

Machine learning-based forecasting of fission gas release in liquid metal-cooled reactor

Authors: Li, Junfeng, Jiawei Yang, Wang, Mr. Mou, Zhao, Dr. Pengcheng, Wang, Mr. Da, Yokoyama, Dr. Ryo, WANG, Dr. KAI, Zhang, Dr. Guangwei, Dr. Wenzhong Zhou, Zhou, Mr. Lijun, WANG, Dr. KAI

Date: 2025-04-26T18:00:09+00:00

Abstract

In this paper, we propose an artificial intelligence (AI)-driven framework for predicting fission gas release (FGR) in liquid metal-cooled reactors. Three AI models are trained on a comprehensive FGR database comprising 10,065 data points generated from multi-physics simulations. The results reveal that Random Forest (RF) achieves superior prediction accuracy, with a mean error (ME) of 3.96% for the optimal parameter combination (p_{lin0} , a_{grain0} , po , and t). While RF excels in database predictions, the Transformer model demonstrates exceptional capability in capturing parameter-effect trends, such as the positive correlation between FGR and fuel linear power or time. This study underscores the potential of AI models, particularly RF and Transformer, in advancing FGR forecasting for reactor safety analysis. This work can provide a replicable methodological paradigm for the FGR prediction of nuclear reactors and reliable technical support for the safety margin assessment of liquid metal-cooled reactors.

Full Text

Machine Learning-Based Forecasting of Fission Gas Release in Liquid Metal-Cooled Reactors

Junfeng Li¹, Jiawei Yang¹, Mou Wang¹, Pengcheng Zhao², Da Wang³, Ryo Yokoyama⁴, Kai Wang^{1,2,4*}, Guangwei Zhang^{5*}, Wenzhong Zhou¹, Lijun Zhou⁶

¹ Sino-French Institute of Nuclear Engineering and Technology, Sun Yat-Sen University, Zhuhai 519082, China

² Key Laboratory of Advanced Nuclear Energy Design and Safety, Ministry of Education, Hengyang 421001, Hunan, China

³ Iwhalecloud Co., Ltd, Building B, Yihua Industrial Park, Yuhuatai District, Nanjing City, China

⁴ Nuclear Professional School, School of Engineering, The University of Tokyo, 2-22 Shirakata, Tokai-mura, Ibaraki, 319-1188, Japan

⁵ School of Advanced Manufacturing, Sun Yat-sen University, Shenzhen 518107, China

⁶ JGC Corporation, Yokohama 220-6001, Japan

Corresponding author. E-mail: wangk326@mail.sysu.edu.cn (K. Wang), zhanggw33@mail.sysu.edu.cn (G. Zhang). Phone number: 18148828510 (K. Wang)

Abstract

In this paper, we propose an artificial intelligence (AI)-driven framework for predicting fission gas release (FGR) in liquid metal-cooled reactors. Three AI models are trained on a comprehensive FGR database comprising 10,065 data points generated from multi-physics simulations. The results reveal that Random Forest (RF) achieves superior prediction accuracy, with a mean error (ME) of 3.96% for the optimal parameter combination (p_{lin0} , a_{grain0} , p_0 , and t). While RF excels in database predictions, the Transformer model demonstrates exceptional capability in capturing parameter-effect trends, such as the positive correlation between FGR and fuel linear power or time. This study underscores the potential of AI models, particularly RF and Transformer, in advancing FGR forecasting for reactor safety analysis. This work can provide a replicable methodological paradigm for the FGR prediction of nuclear reactors and reliable technical support for the safety margin assessment of liquid metal-cooled reactors.

Keywords: Fission gas release, AI models, parameter combination, parameter-effect

1. Introduction

Coal, oil, and natural gas—the traditional three major fossil energy sources—cannot serve as humanity’s mainstream energy sources in the future due to the environmental pollution caused by harmful gases produced during combustion. Nuclear energy, as a clean and efficient large-scale centralized high-quality energy source, is not overly restricted by environmental conditions, and global fuel reserves are substantial, especially when advanced reactors and fuel cycles are employed. After years of development, fourth-generation nuclear reactors have been successfully deployed [1]. Compared with previous generations, the fourth-generation nuclear power system has significantly enhanced its economy, safety, and sustainable utilization through innovative design and technological optimization. Additionally, coolant selection further reduces the occurrence of secondary disasters. For instance, liquid metals (lead/lead-bismuth) or inert gases (helium) are used to replace the water medium in light water reactors,

avoiding explosion risks caused by hydrogen production from zirconium-water reactions.

Compared with traditional water-cooled reactors, liquid metal reactors exhibit superior heat transfer characteristics and higher heat capacity, as well as higher power generation efficiency and uranium resource utilization. However, during reactor operation, fuel generates large quantities of radioactive fission products through fission reactions, which continuously accumulate within the core and may be released into gaps. Under normal circumstances, the fuel cladding prevents fission products from entering the primary circuit. However, when the core cladding is damaged, fission gas accumulated in the cladding from nuclear reactions may be released into the coolant and migrate into the reactor's primary circuit, directly affecting the distribution of radioactive substances and potentially causing hazards to personnel and the surrounding environment. Therefore, evaluating fission gas release (FGR) in liquid metal-cooled reactors during accidents is critically important.

