Adaptive Sampling for Data-efficient AI Prediction in NPP Simulations

Donghee Jung, Junyong Bae, Seung Jun Lee*
Ulsan National Institute of Science and Technology, 50, UNIST-gil, Ulsan, 44911
**Corresponding author: sjlee420@unist.ac.kr*

Keywords : Adaptive sampling, AI prediction model, Nuclear power plant simulation, Data efficiency

1. Introduction

Artificial intelligence (AI) technologies have been increasingly applied in the nuclear engineering field, particularly to improve the reliability and performance of instrumentation and control (I&C) systems. In the main control room, operators are required to monitor hundreds of parameters simultaneously and to make timely decisions under both normal and abnormal conditions. In accident scenarios such as a Loss-of-Coolant Accident (LOCA), the evolution of plant parameters can be highly nonlinear, and small deviations in initial conditions or boundary conditions may result in significantly different system responses. Accurate prediction of key parameters using AI models can therefore provide valuable support to operators, complementing traditional alarm-based systems and enhancing situational awareness [1].

Despite these advantages, the development of robust AI prediction models is often hindered by the cost of data generation. In contrast to typical machine learning applications where large datasets are readily available, nuclear power plant (NPP) simulators must be executed repeatedly to cover a wide range of accident conditions, break sizes, and operator actions. Each simulation run can be computationally expensive and time-consuming. Consequently, training datasets are frequently constructed using uniform grid-based sampling of conditions or random selection from the parameter space. While such approaches ensure coverage of the domain, they often include redundant cases in regions where the model already performs well, while neglecting more challenging scenarios where the prediction accuracy is poor. This inefficiency leads to excessive simulation costs and suboptimal training outcomes [2].

Adaptive sampling, also referred to as active learning in the machine learning community, provides a potential solution to this challenge. The key idea is to allow the prediction model itself to identify regions of high uncertainty or large prediction error, and then to request new data specifically from those regions. This self-reflective loop—training the model, evaluating its weaknesses, and selectively generating new scenarios—enables the model to focus computational resources on the most informative data points. By doing so, adaptive sampling reduces redundancy in the dataset and improves the generalization of the trained model [3], [4].

In the context of nuclear power plants, adaptive sampling has particular relevance for accident scenario simulations. For example, in LOCA scenarios, the severity of the transient and the timing of reactor trip events are highly sensitive to the break size. Some break sizes may lead to rapid trips within a few seconds, while others may result in delayed or no trips over several minutes. Uniform sampling across the entire range of break sizes may under-sample these critical transition regions. Adaptive sampling, on the other hand, can identify and intensively sample such regions based on model performance.

This study explores the application of adaptive sampling to improve the data efficiency of AI prediction models in NPP simulations. The Compact Nuclear Simulator (CNS) is used to generate LOCA scenarios with break sizes ranging from 1 to 500. A long short-term memory (LSTM)-based prediction model is trained to forecast multiple plant variables, and its performance is compared under different data selection strategies: adaptive sampling, grid sampling, random sampling, and full training with all available cases. The objective is not only to evaluate the average prediction error but also to analyze the distribution of errors across the parameter space, particularly in challenging regions.

The remainder of this paper is organized as follows. Section 2 describes the methodology of the proposed adaptive sampling framework, including scenario generation, prediction model structure, and sampling strategies. Section 3 presents the case study results for LOCA scenarios in the CNS. Finally, Section 4 provides discussion and concluding remarks, including implications, limitations, and directions for future research.

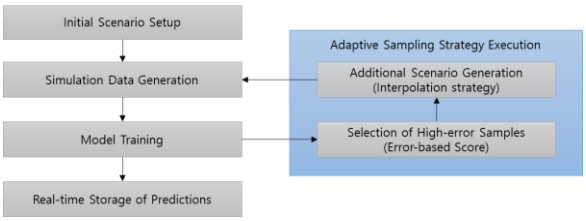

Fig. 1. Workflow of the proposed adaptive sampling framework for AI prediction in nuclear power plant simulations.

2. Methodology

The proposed adaptive sampling framework for data-efficient AI prediction in nuclear power plant (NPP) simulations consists of three main elements: scenario generation, prediction model development, and sampling strategy. Figure 1 illustrates the iterative workflow, in

which the model is trained with an initial dataset, its prediction errors are evaluated, and new scenarios are generated based on the error distribution.

