Optimizing Downsampling Strategies for AI-Based Surrogate Models in Severe Accident Prediction

Taeyeon Min [a], Semin Joo [a], Seok Ho Song [a], Mi Ro Seo [b], Jeong Ik Lee [a*]
[a]*Department of Nuclear and Quantum Engineering, N7-1 KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, Korea 34141*
[b]*KHNP CRI, 70, Yuseong-daero 1312beon-gil, Yuseong-gu, Daejeon, Korea 34101*
[*] *Corresponding author: jeongiklee@kaist.ac.kr*

***Keywords :** surrogate model, downsampling strategy, time series prediction, dynamic time warping*

1. Introduction

The Fukushima Daiichi nuclear power plant accident highlighted the critical need for effective management of severe accidents in nuclear facilities. To support operators in severe accident with high-stress situations, accident management supporting tools (AMSTs) are essential for minimizing human error and improving decision-making during rapidly changing accident conditions.

Recent advancements in artificial intelligence (AI) have demonstrated the potential for integrating AI into AMSTs. The surrogate model based on deep neural networks can operate approximately 4,000 times faster than conventional commercial accident analysis codes [1]. This performance enhancement is largely attributable to the capacity of AI-based surrogate models to substitute the time series generated from the nonlinear and computationally intensive processes of traditional accident analysis codes with outcomes derived from neural network training.

However, when predictions are made with a time resolution higher than 1 hour, error accumulation becomes more pronounced, significantly deteriorating predictive accuracy for longer period [2]. Conversely, employing 1-hour interval time series data results in substantial information loss during the sampling of the original time series for model training. Considering these challenges, the present study proposes a range of sampling strategies aimed at minimizing information loss and enhancing the overall predictive performance of the AI-based system.

2. Methods

2.1 Data collection

The baseline accident scenario for training the AI model is the total loss of component cooling water (TLOCCW), which was selected by the previous studies [1]. A TLOCCW accident is a failure of all seven safety-related components of a reactor (See Table I.). However, for generating the accident scenarios, the following failures of safety components were assumed to occur randomly over time with a few exceptions: the RCP seal LOCA has an 89.2% chance of failure within the first hour, failures of the HPI, LPI, CSS, and charging pump are tied to the depletion of the refueling water storage tank (RWST), which is depleted between 7 and 8 hours in over 80% of cases, leading to failures during this period, and all other component failures were assumed to occur randomly during 72 hours. In addition, three severe accident management guidelines (SAMGs) were randomly initiated within 72 hours.

Moreover, three severe accident management guidelines (SAMGs); SG injection (M1), RCS depressurization (M2), and RCS injection (M3)—are randomly activated within 72 hours with 1 hour-interval. SAMGs are protocols designed to guide operators in mitigating severe accident consequences; they activate when monitored variables meet specified conditions.

Table I. List of components that fail at TLOCCW

Reactor coolant pump (RCP) seal LOCA
Letdown heat exchanger (HX)
High-pressure (HPI) injection pump
Low-pressure (LPI) injection pump
Containment spray system (CSS) pump
Motor-driven auxiliary feedwater (MDAFW) pump
Charging pump

To generate the accident scenario, MAAP 5.03 code was used. The MAAP code calculated the progression of severe accidents over a 72-hour period and returns various thermos-hydraulic (TH) variables. The target TH variables were selected as 10 variables monitored in the main control room (MCR) (see Table II). Therefore, a single accident scenario consists of time series data of 10 TH variables from 0 to 72 hours after the accident occurred. The total dataset is composed of 11,000 accident scenarios with a combination of component failure and SAMG activation timing. The dataset was normalized with a minimum value of 0.2 and a maximum value of 0.8 for the entire scenario for each TH variable. This removes extreme values, which mitigates the problem of slope vanishing at the two extremes of the deep learning model's activation function, improving learning stability and performance

Table II. List of target TH variables for surrogate model

Primary system pressure
Hot leg temperature
Cold leg temperature
Reactor vessel water level (RV WL)
Steam generator pressure (SG P)
Steam generator water level (SG WL)
Maximum core exit temperature (Max CET)
Containment pressure (CTMT P)
Pressurizer pressure (PZR P)
Pressurizer water level (PZR WL)

2.2. Model learning

The surrogate model employed Long Short-Term Memory (LSTM) to capture temporal dependency efficiently (see Fig. 3). Moreover, the surrogate model was trained with two types of datasets: conventional sampling dataset and best representative dataset. These two datasets will be further discussed in the next section. The input layer is composed of the previous 3-time steps with 10 TH variables and binary indicator of 7 component failure and 3 SAMG activation. Based on a rigorous hyperparameter study in previous work [3], the hidden layer was configured with 400 nodes and a batch size of 32 to achieve optimal training performance. Consequently, 10 surrogate models, one for each TH variable, are trained. The loss function was selected to mean absolute error (MAE), and the error metric was selected to root mean error (RMSE).