The migration behavior of fission gas during fuel cladding damage accidents is a complex physical process. First, fission gas is produced in large quantities in the fuel core through fission reactions and accumulates within it. Then the fission gas migrates from the fuel core to the fuel cladding gap. The gas is subsequently transported through the cladding gap, reaching the cladding breach, and finally releases into the coolant through the break. Based on this migration mechanism, Booth [2] first proposed a simplified diffusion model for fission product release from fuel pellets. The model assumes that fission gas diffuses to pellet boundaries driven by concentration gradients and then releases into the interstitial space [3]. Matzke [4] further discovered that below 1000°C, the diffusion coefficient in the Booth model is independent of temperature. For this reason, Turnbull et al. [5] proposed a diffusion coefficient formula including low-temperature effects. Kidson [6] considered the influence of precursor verification on diffusion release of fission products in the fuel matrix and added a correction factor to the Booth model. It must be acknowledged that most current FGR models are modifications based on the Booth model.

Most experimental studies on FGR in liquid metal-cooled reactors focus on fission gas behavior and characteristics after release into the coolant. Chawla and Høglund [7] simulated sudden fission gas release in a sodium-cooled fast reactor and studied the influence of transient flow caused by gas release on fast reactor assemblies. Building on this, Chawla [8] extended the coolant transient model. Van Erp et al. [9,10] conducted experiments on fission gas release due to failure of one or more fuel rods in sodium-cooled fast reactors and studied potential mechanisms leading to propagation of fuel rod failure. Haga et al. [11] performed experimental studies on fission gas release in sodium-cooled fast reactors under blocked fuel sub-channel conditions. Bell et al. [12] analyzed fission product release from lead-cooled fuel under severe accident conditions based on three fuel fragments immersed in liquid lead.

In addition, numerical simulation studies of FGR are emerging continuously.

Simulation methods are roughly divided into computational fluid dynamics (CFD), finite element methods, and multi-physics field coupling methods. Similar experiments such as fuel-coolant interaction jet injection [13], ordinary flash flow [14,15], and liquid flow entrainment [16] demonstrate that three-dimensional CFD can simulate two-phase interaction well and make relevant predictions for the entire physical process. Dong et al. [17] numerically simulated the burst release process of fission gas using CFD software Fluent based on a first-order dynamics model. Previously, the team [18] also used CFD to simulate FGR under defective conditions in pressurized water reactors, obtaining results within the empirical range. Bolotnov et al. [19] conducted numerical simulation analysis of fission gas release behavior during sodium-cooled fast reactor crust rupture using the NPHASE-CMFD program. Zhang et al. [20] simulated fission gas release behavior in fuel pellets based on finite element software ABAQUS, with results verified in [21].

In recent years, data-driven technologies represented by artificial intelligence (AI) have demonstrated unique value in nuclear reactor safety analysis and fuel performance research. Traditional empirical models are prone to strong non-linear effects during coupling of complex multi-physics fields, which not only interferes with simulation convergence but also reduces solution efficiency. In contrast, AI models can mine implicit patterns through data input without requiring understanding of internal physical mechanisms. Kaan et al. [22] developed an evolutionary machine learning correlation algorithm for VVER-1000 nuclear fuel gas release problems that can simultaneously predict fission gas fraction and hydrogen concentration. However, currently there are very few studies predicting FGR using AI methods.

Therefore, research on predicting FGR in fuel elements through AI models is particularly crucial. Based on a liquid metal fuel performance analysis model, this study obtained FGR data by adjusting key parameters and used three AI models to predict FGR and parameter trends. We believe this work can lay a methodological foundation for subsequent AI-based gas release behavior prediction.

This article is divided into five chapters. Chapter 2 introduces the FGR database and three AI models to determine optimal hyperparameter combinations. Chapter 3 investigates the impact of different input parameter combinations on AI models, compares prediction accuracy across models, and explores optimal input parameter combinations. Chapter 4 analyzes the effects of multiple input parameters and comprehensively assesses AI model predictive performance. Chapter 5 presents research conclusions and suggestions for future work.

2.1 Fission Gas Release Database

In this work, we referenced the annular uranium-plutonium mixed oxide (MOX) model for liquid metal-cooled reactors established by Cai et al. [23]. This model includes multiple nuclear technology modules such as fission gas release, poros-

ity change, and grain growth. It simulates key physical phenomena including material creep swelling, cladding corrosion, and Joint Oxide Gain (JOG) layer formation through multi-physical field coupling functions. Most importantly, the simulation results show high consistency with experimental results [24]. Therefore, based on this model, we constructed an FGR database after obtaining results by varying parameters. This database contains 10,065 data sets including seven parameters: initial fuel linear power (p_{lin0}), fuel grain size (a_{grain0}), deviation from stoichiometry of oxygen (Xdev), fuel porosity (po), initial fuel burnup (Bu_0), time (t), and FGR. The first six parameters serve as model inputs, while the output is the predicted FGR value. Among these, Xdev in MOX fuel is a key parameter characterizing deviation from ideal stoichiometric ratio. When Xdev is too small, an oxygen defect exists in the fuel. Table 1 shows the parameter ranges of the fission gas release database.

Table 1 The parameter range of the FGR database. Variable p_{lin0} (kW/m)

2.2.1 Transformer

Since its proposal, the Transformer architecture has demonstrated unique advantages in sequence modeling [25]. Its design breaks through the temporal processing limitations of traditional recurrent neural networks and instead builds a parallel computing framework with self-attention mechanisms at its core. This mechanism dynamically generates correlation weight matrices among elements in the input sequence, enabling the model to efficiently capture long-distance dependencies. Compared with loop structures relying on stepwise recursion, it significantly improves extraction efficiency of complex sequence features. The multi-head attention mechanism further allocates self-attention calculations to multiple independent subspaces. By processing contextual information from different subspaces in parallel, it achieves multi-dimensional representation of sequence semantics and enhances the model's ability to model complex contexts. To compensate for the lack of sequence order information, the Transformer introduces position encoding technology. It injects temporal position features into input embeddings through learnable parameters or trigonometric functions, ensuring the model can still perceive sequential dependence between elements during parallel processing. As shown in Fig. 1 [Figure 1: see original paper], the entire architecture consists of stacked encoder and decoder modules. The encoder extracts high-level features from the input sequence through alternating operations of self-attention layers and feedforward neural networks. The decoder introduces encoder-decoder attention mechanisms based on encoder output to achieve cross-modal alignment between source and target sequences. This hierarchical solution is designed for sequence transformation tasks such as machine translation and text generation. Thanks to the parallel computing characteristics of self-attention mechanisms, the Transformer can fully utilize hardware computing power during training, significantly shortening training cycles compared to mainstream architectures for current cross-modal tasks. Its structure design provides an efficient framework for large-scale models.