2.1 Scenario Generation

Loss-of-Coolant Accident (LOCA) scenarios were generated using the Compact Nuclear Simulator (CNS). The break size was varied from 1 to 500, producing a total of 500 distinct cases. Odd-numbered cases (250) were designated as training data, while even-numbered cases (250) were reserved exclusively for evaluation.

Each scenario began with a malfunction injection at 10 seconds. The subsequent reactor trip timing exhibited strong nonlinearity depending on the break size. For example, break size 50 caused an almost immediate trip, break size 30 produced a trip at 32 seconds, break size 9 resulted in a trip at 75 seconds, and break size 3 delayed the trip until nearly 300 seconds. These variations highlight the need for selective data sampling methods capable of capturing critical transition regions in system dynamics.

2.2 Prediction Model

The prediction model employed in this study was based on a long short-term memory (LSTM) neural network. The architecture included a single LSTM layer with 128 hidden units, followed by a dense output layer. A dropout rate of 0.3 was applied to prevent overfitting.

The model input consisted of two time steps (60 seconds) of 109 plant variables, while the output was defined as the trajectories of 25 key variables over 20 future time steps, corresponding to a 600-second prediction horizon. This configuration enabled the model to forecast medium-term plant behavior following LOCA initiation.

2.3 Adaptive Sampling Strategy

The adaptive sampling procedure began with three initial training cases corresponding to break sizes 1, 251, and 499, chosen to represent the lower, middle, and upper extremes of the parameter space. Two additional candidate cases, with break sizes 127 and 375, were also prepared as midpoints between the initial cases to provide coverage of potentially nonlinear regions.

After initial training, the model was evaluated on the candidate cases, and prediction errors were quantified using mean squared error (MSE). The case with the highest error score was added to the training dataset. To further explore the neighborhood of the selected case, two new break sizes were generated at the midpoints between the selected case and its adjacent training cases. For instance, if break size 127 was identified as the highest-error scenario, additional cases at sizes 65 and 189 were produced and included in the candidate pool. The model was retrained with the expanded dataset, and this loop of evaluation, selection, and retraining was repeated.

This self-reflective loop enabled the model to request new data where it was most needed, thereby improving training efficiency by focusing computational resources on informative regions of the scenario space.

2.4 Comparison Strategies

To evaluate the effectiveness of the adaptive sampling framework, three alternative strategies were also

implemented. The first was grid sampling, in which training cases were selected uniformly across the break size domain to provide evenly spaced coverage of the parameter space. The second was random sampling, where training cases were chosen without consideration of their location or the model's prior performance. The third approach utilized the full training dataset consisting of all 250 odd-numbered break sizes, which represented an upper bound for model accuracy but required substantially more simulation data and training effort. These three strategies, together with adaptive sampling, enabled a balanced comparison of accuracy, robustness, and data efficiency.

2.5 Evaluation Metrics

The performance of each strategy was evaluated using the 250 even-numbered break size cases. The primary metric was the mean squared error (MSE) across all predicted variables and time steps. To capture additional aspects of model reliability, two complementary metrics were also considered: (1) the variance of MSE across evaluation cases, reflecting the stability of model predictions, and (2) the mean error of the top 10% most erroneous cases, highlighting robustness under difficult conditions.

3. Case Study and Results

To evaluate the proposed adaptive sampling framework, a case study was conducted using Loss-of-Coolant Accident (LOCA) scenarios generated from the Compact Nuclear Simulator (CNS). The goal was to compare the predictive performance of the adaptive sampling strategy against grid sampling, random sampling, and full dataset training.

A total of 500 LOCA cases were prepared by varying the break size from 1 to 500. Odd-numbered break sizes (250 cases) were designated as potential training data, while even-numbered break sizes (250 cases) were used exclusively for evaluation. The initial adaptive training dataset consisted of three cases at break sizes 1, 251, and 499, representing the extremes and midpoint of the parameter space. Two additional candidate cases at break sizes 127 and 375 were included, corresponding to the midpoints between the initial cases. After training, the model evaluated these candidates, selected the highest-error case, and expanded the candidate set by generating midpoint scenarios around the selected case. This self-

reflective loop was repeated until 17 training cases were accumulated, matching the size of the grid and random sampling datasets.

