Comparative Study of AI-Based Time Series Prediction Strategies for Nuclear Severe Accident Simulations

Seunghyeok Yang [a*], Daeyoung Lee [a], Suwon Lee [a]
[a]*FNC Technology Co., LTD., 13 Heungdeok 1-ro, 32F, Giheung-gu, Yongin-si, Gyeonggi-do, 16954, Korea*
[*]*Corresponding author: shyang@fnctech.com*

***Keywords :** Severe Accident, Time Series Prediction, Hybrid Strategy, Nuclear Safety

1. Introduction

Nuclear Power Plants, as large-scale energy sources, require both stability and reliability as fundamental conditions. In particular, during severe accidents, the massive release of radioactive materials can occur, making early prediction and effective response critical to ensuring safety. Traditionally, Probabilistic Safety Assessment(PSA) and Deterministic Safety Analysis(DSA)[1] have been employed to support such safety objectives. However, the actual progression of nuclear accidents exhibits high-dimensional and nonlinear characteristics, and conventional methods face limitations in simultaneously considering the complex interactions among diverse operational parameters.

Recently, artificial intelligence(AI)-based research has been actively pursued to utilize measured plant signals for the early detection of accident precursors and the prediction of accident progression. In particular, deep learning techniques for time-series forecasting provide the ability to capture complex patterns hidden within high-dimensional data[2], thereby offering new opportunities for enhancing nuclear power plant operation and accident prevention frameworks.

Previous studies have reported approaches using conventional time-series forecasting models such as LSTM, GRU, and 1D-CNN to predict short-term variations in key reactor parameters (e.g., pressure, temperature, and coolant flow)[3]. However, these models inevitably suffer from recursive forecasting error accumulation, where prediction errors grow as the forecasting horizon extends. To address this issue, the Direct Multi-Horizon(DMH) approach has been proposed[4], in which the model directly learns to predict future states at specific horizons. Nevertheless, DMH methods are still constrained by their limited ability to capture long-term dependencies.

To overcome these limitations, the present study proposes a hybrid forecasting framework that integrates recursive prediction and DMH approaches. By combining the strength of recursive forecasting in capturing cumulative dependencies with the horizon-specific learning capability of DMH, the framework aims to improve both the accuracy and stability of severe accident progression prediction. Furthermore, the proposed multivariate time-series learning structure leverages actual plant measurement data to capture not only individual variable dynamics but also system-level correlations, thereby enabling a more comprehensive representation of accident progression.

The objectives of this study are as follows:

1. To propose a hybrid deep learning framework that combines recursive and Direct Multi-Horizon(DMH) forecasting strategies for nuclear accident progression prediction.
2. To train and evaluate the proposed model using actual nuclear accident simulation data consisting of 72-hour multivariate plant measurement signals.
3. To compare the performance of the proposed method with conventional single-strategy approaches(recursive-only or DMH-only), thereby assessing the advantages, limitations, and practical applicability of the hybrid framework.

2. Methods and Results

2.1 Data Description and Preprocessing

In this study, we utilized safety analysis code simulation data that model severe accidents in nuclear power plants. Each simulation covers a total of 72 hours(3 days), with results recorded at 5-minute intervals, yielding 865 time-series data points per case. The complete dataset consists of 208 scenarios, each differentiated by accident type, initial conditions and operator actions.

The dataset includes the following major variables:

- Input variables (26 types): Observable variables such as pressurizer pressure, RCS flow, steam generator pressure and water level, hot leg and cold leg temperature, and containment temperature.
- Target variables (2 types): Core exit temperature and containment pressure, which is the main variable of interest in the progression of severe accident.

The preprocessing procedure involved the following steps:

1. Missing value check and removal: No NaN values were found in raw data. However, additional masking was applied to prevent computational errors during model input.

