Feasibility Study of XAI Application to Severe Accident Uncertainty Analysis – Insights

Mi Ro Seo [a*], Tae Woo Kim [a], Se Min Joo[b], Jeong Ik Lee[b]
[a]*KOREA Hydro & Nuclear Power Co., Ltd., Central Research Institute, 70, Yuseong-daero 1312beon-gil 70, Yuseong-gu, Daejeon, Korea*
[b]*Dept. Nuclear & Quantum Eng., Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon, Korea*
[*]*Corresponding author: mrseo9710@khnp.co.kr*

***Keywords :** Severe Accident, Artificial Intelligence, Uncertainty, MAAP*

1. Introduction

One unavoidable challenge in analyzing severe accidents and establishing response measures lies in how much uncertainty stemming from the phenomena of severe accidents can be considered. It is known that the sources of uncertainty appearing in severe accident analysis arise mainly in three areas: first, the inherent uncertainty of the severe accident phenomenon itself; second, the uncertainty arising from experiments aimed at predicting the phenomenon; and third, the uncertainty originating from thermal-hydraulic relations established based on experimental results. The phenomena that may occur in nuclear power plants, such as high temperatures, high pressures, and violent phase transitions between molten corium and coolant, fundamentally make it difficult to make an experiment itself for the phenomenological identification. Due to these reasons, until the Fukushima nuclear power plant accident, the severe accident strategies were carried out by the SAMG (Severe Accident Management Guidelines) based on qualitative judgments according to the symptoms. However, since the Fukushima nuclear accident, there has been a demand for more quantitative severe accident analysis, and efforts have been made to clarify the uncertainties in the field of severe accidents.

The numerical analysis using the integrated severe accident analysis code, such as MAAP or MELCOR, is the analytical method widely used for phenomenological analysis, designing of countermeasure equipment and establishing the severe accident management strategies. However, due to the lack of experimental data on severe accident phenomena, the uncertainty of the analysis results remains quite large, so research to clarify this uncertainty is ongoing. In this paper, we evaluated the possibility of providing additional insights into the influence and correlation between variables, which was difficult to find in conventional evaluation methodologies, by utilizing machine learning and AI techniques for uncertainty analysis, which have rapidly developed recently.

2. Methodologies

2.1. The Current SA Uncertainty Analysis

The process of performing uncertainty analysis using MAAP (Modular Accident Analysis Program) codes proceeds as follows [1].

1) Selection of key model variables and their distributions that can affect specific phenomena. This is primarily determined by the qualitative judgment of the model developer or designer.
2) Sampling of variables according to their distribution using random sampling methods. Typically, the Latin Hypercube Sampling (LHS) technique is employed.
3) Execution of the MAAP code with sampled variables to verify specific phenomena. Usually, the code is run with around 100 inputs.

As described above, the current uncertainty analysis in severe accidents relies heavily on random sampling for selecting major model variables. Consequently, it becomes challenging to understand the impact of variables influencing actual phenomena. As a result, sensitivity analysis affecting the uncertainty is additionally performed based on qualitative judgments. However, even this approach struggles to grasp the interconnectivity between variables.

2.2. Application of AI to Uncertainty Analysis

To gain additional insights from severe accident uncertainty analysis results through artificial intelligence, the following steps were taken:

1) Data Augmentation: Typically, uncertainty analysis using computational codes generate analysis results by directly running the code with about 100 samples. However, the amount of data required to train AI models is usually around 10,000 cases, making the output only 0.01% of what's needed. To address this issue, data augmentation methods were applied instead of conducting additional sampling and running the code, which would require substantial analysis time. The method used in this study is called Dirichlet Mix-up, which linearly combines original data using weights obtained from the Dirichlet distribution.
2) Application of Machine Learning Models: The machine learning model utilized in this research is XGBoost, which constructs a powerful predictive model by combining multiple 'weak learners.' This model excels in recognizing rule-based patterns in numerical data with clear rows and columns and

continuously reduces residuals, resulting in highly precise predictions.

3) Implementation of Explainable AI (XAI) Techniques: In this study, the SHAP (SHapley Additive exPlanations) technique, one of the XAI methods, was applied. The principle behind SHAP is to calculate the difference in predicted values when a particular variable is "present" versus "absent", thereby evaluating the influence of input variables on the output

The detailed methodologies for the development and application of AI model will be presented in a separate paper at this conference.

