ARTIFICIAL INTELLIGENCE APPLICATIONS IN NUCLEAR ENERGY

EDITED BY: Xianping Zhong, Jun Wang, Xingang Zhao, Yang Liu and Shripad T. Revankar

PUBLISHED IN: Frontiers in Energy Research

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-708-3
DOI 10.3389/978-2-88976-708-3

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ARTIFICIAL INTELLIGENCE APPLICATIONS IN NUCLEAR ENERGY

Topic Editors:
Xianping Zhong, University of Pittsburgh, United States
Jun Wang, University of Wisconsin-Madison, United States
Xingang Zhao, Oak Ridge National Laboratory (DOE), United States
Yang Liu, Argonne National Laboratory (DOE), United States
Shripad T. Revankar, Purdue University, United States

Citation: Zhong, X., Wang, J., Zhao, X., Liu, Y., Revankar, S. T., eds. (2022). Artificial Intelligence Applications in Nuclear Energy. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-708-3

Table of Contents

frontiers | Frontiers in Energy Research

EDITORIAL
published: 08 July 2022
doi: 10.3389/fenrg.2022.965581

Editorial: Artificial Intelligence Applications in Nuclear Energy

Xianping Zhong[1], Jun Wang[2]*, Xingang Zhao[3], Yang Liu[4] and Shripad T Revankar[5]*

[1]Department of Mechanical Engineering and Materials Science, University of Pittsburgh, Pittsburgh, PA, United States, [2]Department of Engineering Physics, University of Wisconsin, Madison, WI, United States, [3]Nuclear Energy and Fuel Cycle Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [4]Nuclear Science and Engineering Division, Argonne National Laboratory, Lemont, IL, United States, [5]Department of Nuclear Engineering, Purdue University, West Lafayette, IN, United States

Keywords: artificial intelligence—AI, machine learning—ML, software development, equipment prognostics and health management, reactor design optimization, nuclear safety analysis and accident management

Editorial on the Research Topic

Artificial Intelligence Applications in Nuclear Energy

As a highly complex man-machine-network integration system, the nuclear power plant's development, construction and operation are still facing many obstacles and risks. Firstly, plant instruments and equipment may fail during operation, which will affect the performance and safety of nuclear power plants. Secondly, although nuclear power plants have been digitalized after decades of development, most of them still adopt traditional and inefficient operation and control methods. Finally, due to the above reasons and stringent control requirements, human operators are under great pressure. In the past decades, artificial intelligence (AI) and machine learning (ML), especially methods related to deep learning, have made great progress and have been widely used in computer vision, automatic control and other fields (Bakator and Radosav, 2018; Singla et al., 2020; Taskiran et al., 2020; Usuga Cadavid et al., 2020). At present, many researchers have begun to apply AI to the field of nuclear energy to overcome the above obstacles and risks. Potential application scenarios include nuclear power software development (Bao et al., 2019; Liu et al., 2019), equipment prognostics and health management (Zhao et al., 2021; Zhong and Ban, 2022), reactor design optimization (Kumar and Tsvetkov, 2015; Turkmen et al., 2021), reactor autonomous control and operation (Wilson, 2019; Lee et al., 2020; Lin et al., 2021), and nuclear safety analysis and accident management (Zeng et al., 2018; Chung, 2021). This topic explores the application of the latest AI technologies in nuclear energy to promote research, sharing and development.

We have collected two papers on AI for nuclear power software development: Dong et al. and Wu et al. Dong' work proposed a neural network-based data-driven model to predict the bubble departure diameter in subcooled boiling flow. The model is based on mechanistic bubble departure models and takes dimensionless numbers as input, thus demonstrating good generalization capability on a broad range of flow conditions.

We have collected three papers on intelligent prognostics and health management of plant equipment: Fan et al., He et al., and Yao et al. Fan's work focused on the fatigue detection of glass-to-metal seals in nuclear power plants, with the assistance of the spectrum characterization of fiber Bragg grating (FBG) sensors. The spectral response to non-uniform strain distributions in glass-to-metal are reconstructed precisely based on the transfer matrix model, and the asymmetric deformation induced by fatigue conditions is detected efficiently by the variations of Bragg wavelength shift and full width at half maximum.