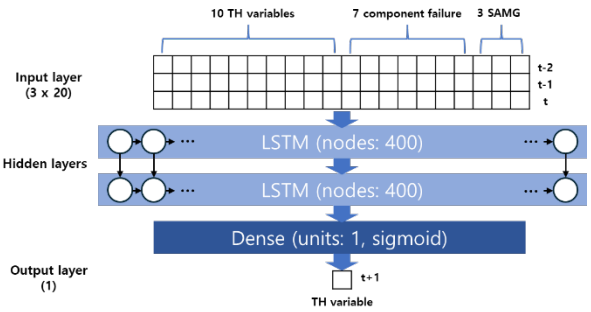

Fig. 2. The overall framework of surrogate model.

2.3 Selection of Representative data for Down sampling

According to the Nyquist theorem, sampling at a frequency lower than twice the maximum frequency present in the original signal induces signal distortion and aliasing during reconstruction. Consequently, conventional downsampling the original time series to a one-hour interval inevitably results in significant information loss. In this study, conventional sampling refers to selecting identical time points for downsampling. To address this issue, various methods were compared, extracting a representative value from the one-hour interval data, with the aim of preserving as much similarity to the original signal as possible. Specifically, for each hour point, an interval spanning n minutes forward and backward (n = 1, 2, 3, …, 30) is defined, and either the mean or the median of the values within this interval is used as the representative value for sampling, as shown in Figure 1. The candidates of representative value selecting methods are summarized in Table III.

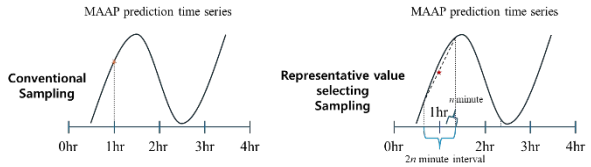

Fig. 1. Schematic illustration of Downsampling method; Conventional (left) and Representative value selecting (right)

Table III. Candidates of representative value selecting methods

	Representative value selecting
Time interval around hour point	*n* minutes forward and backward (n = 1,2,3, … , 30)
Representative value	mean or median

3. Results

3.1 Optimization of Best Selecting Method

The Dynamic time warping (DTW) distance are widely used methods for measuring similarity between data points, particularly in time series. The DTW distance is a technique for measuring similarity between time series by allowing nonlinear time distortions to optimally align two sequences. DTW allows flexible alignment of time series by handling temporal distortions, making it more robust than EUD in pattern recognition (see Fig. 3). Given two sequences X=(x_1,x_2,…,x_N) and Y=(y_1,y_2,…,y_M), DTW computes a cumulative cost matrix D(i, j) using dynamic programming, as defined by the recurrence relation:

$$D(i,j) = d(x_i, y_i) + min\{D(i-1,j) + D(i,j-1) + D(i-1,j-1)\}$$

where d(x_i,y_i) is typically the EUD. The final DTW distance is given by D(N, M), representing the minimum cost of aligning the two sequences.

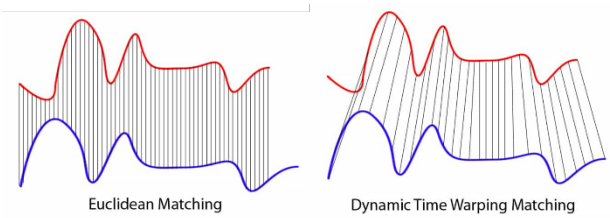

Fig 3. Comparison of Euclidean Distance and Dynamic Time Warping

To optimize the best sampling case for various representative value selecting methods, DTW distance was employed to compare with original time series predicted by the MAAP code with those sampled for conventional and representative value selecting methods. The relative errors of DTW distance were computed (see Fig. 4). The reference value of relative error was used to obtained from conventional sampling, and comparison values were candidates in Table III. Therefore, the negative value means that particular representative value selecting method is more similar to original time series.

For DTW distance, it was found that the relative error increased with longer time intervals, and that the case employing the mean value as the representative exhibited a strong dependency on the time interval. Only three cases yielded a lower DTW compared with the conventional method. The optimal selecting method was identified as the use of the median value over a 4-minute interval, which resulted in a relative error of –0.6%.

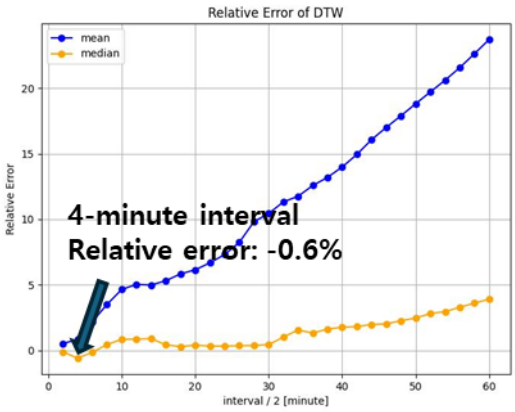

Fig 4. Relative error of DTW distance

Consequently, the optimal metric for downsampling was determined to be the median computed over a 4-minute bin. This configuration is hereafter referred to as the 'best sampling case'.