3. Results

3.1. Results of Accident Progression and Uncertainty Analysis

In this study, the uncertainty analysis results published in the paper titled "Uncertainty Study of RPV Failure and Operator Actions in a MBLOCA Scenario of the OPR1000," presented at the Korean Nuclear Society Spring Meeting in 2025, were utilized [2]. The accident scenario was the 6-inch medium-break loss-of-coolant accident (MB LOCA) ruptured in the cold-leg, coupled with a failure of the safety injection system and no operator mitigation actions. For severe accident analysis and uncertainty analysis, MAAP 5.06 was employed. The major events are listed in Table I, and the 49 variables deemed influential on RPV failure by the model developer, FAI, are listed in Table II [1][3].

Table I: Base case scenario without mitigation operation

Time	Events
0 sec	Initiation of MBLOCA
12 sec	Reactor Scram
172 sec (2.9 min)	RCP Trip
4,901 sec (1.4 hr)	Core Uncover
6,696 sec (1.9 hr)	CET exceeds 1200 F (SAMG entrance)
13,433 sec (3.7 hr)	RPV failure

Table II: Selected uncertain input model parameters

P/M	Uncertain Input Parameter	Num.
TH-PP	FCHFCR, FFRICX, TJBRN, XSTIA, FGBYPA, TAUTO, FWHR, FROUPZ	8
SA-CR	FUPOOL, FDPOOL, FSPOOL, TCLMAX, LMCOL0, LMCOL1, LMCOL2, LMCOL3, EPSCUT, EPSCU2, FZORUP, FACT, FCRDR, FDDP, ENT0, FSGBEN, VFCRCO, FGPOOL, FMOVE, FAOX, IOXIDE, FASSOXID	22
SA-CS	TSPFAL, FPEEL, XDJETO, XLAFALS, FOXBJ, VFENT	6
SA-LP	XGAP0, XGAPLH, IQDPB, XLFALS, FZGAPTOPLH, IOXIDHT, IOCHF	7
SA-LH	ECREPF, ECREPP, EPSPB, FEMISD, FEMISP, FQUEN	6

In that paper, uncertainty analysis was conducted for 52 variables, including the initial 49 variables and an additional three related to Mitigation Actions. A total of 100 inputs were sampled using the LHS technique, and among these, RPV failure were observed in only 39 sample inputs. These findings are detailed in Figure 1. Furthermore, the key variables influencing RPV failure were analyzed as shown in Figure 2.

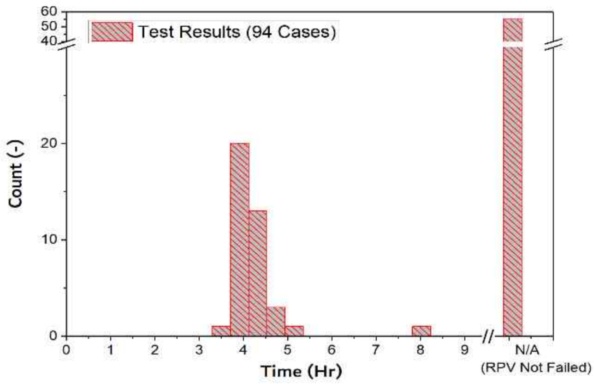
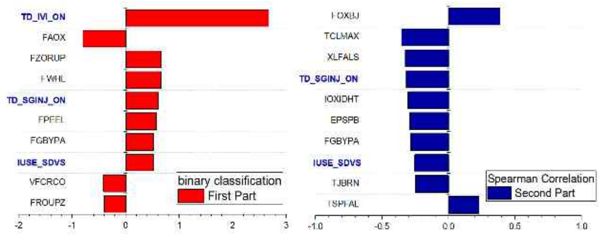


Fig. 1. Histogram of the distribution of RPV Failure times



(a) First Part (b) Second Part
Fig. 2. The Top 10 parameters influencing RPV failure

3.2. Results of XAI Applications

In this study, since the objective was to evaluate the impact of key variables related to RPV failure based on the aforementioned uncertainty analysis results, an additional 100 sampling datasets were re-generated based on the reactor vessel failure case. However, even 100 datasets are quite small for training machine learning models; hence, the previously mentioned Dirichlet mix-up methodology was employed to augment the data. With the augmented dataset, a machine learning model utilizing the XGBoost algorithm was trained, and the XGBoost-specific SHAP computation algorithm was used to assess the influence of variables on the uncertainty analysis results. Detailed modeling incorporating XAI and the computed results derived from XAI is described in a separate paper and will be presented in this conference.