2. Normalization: Z-score normalization was performed to account for differences in units across variables.
3. Data splitting: The dataset was divided into training(60%) validation(20%), and testing(20%) sets. Scenario-level splitting was applied to avoid overlaps between datasets.
4. Sequence construction: Input sequences were set to a length of 24 and 36 (equivalent to 2 and , 3hours), while output horizons were set to 6, 12, 18, 24 (corresponding to 30, 60, 90, and 120 minutes), enabling experiments with multiple forecasting scenarios.

2.2 Model Architecture

In this study, we employed the recently proposed Mamba(a deep learning architecture based on State Space Models)[5]. Compared to traditional RNN and Transformer-based approaches, Mamba is more efficient in handling long sequences and requires lower GPU memory consumption, making it well-suited for large-scale simulation data.

The main components of the model are as follows:

- Input layer: A linear embedding is applied to the 26-dimensional input variables, projecting them into a fixed-dimensional latent space.
- Mamba blocks: These perform state-space operations, capturing both short and long-term dependencies in time series.
- Output layer: Generates multi-step sequence outputs of length m

The training setup was configured as follows:

- Loss function: Mean Squared Error(MSE)
- Optimization algorithm: AdamW with learning rate 10^{-4}, weight decay 10^{-2}
- Batch size: 256
- Epochs: Up to 200 with early stopping
- Hardware environment: Training was conducted on an NVIDIA L40S GPU, 48GB VRAM.

2.3 Forecasting Strategies

To evaluate prediction performance, we compared three forecasting strategies:

1. Recursive One-Step (R1S): The model predicts only a single step at a time, and each prediction is recursively fed back into the input to generate the full m-step sequence. This approach provides high short-term accuracy but suffers from error accumulation as the prediction horizon extends.
2. Direct Multi-Horizon (DMH): The model directly outputs the entire m-step sequence from a single input window. While this method yields stable performance in long-term forecasts, it has limitations in capturing abrupt short-term dynamics (e.g., a sudden drop in core power).
3. Hybrid: To combine the advantages of both approaches, the hybrid strategy applies the R1S method for the initial k-steps and then switches to DMH outputs for the remaining horizon. This design aims to achieve both short-term accuracy and long-term stability.

2.4 Evaluation Metrics

The performance of the proposed models was evaluated using the following metrics:

- RMSE(Root Mean Squared Error): Measures the overall magnitude of prediction errors.
- MAE(Mean Absolute Error): Complements RMSE by reducing the influence of large deviations, providing a more balanced error assessment.
- CE(Conformal Efficiency): Quantifies the efficiency of uncertainty-based prediction intervals, reflecting the reliability of the model's confidence estimates.
- VPH(Variance-Preserving Horizon): Assesses whether the variance characteristics of the original data are preserved in long-horizon forecasts. This metric plays a critical role in validating the physical consistency of predictions against actual nuclear plant measurements.

2.5 Results

2.5.1 Quantitative Results

The performance comparision across differenct forecasting horizons is summarized in Figure 1 and Figure 2. When prediction horizon was relatively short (m=6, equivalent to 30 minutes), the Hybrid strategy did not outperform the other methods and even showed unstable results in terms of short-horizon RMSE. Both R1S and DMH maintained low errors in this regime, with DMH slightly outperforming R1S in average RMSE.

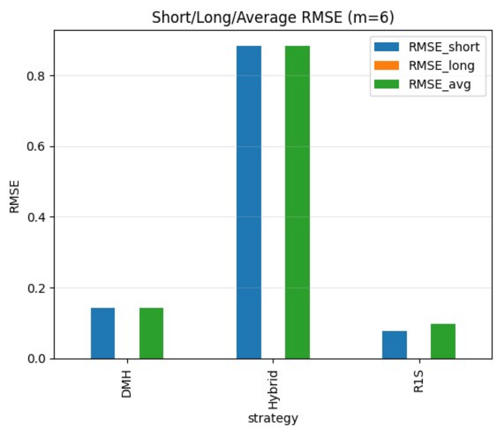

Fig 1. Comparison of short, long and average RMSE for the three forecasting strategies at m=30.

