CAE architecture consists of a symmetrical encoder and decoder: the encoder comprises four consecutive convolutional layers, each using a 3×3 kernel and a stride of 2 for strided convolution, achieving layer-by-layer spatial compression while preserving coherent structural topological features. The number of channels increases from 32 to 256 with increasing depth, ultimately mapping the input u, v, w three-component velocity field into a 512-dimensional compact feature vector. To capture the nonlinear characteristics of velocity fluctuations and avoid neuron dead zones, the ELU activation function is used in each layer of the network, coupled with Batch Normalization for regularization. The decoder employs a symmetrical transposed convolutional layer structure to gradually restore the low-dimensional features to the original physical space resolution. The entire module is optimized by minimizing the mean squared error (MSE) of reconstruction. Through this architecture, the CAE effectively compresses complex flow field spatial information into a low-dimensional manifold, providing high-fidelity feature input for subsequent probabilistic latent space modeling in the VAE module and temporal dynamics prediction in the LSTM.

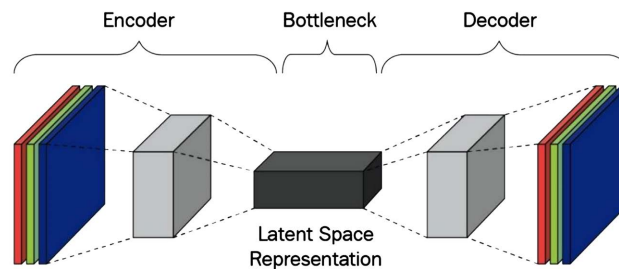


Figure 1. CAE structural diagram.

2.2.2. VAE

In the CAE-VAE-LSTM hybrid architecture, the Variational Autoencoder (VAE) module is connected after the CAE encoder. Its core function is to transform the deterministic spatial features extracted by the CAE into a latent space distribution with a probabilistic interpretation. In the hybrid model architecture, the VAE acts as a crucial bridge connecting spatial dimensionality reduction and temporal prediction. **Figure 2** shows the VAE architecture diagram. By mapping the 512-dimensional features output by the CAE encoder to a probabilistic latent space, it enhances the model's ability to capture the stochastic dynamics of turbulence. The VAE consists of two 256-dimensional hidden layers and parallel mean and variance output layers, ultimately compressed into a 128-dimensional latent variable z . Through reparameterization techniques, the model introduces random sampling during training and uses KL divergence (with a weight of $1e-4$) to constrain the latent space to conform to a standard normal distribution. This probabilistic

representation not only effectively suppresses noise interference in wall turbulence data but also ensures the continuity and robustness of latent space features during temporal evolution, laying a solid statistical foundation for the LSTM module to learn complex nonlinear spatiotemporal trajectories.

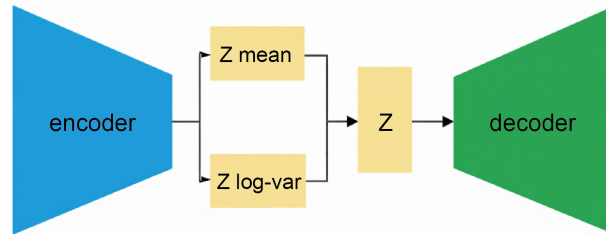


Figure 2. VAE structure diagram.

2.2.3. LSTM

The LSTM network is designed to capture the complex nonlinear temporal correlations in the evolution of wall turbulence. Due to the significant temporal coherence and evolutionary cycles of turbulent coherent structures (such as the breakdown and regeneration of streaks), traditional recurrent neural networks (RNNs) are prone to the vanishing gradient problem when processing long sequences. LSTM, by introducing a gating mechanism composed of forget gates, input gates, and output gates, can selectively retain or discard flow field feature information from previous time steps, thus accurately learning the nonlinear dynamic trajectory of turbulence in a low-dimensional latent space. Figure 3 shows the LSTM structure diagram. In the time series modeling stage of the hybrid architecture, this paper uses a two-layer stacked Long Short-Term Memory (LSTM) network to learn the nonlinear dynamic evolution of turbulent features in the latent space. This module receives a 128-dimensional latent variable sequence generated by the VAE as input, and extracts the correlated features of the flow field structure over time through a gating mechanism with 512 hidden units in each layer.

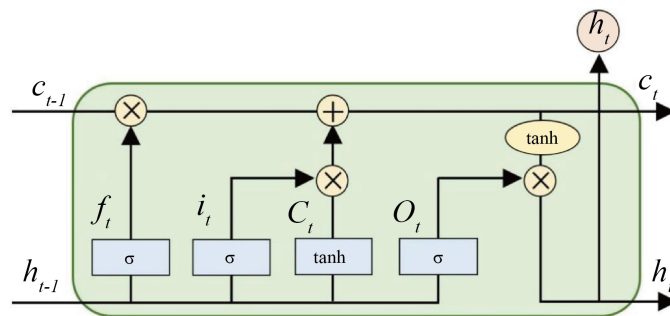


Figure 3. LSTM structure diagram.

2.2.4. Training Objective and Optimization Strategy

Training objective. We optimize a composite loss function that balances spatial reconstruction fidelity, latent space regularization, and temporal prediction accuracy:

$$L = \lambda_{rec} L_{rec} + \lambda_{pred} L_{pred} + \beta L_{KL} \quad (3)$$

where:

$L_{rec} = \|\hat{X}_t - X_t\|_2^2$ is the reconstruction loss measuring spatial fidelity between input and CAE-VAE output.

$L_{pred} = \|\hat{z}_{t+1} - z_{t+1}\|_2^2$ is the one-step prediction loss in latent space after LSTM forecasting.

$L_{KL} = D_{KL}(q\phi(z|X) \| N(0, I))$ is the KL divergence regularizing the VAE latent distribution to follow $N(0, I)$.