3.3. Feasibility of XAI Applications

This paper aims to evaluate whether applying XAI to existing uncertainty analysis can provide additional meaningful insights. To achieve this, we compared the evaluation results of major variables' influences from both XAI evaluations and conventional uncertainty analysis. Figure 3 illustrates a comparison between the main variables analyzed using PCC (Pearson Correlation Coefficient) statistical metrics in traditional uncertainty analysis and those identified by SHAP.

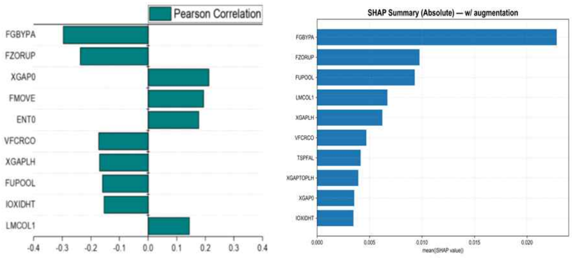


Fig.3. The Top 10 parameters influencing RPV failure

From the figure, it's evident that eight out of the top ten variables match between the two methods. Notably, there are two variables (TSPFAL and XGAPTOPLH) that appear in the top ten according to SHAP analysis but not in the Pearson correlation-based top ten. The meanings of the top five variables relevant here are as follows :

1) FGBYPA : flag to divert gas flows in the core to the bypass channel when an entire axial row in the core is completely blocked
2) FZORUP : minimum fraction of Zr that must be oxidized to keep the cladding intact if the cladding is at TCLMAX
3) FUPPOOL : heat transfer coefficient for core upward convective heat transfer between two molten core nodes or one molten core node and one crust node
4) LMCOL1 : collapse criteria parameter for a Larson-Miller-like functional dependence for a core node below a collapsed core node
5) XGAPLH : initial size of the gap between the debris and the lower head steel wall for the bottom section of the lower head corium pool

According to the results of SHAP, the top 10 parameter pairs with the most powerful interaction effect are listed in Table III.

Table III. Parameter pairs with the top 10 highest interaction

	Parameter 1	Parameter 2	Interaction effect (ϕ_{ij})
1	FUPOOL	LMCOL1	0.002902
2	FGBYPA	XGAPLH	0.002216
3	FZORUP	VFCRCO	0.002211
4	FUPOOL	LMCOL0	0.001627
5	FGBYPA	FZORUP	0.001613

The most influential variables, FUPOOL and LMCOL1, are highlighted in Figure 4.

A clear increase in the SHAP value of FUPOOL is observed when FUPOOL is small and LMCOL1 is large. This indicates that when both variables satisfy specific conditions simultaneously, a significant delay in the RPV failure time occurs. As LMCOL1 increases, the rod bundle geometry is maintained for a longer period, improving structural stability. At the same time, a decrease in FUPOOL reduces upward convective heat transfer, leading to the formation of a thicker crust. This delays the relocation of corium to the lower plenum and consequently postpones RPV failure. A small number of samples with this specific combination appear to have increased both the mean SHAP contribution of FUPOOL and its correlation coefficients.

The second influential variables, FGBYPA and XGAPLH, are highlighted in Figure 5. When FGBYPA is 0, the SHAP value decreases more clearly as XGAPLH increases. This indicates that when gas flow is not diverted to the bypass channel, core cooling capability is relatively maintained, and the RPV failure time becomes more sensitive to XGAPLH, which is the initial size of the gap between the debris and the lower head steel wall. In contrast, when FGBYPA is 1, gas flow diversion reduces core cooling capability, and the relative influence of XGAPLH decreases. This result demonstrates that the influence of XGAPLH on the failure time depends on the state of FGBYPA.

