WILEY

Research Article

Enhancement of Reflood Test Prediction by Integrating Machine Learning and Data Assimilation Technique

Nguyen Huu Tiep [iD],[1,2] **Jong-Sung Kim** [iD],[1] **Hae-Yong Jeong,**[1] **Kyung-Doo Kim,**[3]
Nguyen Xuan-Mung,[4] **Hoai-Nam Tran** [iD],[5,6] **and Nguyen Ngoc Anh** [iD][5,6]

[1]*Department of Quantum and Nuclear Engineering, Sejong University, 209, Neungdong-ro, Gwangjin-gu, Seoul 05006,
Republic of Korea*
[2]*Institute for Nuclear Science and Technology (INST), Vietnam Atomic Energy Institute (VINATOM), 179 Hoang Quoc Viet,
Cau Giay, Hanoi 100000, Vietnam*
[3]*Korea Atomic Energy Research Institute (KAERI), 111, Daedeok-daero 989beon-gil, Yuseong-gu, Daejeon 34057, Republic of Korea*
[4]*Faculty of Mechanical and Aerospace Engineering, Sejong University, 209, Neungdong-ro, Gwangjin-gu, Seoul 05006,
Republic of Korea*
[5]*Phenikaa Institute for Advanced Study (PIAS), Phenikaa University, Hanoi 12116, Vietnam*
[6]*Faculty of Fundamental Science, Phenikaa University, Yen Nghia, Ha Dong, Hanoi 12116, Vietnam*

Correspondence should be addressed to Nguyen Huu Tiep; tiepngh@sejong.ac.kr, Jong-Sung Kim; kimjsbat@sejong.ac.kr,
Hoai-Nam Tran; nam.tranhoai@phenikaa-uni.edu.vn, and Nguyen Ngoc Anh; anh.nguyenngoc1@phenikaa-uni.edu.vn

Data assimilation (DA) was revealed as a highly efficient approach to enhance the prediction's accuracy, demonstrating the
reliability of the simulation through uncertainty quantification analysis. However, in the numerical simulations, most of the
tasks are highly nonlinear relationships between the parameters that directly affect the efficiency of the DA such as the large
search space dimension, resulting in significant computational costs. This limitation can be overcome by implementing
machine learning (ML) predictions that can efficiently expand the DA search space. Therefore, this study is aimed at
enhancing the performance of DA by integrating ML and DA (IMD), proposing a new method to overcome the stand-alone
DA limitations. The approach involved implementing a stand-alone DA using the multidimensional analysis of reactor safety
(MARS) code to enhance the prediction of reflood tests. The output datasets obtained from the stand-alone DA were then
used to train the deep neural networks (DNNs), and the accuracy of the DNN's prediction was thoroughly evaluated with
varying dataset sizes. This DNN subsequently can predict the enormous unobserved samples, enabling the identification and
investigation of the prospective candidates for the subsequent DA process. The results demonstrated that the stand-alone DA
achieved an accuracy enhancement of up to 41.3% in reflood test predictions, while the IMD yielded even more significant
improvements, with a performance enhancement of 47.0%. These findings reveal that the IMD approach outperformed the
stand-alone DA approach, particularly in the case of high flooding rate tests. In this context, the proposed integrating system
can effectively overcome the high computational cost and enhance the performance of stand-alone DA.

1. Introduction

In recent decades, investigations of a postulated accident, the loss of coolant accident (LOCA), using numerical and experimental methods have emerged. Experimentally, LOCAs and the effectiveness of the emergency core cooling system (ECCS) were comprehensively studied for the fuel bundle such as full-length emergency core heat transfer separate effects and systems effects tests (FLECHT SEASET) [1], rod bundle heat transfer (RBHT) [2], and flooding experiment with blocked arrays (FEBA) [3]. They focused on different initial boundary conditions and effects to enhance

the knowledge of postulated conditions after LOCA, a so-called reflooding phenomenon. To predict the complex heat transfer modes in the reflooding phase [4], these reflood test data can be used to evaluate and validate the physical models in various advanced numerical computer codes such as safety and performance analysis code (SPACE) [5], reactor excursion and leak analysis program (RELAP5) [6], multidimensional analysis of reactor safety (MARS) [7], coolant-boiling in rod arrays - two fluids (COBRA-TF) [8], TRAC/RELAP advanced computational engine (TRACE) [9], and computational multifluid dynamics (CUPID) [10, 11].

Moreover, the accuracy of reflood phenomenon prediction significantly depends on the correlated physical models included in the system codes. Therefore, enhancing these physical models can reduce the discrepancies between the predicted and measured values in reflooding simulations [4, 12, 13]. For example, using RELAP5, the effects of physical models and some sensitivity studies were investigated for the FLECHT SEASET experiments [14]. This analysis implied that the turnaround temperatures, turnaround times, and film boiling heat transfer coefficients were underestimated, which needed to be improved in the RELAP5 model. In addition, an assessment of the RELAP5 code against the FLECHT SEASET tests to seek the weaknesses of the reflood models was comprehensively studied [15]. Based on these weaknesses, the physical models of the RELAP5 code were justified to enhance the reflooding prediction. These physical models included the adjustment of droplet size in the dispersed flow regime and the modifications of reflood wall heat transfer. Moreover, a sensitivity analysis was performed for FLECHT SEASET tests using the CUPID code, which revealed that the film boiling was the most sensitive physical model in the reflood simulations, specifically for the peak cladding temperature prediction [16]. Furthermore, the film boiling heat transfer was also identified as the most crucial physical model in the SPACE simulation for reflooding phenomena [17, 18]. In this context, a method that propagates these physical models under uncertainty ranges so-called data assimilation (DA) can enhance the accuracy of the computer code prediction.

In 2022, the sampling method for highly nonlinear system uncertainty analysis (STARU) [17], a DA framework, was developed to deal with the highly nonlinear system and the high-dimensional parameters. STARU can be implemented to propagate the physical models and always find the better candidates, resulting the improvements in the predictions. It can be inferred that STARU enhanced the accuracy of the predictions for the reflood tests and could identify the most sensitive parameters after a dozen thousand computationally expensive samples. DA is consistently an excellent approach to enhance prediction and validate the physical models of simulations [17–22]. Even with the search space being limited in DA execution, there are prospects of obtaining valuable enhancements to the predictions. Moreover, expanding the search space in DA can facilitate better enhancements.

However, expanding the search space increases the computational costs, which is one of the main limitations of the stand-alone DA approach. Therefore, finding an integrated approach that can expand the search space with acceptable accuracy and minor computational costs is essential. In this context, machine learning (ML) predictions can be appropriate to improve DA, because an ML-trained model can predict the enormous unobserved samples while consuming a minimal computation cost and attaining good agreement with the actual values.

Furthermore, the ML predictions are practically revealed as a vital computation model to predict the experiment data within minor dissimilarities. For example, deep neural networks (DNNs) in ML [23–25] are effectively applicable to predict the system behavior of computer code predictions such as LOCA break size prediction using Cascaded fuzzy neural networks [26], reactor water level predictions [27], a minimum departure from nucleate boiling ratio (DNBR) estimation [28], and the wall heat transfer for boiling phenomenon prediction using DNN as well as high-fidelity simulation results [29]. Moreover, utilizing the dataset from the uncertainty quantification process, time series response forecasting (e.g., temperature, pressure, break flow rate, and water level) was adopted using long short-term memory (LSTM) and DNN [30]. DNNs, which typically consist of feedforward networks with multiple hidden layers, can effectively estimate the relationship between independent variables and dependent outputs, especially the highly nonlinear complex system states generated through the propagation of DA parameters or the uncertainty quantification process [30].