We have collected four papers on AI for reactor design optimization: Pevey et al., Zhang et al., Yu et al., and Li et al. Hines' work proposed a convolutional neural network–based surrogate model optimization of fast neutron source configurations. Their new algorithm produced more viable

Edited and reviewed by:
Uwe Schröder,
University of Greifswald, Germany

****Correspondence:***
Xianping Zhong
xianping.zhong@pitt.edu
Jun Wang
jwang564@wisc.edu

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *09 June 2022*
Accepted: *20 June 2022*
Published: *08 July 2022*

Citation:
Zhong X, Wang J, Zhao X, Liu Y and Revankar ST (2022) Editorial: Artificial Intelligence Applications in Nuclear Energy. Front. Energy Res. 10:965581. doi: 10.3389/fenrg.2022.965581

designs that significantly improved the objective function utilizing the same computational resources compared to the standard multi-objective genetic algorithm NSGA-II.

We have collected three papers on AI for nuclear safety analysis and accident management: Gong et al., Dong et al., and Sallehhudin and Diab. With the assistance of a deep learning model called zLSTM, Gong's work focused on the multivariate time series prediction for LOCA development. The zLSTM is constructed by introducing an improved gate function Zigmoid within the original LSTM model, allowing the non-linearity, both short and long-term memory, and multiple system parameters to be fully covered for a more accurate LOCA prediction.

As the development of AI technologies has accelerated in recent years, the nuclear industry has begun to look for the potential of AI for code development, real-time intelligent operation and maintenance, reactor design optimization, and safety analysis and accident management. The industry will follow suit if AI shows strong capabilities in research. In AI research, data as a carrier of knowledge plays a dominant role in the performance of AI and ML models. However, data containing valid information is scarce in the nuclear industry. In the coming period, the focus should be on how to make AI effective in practice under small sample, sample imbalance, and strong noise conditions. This may be a long-term challenge, but in the end all the effort will be worthwhile. In the future, with the popular application of AI technologies, the whole chain of the nuclear industry will become more intelligent.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

Bakator, M., and Radosav, D. (2018). Deep Learning and Medical Diagnosis: A Review of Literature. *Multimodal Technol. Interact.* 2, 47. doi:10.3390/MTI2030047

Bao, H., Dinh, N. T., Lane, J. W., and Youngblood, R. W. (2019). A Data-Driven Framework for Error Estimation and Mesh-Model Optimization in System-Level Thermal-Hydraulic Simulation. *Nucl. Eng. Des.* 349, 27–45. doi:10.1016/j.nucengdes.2019.04.023

Chung, K. Y. (2021). A Machine Learning Strategy with Restricted Sliding Windows for Real-Time Assessment of Accident Conditions in Nuclear Power Plants. *Nucl. Eng. Des.* 378, 111140. doi:10.1016/j.nucengdes.2021.111140

Kumar, A., and Tsvetkov, P. V. (2015). A New Approach to Nuclear Reactor Design Optimization Using Genetic Algorithms and Regression Analysis. *Ann. Nucl. Energy* 85, 27–35. doi:10.1016/j.anucene.2015.04.028

Lee, D., Arigi, A. M., and Kim, J. (2020). Algorithm for Autonomous Power-Increase Operation Using Deep Reinforcement Learning and a Rule-Based System. *IEEE Access* 8, 196727–196746. doi:10.1109/ACCESS.2020.3034218

Lin, L., Athe, P., Rouxelin, P., Avramova, M., Gupta, A., Youngblood, R., et al. (2021). Development and Assessment of a Nearly Autonomous Management and Control System for Advanced Reactors. *Ann. Nucl. Energy* 150, 107861. doi:10.1016/j.anucene.2020.107861