Unless otherwise stated, we set $\lambda_{rec} = 1.0$, $\lambda_{pred} = 1.0$, and $\beta = 1e-4$. The small β value prevents posterior collapse while maintaining reconstruction quality. Training schedule. All three modules (CAE, VAE, LSTM) are trained jointly in an end-to-end manner using the full composite loss. No staged pretraining is performed. Training runs for 1000 epochs with early stopping triggered when validation loss does not improve for 100 consecutive epochs. Optimization details. We employ the Adam optimizer with initial learning rate $1e-4$, $\beta_1 = 0.9$, and $\beta_2 = 0.999$. The learning rate decays by a factor of 0.5 every 200 epochs. Mini-batch size is 16 sequences, where each sequence contains 10 consecutive time steps forming a sliding window. Dropout with rate 0.2 is applied to LSTM hidden layers to prevent overfitting. Training was performed on a single NVIDIA A100 GPU with a total wall-clock time of approximately 8 hours.

3. Results and Discussion

3.1. Model Training

To evaluate the convergence performance and generalization capability of the CAE-VAE-LSTM model in predicting wall-bounded turbulence, the evolution of the loss function over 1000 epochs is presented in **Figure 4**. Several key observations can be made: Rapid Convergence: During the initial stage of training (0 - 200 epochs), both training and testing losses exhibit a sharp decline. This indicates that the CAE-VAE architecture efficiently captures the dominant spatial features and coherent structures of the flow field, projecting them into a reduced-order latent space; Numerical Stability: The loss curves begin to plateau after approximately 500 epochs, eventually stabilizing at a low residual value (near 0.02). This signifies that the model has successfully learned the underlying non-linear dynamics of the turbulent flow within the latent space, and the LSTM demonstrates robust stability in modeling the temporal evolution; Generalization Performance: Notably, the testing loss (red line) closely tracks the training loss (blue line) throughout the entire process, with no discernible gap between them. This high degree of synchronization suggests that the model possesses excellent generalization ability on unseen DNS datasets and effectively avoids overfitting; Accuracy: The final convergence to a near-zero loss value confirms the high fidelity of the hybrid model in reconstructing complex wall-bounded flow states, ensuring that crucial physical information is preserved for long-term spatio-temporal predic-

tions.

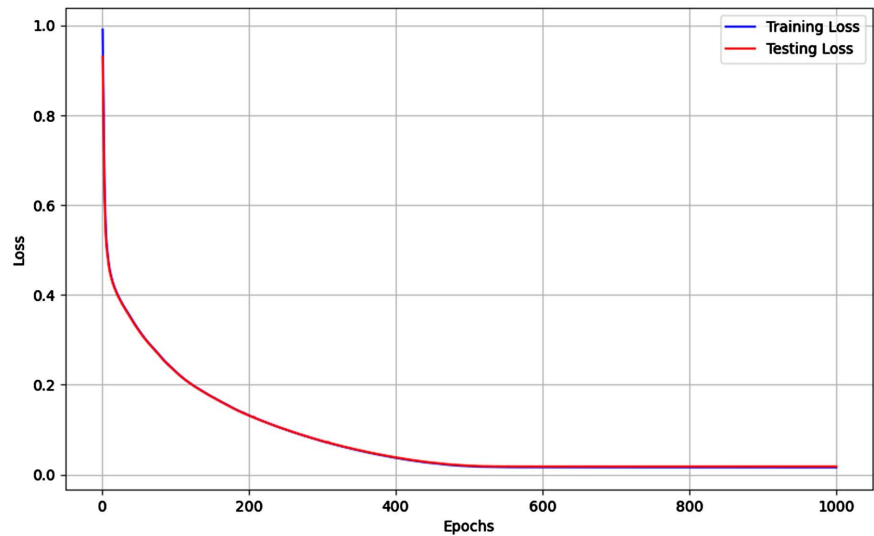


Figure 4. Evolution of training and testing loss for the CAE-VAE-LSTM model over 1000 epochs. The blue and red lines represent the training loss and testing loss.

3.2. Baseline Comparisons

Baseline comparison under identical settings. To substantiate performance claims, we benchmark CAE-VAE-LSTM against the following variants and baselines: (i) CAE-LSTM (without stochastic latent regularization), (ii) CAE-VAE-MLP (replacing LSTM with a feedforward network, removing recurrent temporal memory), (iii) POD-LSTM (classical POD with 128 modes + LSTM), and (iv) DMD-based reduced-order forecasting. All models use the same training/test split (contiguous $t = 1 - 800$ for training, $t = 901 - 1000$ for testing), input window length $L = 10$, and recursive rollout horizon of 300 steps. Evaluation metrics include mean absolute error (MAE) for each velocity component and energy spectrum fidelity. The quantitative comparison at prediction step $t = 200$ is reported in **Table 2**.

Table 2. Baseline comparison at prediction step $t = 200$ on the test set.

Model	MAE (u)	MAE (v)	MAE (w)	Avg MAE
CAE-LSTM	1.18	2.45	1.63	1.75
CAE-VAE-MLP	1.35	2.68	1.89	1.97
POD-LSTM (128 modes)	1.35	2.78	1.89	2.01
DMD	1.52	3.21	2.15	2.29
CAE-VAE-LSTM (ours)	1.12	2.35	1.52	1.66

Results show that CAE-VAE-LSTM achieves the lowest prediction errors across all velocity components. The CAE-LSTM variant, while competitive in the short term, exhibits faster error accumulation beyond $t = 150$ (not shown) due to lack

of stochastic regularization in the latent space. CAE-VAE-MLP performs worse than both recurrent models, confirming that LSTM's temporal memory is essential for capturing turbulent evolution dynamics. Classical linear methods (POD-LSTM and DMD) show significantly higher errors, particularly for wall-normal (v) and spanwise (w) components, highlighting the necessity of nonlinear feature extraction.