In addition, one of the significant efficiencies of DNN is that the computational costs of DNN predictions are meager compared to typical code simulations. This efficiency allows DNNs to effectively expand the search space and approximate nonlinear systems without relying on a deep understanding of the system's fundamental equations [31]. Expanding the search space can facilitate finding the better system states of the DA process [31, 32]. Conversely, high-quality datasets from DA can significantly enhance the training process of DNN models, leading to highly accurate predictions [33]. As a result, based on the DNN predictions for many samples, potential candidates can be identified and subsequently validated through DA.

In this context, the integrating system of ML and DA is expected to enhance DA performance significantly. In this study, we proposed an integrating ML and DA (IMD) system to overcome the limited search space of DA. Therefore, our proposed integrating system can enhance the DA performance and reduce the computation costs for highly nonlinear systems and high-dimensional parameter problems.

In particular, we aimed to implement DNNs for predicting the simulation accuracy of various responses, including cladding temperature, steam temperature, pressure drop, and quenching time, using the MARS code against the FLECHT SEASET reflood tests. Subsequently, the IMD approach was applied to enhance the performance of DA, which involved utilizing real experimental results and stand-alone data assimilation outcomes to evaluate the efficiency of the integrating system.

2. Methods and Materials

2.1. Problem Specification. The proposed solution can be summarized as follows.

TABLE 1: The selected FLECHT SEASET reflood tests [34].

Test no.	Flooding rate (mm/s)	Power (kW/m)	Initial clad temperature at 1.83 m (K)	Pressure (MPa)
F1-31021	38.6	1.3	1153	0.28
F2-31302	76.5	2.3	1142	0.28
F3-31504	24.0	2.3	1136	0.28
F4-33849	25.9	1.9	1018	0.28
F5-34103	38.1	2.4	1158	0.28
F6-34316	25.0	2.4	1162	0.28
F7-34420	38.9	2.4	1392	0.27
F8-34711	17.0	1.4	1161	0.13
F9-35050	25.9	1.6	1031	0.14

(1) We perform a stand-alone DA, in which physical model multipliers are assimilated using the STARU to search for better physical models. The output of this stand-alone DA is a set of different physical model multipliers that correspond to their system states

(2) These high-quality datasets are used to train DNN to predict the system states corresponding to a given physical model. The precision of this ML model is verified using a test dataset excluded from the DNN training process. This ML model is then implemented to predict enormous unobserved samples; the system states are predicted without using the MARS code

(3) Prospective samples can be obtained and implemented in DA again to assimilate the multipliers of the physical model

Therefore, it is expected that IMD could obtain better enhancements, and its computational cost is extensively lower than that of the stand-alone DA. The following two subsections will describe the FLECHT SEASET reflood tests, the reflood model in the MARS code, STARU, and DNN hyperparameters.

2.2. The FLECHT SEASET Reflood Tests and the MARS Model. The FLECHT SEASET test series [1] provides two-phase flow and experimental heat transfer databases of reflooding phenomena, which were crucial for model developers and users in reactor safety analysis. This test examined many effects and two-phase databases, such as rod temperatures, quenching times, heat transfer coefficients, inlet flooding rates, overall mass balance, differential pressures, void fractions, exhaust steam, and liquid carryover mass.

Accordingly, this test consisted of many valid tests such as steam cooling, gravity refill, boiloff, and forced reflood tests. However, in this study, we meticulously neglected the cases with scrammed power, lower power decay than the specified value, bundle distortion occurrence, unknown initial coolant, missing housing temperatures, and suddenly increasing flooding rate. As a result, nine forced reflood test cases were carefully selected in this investigation (see Table 1).
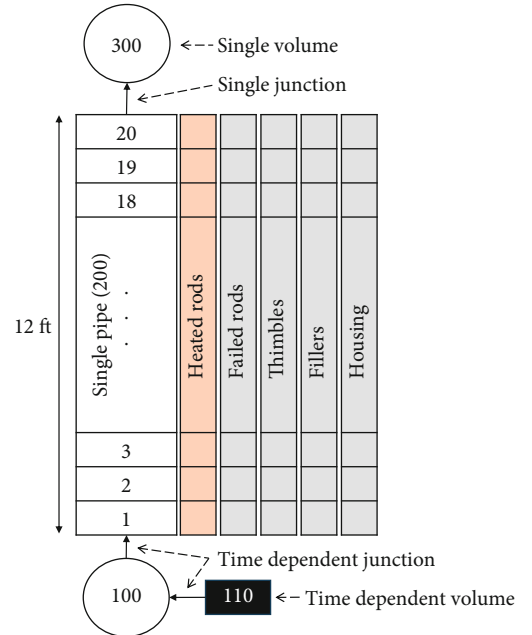

FIGURE 1: The MARS code model for reflood tests. The test section in the MARS code model was divided into 20 nodes. The heat structures included heated rods, failed rods, thimbles, fillers, and housing. The inlet of the test section was modeled using a time-dependent volume and a time-dependent junction, while the outlet used a single volume and junction, respectively.

This study analyzed the reflood tests using the MARS code developed by the Korea Atomic Energy Research Institute and Korea Institute of Nuclear Safety. This code was utilized with multidimensional analysis capacity and updated realistic physical models to treat with the three-dimensional two-fluid three-field formulation (liquid, vapor, and droplet field). Although the MARS code was restructured from the original RELAP5, the legibility, maintenance, and modification ability were enhanced.

The MARS model of the reflood test consisted of a lower plenum (single volume with flow boundary condition), a single pipe with twenty-axial cells, heat structures (the heated rods, failed rods, thimbles, fillers, and housing), and an upper plenum (single volume with pressure boundary condition) (see Figure 1).

TABLE 2: The physical models implemented in STARU for the MARS code.

No.	Type	Parameter descriptions	Uncertainty (%)
1	Entrainment coefficient	Multiplier for droplet We number for reflood	60
2	Dry/wet criteria	Multiplier for dry/wet wall criteria 30°C	60
3	Wall friction factors	Multiplier for the two-phase wall friction for the specified components	60
4	Interfacial drag model	Multiplier for the interfacial drag model for the specified components	30
5	Interfacial heat transfer model	Multiplier for the interfacial heat transfer model for the specified components	60
6		Multiplier for liquid Dittus-Boelter correlation	60
7		Multiplier for Chen nucleate boiling model	60
8		Multiplier for AECL CHF value	60
9		Multiplier for Chen transition boiling model	60
10		Multiplier for Bromley film boiling model	60
11	Correlations	Multiplier for vapor Dittus-Boelter correlation	60
12		Multiplier for Zuber CHF correlation	60
13		Multiplier for modified Weismann correlation	60
14		Multiplier for QF film boiling correlation	60
15		Multiplier for Forslund-Rohsenow correlation	60
16		Multiplier for reflood superheated vapor correlation	60

The investigations of these reflood tests would be verified using the experimental data provided in the FLECHT SEASET data report [1]. In this study, we investigated five responses as follows:

(1) Cladding temperatures (K): we examined the cladding temperatures of the heated rods at two different elevations (1.83 m and 2.44 m linked with cell number 11 and cell number 14 in the MARS model, respectively)

(2) Steam temperature (K): this response is the temperature of the vapor at 2.82 m that is associated with cell number 16 of the MARS model

(3) Pressure drop (MPa): this response is the different pressure between the inlet and outlet of the test section

(4) Quenching time (s): this response displayed the quench front location as a function of time; we implemented the "ZQBOT" reflood quantity to identify the quench front location for all nodes in the test section

