

A Reduced-Order Multi-Variable Surrogate Approach for Cross Flow Dynamics in Helical Coil Steam Generators Using POD and LSTM

Bumjin Cho, Minseop Song*

Departments of Nuclear Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul, 04761, Republic of Korea

*Corresponding author: hysms@hanyang.ac.kr

***Keywords : HCSG, Deep Learning, POD, Long Short-Term Memory, Multi-Variable**

1. Introduction

Small Modular Reactor (SMR) offer distinct advantages in safety, efficiency, and versatility compared to traditional large-scale reactors. The compact design of SMRs requires the development of highly integrated components, which presents the challenge of achieving high heat transfer efficiency within a limited space. In this context, the helical coil steam generator (HCSG) has gained attention as an effective solution. Its helical structure maximizes the heat transfer area, playing a crucial role in the thermal management system of SMRs. However, the tube arrangement, being perpendicular to the direction of flow, inevitably induces cross flow through rod bundles. This cross flow generates complex vortical and turbulent structures between the tube layers, which can significantly impact system stability by causing flow-induced vibration, flow instability, and non-uniform heat transfer, making it a major subject of research [1]. Therefore, precise analysis and control of the primary-side cross flow within HCSGs are essential for optimizing design and operational efficiency.

Within computational fluid dynamics (CFD), large eddy simulation (LES) has been utilized as a powerful tool to capture these complex turbulent phenomena. Previous studies have used LES-based CFD analysis to investigate the characteristics of cross flow and turbulent patterns inside HCSGs, providing deep insights into their thermo-hydraulic properties [2]. Prior research has established that the dynamic interactions within HCSGs are not governed by a single physical quantity but arise from strong physical coupling among flow variables, including velocity and pressure. While high-fidelity transient analysis methods like LES provide accurate and detailed spatio-temporal data on flow characteristics, this accuracy comes at the cost of immense computational resources and long simulation times. Since the analysis is highly time-consuming, high-fidelity CFD becomes an impractical option for applications requiring iterative calculations, such as design optimization or control systems.

To alleviate this computational burden, data-driven surrogate models based on deep learning have recently emerged as a promising alternative. A common methodology involves using a reduced-order model (ROM) to compress high-dimensional CFD data into a low-dimensional space, which is then fed into a time-

series prediction network. As a linear technique for compressing time-series flow fields, proper orthogonal decomposition (POD) has been actively applied [3]. A preceding study has predicted the temporal evolution of the velocity field in a local region of an HCSG by compressing it into low-dimensional modes and coupling them with a Long Short-Term Memory (LSTM) network [4]. However, most existing research has either focused on single-variable prediction or overlooked the stability issues of multi-variable ROMs for complex internal flows. Therefore, extending the existing research framework to multi-variable prediction remains a challenging task that requires ensuring both stability and accuracy.

This study aims to develop a multi-variable surrogate model for the velocity and pressure variables within a local region of an HCSG. To overcome the limitations of single-variable ROMs, we propose a hybrid approach termed Individual POD - Unified LSTM. In this framework, POD is first applied individually to each physical variable (u, v, w , and C_p) to independently extract the high-energy modes. Next, the temporal coefficients of each extracted mode are concatenated and input into a single, unified LSTM network. This strategy enables the LSTM to learn the temporal correlations and interdependencies among the dominant patterns of different physical fields, thereby aiming to achieve the dual goals of the preservation of physical coupling. The ability of the model is evaluated by comparing its reconstruction results with the original LES. While the present study focuses on reconstruction over the observed time window, we also examine autoregressive behavior to characterize long-term stability and delineate actionable directions for predictive deployment.