3.2 Model verification

The performance of the surrogate model for both the conventional sampling and best sampling cases was evaluated. In both cases, the MAE and RMSE values for each model show a difference of around 5%, indicating that there is no significant performance difference in model training (see Table IV). The bold font means lower metrics.

Subsequently, DTW distances were computed between the time series predicted by the surrogate model and the original time series predicted by the MAAP code. Figure 4 presents the DTW values for the surrogate model predictions of each sampling case. The DTW values decrease for all TH variables except for PPS and CTMT P. In the case of training with the best sampling as label data, the predicted time series becomes slightly more similar to the original time series.

Table IV. Error function value and Performance Metric for each TH variable respect to sampling method

	Conventional Sampling		Best Sampling	
	MAE	RMSE	MAE	RMSE
PPS	**0.000934**	**0.005888**	0.001189	0.006574
Cold leg T	**0.002068**	**0.006364**	0.002212	0.006637
Hot leg T	0.001961	**0.005993**	**0.001914**	0.006012
RV WL	0.003922	**0.013692**	**0.003668**	0.014063
SG P	0.001581	0.004982	**0.001525**	**0.00485**
SG WL	0.001731	0.004238	**0.001558**	**0.004014**
MAXCET	0.003664	**0.017171**	**0.003299**	0.017298
CTMT P	**0.000769**	**0.002389**	0.001046	0.002668
PZR P	0.001406	0.00665	**0.000953**	**0.005599**
PZR WL	0.001285	0.006725	**0.001029**	**0.006257**

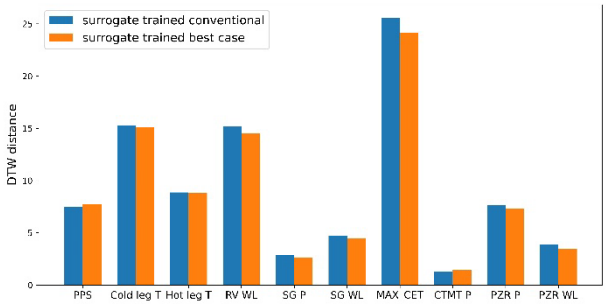

Fig 5. Normalized mean DTW distances of TH variables of surrogate model each sampling case

4. Summary and Conclusions

This study demonstrated the effectiveness of using representative value selection methods to improve the accuracy of downsampling for severe accident prediction using AI based surrogate model. By comparing various representative value selection strategies, it was found that the use of the median value over a 4-minute interval yielded the most optimal results in terms of minimizing information loss while maintaining the accuracy of the surrogate model's predictions.

The comparison of conventional sampling and the best sampling case revealed minimal differences in performance, suggesting that the representative value selection does not significantly affect the overall prediction accuracy, as measured by error metrics such as MAE and root mean square error RMSE.

Further analysis using DTW distance demonstrated that the surrogate model trained with the optimal sampling method exhibited greater similarity to the

original MAAP predictions for most TH variables. However, this method did not yield consistently high similarity across all variables. This trend arises because the optimal sampling case was determined based on the average DTW distance and Euclidean distance across ten TH-variables rather than being optimized for each variable individually.

These findings indicate that selecting appropriate representative values during downsampling can enhance the stability and performance of AI-based surrogate models without sacrificing accuracy. Therefore, the proposed downsampling approach using the median value over a 4-minute interval serves as one possible method to mitigate information loss when using low temporal resolution training data, thereby contributing to the improved predictive capability of AMSTs and supporting operators in managing severe accidents more effectively. However, this approach primarily addresses the limitations of low-resolution data and underscores the fundamental need for developing AI-based surrogate models capable of making predictions at higher temporal resolutions.

ACKNOWLEDGEMENTS

This work was supported by KOREA HYDRO & NUCLEAR POWER CO., LTD (No. 2024-Tech-08).

REFERENCES

[1] Y. Lee, et al. Surrogate model for predicting severe accident progression in nuclear power plant using deep learning methods and Rolling-Window forecast, Annals of Nuclear Energy, Vol. 208, p. 110816, 2024.
[2] S. Joo, *et al*. Accelerated prediction of severe accident progression: Sensitivity of deep neural network performance to time resolution, *Transactions of the Korean Nuclear Society Autumn Meeting, Gyeongju, 2023*.
[3] Joo, S. Development of an Explainable Machine Learning Methodology for Accelerated Prediction of Nuclear Power Plant Severe Accident Scenario, Master's thesis, KAIST, 2024.