2.3. The STARU Framework. It was clarified that the system behaviors of the reflooding phenomenon predictions were highly nonlinear, in which the dependency of prior and posterior parameters was unpredictable [17, 18]. Applying STARU can be a consistent approach for this nonlinear system to obtain prediction enhancements. In the STARU sampling algorithm, uniform and continuous samples can facilitate the convergence of system states. We deployed both the continuous and the uniform samples to sample the new candidates; if $x_i^p = \{x_1^p, x_2^p, \cdots, x_n^p\}$ is the previously accepted sample (with n as the number of parameters), the
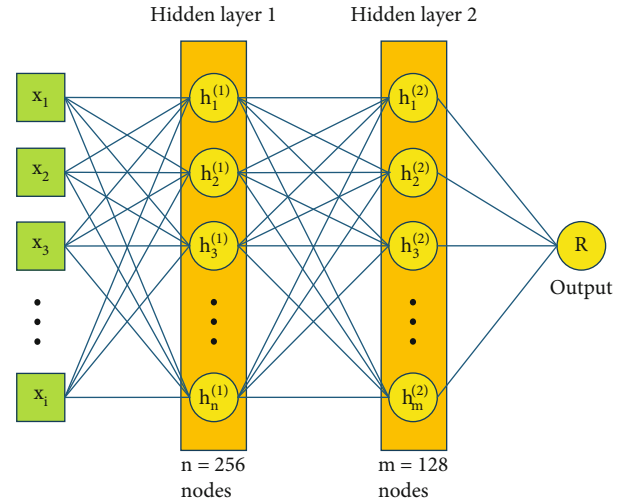

FIGURE 2: The architecture of the DNN to predict the system state (R). There are two hidden layers with 256 and 128 nodes, respectively.

TABLE 3: The selection of the DNN hyperparameters.

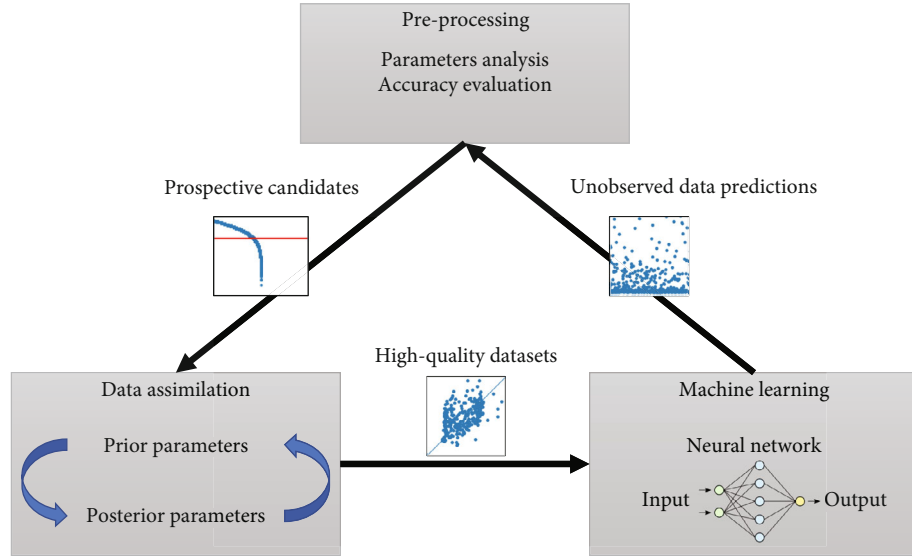| Quantity	Selected values
Number of hidden layers	2
Number of nodes in the hidden layer	256, 128
Loss function	MSE
Activator	ReLU
Optimizer	Adam
Learning rate	0.01
Epochs	200
Validation split	10%

FIGURE 3: The IMD scheme. The prior and posterior parameters in DA are referred to as the multipliers of the physical models before and after the DA process. DA can provide high-quality datasets, including parameter distributions, to train the ML model to generate prospective samples for the following DA process.

continuous sample can be generated as the following equation.

$$x_i^{\text{conti}} = x_i^p + \text{rand}\,(-\varepsilon\,;\varepsilon), \qquad (1)$$

where ε is the step size of the continuous sample and rand $(-\varepsilon\,;\varepsilon)$ is the random number from $-\varepsilon$ to ε. On the other hand, the uniform sample can be defined as

$$x_i^{\text{unif}} = \text{rand}\,(\sigma\,;\tau), \qquad (2)$$

where σ and τ were the uncertainty range's lower and upper bound, respectively. It deduced that the uniform samples in Eq. (2) could be obtained at random values within the uncertainty range, and it was not dependent on the previously accepted sample. Conversely, the continuous samples (see Eq. (1)) can only be obtained from values near the previous state within a small-justified step size ε. In our implementation, the sampling algorithm, the accuracy evaluation method, and the acceptance criteria were adapted from the original STARU [18]. However, minor modifications were made to fit the MARS code to be compatible with STARU.

One of the crucial tasks in the implementation of DA is the selection of the physical models. In this investigation, we selected sixteen physical models typically available in the MARS code for uncertainty quantification analysis [35], such as the multipliers for two-phase wall friction, the interfacial drag models, the interfacial heat transfer models, the entrainment coefficient, the droplet We number, and the multipliers of correlations along with their assumed uncertainty range (see Table 2).

In this selection, we comprehensively investigated the valid multipliers of these physical models before proposing their assumed uncertainties. Thereby, we found that the valid uncertainty range can be 60% for all the physical
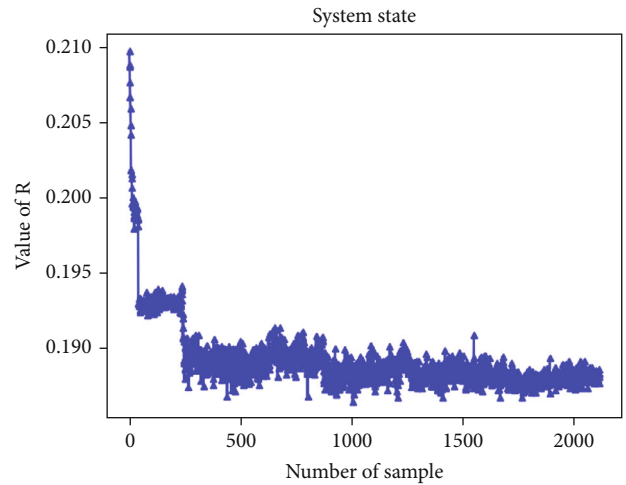

FIGURE 4: Accepted samples in the stand-alone DA using STARU. The system state was referred to as the value of R in Eq. (3).

models except the interfacial drag model, which is only valid for 30% or less. Moreover, this physical model was also confirmed as the most sensitive multiplier to predict the quenching times in reflooding phenomena using the MARS code [36].

Furthermore, the five responses were investigated and compared with the corresponding experimental data using the absolute relative difference (ARD) method in STARU (see Eq. (3)). Accordingly, the system states that the discrepancy between the current prediction and the experimental data can be evaluated as follows.

$$R = \sum_{j=1}^{m}\sum_{i=1}^{n}\left|\frac{S_{C_i}^j - S_{M_i}^j}{\sum_{i=1}^{n} S_{C_i}^j + \sum_{i=1}^{n} S_{M_i}^j}\right| * k^j, \qquad (3)$$

TABLE 4: The best enhancements for all the reflood tests.