2. Numerical Analysis and Datasets

2.1. Geometry and Computational Domain

This study analyzed the local cross flow occurring inside the HCSG via CFD and subsequently generated transient state data as a training dataset. Since simulating the entire steam generator is impractical, a 3×3 array of tubes, based on the design data of the System-integrated Modular Advanced Reactor (SMART) developed by the Korea Atomic Energy Research Institute (KAERI), was established as the analysis domain [5]. When defining the analysis domain, the actual helical structure, with its

rotation and elevation angles, increases the complexity of mesh generation and analysis. Therefore, the tubes were modeled as straight and arranged perpendicular to the flow direction. The dimensions of the entire analysis domain were set to a depth of 0.02 m, a width of 0.056 m, and a height of 0.08 m. A detailed diagram of the structure is provided in Figure 1. Approximately 2.09 million polyhedral cells were generated, with a dense mesh around the tubes to capture vortical and turbulent structures. For the efficient training of the deep learning model, 2D planar data from a focused analysis section were extracted from the 3D simulation results. A total of 46,638 nodes with uniform spacing were defined on the 2D plane, and at each node, velocity and pressure field data were output for every time step. The mesh configuration and node arrangement in the focused analysis section can be seen in Figure 2.

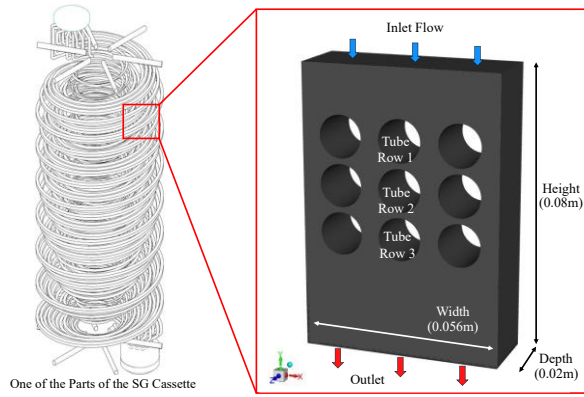


Fig. 1. SG Cassette and Fluid Domain

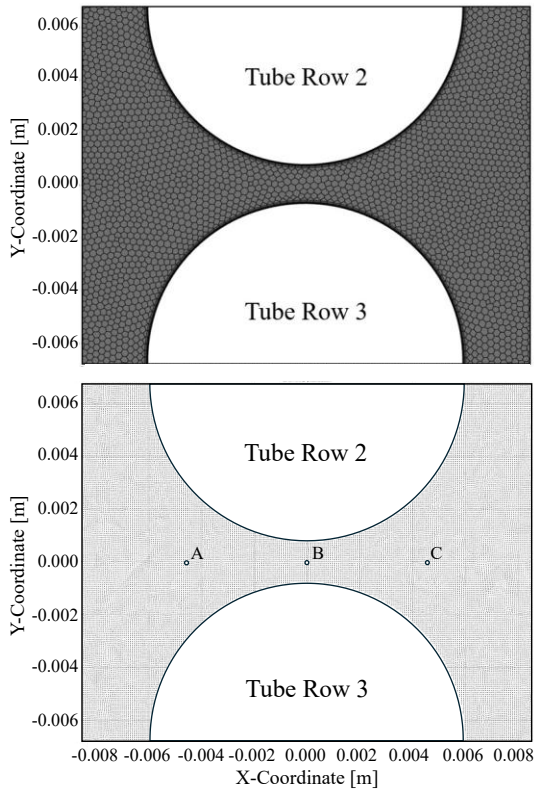


Fig. 2. Computational Mesh of Local 2D Region (top) and Grid Points Distribution (bottom)

2.2. Solver Settings and Data Extraction

Numerical simulation data were generated using ANSYS FLUENT 24R2. The working fluid was assumed to be water with a density (ρ) of 710.78 kg/m³ and a dynamic viscosity (μ) of 8.56×10^{-5} Pa·s. At the inlet, a uniform flow with a velocity (V_{∞}) of 0.5983 m/s was applied. The internal Reynolds number is calculated as $Re = \rho V_{\infty} D_h / \mu \approx 1.4 \times 10^5$.

The outlet boundary was set as a pressure-outlet with a gauge pressure of 0 Pa. The tube walls were modeled as no-slip walls to account for viscous effects, while the side boundaries of the domain were assumed to be free-slip walls. Based on these initial and boundary conditions, a transient analysis of the flow field was performed for a total of 3000-time steps (0.6 s) with a time step size of 2×10^{-4} s. Of the 3,000 snapshots, 80% were used for training and 20% for validation.