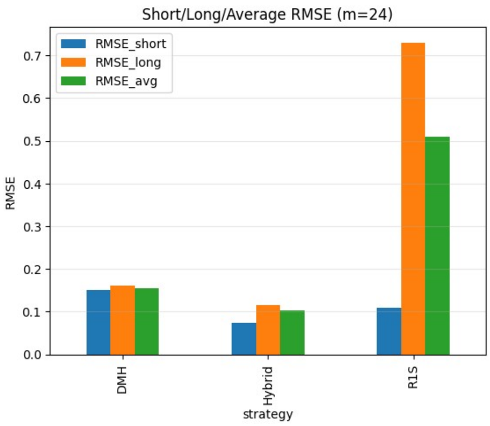

Fig 2. Comparison of short, long and average RMSE for three forecasting strategies at m=24.

However, when the horizon was extended to m=24 (120 minutes), the performance landscape changed significantly. As shown in Figure 2, R1S suffered from severe error accumulation in the long-term region (RMSE_long > 0.7), whereas DMH maintained stable performance with average RMSE around 0.15. Importantly, the Hybrid method demonstrated clear improvement relative to DMH, achieving the lowest error in the long-horizon RMSE segment while keeping competitive performance in short-term predictions. These findings suggest that although Hybrid does not always guarantee superior performance in very short-term horizons, its relative advantage grows as the forecasting horizon increases, highlighting its utility for long-term accident progression prediction.

2.5.2 Qualitative Results

By comparing the time-series prediction results, the following characteristics of each method were identified:

- R1S: While closely matching the ground truth in the initial phase, it exhibited pronounced over and under estimation of core pressure and containment pressure beyond 60 minutes due to accumulated errors.
- DMH: This method consistently captured the overall trend across the entire horizon; however, in regions of rapid change, its responses were relatively smooth, failing to fully follow the steep gradients observed in the actual data.
- Hybrid: Although some sequences displayed unstable patterns during transition phases, its responsiveness was improved in abrupt change intervals (e.g., immediately after accident onset), and it maintained a level of consistency comparable to DMH in long-term stable regions. This suggests potential advantages from an accident-response perspective that may not be apparent when considering average RMSE alone.
- In addition to single-sequence qualitative comparisons, the short/long split analysis further confirmed that Hybrid was particularly effective in maintaining prediction accuracy beyond the 60-120 minute range. This long-term stability, combined with its improved responsiveness to early accident dynamics, makes Hybrid a strong candidate for accident progression forecasting under severe conditions.

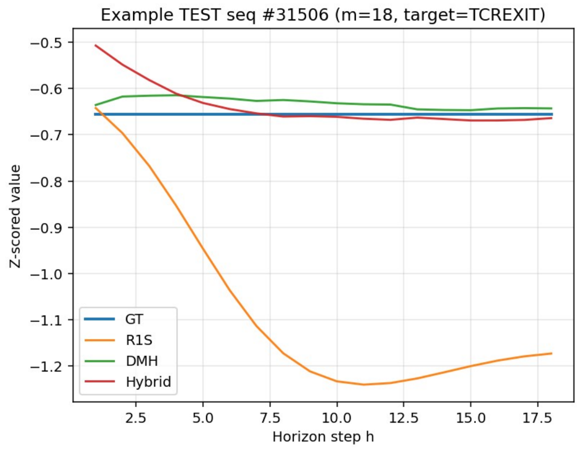

Fig 3. Example test sequence for target *TCREXIT* (m=18)

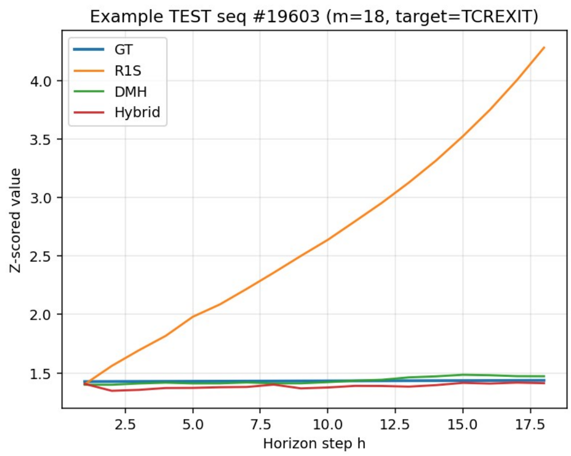

Fig 4. Another Example test sequence for target *TCREXIT* (m=18)