Fig. 1 The architecture diagram of Transformer.

2.2.2 Mamba

The Mamba model [26], as a new sequence modeling architecture in the post-Transformer era, redefines the processing paradigm for long sequence data through dynamic state compression and hardware-aware optimization. As shown in Fig. 2 [Figure 2: see original paper], its core innovation lies in introducing the Selective State Space Mechanism (SSM). By inputting relevant parameters to dynamically adjust the system matrix, it achieves focusing on key information and filtering of redundant information, thereby breaking through static parameter limitations of traditional state space models while maintaining linear computational complexity. This model decomposes recursive computations into parallel scanning operations, enabling reasoning speed to reach more than five times that of the Transformer and capable of handling long sequence data in the millions. Compared with the global attention mechanism of Transformer, Mamba achieves information compression through context-aware hidden state evolution. During the training stage, it adopts convolution mode for full sequence parallel processing, while during inference it switches to recursive mode to complete calculations with constant memory, significantly reducing reliance on video memory resources. Mamba has demonstrated outstanding performance in tasks such as language modeling, genomic analysis, and video timing modeling, marking a paradigm shift in sequence modeling from explicit attention-dependence to implicit state evolution.

Fig. 2 The architecture diagram of Mamba.

2.2.3 Random Forest

Random Forest is an ensemble algorithm widely used in traditional machine learning [27]. Its core feature is improving model stability and prediction performance by constructing a set of multiple decision trees, as shown in Fig. 3 [Figure 3: see original paper]. Unlike traditional single decision trees, RF reduces model variance and overfitting risk by introducing randomness. This randomness is reflected in two aspects: First, each tree's training dataset is generated from original data through sampling with replacement, i.e., the Bootstrap method; Second, when splitting each tree node, only a subset of features is randomly selected as candidate split attributes, thereby increasing differences between trees. Precisely this dual randomness enables RF to perform well when handling high-dimensional data while maintaining strong robustness against noise and outliers. Additionally, RF makes predictions through majority voting in classification tasks, while in regression tasks it calculates the average output of multiple trees, achieving flexible adaptation to different problem types. It also has high computational efficiency because each tree can be generated independently, supporting parallel training. This enables RF to run efficiently

on large-scale datasets and provides the ability to evaluate feature importance, offering a powerful tool for feature selection.

Fig. 3 The architecture diagram of the Random Forest.

2.3 Hyperparameter Analysis

To demonstrate predictive performance of the three AI models, we developed corresponding evaluation criteria: root mean square percentage error (RMSPE) and mean absolute percentage error (MAPE), represented by equations (1) and (2):

$$\text{RMSPE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_{\text{true}} - y_{\text{pred}}}{y_{\text{true}}} \right)^2} \times 100\%$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{\text{true}} - y_{\text{pred}}}{y_{\text{true}}} \right| \times 100\%$$

where n is the total amount of data, y_{true} is the true value, and y_{pred} is the predicted value. Compared with root mean square error (RMSE) and mean absolute error (MAE), the core advantages of RMSPE and MAPE lie in converting errors into relative proportions rather than absolute values, thereby eliminating dimensional influence and facilitating cross-scale comparisons. To present error magnitude more intuitively, we defined mean error (ME) as the final evaluation criterion, representing the average of RMSPE and MAPE:

$$\text{ME} = \frac{\text{RMSPE} + \text{MAPE}}{2}$$

Before training an AI model, hyperparameter setting is extremely important as it directly affects training speed, convergence, generalization ability, and final performance. Hyperparameters are not learned from training data but are preset by researchers based on experience and domain knowledge. Six hyperparameters were selected for this work: `batch_size`, `d_model`, `nhead`, `num_encoder_layers`, `dim_feedforward`, and `dropout`. In previous studies, the Transformer has demonstrated excellent predictive performance [28,29]. Therefore, we adopted the Transformer as the guiding model for this study. Regarding the Transformer architecture, we referenced previous settings and experiences, screening possible values for each hyperparameter as shown in Table 2.

Table 2 The value range of hyperparameters.

| Hyperparameters | Scope |
|-------------------------------------------------------------------|--------------|
| $\text{batch_}\{\text{size}\}$ | [256,512] |
| $\text{d_}\{\text{model}\}$ | [64,128,256] |
| nhead | [16,32,64] |
| $\text{num_}\{\{\{\text{encoder}\}\}_{\{\{\text{layers}\}\}\}}$ | [8,10] |
| $\text{dim_}\{\text{feedforward}\}$ | [4096] |
| dropout | [0.008,0.01] |

We preprocessed the FGR database, using 60% of data as the training set and 20% as the cross-validation set during the training stage. The remaining 20% served as the final test set to evaluate AI model predictive performance. Compared with L2 loss function, L1 loss function has stronger stability. Therefore, we selected L1 loss function as the basis for judging model convergence. Additionally, the learning rate was set at 0.0001. Subsequently, 72 permutations and combinations of hyperparameters were performed, obtaining different error results. We found that when $\text{batch_}\{\text{size}\}$ is 256, ME is relatively small. Therefore, Table 3 only shows 36 groups of hyperparameter combinations.