3. Methodologies

This study aims to predict flow variables by compressing the dimensionality of a transient flow dataset, obtained through numerical analysis, using POD and then utilizing it as training data for an LSTM network. The flow variables (u, v, w , and C_p) are used to analyze the vortex and turbulent phenomena occurring between the tubes within the HCSG. In addition to the issue that static pressure suffers from scale disparities, which cause the learning process to be dominated by large absolute errors, it also exhibits weak correlations with velocity variables. Therefore, the pressure variable in the multi-variable model was represented by the pressure coefficient (C_p).

$$C_p = \frac{p - p_{ref}}{\frac{1}{2} \rho U_{ref}^2} \quad (1)$$

The C_p reflects relative distributions, which stabilizes the data range and facilitates training. Moreover, since it is directly linked to the velocity term in the momentum equation, it serves as a physically more consistent target.

This section details the dimension reduction procedure using POD and the architecture of the LSTM network. It includes an explanation of the hybrid approach, in which the individual POD results for each flow variable are fed into a single LSTM network.

3.1. Proper Orthogonal Decomposition (POD)

POD was applied to extract dominant spatial structures from high-dimensional flow data. The snapshot matrix U was constructed using the fluctuation components obtained by subtracting the mean flow from the instantaneous flow fields. Singular value decomposition (SVD) was then performed as follows:

$$U = \Phi \Sigma V^T \quad (2)$$

Here, Φ represents the spatial modes, Σ contains the singular values indicated the energy of each mode, and V is the temporal coefficient matrix. By selecting the top K modes, most of the total energy can be preserved, and the flow field can be approximated as:

$$u(t) \approx \sum_{i=1}^K a_i(t) \phi_i(x) \quad (3)$$

where $a_i(t)$ denotes the time coefficients and $\phi_i(x)$ the spatial modes. POD is particularly effective in extracting large-scale vortex structures from turbulent flows and was employed for reduced-order modeling in this paper.

3.2. Long Short-Term Memory Network (LSTM)

LSTM networks, a variant of recurrent neural networks (RNNs), are designed to effectively learn and retain patterns in sequential data by incorporating memory cells and gating mechanisms. The architecture of the network is illustrated schematically in Figure 3. The input gate controls the inclusion of new information, the forget gate discards irrelevant past information, and the output gate selects the relevant cell state for subsequent layers. LSTM is trained on LES data from 0 to 0.6 s and used to predict the flow variables (u, v, w , and C_p), enabling accurate turbulence prediction by leveraging its capability to capture complex temporal dynamics.

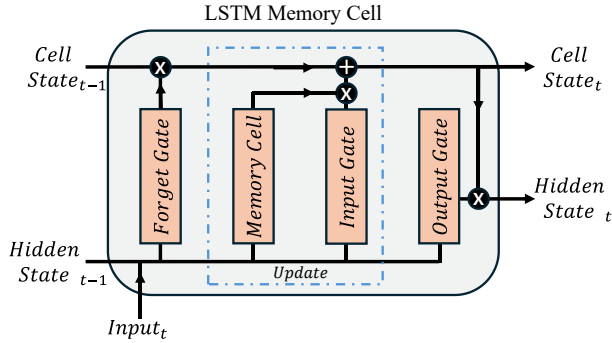


Fig. 3. Schematic of the LSTM Networks

3.3. Multiple-Variable POD-LSTM Framework

This study proposes a ROM framework that combines the linear dimensionality reduction technique POD with