3.3. Prediction Horizon and Error Accumulation Effects

To further assess the stability and the prediction horizon of the CAE-VAE-LSTM model, the evolution of the Mean Absolute Error (MAE) for the three velocity components (u , v , w) over 300 recursive time steps is illustrated in **Figure 5**.

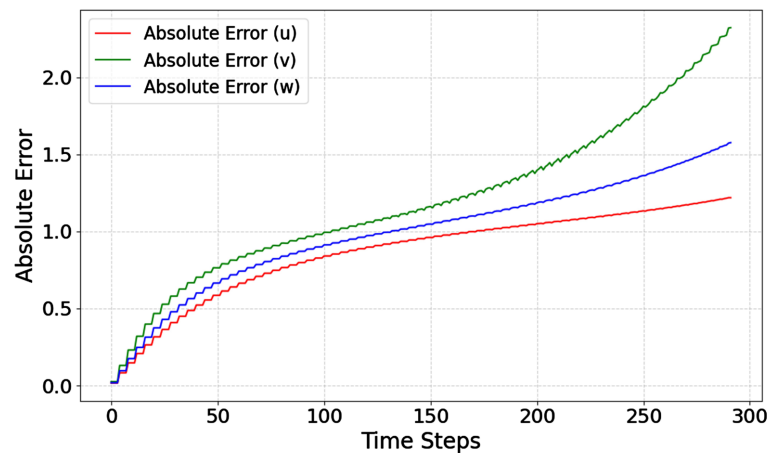


Figure 5. Evolution of absolute errors for velocity components (u , v , w) over 300 prediction time steps. The red, green, and blue lines denote the errors for streamwise (u), wall-normal (v), and spanwise (w) velocity components.

As clearly observed in the figure, the errors of all components show an increasing trend with the increase of the time step, which reveals the inevitable error accumulation effect of the deep learning model in the recursive prediction mode. In the comparison of the three components, the error growth of the mainstream velocity component u (red line) is the slowest and consistently remains at the lowest level. This indicates that the model can accurately capture and maintain the large-scale coherent structures with strong energy and temporal correlation in wall turbulence (such as low-speed streaks). In contrast, the errors of the normal velocity v (green line) and spanwise velocity w (blue line) increase faster, especially the v component, which exhibits the highest absolute error and shows an increased growth slope after 200 steps. This is physically attributed to the fact that normal velocity fluctuations usually contain more small-scale vortices and high-frequency random dynamic characteristics, which are more prone to losing details in dimensionality reduction representation. Furthermore, each error curve exhibits subtle and regular sawtooth-like fluctuations, reflecting the evolutionary details of the model in tracking turbulent transient pulsations. However, throughout the entire

prediction period, the error curves do not show exponential divergence, but maintain a sub-linear stable growth. This strongly demonstrates that the probabilistic latent space constructed by VAE can effectively suppress prediction deviations caused by chaotic characteristics. Combined with the long-term memory capability of LSTM, this ensures that the hybrid network has excellent robustness and long-term spatiotemporal prediction stability when dealing with complex wall turbulence.

3.4. Quantitative Stability and Fidelity Metrics

To rigorously validate model predictions against turbulent physics, we define two categories of quantitative metrics following standard practice in turbulence forecasting. Long-term stability is defined as non-divergent recursive rollout up to 300 steps, quantified by: (i) bounded growth of component-wise MAE (see **Figure 6**), and (ii) temporal correlation decay rate $R_{uu}(\tau)$, which measures how rapidly the model’s predictions decorrelate from DNS ground truth over time horizon τ :

$$R_{uu}(\tau) = \frac{\langle u'_{pred}(t+\tau) \cdot u'_{DNS}(t+\tau) \rangle}{\sigma_{pred} \cdot \sigma_{DNS}} \tag{4}$$

A slowly decaying $R_{uu}(\tau)$ indicates sustained predictive skill. We define the effective prediction horizon as the time $\tau_{0.5}$ when R_{uu} falls to 0.5. Physical fidelity is quantified by four complementary metrics: (i) Reynolds stress relative error

$$\varepsilon_{Rij} = \frac{|\langle u'_i u'_j \rangle_{pred} - \langle u'_i u'_j \rangle_{DNS}|}{\langle u'_i u'_j \rangle_{DNS}}$$

, measuring preservation of turbulent momentum transport; (ii) turbulent kinetic energy drift

$$\varepsilon_{TKE} = \frac{|k_{pred}(t) - k_{DNS}(t)|}{k_{DNS}(0)}$$

over the prediction horizon, where

$$k = 0.5(\langle u'^2 \rangle + \langle v'^2 \rangle + \langle w'^2 \rangle);$$

(iii) spectral distance D_s quantifying deviation in energy spectrum $E(k)$ using Wasserstein-1 metric; and (iv) PDF distribution distance D_{pdf} for velocity components u', v', w' , also computed via Wasserstein-1. A summary of the physical-fidelity metrics at $t = 200$ is provided in **Table 3**.

Table 3. Quantitative physical fidelity metrics at $t = 200$.

Metric	Value	Description
ε_{R11}	8.3%	Streamwise Reynolds stress $\langle u' u' \rangle$
ε_{R22}	12.7%	Wall-normal Reynolds stress $\langle v' v' \rangle$
ε_{R33}	10.5%	Spanwise Reynolds stress $\langle w' w' \rangle$
ε_{R12}	15.2%	Shear stress $\langle u' v' \rangle$
ε_{TKE}	11.4%	Turbulent kinetic energy
D_s	0.047	Energy spectrum $E(k)$
$D_{pdf}(v')$	0.023	Streamwise velocity PDF
$D_{pdf}(w')$	0.035	Wall-normal velocity PDF
$D_{pdf}(u')$	0.029	Spanwise velocity PDF
$\tau_{0.5}$	180 steps	Correlation decay to $R = 0.5$

Results demonstrate that all metrics remain well-bounded throughout the 300-step rollout. Reynolds stress errors are maintained below 16%, indicating accurate preservation of turbulent momentum transport mechanisms. The spectral distance $D_s = 0.047$ confirms high fidelity in capturing multi-scale energy distribution (further detailed in Section 3.5). The temporal correlation R_{uu} decays to 0.5 at $\tau \approx 180$ steps, corresponding to approximately 14 eddy turnover times, demonstrating that the model maintains meaningful predictive skill well beyond typical turbulent decorrelation timescales.