Table 3 The error result of hyperparameter combinations when $\text{batch_}\{\text{size}\}$ is 256.

The optimal hyperparameter combination obtained from Table 3 is: $\text{batch_}\{\text{size}\} = 256$, $\text{d_}\{\text{model}\} = 256$, $\text{nhead} = 16$, $\text{num_}\{\{\{\text{encoder}\}\}_{\{\{\text{layers}\}\}\}} = 8$, $\text{dim}\{\text{feedforward}\} = 4096$, and $\text{dropout} = 0.008$. Moreover, during hyperparameter debugging, we found that when epoch reached 100, the AI model gradually approached convergence. We set the final epoch at 300 and present the loss function curve for the optimal hyperparameter combination (Fig. 4 [Figure 4: see original paper]).

Fig. 4 The loss function curve graph of the optimal hyperparameter combination.

3. Influence of Input Parameters Combination

As mentioned in Wang et al. [29], sometimes AI model error decreases when input parameters decrease, indicating that model accuracy does not simply depend on the number of input parameters. This phenomenon can be attributed to the balance between model complexity and data quality: excessive input parameters may introduce redundant information or noise, leading to model overfitting, while reducing parameters can help the model focus on more relevant data features, thereby improving prediction stability and accuracy. However, due to the “black-box” nature of AI models, the specific mechanisms behind this performance are often difficult to explain intuitively, especially in deep learning models where this opacity stems from interactions among large numbers of parameters and nonlinear structures.

We reviewed relevant literature [30,31] and found that linear power and grain size have non-negligible influences on FGR. Additionally, since each dataset is temporal, time is also a necessary parameter. Finally, we set the initial input parameters as p_{lin0} , a_{grain0} , and t to study the influence of different input parameters on AI prediction performance.

3.1 Input Parameters Combination of Transformer Model

Since the Transformer is the guiding model of this work, we first studied input parameter combinations for the Transformer, with results shown in Table 4. When the number of input parameters is only 3 (Case1), the resulting error is relatively large. Therefore, we added one additional input parameter based on Case1, obtaining Cases 2-4. However, error for four-parameter combinations did not decrease significantly. In fact, ME for Case4 is even larger than for Case1. So we added two additional input parameters, obtaining Cases 5-7. Compared with Case1, RMSPE for Case6 decreased significantly. However, MAPE for the two cases is very close, so ME for Case6 is only about 3% lower than for Case1. Finally, when we increased input parameters to 6 (full parameter combination), ME dropped to 20.36%. This indicates that for this FGR database, only by accepting full parameter input can the Transformer model achieve optimal prediction performance.

Table 4 The error of different combinations of input parameters in the Transformer.

We selected the parameter combination with smallest error in each size and plotted a comparison graph between predicted results and true values (Fig. 5 [Figure 5: see original paper]). From a qualitative perspective, images for each case do not differ much. However, when quantified in table form (Table 4), it can be clearly seen that Case8 prediction performance is best. It is worth mentioning that when FGR exceeds 0.1, predicted values for all cases in Fig. 5 show good consistency with true values. Conversely, at low FGR, model prediction accuracy is relatively low. This might be because data is concentrated in the low FGR range, and data redundancy has significant impact on the Transformer.

Fig. 5 The comparison graph of predicted results and actual values in the Transformer. (a) Case 1, (b) Case2, (c) Case6, (d) Case8.

3.2 The Combination of Input Parameters of Mamba and RF Models

Based on Transformer results, we selected parameter combinations from Cases 1, 2, 6, and 8 as research objects for Mamba and RF models. Although in the Transformer, errors for these cases decrease as input size increases, we were curious whether this phenomenon would occur in Mamba and RF models. Therefore, we conducted the same research in Mamba and RF models, obtaining results in Table 5.

Table 5 The errors of different combinations of input parameters in the Mamba and RF models.

As shown in Table 5, prediction accuracy of the Mamba model does decrease as input size increases. However, the RF model deviates somewhat from our speculation. The error of the four-parameter combination for RF is larger than that of the three-parameter combination, although only by 0.01%. It should also be noted that ME for the RF model does not change much. This indicates that RF, as a traditional machine learning model, may experience slight overfitting when certain parameters increase. Additionally, although we mentioned in Chapter 2 that the Mamba model is more suitable for predicting time series data, Table 5 shows that the Mamba model has extremely high errors. Even for the full-parameter combination with highest prediction accuracy (Case12), the error is larger than all cases in the Transformer in Table 4.

Surprisingly, prediction accuracy of the RF model is the best among the three models, and its error is much smaller than that of the Transformer. Even with a three-parameter combination, ME for Case13 can still reach around 4%. Just as mentioned in [29], RF has very good predictive performance for a single database. We selected cases with the lowest ME in Mamba and RF models and plotted the result comparison chart (Fig. 6 [Figure 6: see original paper]). It can be found that prediction accuracy of the Mamba model is much lower than other models. More importantly, predicted values from Mamba tend to be negative, which violates physical rules. Predicted values from the RF model highly overlap with true values, demonstrating good stability and adaptability.

Fig. 6 Comparison chart of predicted results and actual values in Mamba and RF. (a) Case12, (b) Case16.