In addition to the average performance metrics, individual test sequences revealed important qualitative differences among the forecasting strategies. Figure 3 and Figure 4 illustrate two representative cases for the target variable *TCREXIT*. In both cases, the R1S approach initially followed the ground truth but diverged rapidly as the horizon extended, leading to large deviations that are physically unrealistic. By contrast, DMH maintained stable trends but occasionally under-represented abrupt dynamics.

Notably, the Hybrid method combined the stability of DMH with improved alignment to the ground truth over long horizons. While Hybrid did not perfectly eliminate errors, it consistently prevented the severe drift observed in R1S, thereby preserving the physical plausibility of the predicted trajectory. These qualitative examples reinforce the quantitative findings that Hybrid becomes increasingly advantageous as the prediction horizon lengthens. This property is particularly critical for accident management, where sustained accuracy over one to two hours is essential for reliable operator decision support.

2.5.3 Feature Importance Analysis

In this study, we applied the Permutation Importance method[6] to quantify the influence of input variables across the three forecasting strategies, R1S, DMH and Hybrid. The analysis was conducted for both short-term prediction horizons (m=6, 30 minutes) and long-term prediction horizons (m=24, 120 minutes).

Table I Permutation Importance of top-3 variables for short-term and long-term horizons.

Horizon	Strategy	Var1	Var2	Var3
m=6	R1S	PPZ (0.631)	TCREXIT (0.224)	PEX(0)9 (0.133)
	DMH	PEX0(9) (0.345)	TCREXIT (0.272)	PPZ (0.258)
	Hybrid	---	---	---
m=24	R1S	PPZ (0.424)	ZWV (0.163)	RCSINFLOW (0.058)
	DMH	PEX0(9) (0.331)	TCREXIT (0.299)	PPZ (0.258)
	Hybrid	PEX(0)9 (0.393)	TCREXIT (0.349)	PPZ (0.290)

The results revealed that the R1S model exhibited excessive dependence on the pressurizer pressure variable across both horizons. In particular, ΔRMSE reached 0.63 at m=6 and 0.42 at m=24, far exceeding the contributions of other variables. This indicates that while R1S may be advantageous for capturing short-term fluctuations, its over-reliance on a single variable contributes to error accumulation and degraded performance in long-term forecasts.

The DMH model consistently highlighted containment pressure, core exit temperature, and pressurizer pressure as the most influential variables. These physically meaningful thermal-hydraulic parameters were identified as critical for both short- and long-term predictions, underscoring DMH's ability to incorporate stable and relevant signals across horizons.

The Hybrid model showed negligible sensitivity at m=6, where permutation resulted in nearly zero ΔRMSE across all variables. However, at m=24, the Hybrid model closely resembled the DMH pattern, emphasizing containment pressure, core exit temperature, pressurizer pressure, and containment spray flow rate as the most important variables. This suggests that the Hybrid model may not leverage variable-level distinctions in very short horizons, but demonstrates enhanced sensitivity to physically meaningful variables in long horizons.

These findings indicate that the Hybrid approach not only reduces numerical errors but also adapts to emphasize key safety-related variables as the prediction horizon increases.

2.5.4 Discussion

The results of this study indicate that AI-based prediction models can capture critical physical characteristics of nuclear accident simulations. R1S and DMH each demonstrated clear advantages in short-term and long-term horizons, respectively. However, the comparative analysis across different horizons revealed an important trend: while Hybrid did not outperform the other strategies at very short horizons (m=6, 30 minutes), its relative performance improved as the horizon length increased (m=24, 120 minutes). This horizon-dependent advantage suggests that the Hybrid approach becomes increasingly effective for long-term accident progression prediction, which is of high relevance to nuclear safety decision-making.