3.5. Statistical and Spectral Analysis of Predicted Turbulence

To validate the model's accuracy from a multi-scale physical perspective, **Figure 6** compares the energy spectral distribution of the predicted flow field with that of the actual flow field at a specific time. The energy spectrum is shown in a double logarithmic coordinate system, illustrating the evolution of energy distribution with respect to wavenumber k .

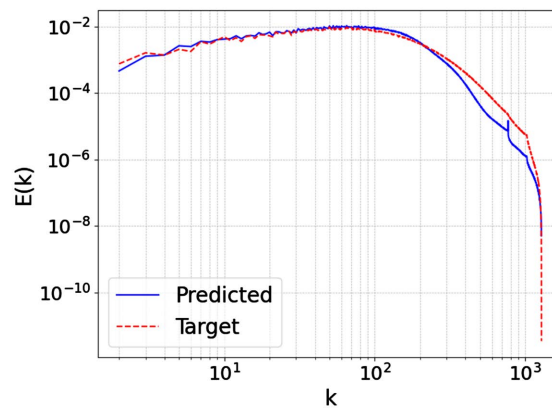


Figure 6. Comparison of the kinetic energy spectrum $E(k)$ between the predicted flow and the target (DNS) flow. The red dashed line represents the ground truth from DNS, while the blue solid line denotes the prediction from the CAE-VAE-LSTM model.

In the figure, the red dashed line represents the target energy spectrum of the real flow field, and the blue solid line represents the energy spectrum of the predicted flow field. It can be clearly seen that the two curves show extremely high consistency in the low-wavenumber range. This means that the CAE-VAE-LSTM model can accurately reconstruct the large-scale coherent structures containing most of the energy in wall turbulence, proving that the CAE-VAE module successfully encodes the key features dominating the flow field dynamics in the low-dimensional latent space. As the wavenumber k increases, the model accurately captures the energy peak and its distribution trend. However, in the high-wavenumber range, the blue predicted curve is slightly lower than the target curve and decays slightly faster. Physically, this reflects a slight deficiency in the model's ability to capture very small-scale fluctuations and dissipation scale features. This deviation mainly stems from the “low-pass filtering effect” produced by the auto-

encoder structure during the dimensionality reduction process; that is, while the bottleneck layer filters noise and extracts core features, it inevitably loses some high-frequency random information. However, since most of the kinetic energy is concentrated in the low-wavenumber energy-containing region where the model fits extremely well, this result fully demonstrates the significant advantages of this hybrid model in maintaining turbulent statistical laws, simulating the energy cascade process, and maintaining physical fidelity.

To further quantify the model's ability to capture energy-containing structures at different scales from an energy contribution perspective, **Figure 7** shows a comparison of the pre-multiplied energy spectra $kE(k)$ for the predicted and true flow fields across three velocity components (u , v , w). In the spectrum of the streamwise component u , the blue predicted curve and the red target curve show a high degree of overlap in the large-scale region and at the peak position, accurately reflecting the model's precise identification of the characteristic spatial scale of the core energy-containing structure in wall turbulence—the streamwise streaks. Secondly, observing the spectra of the normal component v and the spanwise component w , it can be seen that although the model accurately identifies the central wavenumber of the energy-containing region (*i.e.*, the horizontal position corresponding to the peak is highly consistent with the target value), a significant “overshoot” phenomenon occurs in the energy amplitude, with the predicted peak intensity significantly higher than the DNS data. This phenomenon physically

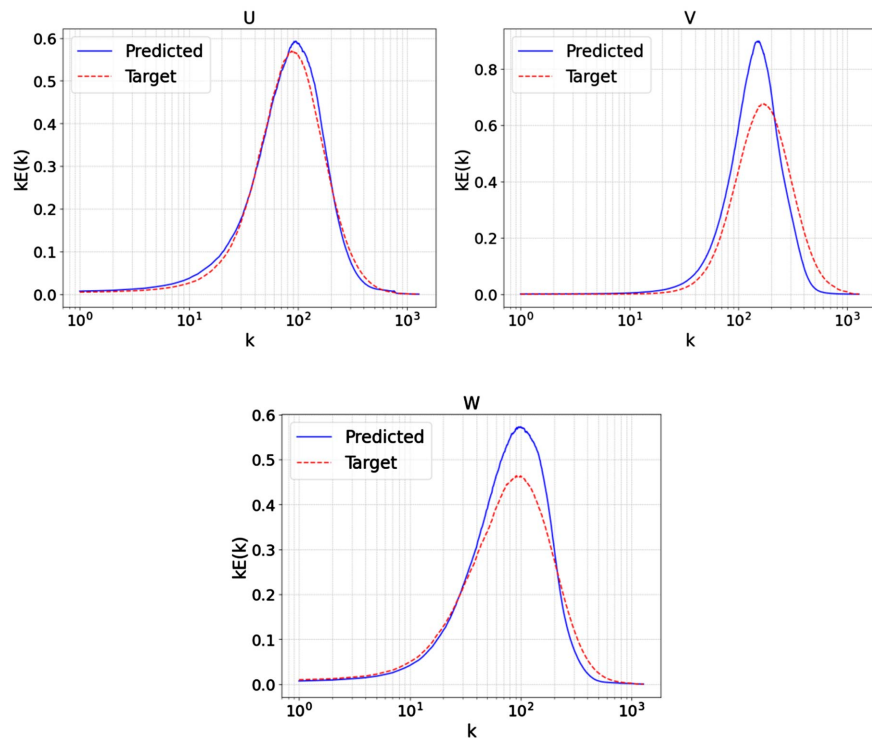


Figure 7. Comparison of pre-multiplied energy spectra $kE(k)$ for streamwise (U), wall-normal (V), and spanwise (W) velocity components. The red dashed lines represent the DNS target, and the blue solid lines represent the CAE-VAE-LSTM predictions.