Table 1 summarizes the performance of each strategy in terms of the overall mean squared error (MSE), the average error of the worst 10% evaluation cases, and the variance of errors across all evaluation cases.

Table I: Performance comparison of sampling strategies

Training Strategy	No. of Training Cases	Evaluation Mean MSE	Worst 10% Mean MSE	MSE Variance
Adaptive Sampling	17	0.001749	0.007375	0.000010
Grid Sampling	17	0.001521	0.008913	0.000064
Random Sampling	17	0.002002	0.013140	0.000160
Full Dataset	250	0.000834	0.004830	0.000014

The results indicate that adaptive sampling achieved a mean error of 0.001749, slightly higher than the grid-based approach (0.001521) but substantially lower than random sampling (0.002002). More importantly, adaptive sampling outperformed grid and random sampling in terms of worst-case robustness, with a worst 10% error of 0.007375 compared to 0.008913 for grid sampling and 0.013140 for random sampling. The variance of prediction errors was also lowest for adaptive sampling (1.0×10^{-5}), indicating stable performance across different break sizes.

Figure 2 illustrates the distribution of MSE across all 250 evaluation cases for the four training strategies. To facilitate direct comparison, the vertical axis limits were fixed across all subplots. As shown, adaptive sampling mitigated extreme error spikes observed in random sampling, while maintaining relatively stable performance across the parameter space. Grid sampling produced low average errors but exhibited localized peaks at specific break sizes, suggesting that uniform spacing does not always capture highly nonlinear transitions. The full dataset unsurprisingly yielded the best accuracy and stability, but at the cost of generating fifteen times more training cases than the reduced-data strategies.

These findings demonstrate that adaptive sampling can provide a balanced compromise between accuracy and robustness. While grid sampling was slightly superior in terms of mean error, adaptive sampling was more effective in reducing severe errors, which is critical for accident scenarios where outlier cases can significantly impact operator decision support. Random sampling, by contrast, consistently underperformed, underscoring the inefficiency of unguided data selection.

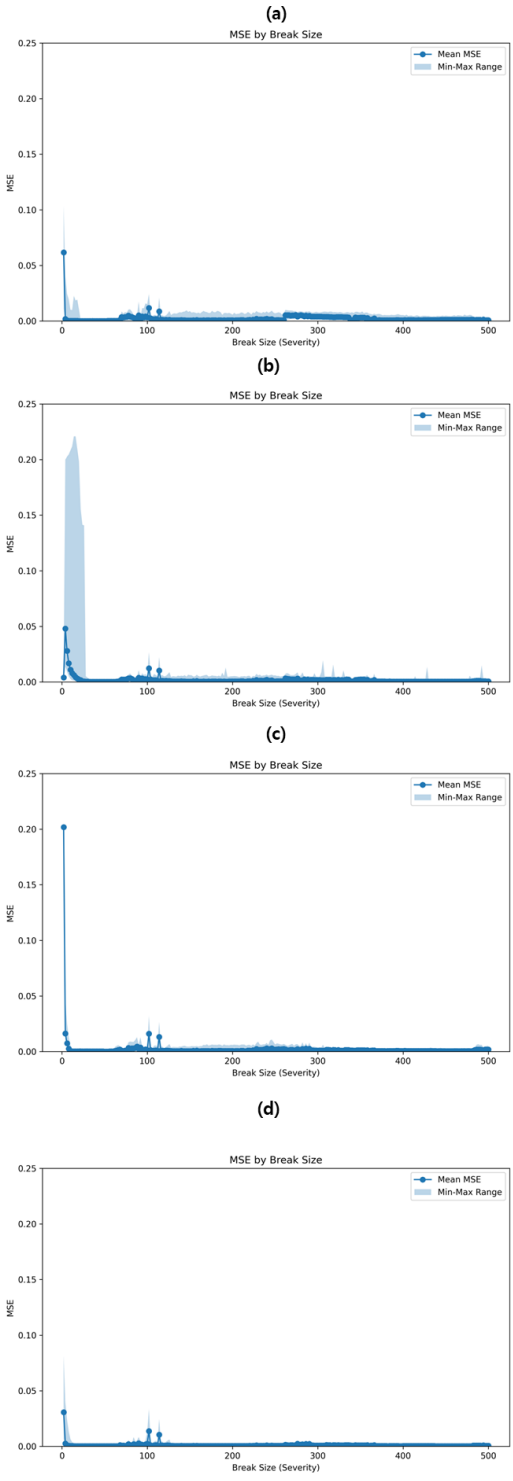

Fig. 2. Distribution of mean squared error (MSE) across break sizes for different sampling strategies: (a) adaptive sampling, (b) grid sampling, (c) random sampling, and (d) full dataset training. The shaded region represents the min–max range across predicted variables, and the solid line indicates the mean MSE.