Test case	System state (R value)								
	F1	F2	F3	F4	F5	F6	F7	F8	F9
Nominal	0.225	0.210	0.150	0.244	0.235	0.236	0.194	0.337	0.207
Stand-alone DA	0.188	0.192	0.088	0.201	0.197	0.156	0.179	0.301	0.189
IMD	0.180	0.111	0.116	0.185	0.211	0.100	0.186	0.294	0.199
Enhancement (%)	1-(Nominal)/(R^*)								
Stand-alone DA	16.7	8.2	41.3	17.5	15.9	33.9	7.7	10.8	8.5
IMD	20.2	47.0	22.6	24.1	10.3	57.7	4.2	13.0	3.6

*R value can be of either the stand-alone DA or the IMD.

where R is the system state; n is the total number of data; $S_{C_i}^j$ and $S_{M_i}^j$ are the calculated and measured values of the response j and at data i^{th}, respectively; j is the response; and k^j is the weighting factors associated with the response j.

2.4. DNNs. In this section, we described the generation and structure of the training dataset for the ML model. Furthermore, we provide a clarification of the deep learning method that was implemented and the selected hyperparameters for the DNNs.

As depicted in Table 2, we employed sixteen adjustable input parameters under their uncertainty range for the stand-alone DA using the MARS code. These adjusted input parameters and their corresponding system states were subsequently incorporated into the DNNs as training datasets. Specifically, in the stand-alone DA approach, system states reflecting the accuracy of the MARS code predictions were computed for 23,000 uniform samples. Each sample was generated using the uniform sampling approach in STARU and corresponded to a unique combination of values for sixteen physical model multipliers within their uncertainty range and their corresponding system state. It is important to reveal that the computational cost associated with these system states is very expensive, which has limited the improvements achieved by the stand-alone DA approach, presenting a significant challenge.

We proposed constructing a DNN model to address this issue to replace the computationally expensive MARS code calculations. DNNs are fundamental models in ML that comprise multiple hidden layers between the input and output layers. These models effectively estimate unknown underlying functions, mapping inputs to outputs based on provided datasets. By utilizing this DNN model, we can predict the accuracy of the MARS code simulation based on the sixteen parameters. This approach enables us to overcome the computational burden and enhance the efficiency of the DA process.

Figure 2 illustrates the architecture of the DNNs to predict the value of the system state using 16 input parameters as the input features to train. These DNNs comprised two hidden layers of 256 and 128 neurons, respectively. The output of these two hidden layers and the R can be determined as the following equations.

$$h_1 = f_A\left(W_1^T X + b_1\right),$$
$$h_2 = f_A\left(W_2^T h_1 + b_2\right), \quad (4)$$
$$R = f_B\left(W_3^T h_2 + b_3\right),$$

where X is the input layer; f_A and f_B are the activation functions; W_1^T, W_2^T, and W_3^T are the weight matrixes whose sizes are 16×256, 256×128, and 128×1, respectively; and b_1, b_2, and b_3 are biases.

For this activation function, we used the rectified linear unit (ReLU) [37] as the nonlinear activation function between hidden layers due to the consistency and simplicity of the computation compared to the sigmoid or tanh function for the regression problems.

In addition, the accuracy of the DNN predictions relies significantly on the size and quality of the datasets used, the coverage range, and the selection of appropriate hyperparameters [28–30, 33, 38]. Moreover, it has been observed that the optimal values of hyperparameters are highly dependent on the specific problem and dataset. Therefore, this topic remains an active area of research within the deep learning community, necessitating further investigations [29, 30, 38].

In this study, to ensure reliable DNN predictions, we conducted numerous trial-and-error tests to determine appropriate hyperparameters for our problem, as outlined in Table 3. For instance, achieving good accuracy requires a training dataset of over 20,000 samples. Furthermore, the Adam optimizer was employed in the DNN to adjust the weights and biases of the neurons in the hidden layers. The Adam optimizer was selected based on its computational efficiency, low memory requirements, and suitability for handling complex systems with large amounts of data and parameters [39]. Additionally, we observed that two hidden layers outperformed three or four layers when the number of nodes in the additional layers remained constant. Accordingly, we selected 256 nodes for the first hidden layer and 128 nodes for the second hidden layer. The training and testing datasets were divided with a ratio of 95 : 5, while the validating dataset was imposed to be 10% of the training datasets.

To prepare the training data for the neural network more efficiently, we utilized standardization to transform the value of the features so that they maintain similar scales. Standardizing the training data can enhance the model's generalization
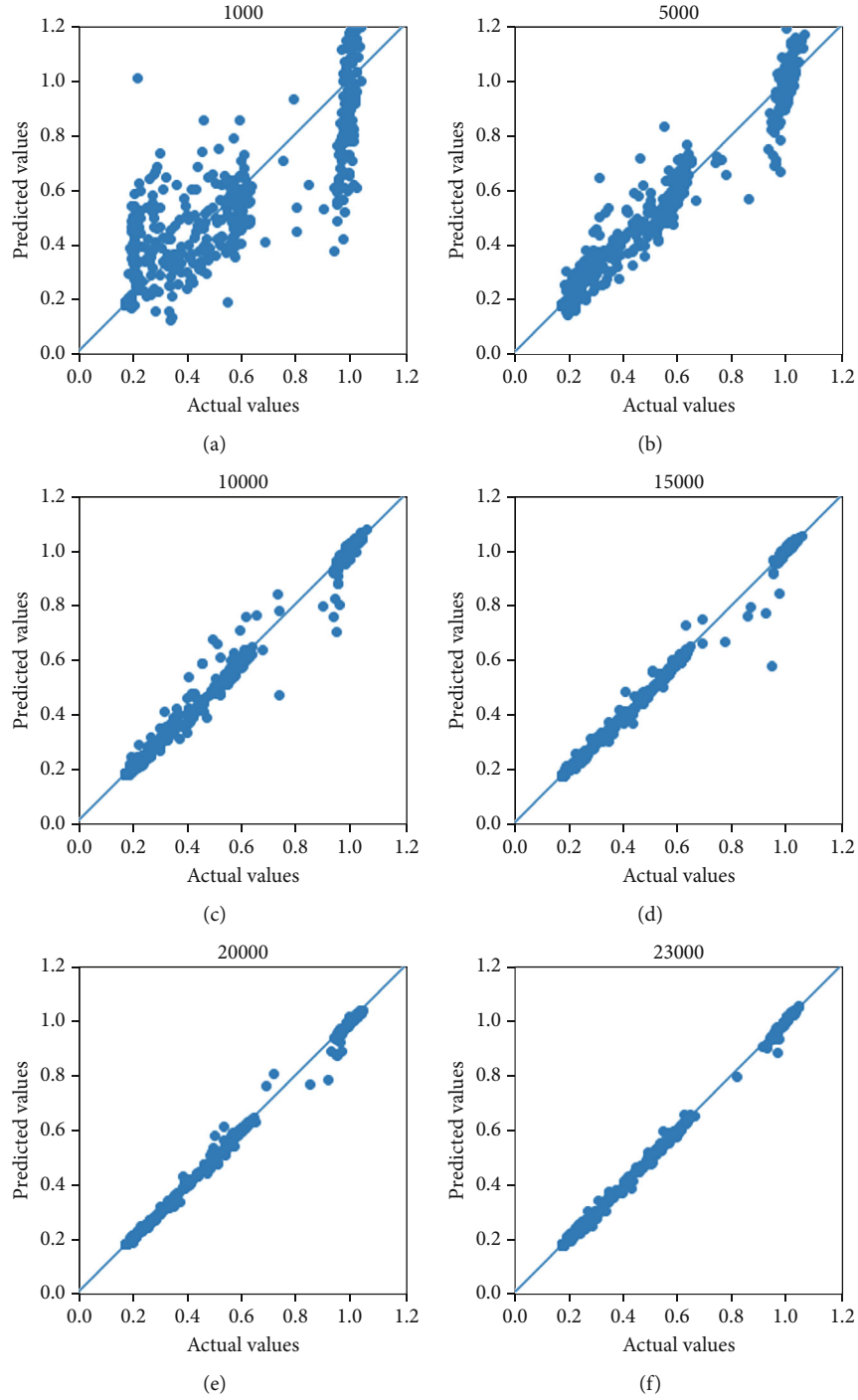
FIGURE 5: The ML predictions for the system state concerning the size of trained datasets. The actual values refer to the DA output values of R, computed by the MARS code, which are not included in the trained datasets. The numbers 1,000; 5,000; 10,000; 15,000; 20,000; and 23,000 are the sizes of the samples in the training dataset on (a)–(f), respectively.