Liu, Y., Sun, X., and Dinh, N. T. (2019). Validation and Uncertainty Quantification of Multiphase-CFD Solvers: A Data-Driven Bayesian Framework Supported by High-Resolution Experiments. *Nucl. Eng. Des.* 354, 110200. doi:10.1016/j.nucengdes.2019.110200

Singla, N., Kaur, M., and Sofat, S. (2020). Automated Latent Fingerprint Identification System: A Review. *Forensic Sci. Int.* 309, 110187. doi:10.1016/j.forsciint.2020.110187

Taskiran, M., Kahraman, N., and Erdem, C. E. (2020). Face Recognition: Past, Present and Future (A Review). *Digit. Signal Process.* 106, 102809. doi:10.1016/J.DSP.2020.102809

Turkmen, M., Chee, G. J. Y., and Huff, K. D. (2021). Machine Learning Application to Single Channel Design of Molten Salt Reactor. *Ann. Nucl. Energy* 161, 108409. doi:10.1016/j.anucene.2021.108409

Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R., and Fortin, A. (2020). Machine Learning Applied in Production Planning and Control: a State-Of-The-Art in the Era of Industry 4.0. *J. Intell. Manuf.* 31, 1531–1558. doi:10.1007/s10845-019-01531-7

Wilson, J. (2019). *Machine Learning for Nuclear Fission Systems: Preliminary Investigation of an Autonomous Control System for the MGEP*. Cambridge, MA: Massachusetts Institute of Technology.

Zeng, Y., Liu, J., Sun, K., and Hu, L.-W. (2018). Machine Learning Based System Performance Prediction Model for Reactor Control. *Ann. Nucl. Energy* 113, 270–278. doi:10.1016/j.anucene.2017.11.014

Zhao, X., Kim, J., Warns, K., Wang, X., Ramuhalli, P., Cetiner, S., et al. (2021). Prognostics and Health Management in Nuclear Power Plants: An Updated Method-Centric Review with Special Focus on Data-Driven Methods. *Front. Energy Res.* 9, 696785. doi:10.3389/fenrg.2021.696785

Zhong, X., and Ban, H. (2022). Pre-Trained Network-Based Transfer Learning: A Small-Sample Machine Learning Approach to Nuclear Power Plant Classification Problem. *Ann. Nucl. Energy* 175, 109201. doi:10.1016/j.anucene.2022.109201

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Energy Research

ORIGINAL RESEARCH
published: 08 October 2021
doi: 10.3389/fenrg.2021.755638

Using Machine Learning to Predict the Fuel Peak Cladding Temperature for a Large Break Loss of Coolant Accident

Wazif Sallehhudin[1] *and Aya Diab*[1,2]*

[1]Nuclear Power Plant Engineering Department, KEPCO International Nuclear Graduate School (KINGS), Ulsan, South Korea, [2]Mechanical Power Engineering Department, Faculty of Engineering, Ain Shams University, Cairo, Egypt

In this paper the use of machine learning (ML) is explored as an efficient tool for uncertainty quantification. A machine learning algorithm is developed to predict the peak cladding temperature (PCT) under the conditions of a large break loss of coolant accident given the various underlying uncertainties. The best estimate approach is used to simulate the thermal-hydraulic system of APR1400 large break loss of coolant accident (LBLOCA) scenario using the multidimensional reactor safety analysis code (MARS-KS) lumped parameter system code developed by Korea Atomic Energy Research Institute (KAERI). To generate the database necessary to train the ML model, a set of uncertainty parameters derived from the phenomena identification and ranking table (PIRT) is propagated through the thermal hydraulic model using the Dakota-MARS uncertainty quantification framework. The developed ML model uses the database created by the uncertainty quantification framework along with Keras library and Talos optimization to construct the artificial neural network (ANN). After learning and validation, the ML model can predict the peak cladding temperature (PCT) reasonably well with a mean squared error (MSE) of ~0.002 and R^2 of ~0.9 with 9 to 11 key uncertain parameters. As a bounding accident scenario analysis of the LBLOCA case paves the way to using machine learning as a decision making tool for design extension conditions as well as severe accidents.