3.3 Optimal Parameter Combination Analysis

Error magnitudes of representative parameter combinations were presented previously. However, for AI models, parameter sensitivity varies, and prediction results differ significantly. Therefore, in this section, we used the RF model to perform full permutation and combination of the six parameters, with results shown in Appendix A. We plotted ME for 63 groups of parameter combinations as a graph (Fig. 7 [Figure 7: see original paper]) and obtained the following results: (1) When input size is 4, the RF model achieves the lowest error, with $ME = 3.95\%$. The corresponding optimal parameter combination is $p_{\{lin0\}}$, $a_{\{grain0\}}$, po , and t . (2) When the parameter combination is $p_{\{lin0\}}$ and t , ME for the RF model is 16.62%, which is even lower than ME for optimal parameter combinations of Transformer and Mamba in Chapter 3. This indicates that for prediction of a single database, RF has unparalleled advantages. (3) When only single parameter input is discussed, $p_{\{lin0\}}$ is more helpful for RF model prediction than t . Although RF achieved 313.5% error when input size was 1, $p_{\{lin0\}}$ was the optimal input parameter. (4) Although fuel burnup is a key parameter for FGR in simulation processes, in RF model prediction, ME

for parameter combinations with Bu_0 added is relatively large.

Fig. 7 The relationship diagram between ME and input size.

3.4 Summary

The work of this chapter studied and analyzed different combinations of input parameters. Based on Transformer error results, representative parameter combinations were applied to Mamba and RF. Moreover, full permutation and combination of RF model parameters yielded the following results:

1. When input parameters increase, ME for both Transformer and Mamba basically decreases. However, ME for the four-parameter combination of RF (Case14) is larger than for the three-parameter combination (Case13).
2. Among these three models, RF has the best prediction accuracy, with ME being 3.96%. Transformer ranks second. Mamba model prediction accuracy is the lowest, with even its lowest ME showing only 28.83% prediction accuracy.
3. Regardless of the model, prediction accuracy for low FGR is lower than for high FGR. This might be because data is concentrated in the low FGR range, and uneven data distribution has considerable impact on AI model predictions.
4. For the FGR database, the minimum ME of the RF model is 3.95%, and the corresponding optimal input parameter combination is $p_{\{lin0\}}$, $a_{\{grain0\}}$, p_o , and t .

4. Input Parameter Effects

This chapter studies the effects of different parameters on FGR based on error analysis from Chapter 3. Analyzing the influence of a single parameter on results is very necessary. This can not only eliminate interference from other variables but also reveal inherent logical relationships between parameters and results. Carroll et al. [32] conducted irradiation experiments on specimens with different microstructures to study effects of parameters such as fission density and fuel burnup on FGR. Bernard et al. [33] established a MOX fuel model for pressurized water reactors and evaluated FGR output under different fuel burnup and power conditions. All these works were accomplished using traditional research methods. Therefore, we innovatively used AI models to predict and analyze effects of single parameters. On one hand, due to limitations of traditional methods, it is difficult to obtain data in extreme cases. On the other hand, by studying parameter effects, predictive performance of different models can be comprehensively evaluated.

We produced different data slices. In each group of slices, the effect of a single parameter on FGR was studied using control variables, as shown in Table 6. The first line of each group represents parameter values during model prediction, where only one input parameter changes while others remain constant. The second line represents true values from the FGR database. Since simulation results

may have small offsets, to eliminate influence from result oscillation, we set an independent parameter reference range. It is particularly worth mentioning that since parameter values of p_0 and Bu_0 in the database are basically maintained around 0.05 and 0.001 respectively, parameter values for each group remain unchanged. In this chapter, we adopt the best parameter combination for each model—full parameter input—to study the influence of $p_{\text{lin}0}$, $a_{\text{grain}0}$, X_{dev} , and t on FGR.

Table 6 FGR data slice.

4.1 Fuel Initial Line Power Effect

As shown in Fig. 8 [Figure 8: see original paper], predicted values from all three models basically increase as $p_{\text{lin}0}$ increases, and FGR also increases. This trend matches the true value. Mogensen et al. [34] also mentioned this trend. In their study, as line power increased, FGR could increase from 0.25 to 0.4. This is because increased linear power raises core block temperature, and increased temperature promotes diffusion of fission gas atoms in fuel and formation of lattice defects, causing gas to accumulate more quickly on grain boundary surfaces. When saturation concentration is reached, release begins. Therefore, gas release at high power will be significantly advanced, and the fission reaction will be more intense at high power, corresponding to higher gas release.

Additionally, according to Fig. 8, although trends of the three models are the same, predicted values from Transformer and Mamba models are more consistent with true values, while RF predicted values are larger in most cases. This seems different from previous results where RF error was much lower than Transformer and Mamba. It is notable that when $p_{\text{lin}0}$ exceeds 40 kW/m, predicted values from all models decrease. This might imply that when $p_{\text{lin}0}$ reaches a certain threshold, its promoting effect on FGR begins to weaken. We believe this turning point may be related to computational characteristics of the model when dealing with high parameter values or changes in data distribution. Subsequent research can optimize model structure or parameter settings by deeply investigating the internal mechanism of this turning point.

Fig. 8 The influence of the initial line power on FGR.

4.2 Fuel Grain Size Effect

The relationship between $a_{\text{grain}0}$ and FGR is shown in Fig. 9 [Figure 9: see original paper]. Predicted values from both Transformer and Mamba models decrease as $a_{\text{grain}0}$ increases, consistent with true value trends. However, the prediction curve presented by RF was much as expected. Only when $a_{\text{grain}0}$ changed from 1.1 to 1.3 did predicted values show a relatively large downward slope. At other times, they remained stable. This indicates that the RF model is not competent for studying grain size effects.

Turnbull [35] conducted irradiation experiments on samples of different sizes and obtained the same result: as grain size increased, FGR would decrease. Meanwhile, Brown et al. [36] also mentioned there is a very strong negative correlation between grain size and FGR. This is because small grains have relatively high grain boundary density. These grain boundaries become rapid diffusion channels for fission gas atoms, promoting gas migration and release to the material surface. However, large-grain samples have significantly fewer grain boundaries, gas diffusion paths are blocked, and more fission gases are confined within grains, making them difficult to escape. Additionally, the crystal structure of large grains is relatively stable with fewer internal defects, which further reduces fission gas diffusion rate.