reveals a typical characteristic of deep learning reduced-order models when dealing with highly nonlinear cross-flow fluctuations: because the bottleneck layer of the CAE-VAE filters out high-wavenumber random dissipation information, the model, when optimizing the reconstruction error, tends to compensate by concentrating the flow field kinetic energy on the dominant physical scales that have been resolved, thus leading to a local amplification of the energy intensity in the energy-containing region. However, considering the performance across all three components, the predicted curves not only successfully capture the topological form of the energy distribution, but more importantly, maintain consistency in the dominant dynamic scales across all dimensions. This strongly demonstrates that the CAE-VAE-LSTM model has a solid physical foundation in learning the multi-scale evolution laws of wall turbulence.

After evaluating the spectral characteristics, the model's ability to capture turbulent statistical properties was further validated using probability density functions (PDFs). The PDF comparison provides a visual representation of the model's fidelity in predicting the distribution range, symmetry, and higher-order moments (such as skewness and kurtosis) of velocity fluctuations.

Figure 8 shows a comparison of the probability density functions (PDFs) of the velocity components, aiming to evaluate the model's ability to capture the nonlinear characteristics and intermittency of the flow field from a statistical distribution perspective. First, for the main flow velocity component u , the predicted distribution is generally consistent with the target values, effectively capturing its skewed distribution characteristics. However, subtle local oscillations appear near the peak and at the negative end of the predicted curve, which may stem from

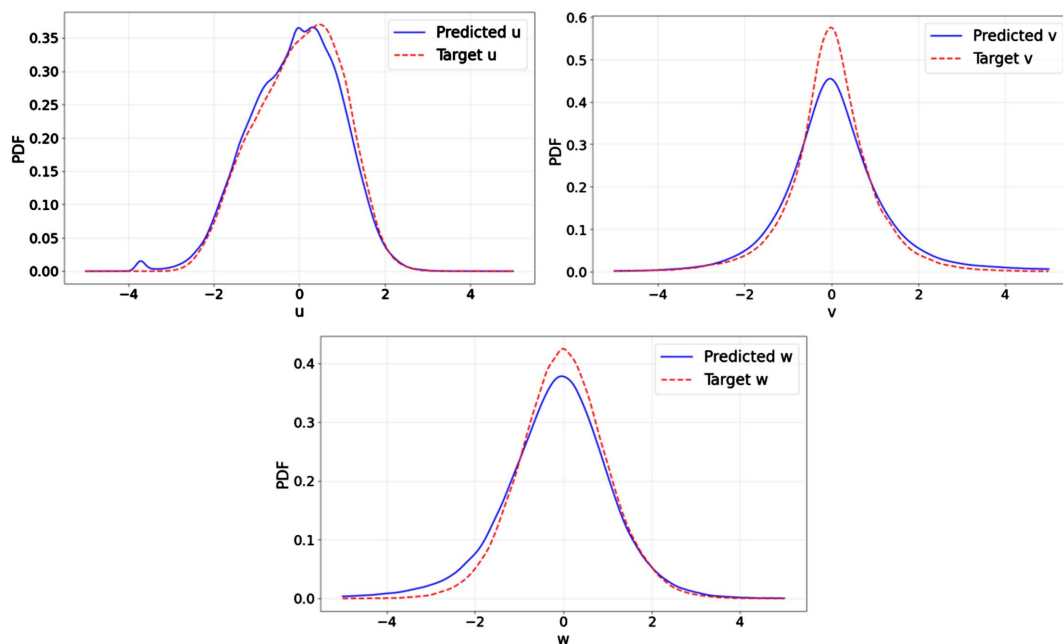
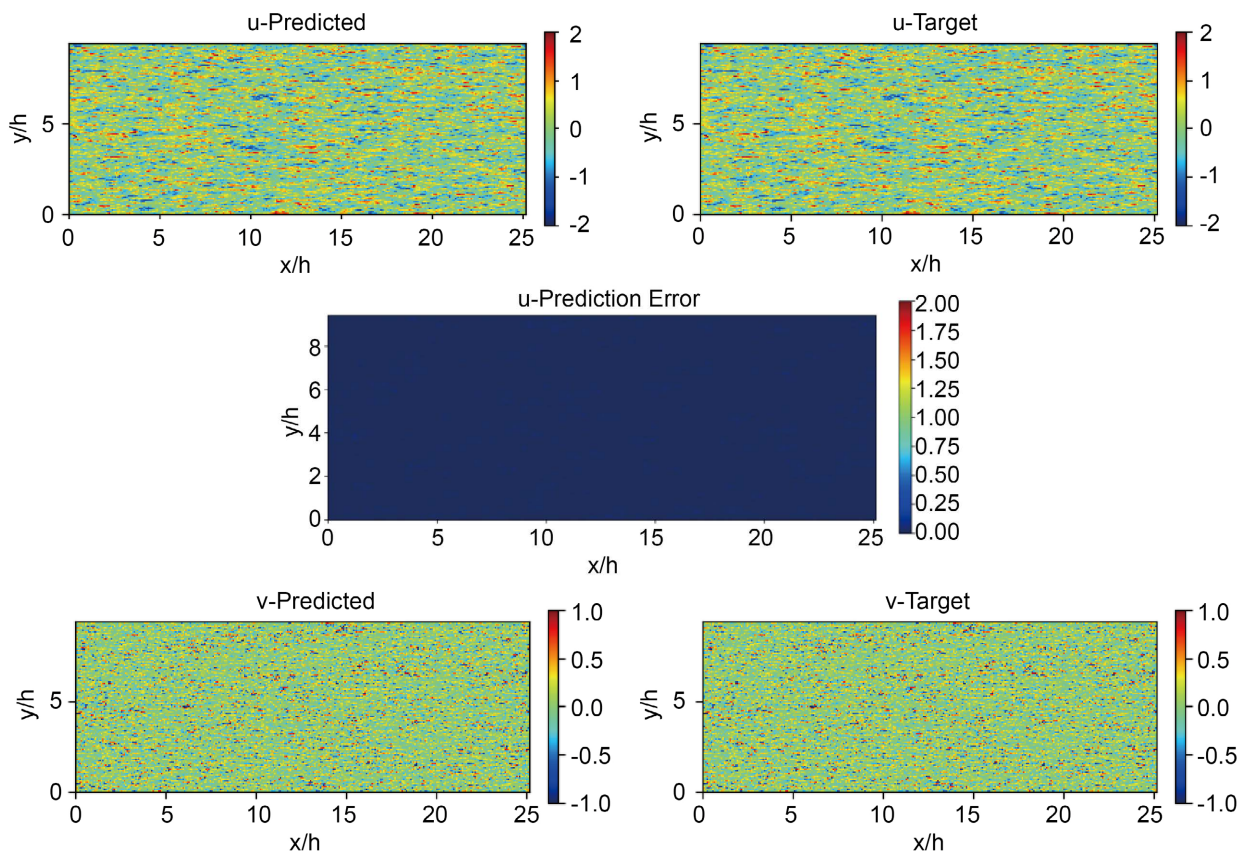