Keywords: nuclear safety, large break LOCA, artificial neural network, machine learning, uncertainty quantification, peak cladding temperature

OPEN ACCESS

Edited by:
Xingang Zhao,
Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Yue Jin,
Massachusetts Institute of Technology, United States
Anil Gurgen,
National Institute of Standards and Technology (NIST), United States
Zhegang Ma,
Idaho National Laboratory (DOE), United States

***Correspondence:**
Aya Diab
aya.diab@kings.ac.kr
aya.diab@eng.asu.edu.eg

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *09 August 2021*
Accepted: *21 September 2021*
Published: *08 October 2021*

Citation:
Sallehhudin W and Diab A (2021) Using Machine Learning to Predict the Fuel Peak Cladding Temperature for a Large Break Loss of Coolant Accident. Front. Energy Res. 9:755638. doi: 10.3389/fenrg.2021.755638

INTRODUCTION

Deterministic safety analysis has traditionally been utilized to demonstrate the robustness of nuclear power plants, usually adopting a conservative approach. However, the conservative approach relies on a number of assumptions that do not necessarily reflect the real plant performance (Queral et al., 2015). On the other hand, the best estimate (BE) approach provides a more realistic system response based on detailed thermal-hydraulic mechanistic models provided it is accompanied with uncertainty quantification (UQ). The integration of BE and UQ is known as best estimate plus uncertainty (BEPU) and is built upon a statistical foundation to provide a more realistic estimation of the safety margin and hence ensure that the safety limit is met.

Utilities were given an ample opportunity to apply the best estimate plus uncertainty (BEPU) methodology following the United States Regulatory Commission (USNRC) amendment of 10CFR50.46 Appendix-K in 1988. Accordingly, the USNRC assisted in the steady transition from the conservative to BEPU methodology by introducing the USNRC Regulatory Guide

1.157, "Best Estimate Calculations of Emergency Core Cooling System Performance" and the demonstration of code scaling, applicability and uncertainty (CSAU) methodology in 1989 aiming to quantify the uncertainty parameters (USNRC, 1989). In addition to the CSAU methodology, various international collaboration projects had been undertaken to propose and validate other uncertainty quantification methodologies such as the uncertainty method study, UMS, (OECD, 1998), the BEMUSE project (OECD, 2007a) and the SM2A study within the SMAP framework (OECD, 2007b).

The BEPU methodology has been used to predict key safety parameters such as the peak cladding temperature (PCT), departure from nucleate boiling ratio (DNBR), etc. for critical accident scenarios. In BEPU analysis, a BE code is used to simulate the plant response given the variations in a multitude of uncertain parameters (UPs) that can be propagated within the thermal-hydraulic system code. The process of uncertainty propagation is however lengthy and hence BEPU analysis has so far been limited to the analysis of bounding design basis accident (DBA) scenarios, e.g. large break loss of coolant accident (LBLOCA) (Chang et al., 2020) and only recently to the analysis of a station blackout (SBO) (Musoiu et al., 2019) and to the main steam line break (MSLB) (Petruzzi et al., 2016).

This concern can be addressed by using data-driven approaches that provide a prediction based only on the database previously obtained from experimental, or simulation results. A data-driven model tries to learn the salient characteristics embedded within the system by developing a mathematical relationship between the system parameters rather than solving the physics-based models to describe the system performance. This process is known as machine learning (ML).

ML is one of the branches of artificial intelligence (AI). Currently, there are many machine learning tools that can be used for prediction or classification such as artificial neural network (ANN), support vector machine (SVM), Naïve-Bayes algorithm, random forest, decision tree, logistic regression (LR), K- nearest neighbors (KNN), etc. Any of these tools may be used to develop a machine learning algorithm. Each algorithm is based on its unique strategy in making predictions. Generally speaking, ML algorithms learn from the datasets and try to decipher the salient characteristics within the data that reflect the relationship between the inputs and outputs. Based on the datasets, a mathematical relationship can be generated between the input vector variables and the scalar output variables. The learning process helps improve the relationship by constantly changing the learning parameters to tune the model until the objective function is optimized. The objective functions for each machine learning algorithm is different and needed to be specified accordingly.