Fig. 9 The influence of fuel grain size on FGR.

4.3 Deviation from Stoichiometry of Oxygen Effect

As shown in Fig. 10 [Figure 10: see original paper], both predicted values from Transformer and real FGR increase with Xdev, although the increment of the former is relatively small. For parameter Xdev, the Mamba model did not demonstrate excellent predictive performance. The general trend was to decrease first and then increase. Similarly, the RF model still fails to provide satisfactory results. From this perspective, the RF model has a considerable gap compared with other models in parameter trend prediction. The conclusion of Cai et al. [23] is the same as the prediction result of Transformer, and FGR shows a positive correlation with Xdev. They mentioned that when Xdev becomes larger, fuel thermal conductivity decreases, causing fuel temperature to increase and eventually leading to increased FGR. Additionally, increased fuel temperature leads to intensified changes in fuel-cladding clearance size. At high temperatures, fuel swelling becomes more severe, resulting in faster reduction in gap size.

Fig. 10 The influence of deviation from stoichiometry of oxygen on FGR.

4.4 Time Effect

Fig. 11 [Figure 11: see original paper] shows the time effect on FGR. Prediction results from all three models show high consistency with true values. As time progresses, FGR increases, which aligns with our understanding. The essence of this time dependence stems from coupling effects between nuclear fuel microstructure evolution and fission gas behavior during irradiation: In the initial stage, fission gas atoms (such as krypton, xenon, etc.) are mainly captured in the fuel lattice as point defects or small bubbles, with relatively low release rates. With extended irradiation time, fission gas accumulates, bubbles gradually agglomerate and grow, and grain boundaries are activated as fast channels for gas migration, causing more gas to diffuse to the fuel surface through grain boundaries and release.

From the perspective of model prediction characteristics, accurate capture of

time effects by the three models reflects their ability to learn complex temporal features. Whether Transformer and Mamba models based on sequence modeling or the RF model based on ensemble learning, they can all effectively identify implicit cumulative effect laws in time series. It is worth mentioning that among these three models, predicted values from the Transformer are closer to true values.

Fig. 11 The influence of time on FGR.

4.5 Summary

This chapter studied and analyzed the effects of four parameters: initial line power, grain size, oxidation chemical coefficient offset, and time, obtaining the following conclusions:

1. Model prediction results and true values indicate that FGR increases as p_{lin0} increases. However, when p_{lin0} is 45 kW/m, predicted values from all three models decrease. Perhaps this is because the AI model considers that p_{lin0} has reached a threshold at 40 kW/m, and FGR begins to decrease.
2. As a_{grain0} increases, FGR gradually decreases. Both Transformer and Mamba models can conform to real trends, while the RF model cannot demonstrate good predictive performance.
3. For Xdev prediction, only the Transformer provides the correct trend curve, and there is a positive correlation between FGR and Xdev. Neither Mamba nor RF models can capture key features.
4. Prediction curves from all three models indicate that FGR increases over time, consistent with true values. This reflects their high learning ability for complex temporal series features.
5. Results of the four parameter effects indicate that the Transformer has the best prediction performance for parameter trends. Although the RF model has the best adaptability to the FGR database, in parameter effect studies, RF is far inferior to other models. This might be because the RF model is prone to overfitting when dealing with complex nonlinear relationships, resulting in systematic deviations in parameter effect predictions.

5. Conclusion and Prospect

This study predicts FGR based on AI methods, discusses the influence of multiple combinations of input parameters on AI model prediction accuracy, and evaluates model applicability to parameter effects. Through this work, we have drawn several important conclusions:

1. For the FGR database, model prediction accuracy ranks from high to low: RF > Transformer > Mamba. The optimal input parameter combination for RF is p_{lin0} , a_{grain0} , p_0 , and t , with corresponding ME of only 3.95%.

2. When RF model input size is 1, $p_{\{lin0\}}$ has the smallest error. Although fuel burnup is a key parameter for FGR, in the RF model, ME for parameter combinations with Bu_0 is relatively large.
3. For all three models, they exhibit relatively low prediction performance in the low FGR interval, which may stem from training data distribution characteristics. Due to highly concentrated and heterogeneous distribution characteristics of sample data within the low FGR range, this unbalanced data structure may lead to insufficient feature learning by the model in this region.
4. When studying input parameter effects, the RF model obtained the worst prediction results. This might be because in high-dimensional covariate space, RF has inherent limitations in capturing mechanisms of nonlinear parameter interactions. Especially when parameter effects show non-stationarity characteristics, the splitting criterion of decision trees is difficult to effectively identify complex action paths.
5. Under comprehensive evaluation, the Transformer model has the best overall prediction performance. Although its error in FGR database prediction is larger than that of the RF model, the Transformer can demonstrate accurate parameter trends, so this model has the greatest application value. Its error is still within acceptable range.

This work demonstrates the great potential of Transformer and RF models in FGR prediction, indicating that AI methods, after training and optimization, can provide auxiliary prediction for gas release behavior under reactor accident conditions. However, research always has limitations. We have put forward some suggestions for future work:

1. Due to current scarcity of FGR experimental data, it is impossible to form a unified database. Therefore, researchers are encouraged to focus on multi-dimensional data collection in high-temperature irradiation environments of nuclear fuel, and expand accumulation of effective data through collaborative experimental designs across reactor platforms and irradiation conditions.
2. Model input parameters have significant impact on prediction accuracy. Therefore, future work can add parameters relatively important for FGR, such as temperature and creep parameters. Through quantitative analysis of parameter interactions, model performance under complex working conditions can be significantly improved.
3. Although AI models have relatively good predictive performance, we have no way of knowing their prediction methods. The “black box” nature of AI methods has always been a key bottleneck restricting their in-depth application in nuclear engineering safety. Future research needs to focus on deep integration of interpretable AI technology and prior knowledge of nuclear physics, transforming model weight matrices into parameter correlation graphs with physical significance.
4. As both Transformer and RF have their own advantages, establishing an integrated model is crucial for future research. This method can achieve

complementary advantages, improve prediction accuracy, and conduct efficient assessment of FGR risks under complex working conditions.