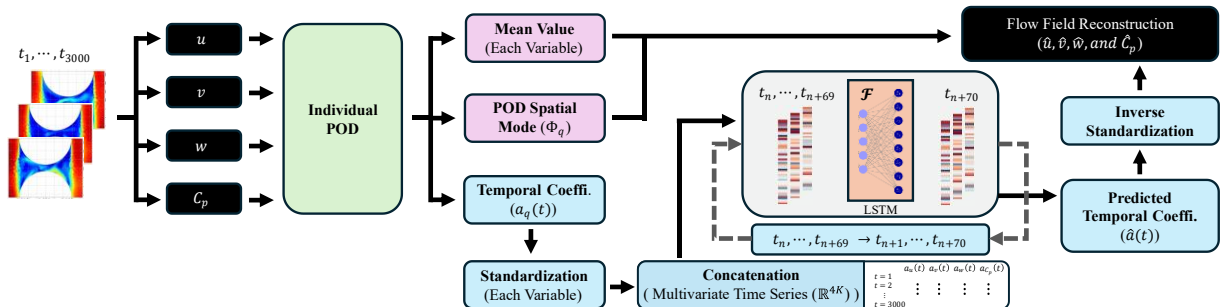


Fig. 4. Schematic of the Individual POD – Unified LSTM Networks Architecture

the nonlinear time-series model LSTM to efficiently reconstruct flow fields from 3,000 time-step snapshots. The overall framework is illustrated in Figure 4. All models were implemented in TensorFlow/Keras, using the Adam optimizer (learning rate 10^{-3}) with mean squared error (MSE) as the loss function. Overfitting was mitigated with EarlyStopping and ReduceLROnPlateau.

For each variable, POD was applied to extract the leading K spatial modes $\Phi \in \mathbb{R}^{N \times K}$ and the corresponding temporal coefficients $a(t) \in \mathbb{R}^K$. The number of modes K was tested under four settings, 25, 50, 100, and 150. To stabilize training, coefficients were standardized mode-wise (z-score) using statistics computed over the available snapshots, and then concatenated across the four variables to form a multivariate time series in $\mathbb{R}^{4K}$.

Input sequences were generated with a sliding window of length $L = 70$. Thus, the input to the LSTM is the past coefficient sequence $[a(t-L), \dots, a(t-1)] \in \mathbb{R}^{L \times 4K}$. And the output is the next-step coefficient vector $\hat{a}(t) \in \mathbb{R}^{4K}$. This constitutes a one-step autoregressive prediction, since the model learns to predict the next coefficient vector from its immediate history.

After training, the reconstruction process is carried out in successive steps. First, the LSTM network $\mathcal{F}_\theta$ takes as input the past sequence of coefficients $[a(t-L), \dots, a(t-1)]$ and predicts the next-step coefficient vector,

$$\hat{a}(t) = \mathcal{F}_\theta(a(t-L), \dots, a(t-1)). \quad (4a)$$

This step constitutes a one-step autoregressive update, because the model learns to infer the next coefficient from its immediate history and then rolls it out sequentially within the training interval. The predicted coefficients are then projected back onto the spatial modes to reconstruct the fluctuation field as

$$\tilde{q}(t) = \hat{a}(t) \Phi^T. \quad (4b)$$

This field represents the deviations from the mean flow that were extracted during decomposition. Finally, the global temporal mean field $\bar{q}$, computed over the entire training dataset and subtracted during preprocessing, is added back to obtain the absolute flow field,

$$\hat{q}(t) = \tilde{q}(t) + \bar{q}, \quad q \in \{u, v, w, C_p\}. \quad (4c)$$

3.4. Performance Evaluation Metrics

The performance of the model was evaluated using three metrics. First, the coefficient of determination (R^2) quantifies how well the predicted values fit the actual data:

$$R^2 = 1 - \frac{\sum_{i=1}^N (r_i - \hat{r}_i)^2}{\sum_{i=1}^N (r_i - \bar{r}_i)^2} \quad (5)$$

Second, the root mean square error (RMSE) measure the overall prediction error:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_i - \hat{r}_i)^2} \quad (6)$$

Finally, the variable-specific RMSE (VS-RMSE) is computed for each physical variable m , allowing independent evaluation of accuracy:

$$VS - RMSE_m = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_{i,m} - \hat{r}_{i,m})^2} \quad (7)$$

4. Results and Discussion

A CFD analysis was conducted to confirm flow characteristics such as vortex formation between the tube layers induced by cross flow. A snapshot dataset of 3000-time steps, spanning from 0.0002 s to 0.6 s, was obtained. The analysis confirmed that this dataset includes vortex stagnation points and the non-periodic generation, dissipation, and oscillation of vortices. No dominant, specific phase was identified. Therefore, it can be concluded that applying a dimensionality reduction technique is crucial for accurately capturing multi-scale structures and identifying key patterns in the current chaotic dataset.