performance, and the training process converges faster. During the training process, the mean squared error (MSE) was utilized as the loss function, defined as the average squared difference between the predicted and actual values (see Eq. (5)). One advantage of the MSE loss was that it was differentiable and could be used with gradient-based optimization algorithms such as Stochastic gradient descent or Adam. Accordingly, the model parameters can be easily optimized

by minimizing the MSE loss. The MSE loss for a single sample is given by

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \tilde{y}_i)^2, \tag{5}$$

where y is the predicted value, $\tilde{y}$ is the actual value, and N is

the number of datasets. After the training process, the mean absolute error (MAE) was also calculated as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \tilde{y}_i|. \tag{6}$$

The MAE is also used to evaluate the accuracy of the DNN in Eq. (6). The lower MSE loss or MAE indicated the more accurate prediction of the trained ML model. The selected DNN hyperparameters are listed in Table 3.

2.5. The IMD Scheme. In this implementation, "stand-alone DA" referred to the typical STARU process using the MARS code to enhance the FLECHT SEASET reflood test predictions. The high-quality output datasets from this stand-alone DA can be effectively used to train DNNs to evaluate unobserved samples. Thereby, this ML model can predict enormous unobserved samples with acceptable accuracy; this process refers to the expanding DA search space, which can facilitate the discovery of new-better system states. Therefore, we aimed to integrate DA and ML to find better system states than the stand-alone DA computations. In this integration, the first step will be the implementation of DA for the MARS code against the FLECHT SEASET reflood tests to generate high-quality datasets with the sixteen-dimensional parameters (associated with the sixteen physical models in Table 2) and their corresponding system state values. In the second step, we utilized standardization to justify the data scales of the datasets to train DNNs effectively. In the next step, after optimizing the ML model's accuracy, it can be used to predict the enormous unobserved samples generated by the uniform sampling technique to avoid bias in the value of the samples. Subsequently, the prospective samples were rigorously selected based on the upper limit value observed by the stand-alone DA process.

Accordingly, the performances of these prospective candidates were evaluated in the subsequent DA process (see Figure 3), in which the implementation of continuous samples was concentrated. For instance, a prospective candidate was examined in DA as the first accepted system state, and we used five continuous samples to enforce a high probability of finding the better system states. This justification can facilitate finding better candidates than a stand-alone DA.

3. Results and Discussions

3.1. The Stand-Alone DA Result. The MARS model for the reflood simulation, constructed in Section 2.1, was employed to predict responses and provide information for evaluating the system state using the ARD method in STARU (see Eq. (3)). We used 5,000 iterations in this computation, and about 2,100 samples were accepted. The accepted samples are better candidates than the current system states, satisfying the STARU framework's alpha condition [18]. Note that using the STARU sampling approach, the total samples will be 25,000 in 5,000 iterations. However, due to the computation error of the MARS code, some of these samples were not successfully computed; therefore, only about 23,000 samples were validly obtained. We found that the system state strongly
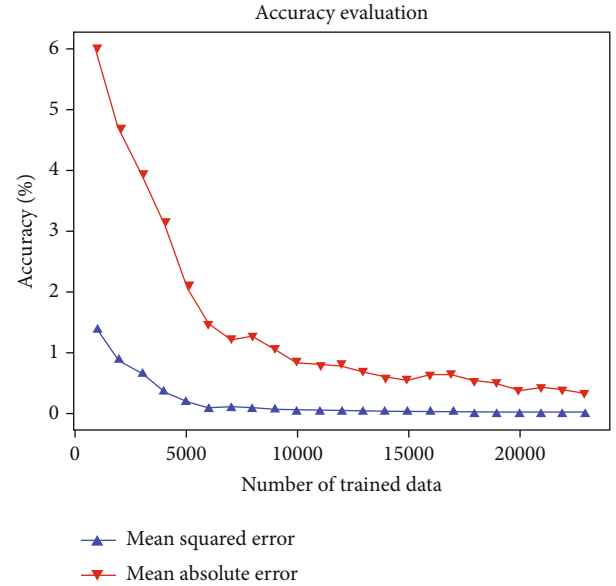

Figure 6: The MAE and MSE of the ML predictions regarding the size of the training datasets. The MAE and MSE values were calculated using Eqs. (5) and (6), respectively.
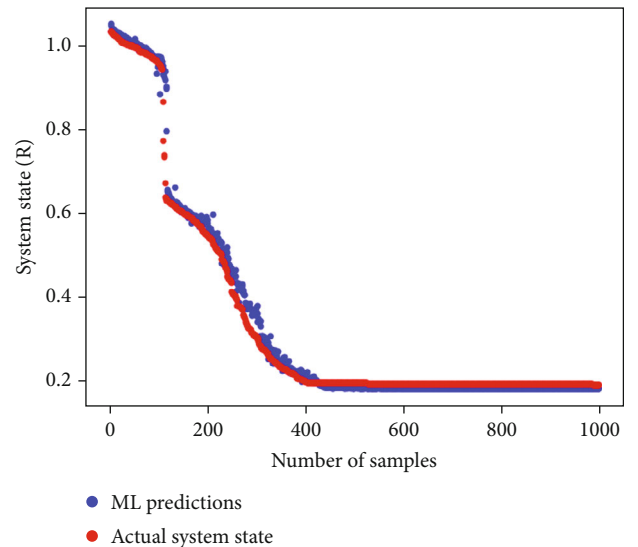

Figure 7: Comparison between the ML predictions and the actual system states. A thousand data points for ML predictions are compared with the actual values of the system state.

decreased after the first few iterations and then typically converged after about 300 iterations. It revealed that the enhancements of the MARS code to predict the reflooding phenomena were achieved (see Figure 4). These enhancements are quantitatively displayed in Table 4. However, we found that the enhancements of the steam temperature and pressure drop predictions could have been better if their predicted results had not fluctuated. The previous studies [17, 18] also indicated the same consequence that the enhancements of the steam temperature and pressure drop predictions were minor. Eventually, these stand-alone DA high-quality datasets were crucial
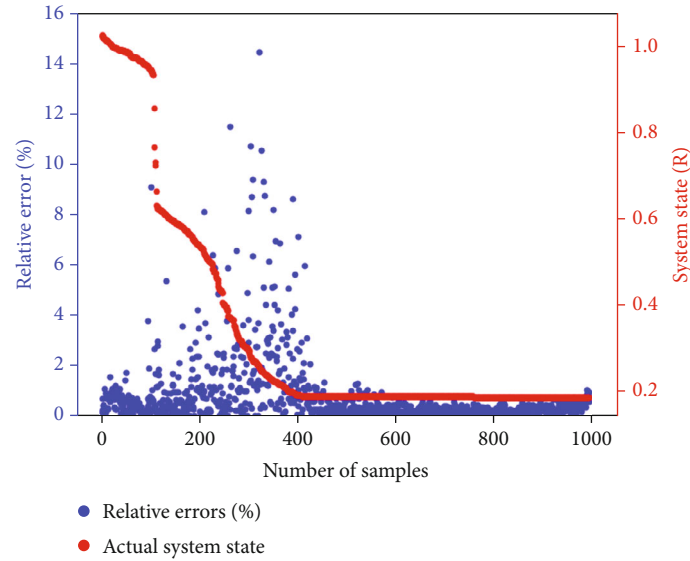
FIGURE 8: Relative errors in the system states predicted by ML. Relative errors were calculated using $(C - E)/C$, where E represented the actual values of the system state and C reflected the predicted ML values.

to training the DNN, especially in determining its minimum system state, which was used to select the prospective candidates introduced in the next section.