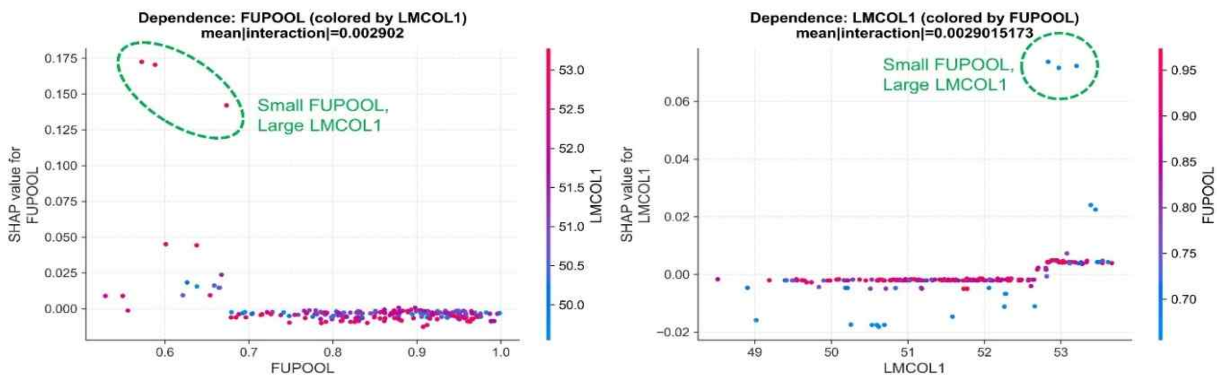


Fig. 4. Interaction effect between FUPOOL and LMCOL1 (Ranked first)

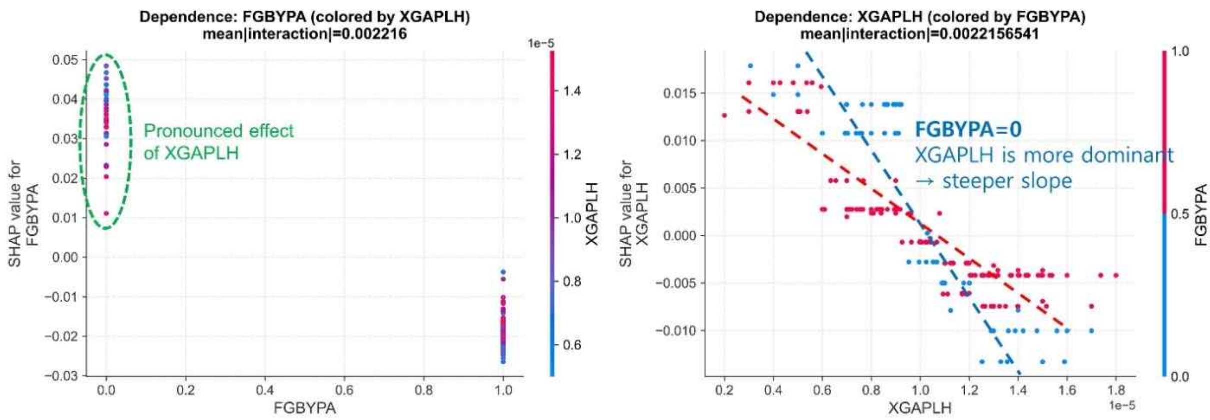


Fig.5. Interaction effect between FGBYPA and XGAPLH (Ranked second)

These results demonstrate that not only the individual effects of variables but also their interaction effects play an important role in determining the RPV failure time.

4. Conclusions

This study has demonstrated the potential of leveraging machine learning combined with XAI to analyze the influence of key variables that were challenging to predict using traditional uncertainty analysis derived from computational codes for severe accidents. While this research does not suggest that artificial intelligence can fully replace existing uncertainty evaluations, it highlights the necessity of interpreting and utilizing historical severe accident uncertainty analysis results as reference data for AI training. Importantly, the findings indicate that by employing data augmentation techniques on time-consuming thermal-hydraulic data, AI can indeed be trained and utilized to unearth hidden insights within the analysis results - signaling a positive outlook for the expansion of AI applications.

Looking ahead, securing the reliability of AI necessitates further investigation into comparing the actual effects of major variables, including the use of real-world data and the trustworthiness of data obtained through data augmentation methodologies. Additionally, there is a need for research on AI training and analysis methodologies covering a broader range of accident scenarios and phenomena. This holistic approach will ensure that AI technologies continue to evolve as a robust tool for enhancing the safety and resilience of nuclear power plants, providing valuable insights that complement and augment existing engineering practices and risk assessment frameworks.

REFERENCES

[1] EPRI, Severe Accident Uncertainty Quantification and Analysis Using the Modular Accident Analysis Program (MAAP), 2021 Technical Report, 2021

[2] T.-W. Kim, S. Shin, M. Seo, "Uncertainty Study of RPV Failure and Operator Actions in an MBLOCA Scenario of the OPR1000," Transactions of the Korean Nuclear Society Spring Meeting, May 2025.

[3] EPRI, Computer Code Manual for MAAP5 – Modular Accident Analysis Program for LWR Power Plants, 2021.

.