Acknowledgement

This work is supported by open fund from Key Laboratory of Advanced Nuclear Energy Design and Safety, Ministry of Education (KLANEDS202308), and the Science and Technology Program of Guangdong Province (Grant No. 2023A1515110608, 2024A1515011614).

References

- [1] B. Zohuri, 6 - Generation IV nuclear reactors, in: S.U.-D. Khan, A. Nakhabov (Eds.), *Nuclear Reactor Technology Development and Utilization*, Woodhead Publishing, 2020: pp. 213–246. <https://doi.org/10.1016/B978-0-12-818483-7.00006-8>.
- [2] A.H. Booth, C.R. Atomic Energy of Canada Limited, A method of calculating fission gas diffusion from UO₂ fuel and its application to the X-2-f loop test, Atomic Energy of Canada Limited, Chalk River, Ontario (Canada), 1957. <https://inis.iaea.org/records/c9rqa-e4713> (accessed April 24, 2025).
- [3] A.H. Booth, A suggested method for calculating the diffusion of radioactive rare gas fission products from UO₂ fuel elements and a discussion of proposed in-reactor experiments that may be used to test its validity, Atomic Energy of Canada Limited, 1957.
- [4] H. Matzke, Radiation enhanced diffusion in UO₂ and (U, Pu)O₂, *Radiation Effects* (1983). <https://doi.org/10.1080/00337578308224715>.
- [5] J.A. Turnbull, C.A. Friskney, J.R. Findlay, F.A. Johnson, A.J. Walter, The diffusion coefficients of gaseous and volatile species during the irradiation of uranium dioxide, *Journal of Nuclear Materials* 107 (1982) 168–184. [https://doi.org/10.1016/0022-3115\(82\)90419-6](https://doi.org/10.1016/0022-3115(82)90419-6).
- [6] G.V. Kidson, A generalized analysis of the cumulative diffusional release of fission product gases from an “Equivalent Sphere” of UO₂, *Journal of Nuclear Materials* 88 (1980) 299–308. [https://doi.org/10.1016/0022-3115\(80\)90286-X](https://doi.org/10.1016/0022-3115(80)90286-X).
- [7] T.C. Chawla, B.M. Høglund, A Study of Coolant Transients During a Rapid Fission Gas Release in a Fast Reactor Subassembly, *Nuclear Science and Engineering* (1971). <https://doi.org/10.13182/NSE71-A20165>.
- [8] T.C. Chawla, Analysis of Inertial Contribution of the Fluid in a Plenum to a Rapid Coolant-Expulsion Process in a Liquid Metal Fast Breeder Reactor Single-Subassembly Accident, *Nuclear Science and Engineering* (1972). <https://doi.org/10.13182/NSE72-A22507>.
- [9] J.B. Van Erp, T.C. Chawla, R.E. Wilson, POTENTIAL FUEL FAILURE PROPAGATION DUE TO FISSION-GAS RELEASE IN LMFBR

SUBASSEMBLIES., 1972. <https://doi.org/10.2172/4599629>.

[10] J.B. van Erp, T.C. Chawla, H.K. Fauske, An evaluation of pin-to-pin failure propagation due to fission gas release in fuel subassemblies of liquid-metal-cooled fast breeder reactors, *Nuclear Engineering and Design* (1974). [https://doi.org/10.1016/0029-5493\(74\)90142-3](https://doi.org/10.1016/0029-5493(74)90142-3).

[11] K. Haga, K. Yamaguchi, F. Namekawa, Experimental study on the effect of fission product gas release into blockage wake region using a simulated LMFBR fuel subassembly, *Nuclear Engineering and Design* (1984). [https://doi.org/10.1016/0029-5493\(84\)90219-X](https://doi.org/10.1016/0029-5493(84)90219-X).

[12] J.S. Bell, R.S. Dickson, J. Sheedy, R.T. Peplinskie, M.D. Gauthier, J. Ruhnke, K. Stoll, Fission product release tests with uranium dioxide fuel and lead under conditions related to reactor accidents for lead-cooled reactors, *Annals of Nuclear Energy* 148 (2020) 107683. <https://doi.org/10.1016/j.anucene.2020.107683>.

[13] Y. Zhou, M. Zhong, X. Fan, J. Chen, M. Lv, A numerical simulation of water jet injection behavior in fuel-coolant interaction, *Journal of Nuclear Science and Technology* (2017). <https://www.tandfonline.com/doi/abs/10.1080/00223131.2016.1224740> (accessed April 24, 2025).

[14] Y. Liao, E. Krepper, D. Lucas, A baseline closure concept for simulating bubbly flow with phase change: A mechanistic model for interphase heat transfer coefficient, *Nuclear Engineering and Design* 348 (2019) 1–13. <https://doi.org/10.1016/j.nucengdes.2019.04.007>.

[15] Y. Liao, D. Lucas, 3D CFD simulation of flashing flows in a converging-diverging nozzle, *Nuclear Engineering and Design* (2015). <https://doi.org/10.1016/j.nucengdes.2015.06.015>.