3.2. The ML Prediction and Validation. In this section, we first examined the data size effect to prepare an effective ML model to predict our system states because the size of the datasets considerably influenced the estimated target function in the ML model. To examine the effect of data size, we assessed the predictability of the trained ML model by tracking the MAE and the MSE regarding the different sizes of the training dataset (see Figures 5 and 6). Consequently, we confirmed that for the reflooding simulations using the MARS code, the ML model predictions agreed reasonably with the actual values when the number of trained data was greater than 20,000 samples. Furthermore, the machine learning-predicted values and their relative errors against the actual system state were evaluated using 23,000 trained samples and 1,000 test samples (refer to Figures 7 and 8). In this study, this amount of training and test datasets was subsequently implemented for further investigations.

Figure 8 illustrates the actual value of the system states corresponding to their relative errors in the ML predictions. It is found that the accuracy of the ML prediction for low system state is in good agreement (relative error less than 1.5%). These precise predictions allowed us to implement the ML model to predict the system state for highly nonlinear simulation results. Furthermore, this outcome facilitated seeking the better system states described by the low values in DA. Focusing on low system states is crucial in DA because it reveals minor discrepancies between the simulations and the experimental data.

Consequently, we implemented this ML model to predict one million uniform random samples, which consumed approximately two years of the stand-alone STARU
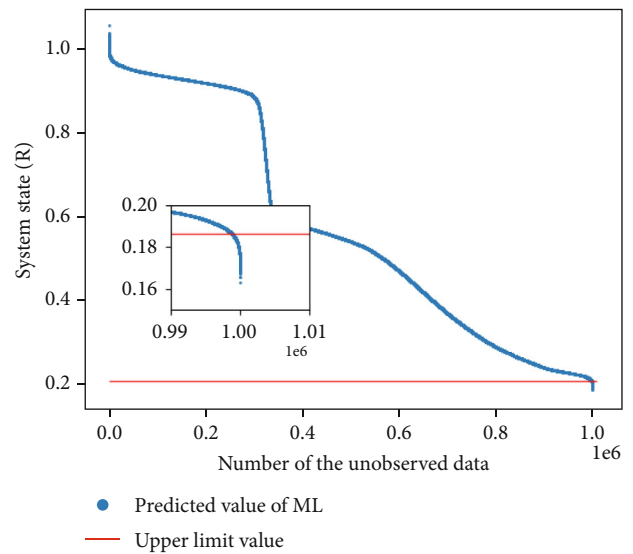

FIGURE 9: Selection of the prospective candidates. The predicted values of ML are referred to as the ML predictions for a million uniform sample in which their system states were rearranged. The red line indicated the upper limit value of a prospective candidate that corresponded to the minimum system state of the stand-alone DA in Section 3.1.

computation but only around two minutes using the ML model prediction. These prediction values of the enormous unobserved dataset were rearranged and displayed in Figure 9. To select the suitable candidates effectively, we assumed the upper limit value for the selected candidates must be lower than the current minimum system state of the stand-alone DA result. All these prospective candidates were subsequently implemented in the following DA process to find their best performance.
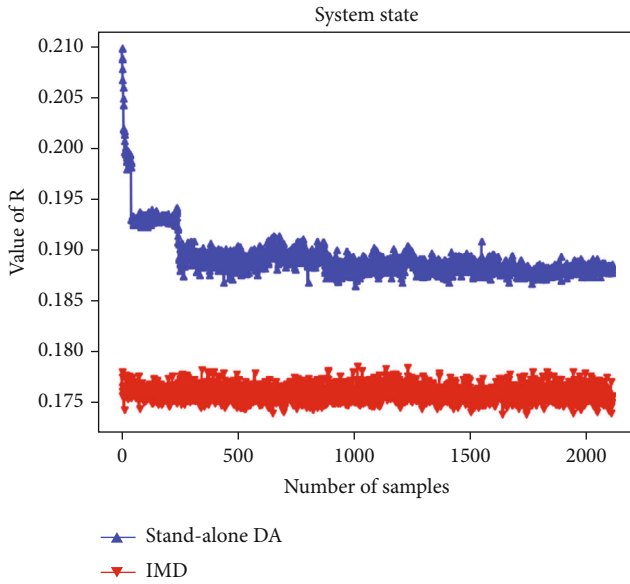
FIGURE 10: Comparison of the evolution of the system states (*R* values) between the stand-alone DA and the IMD.
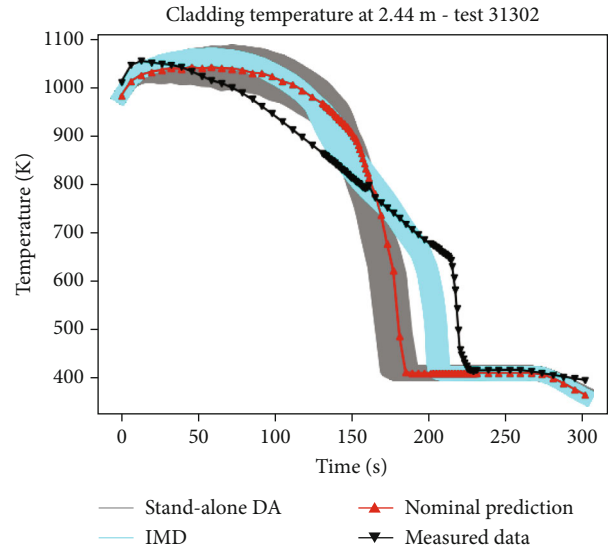

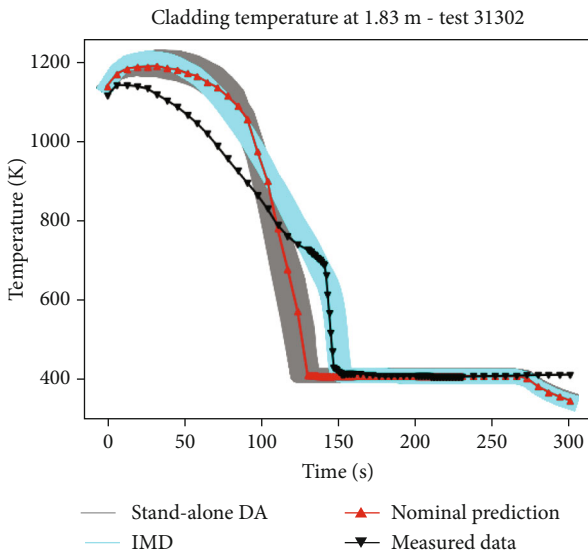FIGURE 12: The cladding temperature at 2.44 m comparison—test 31302.