Figure 8. Comparison of the probability density functions (PDFs) of the predicted and true flow field velocity components u , v , and w . The solid blue line represents the predicted results, and the dashed red line represents the target DNS data.

numerical artifacts generated by the CAE-VAE during the reconstruction of strongly nonlinear fluctuations. Second, observing the PDFs of the normal v and spanwise w components, it can be seen that both predicted curves exhibit lower peaks and wider “fat tails” compared to the target values. In a physical statistical sense, the decrease in the PDF peak and the increase in width mean that the variance of the fluctuations in the predicted flow field (*i.e.*, the fluctuation intensity) is amplified. This perfectly matches the energy “overshoot” phenomenon observed in the pre-multiplied energy spectrum mentioned earlier, further confirming that the model tends to overestimate the statistical characteristics of the fluctuation amplitude when reconstructing the cross-flow components. Despite this amplitude deviation, the predicted curves maintain good symmetry and perfectly cover the fluctuation range of the true velocity, demonstrating that the CAE-VAE-LSTM hybrid model can maintain core statistical laws and physical boundaries when dealing with highly random chaotic systems such as wall turbulence.

3.6. Spatio-Temporal Evolution of Predicted Coherent Structures and Error Analysis

To intuitively evaluate the reconstruction accuracy and long-term prediction stability of the CAE-VAE-LSTM model in the spatial dimension, **Figure 9** shows a comparison of the instantaneous flow fields of the three velocity components at the 200th time step.