[16] M. Lu, H. Xie, An investigation of pool entrainment based on the method of Volume of Fluid, *Nuclear Engineering and Design* (2017). <https://doi.org/10.1016/j.nucengdes.2017.04.006>.

[17] B. Dong, C. Li, L. Li, J. Yin, D. Wang, CFD study on mechanisms of fission gas burst release from defective fuel rods of a typical PWR, *Annals of Nuclear Energy* 140 (2020) 107089. <https://doi.org/10.1016/j.anucene.2019.107089>.

[18] B. Dong, L. Li, C. Li, J. Yin, D. Wang, CFD Simulation of Fission Gas Release Under Defective Conditions in Pressurized Water Reactor, *ASME* 2018. <https://doi.org/10.1115/ICONE26-81398>.

[19] I.A. Bolotnov, S.P. Antal, K.E. Jansen, M.Z. Podowski, Multidimensional analysis of fission gas transport in reactor fuel rods, *Nuclear Engineering and Design* (2012). <https://doi.org/10.1016/j.nucengdes.2012.02.004>.

[20] Z. Xiaodi, H. Meng, L. Long, F. Yixiang, W. Peng, H. Ning, The research on the fission gas behavior of nuclear fuel pellets based on ABAQUS, *Nuclear Engineering and Design* 36 (2016) 1200–1204.

- [21] R.L. Williamson, Enhancing the ABAQUS thermomechanics code to simulate multipellet steady and transient LWR fuel rod behavior, *Journal of Nuclear Materials* 415 (2011) 74–83. <https://doi.org/10.1016/j.jnucmat.2011.05.044>.
- [22] Y.I. Kaan, K. Sadeghi, S.H. Ghazaie, E. Sokolova, V. Modestov, V. Sergeev, P. Gao, Machine learning-assisted correlations for prediction of fission gas fractions and hydrogen concentration in VVER-1000 fuel, *Annals of Nuclear Energy* 212 (2025) 111073. <https://doi.org/10.1016/j.anucene.2024.111073>.
- [23] Y. Cai, K. Chen, W. Zhou, L. Gu, Multiphysics thermo-mechanical behavior modeling for annular uranium-plutonium mixed oxide fuels in a lead-cooled fast reactor, *Front. Energy Res.* 10 (2022). <https://doi.org/10.3389/fenrg.2022.993383>.
- [24] A. Karahan, J. Buongiorno, Modeling of thermo-mechanical and irradiation behavior of mixed oxide fuel for sodium fast reactors, *Journal of Nuclear Materials* 396 (2010) 272–282. <https://doi.org/10.1016/j.jnucmat.2009.11.021>.
- [25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is All you Need, in: *Advances in Neural Information Processing Systems*, 2017. <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a> Abstract.html (accessed February 26, 2025).
- [26] A. Gu, T. Dao, Mamba: Linear-Time Sequence Modeling with Selective State Spaces, (2024). <https://doi.org/10.48550/arXiv.2312.00752>.
- [27] L. Breiman, Random Forests, *Machine Learning* (2001). <https://doi.org/10.1023/A:1010933404324>.
- [28] Xu B., Zhou Y., Bian X., Self-supervised learning based on Transformer for flow reconstruction and prediction, *Physics of Fluids* 36 (2024). <https://doi.org/10.1063/5.0188998>.
- [29] K. Wang, D. Wang, X. Liu, S. Cheng, S. Wang, W. Zhou, S. Miwa, K. Okamoto, Re-examining the input-parameters and AI strategies for Critical Heat Flux prediction, *Energy* 318 (2025) 134606. <https://doi.org/10.1016/j.energy.2025.134606>.
- [30] B.J. Lewis, F.C. Iglesias, D.S. Cox, E. Gheorghiu, A Model for Fission Gas Release and Fuel Oxidation Behavior for Defected UO₂ Fuel Elements, *Nuclear Technology* (1990). <https://doi.org/10.13182/NT90-A16236>.
- [31] U.M. El-Saied, D.R. Olander, Fission gas release during grain growth in a microstructure with a distribution of grain sizes, *Journal of Nuclear Materials* 207 (1993) 313–326. [https://doi.org/10.1016/0022-3115\(93\)90274-3](https://doi.org/10.1016/0022-3115(93)90274-3).
- [32] R.M. Carroll, J.G. Morgan, R.B. Perez, O. Sisman, Fission Density, Burnup, and Temperature Effects on Fission-Gas Release From UO₂, *Nuclear Science and Engineering* (1969). <https://doi.org/10.13182/NSE69-A19519>.
- [33] L.C. Bernard, J.L. Jacoud, P. Vesco, An efficient model for the analysis of fission gas release, *Journal of Nuclear Materials* (2002). [https://doi.org/10.1016/S0022-3115\(02\)00793-6](https://doi.org/10.1016/S0022-3115(02)00793-6).

- [34] M. Mogensen, C.T. Walker, I.L.F. Ray, M. Coquerelle, Local fission gas release and swelling in water reactor fuel during slow power transients, *Journal of Nuclear Materials* 131 (1985) 162–171. [https://doi.org/10.1016/0022-3115\(85\)90455-6](https://doi.org/10.1016/0022-3115(85)90455-6).
- [35] J.A. Turnbull, The effect of grain size on the swelling and gas release properties of uo₂ during irradiation, *Journal of Nuclear Materials* (1974). [https://doi.org/10.1016/0022-3115\(74\)90061-0](https://doi.org/10.1016/0022-3115(74)90061-0).
- [36] C. Brown, K.W. Hesketh, I.D. Palmer, BNFL assessment of methods of attaining high burnup MOX fuel, (1998). <https://www.osti.gov/etdeweb/biblio/654939> (accessed April 23, 2025).

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.