FIGURE 11: The cladding temperature at 1.83 m comparison—test 31302. There are more than two thousand lines for the results of either the stand-alone DA or the IMD.
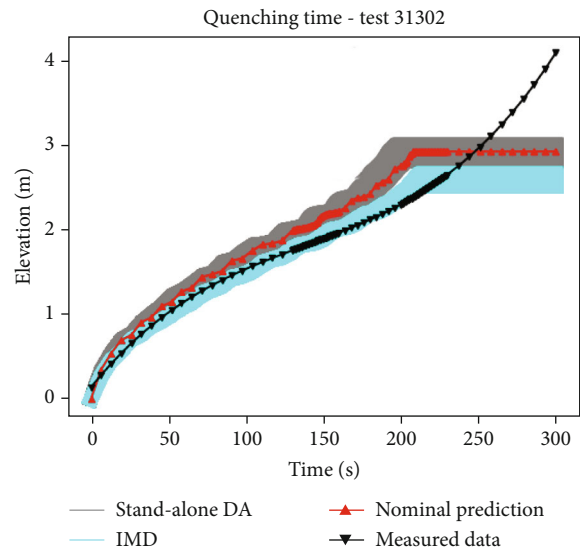

FIGURE 13: The quenching time comparison—test 31302.

3.3. The IMD Results.

The comparison of the system state for the IMD with the stand-alone STARU is illustrated in Figure 10. Furthermore, their enhancements for all FLECHT SEASET reflood tests are compared and displayed in Table 4. It indicates that the IMD enhancements are better than the stand-alone DA in finding the smaller system states. Even with smaller system states, IMD enhancements were sometimes not better than stand-alone DA.

In the case of stand-alone DA results, we also found that the enhancements for the high flooding rate reflood tests were challenging; this outcome was also indicated in the previous study [16–18]. Notwithstanding this situation, the IMD results indicated a better enhancement of the high flooding rate test F2-31302 (see Table 4). Specifically, we confirmed this enhancement by comparing the evolutions in the cladding temperatures, quenching time, and steam temperature prediction for the test F2-31302 (see Figures 11, 12, 13, and 14) in which the posterior responses of the stand-alone DA were compared with the IMD results, as well as the experiment data and nominal prediction of MARS code. Crucially, these results demonstrated that the IMD was an excellent approach to enhance the predictions of reflood tests.
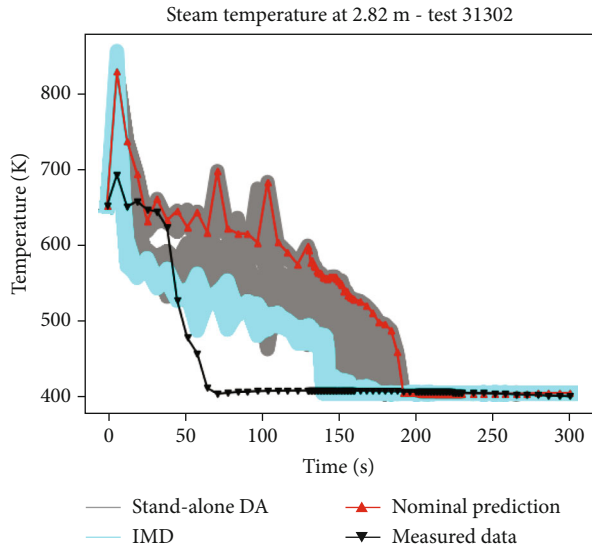
FIGURE 14: The steam temperature comparison—test 31302.

4. Conclusions

Our study is aimed at improving the prediction of reflooding phenomena in thermal hydraulics simulations through the integration of DA with ML, specifically employing DNNs.

Initially, we utilized stand-alone DA to enhance the predictions of the MARS code. However, due to the limitations of the search space, the effectiveness of stand-alone DA was somewhat restricted, necessitating further improvement.

To overcome this limitation, we utilized the datasets generated by stand-alone DA to train a DNN-based ML model. By expanding the search space, the ML model successfully identified better system states, as demonstrated by the accurate predictions which showed the effectiveness of the ML model, particularly with training datasets exceeding 20,000 samples.

Additionally, it was found that the ML predictions of the low system state value closely matched the actual system state, with a relative error of 1.5% or less. This outcome facilitated the identification of better system states by the ML model. Consequently, we employed this ML model to predict the large unobserved samples and carefully selected the prospective candidates for the subsequent DA process.

As a result, the IMD enhancements generally outperformed stand-alone DA, particularly in predicting the high flooding rate test. Specifically, the stand-alone DA achieved an accuracy enhancement of up to 41.3% in reflood test predictions, while the IMD yielded even more significant improvements, with a performance enhancement of 47.0%. This finding holds significant promise for the application and refinement of reflooding test predictions—an inherently challenging task in thermal hydraulics simulations. Moreover, future studies should concentrate on improving the hyperparameters of the DNN model to enhance its prediction accuracy for a limited training dataset.

Nomenclature

Acronyms

ARD:	Absolute relative difference
COBRA-TF:	Coolant-boiling in rod arrays - two fluids
CUPID:	Computational multifluid dynamics computer code developed by Korea Atomic Energy Research Institute
DA:	Data assimilation
DNBR:	Departure from nucleate boiling ratio
DNN:	Deep neural network
ECCS:	Emergency core cooling system
FEBA:	Flooding experiment with blocked arrays
FLECHT SEASET:	Full-length emergency core heat transfer separate effects and systems effects tests
IMD:	Integrating machine learning and data assimilation
LOCA:	Loss of coolant accident
LSTM:	Long short-term memory
MAE:	Mean absolute error
MARS:	Multidimensional analysis of reactor safety
ML:	Machine learning
MSE:	Mean squared error
RBHT:	Rod bundle heat transfer
RELAP5:	Reactor excursion and leak analysis program
SPACE:	Safety and performance analysis code for nuclear power plants
STARU:	Sampling method for highly nonlinear system uncertainty analysis
TRACE:	TRAC/RELAP advanced computational engine.

Symbols

x_i^p:	Previous accepted sample i^{th}
x_i^{conti}:	Continuous sample i^{th}
x_i^{unif}:	Uniform sample i^{th}
ε:	Step size of the continuous samples
σ:	Lower bound of the parameter uncertainty range
τ:	Upper bound of the parameter uncertainty range
R:	System state
$S_{C_i}^j$:	Calculated value of response j at data i^{th}
$S_{M_i}^j$:	Measured value of response j at data i^{th}
h_1, h_2:	Output of hidden layers
W_1^T, W_2^T, W_3^T:	Weight matrixes
b_1, b_2, b_3:	Biases
$\boldsymbol{X}$:	Input layer
f_A, f_B:	Activation functions
x:	Input to the activation function
y:	Predicted value of machine learning model
$\tilde{y}$:	Actual value of the test data.

Data Availability

The data used to support the findings of this study are available from the corresponding authors upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

[1] M. J. Loftus, L. E. Hochreiter, C. E. Conway et al., "PWR FLECHT SEASET unblocked bundle, forced and gravity reflood task data report," in *Nuclear Energy Systems Div*, vol. 1, No. EPRI-NP-1459, Westinghouse Electric Corp., Pittsburgh, PA (USA), 1981.

[2] L. E. Hochreiter, F. B. Cheung, T. F. Lin et al., *RBHT reflood heat transfer experiments data and analysis*, NUREG/CR-6980, U.S. Nuclear Regulatory Commission, Washington, DC, 2012.

[3] P. Ihle and K. Rust, *FEBA-Flooding Experiments with Blocked Arrays: Evaluation Report*, Kernforschungszentrum Karlsruhe, 1984.

[4] R. Mendizábal, E. de Alfonso, J. Freixa, and F. Reventós, "OECD/NEA PREMIUM benchmark final report," *NEA/CSNI/R*, vol. 18, 2016.