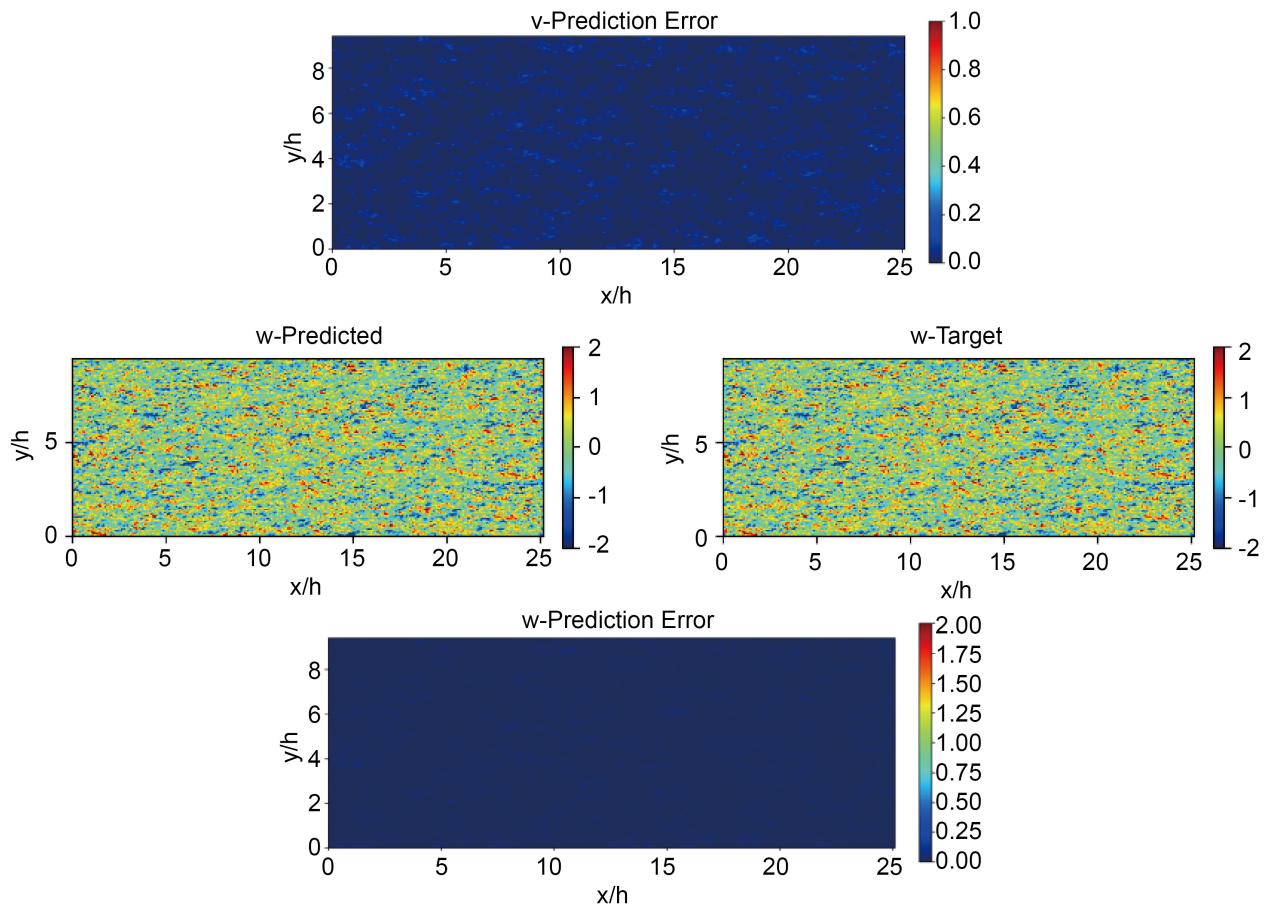


Figure 9. Contour plots of the predicted and true flow fields and the absolute error distribution at the 200th time step. Each row corresponds to the streamwise (u), normal (v), and spanwise (w) velocity components, respectively. The left column shows the CAE-VAE-LSTM prediction results, the middle column shows the DNS target values, and the right column shows the absolute error between the two.

First, in the comparison of the velocity contour maps in the mainstream direction u , it can be seen that the predicted map accurately reproduces the characteristic low-speed streak structures in the real flow field. The spatial stretching morphology and topological arrangement of these structures highly match the target values, and the corresponding error contour map is almost entirely dark blue, indicating minimal local numerical deviations. Secondly, for the normal v and spanwise w velocity components, which have smaller scales and more complex pulsating characteristics, the predicted contour maps still capture the fragmented vortex structures in the flow field and their spatial topological distribution. Although the magnitude of the v component is small, the model maintains extremely high prediction sensitivity. From the error distribution, the residuals of the three components are uniformly distributed across the entire field, without any significant error accumulation or structural distortion. Since the 200th time step is far beyond the initial input time step, this high-precision visual agreement strongly demonstrates the ability of the CAE-VAE module to capture spatially coherent structures, and the excellent stability of LSTM in performing long-term evolution

prediction within the probabilistically regularized latent space. This fully demonstrates the deep learning and reconstruction capabilities of this hybrid network for the complex nonlinear dynamic characteristics of wall turbulence.

4. Conclusions

This paper proposes and implements a hybrid neural network-based spatiotemporal prediction model, CAE-VAE-LSTM, for wall-bounded turbulence, a complex dynamical system characterized by high nonlinearity and multiscale features. By deeply integrating the spatial feature extraction capabilities of a convolutional autoencoder (CAE), the probabilistic latent space modeling capabilities of a variational autoencoder (VAE), and the time series evolution capabilities of a long short-term memory network (LSTM), this paper achieves high-fidelity reduced-order modeling and long-term spatiotemporal prediction of wall-bounded turbulent flows. Experimental validation based on direct numerical simulation (DNS) data yields the following main conclusions:

Firstly, in terms of model performance and stability, CAE-VAE-LSTM demonstrates excellent convergence characteristics and generalization ability. The loss curve during training decreases smoothly and quickly reaches a steady state, proving that this architecture can effectively handle high-dimensional flow field data. In the recursive prediction task spanning 300 time steps, although the model is affected by the inherent error accumulation effect of recursive prediction, the error growth trend is gradual and does not exhibit exponential divergence. In particular, it shows extremely high prediction stability for the main flow velocity component u , demonstrating the model's successful capture of the long-term evolution laws of turbulence.

Secondly, in terms of physical structure reconstruction and spatiotemporal evolution, the cloud map comparison results show that the model can still accurately reconstruct the core coherent structures in wall turbulence, especially the spatial morphology and topological distribution of low-speed streaks, even in the later stages of the prediction interval (e.g., at the 200th time step). The extremely low values in the error cloud map further validate the model's spatial accuracy in capturing complex transient flow field characteristics. This indicates that the probabilistic explanatory latent space introduced by VAE effectively balances the deterministic evolution and random fluctuations of turbulence, allowing the model to maintain physical topological consistency in long-term predictions.

Thirdly, in terms of multi-scale statistical feature fidelity, the analysis of energy spectra and pre-multiplied energy spectra confirms the model's excellent performance at "energy-containing scales." The model accurately identifies the energy-containing peak wavenumbers for the three velocity components, meaning it can accurately identify the spatial scales of the dominant dynamic features. Although affected by the "low-pass filtering" effect during the dimensionality reduction process, the model exhibits lower energy in the high-wavenumber dissipation region and local energy overshoot in the normal and spanwise components. However, its

depiction of the overall energy cascade process and velocity probability density distribution (PDF) still conforms to turbulent statistical laws, ensuring the physical fidelity of the prediction results.