4. Discussion and Conclusion

The results of this study demonstrate that adaptive sampling can serve as an effective strategy for improving

the data efficiency of AI prediction models in nuclear power plant (NPP) simulations. By iteratively identifying error-prone regions of the scenario space and selectively generating additional training cases, the framework reduced redundant data and enhanced robustness in difficult scenarios.

In terms of overall accuracy, grid sampling slightly outperformed adaptive sampling, achieving the lowest mean squared error (MSE). However, the adaptive strategy provided superior robustness by reducing the average error of the worst 10% evaluation cases from 0.008913 in the grid sampling approach to 0.007375. This indicates that adaptive sampling is particularly effective in mitigating extreme prediction errors, which are of critical importance in accident scenarios such as Loss-of-Coolant Accidents (LOCAs). In addition, adaptive sampling exhibited the lowest variance of prediction errors among the reduced-data strategies, suggesting stable performance across the evaluation domain. These results highlight the trade-off between mean accuracy and robustness, with adaptive sampling favoring balanced performance over localized optimization.

From a practical standpoint, the comparison with the full dataset underscores the significance of data efficiency. Training with all 250 cases yielded the best results in terms of both mean and worst-case errors, but at a prohibitive computational cost. In contrast, adaptive sampling achieved competitive performance using only 17 training cases, representing a fifteen-fold reduction in data requirements. This efficiency is particularly relevant in nuclear engineering, where generating large numbers of simulation cases can be expensive and time-consuming.

Despite these advantages, several limitations of the present study must be acknowledged. First, the adaptive sampling strategy relied solely on mean squared error as the performance criterion. While effective for this pilot study, more sophisticated measures of model uncertainty, such as ensemble variance or Bayesian approximations, may provide a richer basis for sample selection. Second, the midpoint-based expansion method, though simple, risks concentrating training data in narrow regions of the parameter space, potentially neglecting other important areas. Finally, the iterative retraining process increases computational cost, which may limit scalability when applied to broader sets of accident scenarios.

Future work will address these limitations by exploring alternative scoring metrics, incorporating multiple accident types beyond LOCA, and extending the approach to diverse boundary conditions and operator actions. Another promising direction is the integration of adaptive sampling with real-time I&C applications, where robust and data-efficient AI prediction models could enhance operator decision support and accident management strategies.

In conclusion, this study presented a pilot application of adaptive sampling for AI prediction in nuclear power plant simulations, focusing on LOCA scenarios generated with the Compact Nuclear Simulator (CNS). The findings demonstrate that adaptive sampling can effectively reduce redundant training data while improving robustness against extreme cases, offering a practical balance between accuracy and efficiency. These results highlight the potential of adaptive sampling as a valuable tool in the development of AI-assisted instrumentation and control systems in nuclear power plants.

ACKNOWLEDGEMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. RS-2024-00401407).

This research was supported by the National Research Council of Science & Technology(NST) grant by the Korea government (MSIT) (No. GTL24031-400)

REFERENCES

[1] J. Bae, G. Kim, and S. J. Lee, “Real-time prediction of nuclear power plant parameter trends following operator actions,” *Annals of Nuclear Energy*, Vol. 156, 108191, 2021.

[2] T. Nguyen and H. Diab, “Uncertainty-aware prediction of peak cladding temperature during extended station blackout using Transformer-based machine learning,” *Nuclear Engineering and Design*, Vol. 438, 113563, 2025.

[3] Y. Gal and Z. Ghahramani, “Dropout as a Bayesian approximation: Representing model uncertainty in deep learning,” *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, pp. 1050–1059, 2016.

[4] B. Settles, “Active Learning Literature Survey,” Computer Sciences Technical Report 1648, University of Wisconsin-Madison, 2010.