4.1. Dimension Reduction with POD

In the process of ROM, data loss inevitably occurs during the remapping from the compressed dimension back to the high dimensional space. Therefore, dimensionality reduction must be considered within an acceptable range that accounts for this data loss. A POD analysis was performed on the training dataset. Table I summarizes the cumulative energy ratio of the spatial modes after applying POD to each of the four variables. In this context, the energy represents the total variance contained in the data, corresponding to the direction with the highest energy. POD was conducted with up to 150 modes, where it was observed that for all four variables, over 90% of the energy is captured with 50 or more modes.

Table I: Cumulative Energy and Percentage at each Variables

Modes	u	v	w	C_p
1	0.2031	0.5145	0.1514	0.9816
25	0.8624	0.9860	0.7809	0.9983
50	0.9470	0.9952	0.9080	0.9995
100	0.9901	0.9992	0.9832	0.9999
150	1	1	1	1

4.2. Comparison between Real and Reconstruction

Table II shows the error rates of the flow variables reconstructed by the LSTM-based models over time steps 71-3000, for different reduction dimensions: 25, 50, 100, and 150. As the dimensionality of the reduced space increased, the error rate decreased. In this study, the comparison is conducted based on a dimension of 100, at which the R^2 value for all variables exceeded 0.9. The error rates of the reconstructed data indicate that, at this stage, the models are not yet mature enough to fully substitute CFD, highlighting the need for further improvement.

Table II: Reconstruction Performance Metrics

Modes	R^2	RMSE	CV-RMSE
25	u	0.8487	0.0580
	v	0.9878	0.0894
	w	0.7443	0.0728
	C_p	0.8514	52.9872
50	u	0.9340	0.0382
	v	0.9932	0.0667
	w	0.8756	0.0494
	C_p	0.9137	39.0360
100	u	0.9750	0.0238
	v	0.9932	0.0667
	w	0.9543	0.0287
	C_p	0.9386	33.5282
150	u	0.9916	0.0204
	v	0.9916	0.0743
	w	0.9742	0.0226
	C_p	0.9348	34.4005

Figure 5 displays the temporal variation of velocity at points A, B, and C, which are marked at the bottom of Figure 2. For the comparative analysis of the velocity-magnitude, the actual data was taken from the values output by the CFD, while the predicted values were calculated from the u, v, w components predicted by the model. The model accurately captures the overall trends of the flow, which is characterized by non-periodicity and dissipation of vortices. Figure 6 shows the velocity-magnitude contours at three training timesteps: 500th, 1,500th, and 2,500th. It was confirmed that with only 100 modes, not only the core vortex structures but also the fine features are captured. However, an examination of the loss curve during training revealed that the training rollout loss gradually increases. This suggests a lack of long-term stability in rollout predictions. In other words, when the current network model predicts data

autoregressively, errors can accumulate, leading to a sharp decline in performance.