[5] S. J. Ha, C. E. Park, K. D. Kim, and C. H. Ban, "Development of the SPACE code for nuclear power plants," *Nuclear Engineering and Technology*, vol. 43, no. 1, pp. 45–62, 2011.

[6] C. D. Fletcher and R. R. Schultz, *RELAP5/MOD3 code manual*, Nuclear Regulatory Commission, Washington, DC (United States). Div. of Systems Research; EG and G Idaho, Inc., Idaho Falls, ID (United States), 1992, No. NUREG/CR-5535-Vol. 5; EGG-2596-Vol. 5.

[7] J. J. Jeong, K. S. Ha, B. D. Chung, and W. J. Lee, "Development of a multi-dimensional thermal-hydraulic system code, MARS 1.3.1," *Annals of Nuclear Energy*, vol. 26, no. 18, pp. 1611–1642, 1999.

[8] R. Salko Jr., M. Avramova, A. Wysocki et al., *CTF 4.0 Theory Manual*, Oak Ridge, TN (United States), No. ORNL/TM-2019/1145. Oak Ridge National Lab. (ORNL), 2019.

[9] US NRC, *TRACE V5. 0 Theory Manual*, ML120060218, US Nuclear Regulatory Commission, 2010.

[10] J. J. Jeong, H. Y. Yoon, I. K. Park, and H. K. Cho, "The CUPID code development and assessment strategy," *Nuclear Engineering and Technology*, vol. 42, no. 6, pp. 636–655, 2010.

[11] H. Y. Yoon, J. R. Lee, H. Kim et al., "Recent improvements in the CUPID code for a multi-dimensional two-phase flow anal-ysis of nuclear reactor components," *Nuclear Engineering and Technology*, vol. 46, no. 5, pp. 655–666, 2014.

[12] H. Akimoto and Y. Murao, "Development of reflood model for two fluid model code based on physical models used in REFLA code," *Journal of Nuclear Science and Technology*, vol. 29, no. 7, pp. 642–655, 1992.

[13] Y. Sudo, "Film boiling heat transfer during reflood phase in postulated PWR loss-of-coolant accident," *Journal of Nuclear Science and Technology*, vol. 17, no. 7, pp. 516–530, 1980.

[14] J. S. Baek, W. J. Lee, S. Y. Lee, and J. E. Kuh, "Assessments of FLECHT SEASET unblocked forced reflood tests using RELAP5/MOD3," *Nuclear Engineering and Technology*, vol. 24, no. 3, pp. 297–310, 1992.

[15] B. D. Chung, Y. L. Lee, C. E. Park, and S. Y. Lee, *Improvements to the RELAP5/MOD3 reflood model and uncertainty quantification of reflood peak clad temperature*, US Nuclear Regulatory Commission (NRC), Washington, DC (United States). Office of Nuclear Regulatory Research; Korea Inst. of Nuclear Safety, Taejon (Korea, Republic of), No. NUREG/IA-0132, 1996.

[16] J. M. Yoo, Y. J. Cho, H. Y. Yoon, and J. J. Jeong, "Analysis of the effect of liquid droplet models on the reflood heat transfer using the CUPID code," *Nuclear Engineering and Design*, vol. 354, article 110148, 2019.

[17] N. H. Tiep, K. D. Kim, and J. Heo, "Improvement in the accuracy of SPACE prediction for the unblocked FLECHT SEASET reflood tests by data assimilation," *Annals of Nuclear Energy*, vol. 161, article 108462, 2021.

[18] N. H. Tiep, K. D. Kim, J. Heo, C. W. Choi, and H. Y. Jeong, "A newly proposed data assimilation framework to enhance predictions for reflood tests," *Nuclear Engineering and Design*, vol. 390, article 111724, 2022.

[19] M. Asch, M. Bocquet, and M. Nodet, *Data Assimilation: Methods, Algorithms, and Applications*, Society for Industrial and Applied Mathematics, 2016.

[20] J. Heo and K. D. Kim, "PAPIRUS, a parallel computing framework for sensitivity analysis, uncertainty propagation, and estimation of parameter distribution," *Nuclear Engineering and Design*, vol. 292, pp. 237–247, 2015.

[21] D. G. Cacuci and M. Ionescu-Bujor, "Best-estimate model calibration and prediction through experimental data assimilation-I: mathematical framework," *Nuclear Science and Engineering*, vol. 165, no. 1, pp. 18–44, 2010.

[22] H. Abdel-Khalik, P. Turinsky, M. Jessee, J. Elkins, T. Stover, and M. Iqbal, "Uncertainty quantification, sensitivity analysis, and data assimilation for nuclear systems simulation," *Nuclear Data Sheets*, vol. 109, no. 12, pp. 2785–2790, 2008.

[23] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[24] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.

[25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[26] G. P. Choi, K. H. Yoo, J. H. Back, and M. G. Na, "Estimation of LOCA break size using cascaded fuzzy neural networks," *Nuclear Engineering and Technology*, vol. 49, no. 3, pp. 495–503, 2017.

[27] D. Y. Kim, K. H. Yoo, G. P. Choi, J. H. Back, and M. G. Na, "Reactor vessel water level estimation during severe accidents using cascaded fuzzy neural networks," *Nuclear Engineering and Technology*, vol. 48, no. 3, pp. 702–710, 2016.

[28] D. Y. Kim, K. H. Yoo, and M. G. Na, "Estimation of minimum DNBR using cascaded fuzzy neural networks," *IEEE Transactions on Nuclear Science*, vol. 62, no. 4, pp. 1849–1856, 2015.

[29] Y. Liu, N. Dinh, Y. Sato, and B. Niceno, "Data-driven modeling for boiling heat transfer: using deep neural networks and high-fidelity simulation results," *Applied Thermal Engineering*, vol. 144, pp. 305–320, 2018.

[30] M. I. Radaideh, C. Pigg, T. Kozlowski, Y. Deng, and A. Qu, "Neural-based time series forecasting of loss of coolant accidents in nuclear power plants," *Expert Systems with Applications*, vol. 160, article 113699, 2020.

[31] C. Buizza, C. Q. Casas, P. Nadler et al., "Data learning: integrating data assimilation and machine learning," *Journal of Computational Science*, vol. 58, article 101525, 2022.

[32] R. Arcucci, J. Zhu, S. Hu, and Y.-K. Guo, "Deep data assimilation: integrating deep learning with data assimilation," *Applied Sciences*, vol. 11, no. 3, p. 1114, 2021.

[33] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning (adaptive computation and machine learning series)*, Cambridge Massachusetts, 2017.

[34] N. H. Tiep, K. D. Kim, H. Y. Jeong et al., "Machine learning applications and uncertainty quantification analysis for reflood tests," *Applied Sciences*, vol. 14, no. 1, p. 324, 2024.

[35] B. D. Chung, K. D. Kim, S. W. Bae et al., *MARS code manual volume I: code structure, system models, and solution methods*, Korea Atomic Energy Research Institute, No. KAERI/TR-2812/2004, 2010.

[36] M. Hwang and B. D. Chung, "Improvement of MARS code reflood model," in *Proceedings of the ICONE-19 The 19th international conference on nuclear engineering*, Japan, 2011.

[37] V. Nair and G. E. Hinton, *Rectified linear units improve restricted Boltzmann machines*, International Conference on Machine Learning (ICML), 2010.

[38] S. Shin, Y. Lee, M. Kim, J. Park, S. Lee, and K. Min, "Deep neural network model with Bayesian hyperparameter optimization for prediction of NOx at transient conditions in a diesel engine," *Engineering Applications of Artificial Intelligence*, vol. 94, article 103761, 2020.

[39] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, International Conference on Learning Representations (ICLR), 2014.