From a physical perspective, this improvement may be attributed to the Hybrid framework's ability to balance recursive error correction in the early stages with direct horizon-specific forecasting in the later stages, thereby mitigating cumulative error propagation. In practice, this property is particularly valuable for nuclear accident management, where operators require both short-term responsiveness and reliable long-term foresight.

Beyond error metrics, the variable importance analysis provided additional insights into the interpretability of the models. R1S showed disproportionate reliance on a single input pressurizer pressure, which explains its vulnerability to error accumulation in long horizons. In contrast, DMH and Hybrid consistently emphasized core thermal-hydraulic variables such as core exit temperature, containment pressure, which are physically meaningful indicators of accident progression. Notably, the Hybrid model, although insensitive at very short horizons, increasingly prioritized these critical variables as the horizon lengthened. This horizon-dependent adaptation highlights Hybrid's potential not only to achieve balanced accuracy but also to preserve physical relevance, reinforcing its value as a trustworthy tool for early warning and operator support.

Although the Hybrid strategy exhibited variability in average error, its growing advantage at longer horizons highlights its potential as a robust early-warning and operator-support tool under severe accident conditions. Furthermore, when integrated with advanced evaluation metrics such as conformal efficiency (CE) and variance-preserving horizon (VPH), the proposed framework can be extended into next-generation digital safety systems. These systems have the potential to reinforce situational awareness, provide more reliable forecasts over operationally critical timescales, and enhance decision-making reliability in nuclear power plant operations.

3. Conclusions

In this study, we developed AI-based time series forecasting models using severe accident simulation data from nuclear power plants and compared the performance of three prediction strategies: Recursive One-Step(R1S), Direct Multi-Horizon(DMH), and Hybrid. The experimental results demonstrated that the R1S method achieved superior accuracy in short-term horizons, while the DMH approach provided more stable performance in long-term forecasts. The hybrid strategy, which combines the strengths of both methods, yielded the most balanced and reliable outcomes overall.

These findings highlight the potential of AI-based forecasting techniques to support early warning and accident progression prediction in nuclear emergency decision-making. In particular, the Hybrid approach shows promise as a supplementary tool for enhancing operator situational awareness and strengthening Severe Accident Management Guidelines(SAMG).

Nevertheless, as this study relies on simulation data, further work is required to ensure applicability to real nuclear plant measurements. Key challenges include verifying data homogeneity, handling sensor noise, and optimizing real-time inference performance. Future research will focus on developing hybrid frameworks that integrate physics-based models, validating generalizability across different reactor types, and advancing uncertainty quantification methods to reinforce the practical utility of this approach.

ACKNOWLEGDEMENTS

This work was supported by the Ministry of Trade, Industry and Energy (MOTIE) and the Korea Institute of Energy Technology Evaluation and Planning (KETEP) (No. 20224B10300050).

REFERENCES

[1] US NRC. PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessment for Nuclear Power Plants. NUREG/CR-2300, 1983.
[2] B. Lim, S. Zohren, Time-series forecasting with deep learning: A survey. Philosophical Transactions of the Royal Society A, 379:20200209, 2021.
[3] J. Song, S. Kim, A machine learning informed prediction of severe accident progressions in nuclear power plants, Nuclear Engineering and Technology, 56(6), 2024.
[4] B. Lim, S. O. Arik, N, Loeff, T. Pfister, Temporal Fusion Transformer for Interpretable Multi-horizon Time Series Forecasting. International Journal of Forecasting, 37(4), 2021.
[5] A. Gu, T. Dao, Mamba: Linear-Time Sequence Model with Selective State Space. arXiv:2312.00752, 2023.
[6] A. Altmann, L. Tolosi, O. Sander, T. Lengauer, Permutation importance: a corrected feature importance measure. Bioinformatics, 26(10), 1340-1347, 2010.