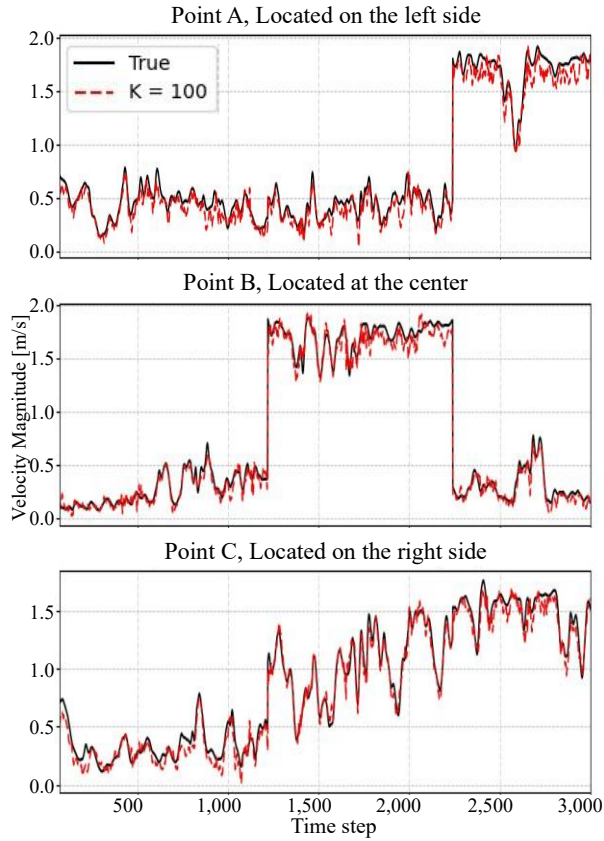


Fig. 5. Time Evolution of the Velocity Magnitude at A, B, C; Predictions against the Real Data

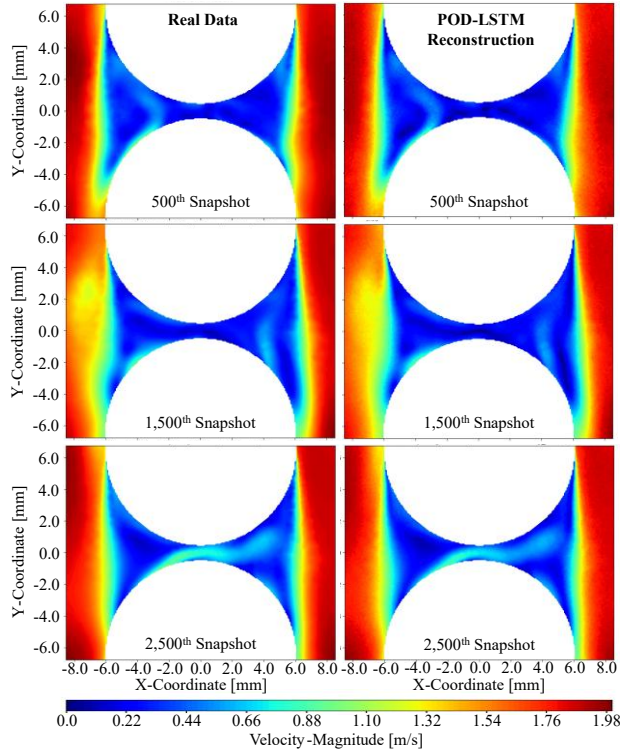


Fig. 6. Velocity Magnitude Contours at 500th, 1,500th, 2,500th

For a more detailed component-wise analysis, a Power Spectral Density (PSD) analysis was conducted. The time-series data of the u, v, w components at Point C were compared in the frequency domain. Figure 7 presents the PSD of the true data and the reconstructed data. For all three components, the model was able to predict the dominant energy distribution in the low- and mid-frequency ranges with reasonable accuracy. However, in the high-frequency region, which contains fine turbulent structures, the reconstructed results showed weaker performance. This limitation is attributed to the dimensionality reduction process of the learning model, which tends to filter out small-scale turbulence. In addition, the u, v components exhibited an overestimation tendency.

Finally, the distribution of C_p was examined at the same time steps where the velocity magnitude was evaluated. Figure 8 shows that the model's sole pressure variable was also found to be reconstructed with high fidelity to the real data. The accuracy of C_p indicates that the physical coupling with the velocity variables has been successfully captured.