Recently, the International Atomic Energy Agency (IAEA) has urged the nuclear community to integrate ML in the industry within the framework of emerging technologies, given its superior capability in handling big-data (IAEA, 2020). In fact, the potential of using ML technology has been explored to estimate some key figures of merit such as the power pin peaking factor (Bae et al., 2008), the wall temperature at critical heat flux (Park et al., 2020), the flow pattern identification (Lin, 2020), to detect anomalies and warn of equipment failure (Ahsan and Hassan, 2013; Chen and Jahanshahi, 2018; Devereux et al., 2019); to determine core configuration and core loading pattern optimization (Siegelmann et al., 1997; Faria and Pereira, 2003; Erdogan and Gekinli, 2003; Zamer et al., 2014; Nissan, 2019), to identify initiating events and categorize accidents (Santosh et al., 2003; Na et al., 2004; Lee and Lee, 2006; Ma and Jiang, 2011; Pinheiro et al., 2020; Farber and Cole, 2020) and to determine of key performance metrics and safety parameters (Ridlluan et al., 2009; Montes et al., 2009; Farshad Faghihi and Seyed, 2011; Patra et al., 2012; Young, 2019; Park et al., 2020; Alketbi and Diab, 2021), and in radiation protection for isotope identification and classification (Keller and Kouzes, 1994; Abdel-Aal and Al-Haddad, 1997; Chen, 2009; Kamuda and Sullivan, 2019), etc. However, it is worth noting that the application of ML in nuclear safety is still limited despite its potential to enhance performance, safety, as well as economics of plant operation (Chai et al., 2003) which warrants further research (Gomez Fernandez et al., 2017). For a more comprehensive review of the status and development efforts utilizing data-driven approaches in nuclear industry, the reader may consult (Gomez Fernandez et al., 2017; Gomez Fernandez et al., 2020).

In this study, an artificial neural network (ANN) is developed to predict the PCT under LBLOCA conditions as a bounding accident scenario. The goal is to develop a fast and cost-effective tool for uncertainty quantification of PCT under LBLOCA conditions using ML. This is achieved by using a database to train the ML algorithm, and once trained and tested, the meta-model can be used as a predictive tool. The database required to train and test the model is generated via the thermal hydraulic system code MARS-KS (KAERI, 2004) within an uncertainty quantification framework using Dakota (Adams et al., 2020). Once proven, the ML technology may be used to help the nuclear designers and/or operators to expedite the decision making process particularly in those situations that involve complex interconnected phenomena during design optimization or in the event of a nuclear accident.

ARTIFICIAL NEURAL NETWORK

ANN is a machine learning model inspired by the biological network of the nerve cells that make up the brain. Fundamentally, the ANN behaves in a way similar to the nerve cells. However, each biological structure is replaced with layers of neurons with a pre-defined architecture that communicates data between the input signals and output signals via weights, biases, and activation functions to find the best weight matrix that best describes the relationship between the inputs and outputs. The ANN structure can be split into three different classes; artificial neural network (ANN), convolutional neural network (CNN) and recurrent neural network (RNN). This research focuses only on ANN.

The ANN generally refers to the modelling of the data through a stack of computational layers. The ANN utilizes the back propagation based on the stochastic gradient descent (SGD) technique that approximates the loss function optimal points which guarantees convergence and terminate at the optimal

TABLE 1 | Common hyper-parameters for ANN.