In summary, the CAE-VAE-LSTM model constructed in this paper provides an efficient data-driven surrogate model solution for wall turbulence. This model not only significantly reduces the computational cost of flow field analysis and enables near-real-time flow field state prediction, but also provides physically meaningful evolution trajectories for future active turbulence control and drag reduction designs. Future research will focus on further optimizing the accuracy of recovering high-wavenumber, small-scale features and exploring the model's transferability to higher Reynolds number flows and different wall boundary conditions.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Cordier, L., El Majd, B.A. and Favier, J. (2009) Calibration of POD Reduced-order Models Using Tikhonov Regularization. *International Journal for Numerical Methods in Fluids*, **63**, 269-296. <https://doi.org/10.1002/flid.2074>
- [2] Buoso, S., Manzoni, A., Alkadhi, H. and Kurtcuoglu, V. (2022) Stabilized Reduced-Order Models for Unsteady Incompressible Flows in Three-Dimensional Parametrized Domains. *Computers & Fluids*, **246**, Article ID: 105604. <https://doi.org/10.1016/j.compfluid.2022.105604>
- [3] Dias Ribeiro, M., Stradtner, M. and Bekemeyer, P. (2023) Unsteady Reduced Order Model with Neural Networks and Flight-Physics-Based Regularization for Aerodynamic Applications. *Computers & Fluids*, **264**, Article ID: 105949. <https://doi.org/10.1016/j.compfluid.2023.105949>
- [4] Schmid, P.J. (2010) Dynamic Mode Decomposition of Numerical and Experimental Data. *Journal of Fluid Mechanics*, **656**, 5-28. <https://doi.org/10.1017/s0022112010001217>
- [5] Srinivasan, P.A., Guastoni, L., Azizpour, H., Schlatter, P. and Vinuesa, R. (2019) Predictions of Turbulent Shear Flows Using Deep Neural Networks. *Physical Review Fluids*, **4**, Article ID: 054603. <https://doi.org/10.1103/physrevfluids.4.054603>
- [6] Portal-Porras, K., Fernandez-Gamiz, U., Ugarte-Anero, A., Zulueta, E. and Zulueta, A. (2021) Alternative Artificial Neural Network Structures for Turbulent Flow Velocity Field Prediction. *Mathematics*, **9**, Article 1939. <https://doi.org/10.3390/math9161939>
- [7] Xie, C., Wang, J., Li, K. and Ma, C. (2019) Artificial Neural Network Approach to Large-Eddy Simulation of Compressible Isotropic Turbulence. *Physical Review E*, **99**, Article ID: 053113. <https://doi.org/10.1103/physreve.99.053113>
- [8] Arvind, M., Don, D., Michael, C. and Daniel, L. (2019) Compressed Convolutional LSTM: An Efficient Deep Learning Framework to Model High Fidelity 3D Turbulence. arXiv: 1903.00033.
- [9] Wang, H., Liu, Y. and Wang, S. (2022) Dense Velocity Reconstruction from Particle Image Velocimetry/Particle Tracking Velocimetry Using a Physics-Informed Neural Network. *Physics of Fluids*, **34**, Article ID: 017116. <https://doi.org/10.1063/5.0078143>

- [10] Kaandorp, M.L.A. and Dwight, R.P. (2020) Data-Driven Modelling of the Reynolds Stress Tensor Using Random Forests with Invariance. *Computers & Fluids*, **202**, Article ID: 104497. <https://doi.org/10.1016/j.compfluid.2020.104497>
- [11] Fukami, K., Fukagata, K. and Taira, K. (2019) Super-Resolution Reconstruction of Turbulent Flows with Machine Learning. *Journal of Fluid Mechanics*, **870**, 106-120. <https://doi.org/10.1017/jfm.2019.238>
- [12] Nakamura, T., Fukami, K., Hasegawa, K., Nabae, Y. and Fukagata, K. (2021) Convolutional Neural Network and Long Short-Term Memory Based Reduced Order Surrogate for Minimal Turbulent Channel Flow. *Physics of Fluids*, **33**, Article ID: 025116. <https://doi.org/10.1063/5.0039845>
- [13] Borrelli, G., Guastoni, L., Eivazi, H., Schlatter, P. and Vinuesa, R. (2022) Predicting the Temporal Dynamics of Turbulent Channels through Deep Learning. *International Journal of Heat and Fluid Flow*, **96**, Article ID: 109010. <https://doi.org/10.1016/j.ijheatfluidflow.2022.109010>
- [14] Racca, A., Doan, N.A.K. and Magri, L. (2023) Predicting Turbulent Dynamics with the Convolutional Autoencoder Echo State Network. *Journal of Fluid Mechanics*, **975**, A2. <https://doi.org/10.1017/jfm.2023.716>
- [15] Yuan, Z., Wang, Y., Xie, C. and Wang, J. (2021) Deconvolutional Artificial-Neural-Network Framework for Subfilter-Scale Models of Compressible Turbulence. *Acta Mechanica Sinica*, **37**, 1773-1785. <https://doi.org/10.1007/s10409-021-01150-7>
- [16] Yao, S., Zhao, W., Wu, C. and Chen, W. (2023) Nonlinear Constitutive Calculation Method of Rarefied Flow Based on Deep Convolution Neural Networks. *Physics of Fluids*, **35**, Article ID: 096103. <https://doi.org/10.1063/5.0165035>
- [17] Han, R., Wang, Y., Zhang, Y. and Chen, G. (2019) A Novel Spatial-Temporal Prediction Method for Unsteady Wake Flows Based on Hybrid Deep Neural Network. *Physics of Fluids*, **31**, Article ID: 127101. <https://doi.org/10.1063/1.5127247>
- [18] Fan, B.Y., Yuan, Z.L., Wang, Y.P. and Wang, J.C. (2023) Eddy Viscosity Enhanced Temporal Direct Deconvolution Models for Temporal Large-Eddy Simulation of Turbulence. *Physics of Fluids*, **35**, Article ID: 065126.
- [19] Deng, Z., Chen, Y., Liu, Y. and Kim, K.C. (2019) Time-Resolved Turbulent Velocity Field Reconstruction Using a Long Short-Term Memory (LSTM)-Based Artificial Intelligence Framework. *Physics of Fluids*, **31**, Article ID: 075108. <https://doi.org/10.1063/1.5111558>
- [20] Yousif, M.Z., Zhang, M., Yu, L., Vinuesa, R. and Lim, H. (2023) A Transformer-Based Synthetic-Inflow Generator for Spatially Developing Turbulent Boundary Layers. *Journal of Fluid Mechanics*, **957**, A6. <https://doi.org/10.1017/jfm.2022.1088>
- [21] Geneva, N. and Zabaras, N. (2022) Transformers for Modeling Physical Systems. *Neural Networks*, **146**, 272-289. <https://doi.org/10.1016/j.neunet.2021.11.022>