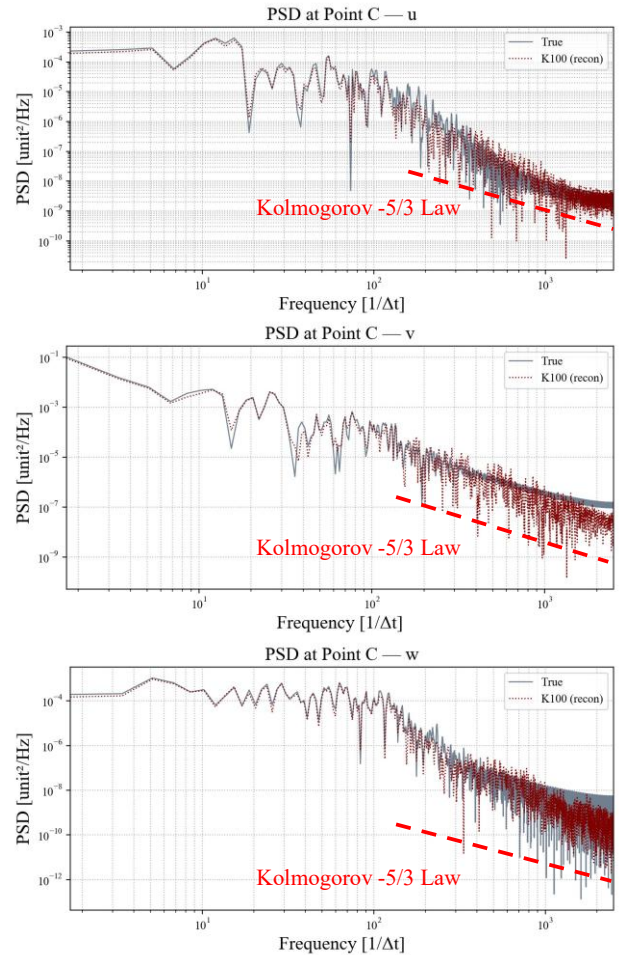


Fig. 7. PSD comparison of velocity-fluctuation signals at C

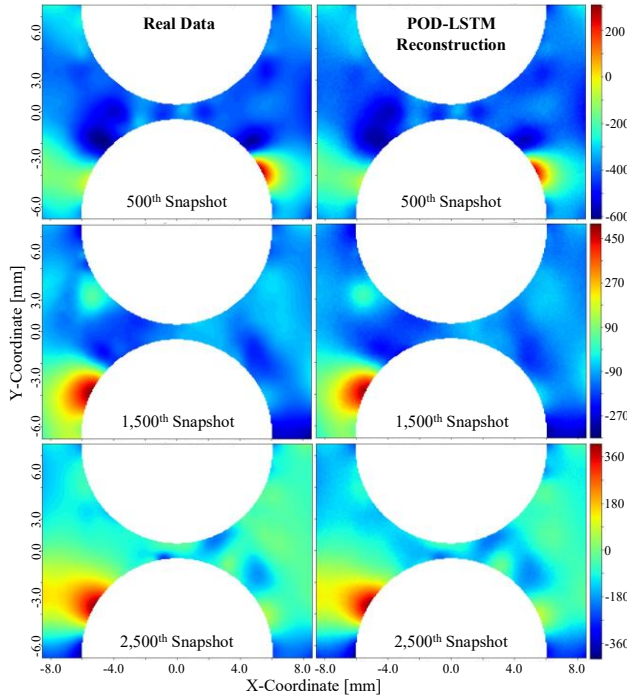


Fig. 8. Pressure Coefficient Contours at 500th, 1,500th, 2,500th

4.3. Computational Cost

The computational efficiency of the POD-LSTM models with different reduced dimensions is compared against CFD (LES) results. Table III summarizes the training and reconstruction times of the POD-LSTM models with 25, 50, 100, and 150 modes, expressed in a non-dimensional form as t_{POD}/t_{CFD} with respect to the CFD runtime required for dataset generation. Here, t_{POD} denotes the sum of the training and reconstruction times of the POD-LSTM models. Once the issues of long-term extrapolation and error accumulation are addressed, the POD-LSTM approach demonstrates a significant advantage in computational cost reduction.

Based on this advantage, deep learning-based predictive models can enable rapid evaluation of thousands of operating conditions and geometric configurations during the conceptual design and requirement definition stages of nuclear reactors. Furthermore, such models are expected to provide reasonable initial conditions for CFD analyses in the preparation stage, and to be applied in various areas such as optimizing geometry and operating conditions for experimental facility setups.

Table III: Computational Cost for Training (Offline) and Reconstruction (Online)

Modes	Offline [second]	Online [second]	t_{POD}/t_{CFD}
25	1338.77	15.31	5.256×10^{-3}
50	1362.312	14.30	5.343×10^{-3}
100	1410.061	17.16	5.540×10^{-3}
150	1470.964	16.92	5.775×10^{-3}
CFD	257,640		1

5. Conclusion

This study presents a framework that integrates POD-based reduced-order modeling with LSTM to reconstruct complex, unsteady flow dynamics in helical coil steam generators from LES-based CFD data. Dimensionality reduction was performed on four flow variables, and the standardized POD coefficients were concatenated and processed by a unified LSTM to enable temporal reconstruction within the observed interval.

Using 100 out of 150 POD modes, the POD-LSTM reconstructed the fundamental vortex patterns with high fidelity. PSD analysis confirmed that the reconstructed coefficients remained within acceptable error bounds in the low- and mid-frequency ranges. The pressure coefficient C_p , the only pressure-related variable among the four, was also well reconstructed, indicating that the physical coupling among variables was effectively captured. Overall, the model demonstrated excellent reconstruction performance over the training window (0–0.6 s).

Despite these strengths, PSDs for the u and v components show overestimation at high frequencies, suggesting the loss of fine-scale turbulence through dimensionality reduction and sequence modeling. In addition, long-horizon autoregressive rollouts exhibit error accumulation, indicating limited stability beyond the training window. The present validation is further constrained by a single geometry and operating condition, so broader generalization was not established.

As a foundational multivariable reconstruction study, our primary objective was to establish feasibility and quantify reconstruction quality under a controlled setting. Building on these findings, future work will (i) construct an independent test dataset to evaluate generalization beyond the training/validation split, (ii) pursue adaptive mode selection and larger truncation levels guided by energy and PSD criteria to balance compression with fidelity, and (iii) improve long-horizon stability via multi-step rollout losses and scheduled sampling. To better capture fine-scale/high-frequency content, we will investigate neural operator formulations (e.g., FNO/TFNO) and hybrid ROM-operator couplings. We will quantify uncertainty via multi-seed ensembles, bootstrap confidence intervals, and conformal prediction, and assess generalization across operating conditions and geometries. Ultimately, this line of research aims to provide a viable strategy for addressing the challenges of high-dimensional, large-scale LES datasets in the analysis and design of HCSG for SMRs.

6. Acknowledgements

This research was supported by the National Research Council of Science & Technology (NST) grant by the Korea government (MSIT) (No. GTL24031-000).

REFERENCES

- [1] S.J. Lee and Y.A. Hassan, “Numerical investigation of helical coil tube bundle in turbulent cross flow using large eddy simulation,” *Int. J. Heat Fluid Flow*, Vol. 81, pp. 108-122, 2020.
- [2] S. Lee and Y.A. Hassan, “Flow structures of the cross-flow over a five-layer helically coiled steam generator geometry,” *Int. J. Heat Fluid Flow*, Vol. 95, 108950, 2022.
- [3] H. Eivazi, H. Veisi, M.H. Naderi, and V. Esfahanian, “Deep neural networks for nonlinear model order reduction of unsteady flows,” *Phys. Fluids*, Vol. 32(10), pp. 105104, 2020
- [4] B. Cho and M. Song, “Efficient prediction of turbulent cross flow in helical coil steam generators of SMR via deep learning–driven reduced-order models,” *NURETH-21*, Busan, Korea, Aug. 31–Sep. 5, 2025
- [5] Korea Atomic Energy Research Institute (KAERI), “Basic Design Report of SMART,” KAERI/TR-2142/2002, Daejeon, Korea (2002)