Hyper parameters	Typical values
Number of input neurons	One per input feature
Number of hidden layers	Problem-dependent, usually 1–5
Neurons per hidden layers	Problem-dependent, usually 10–100
Number of output neurons	One per prediction value
Activation functions	ReLU, SELU, Softplus, Logistic, tanh
Loss functions	MSE, MAE, Huber loss

solution (Dawani, 2020). Currently, the improvised gradient descent techniques such as the adaptive moment estimate (Adam) is being widely used for many ANN applications. The SGD teaches the ANN how to tweak the connection weights and biases in order to converge to the closest mathematical representation of the data at hand. A number of activation functions, such as the hyperbolic tangent function and the rectifier linear unit function (ReLU), may be used to provide signal transformations for each input layer and hence provide better representation of the underlying non-linearity of complex systems.

The ANN is based on a multi-layer perceptron model and can be used for both regression and classification problems. The training process of an ANN is a two-step process. The first is a forward propagation step and involves evaluating the error or loss function. In the second step, the resulting error is propagated backwards through the network to adjust the weights and biases. This process is repeated until no further improvement in the error between the predicted outputs and desired values is achieved. The complexity of the network is determined by the number of hidden layers, the number of nodes in each layer, the type of activation function. The output is predicted by summing the functions within the hidden layers to produce a net input function.

The goal of the training process is to tune the model hyper-parameters for better prediction. Those hyper-parameters include: the number of neurons, number of hidden layers, network structure, activation function, type of optimizer, loss function, etc. Common hyper-parameters associated with the regression problems are described in **Table 1**.

During the optimization process, provisions should be made to ensure a global minimum is achieved rather than a local minimum or saddle point, for which small neural networks are prone. A balance between generalization and fitness to the training data should be achieved to ensure that over-fitting or retaining redundant features in the neural network is avoided.

METHODOLOGY

To achieve the goal of this paper, three main objectives can be identified which are, 1) the thermal hydraulic model development, 2) the uncertainty quantification and database creation and 3) machine learning model development. Each will be delineated in the next subsections.

Thermal-Hydraulic Model Development

This section focuses on the details of the thermal hydraulic model development. In this investigation, the best estimate system code, MARS-KS version 1.4, is used to simulate the nuclear power plant response under LBLOCA conditions. MARS-KS is a multi-dimensional two-phase thermal hydraulic system code developed by KAERI (2009). The model representation in MARS-KS, including nodalization, boundary and initial conditions as well as the main assumptions will be presented next.

APR1400 Nodalization

First, the details of the APR1400 reactor are described using the system nodalization shown in **Figure 1** to represent the key systems and components which includes a reactor pressure vessel, a pressurizer, two loops with four cold legs and two hot legs. The nodalization also includes detailed description of the two steam generators (SG), the main steam lines and associated valves (MSSV, MSIV, ADV, etc). The turbine however, is modeled as a boundary condition. The safety injection system (SIS) is modeled to represent the emergency core cooling system (ECCS) of the APR1400. The SIS modelling is necessary to understand the reactor dynamic behavior in dealing with the LBLOCA scenario and to ensure the safety parameters stay within the safety limits during the different phases of the accident.

The containment building is modelled as a boundary condition and the containment spray system (CSS) are excluded from the nodalization process as it is not required in the analysis of the reactor response under the LBLOCA scenario. The core is divided into a hot channel and an average channel. The hot channel represents the hottest fuel assembly; while the average channel represents the remaining 241 fuel assemblies in the APR1400 nuclear power plant core configuration.

The core channels are modeled in MARS using a pipe hydrodynamic component with a vertical orientation using 20 axial nodes. A solid structural element is attached to the hydrodynamic component to represent the fuel assembly with 20 axial nodes and nine radial nodes. For the downcomer, an annulus hydrodynamic component is used. The annulus is divided circumferentially into six channels at different angle (0°, 60°, 120°, 180°, 240°, 300°, 360°), each having five axial nodes with a vertical orientation.

The same nodalization scheme is used for the upper annulus region, which hosts the entry location of the emergency coolant from the SIS that should be directed to the core. The general pathway of emergency coolant should start from the upper annulus to the downcomer, to the lower plenum, to the bottom of the core and finally to the core itself.

The inlet and the outlet of the core are nodalized using branch components. The core inlet hosts the lower support structure (LSS) and in-core instrumentation while, the core outlet hosts the

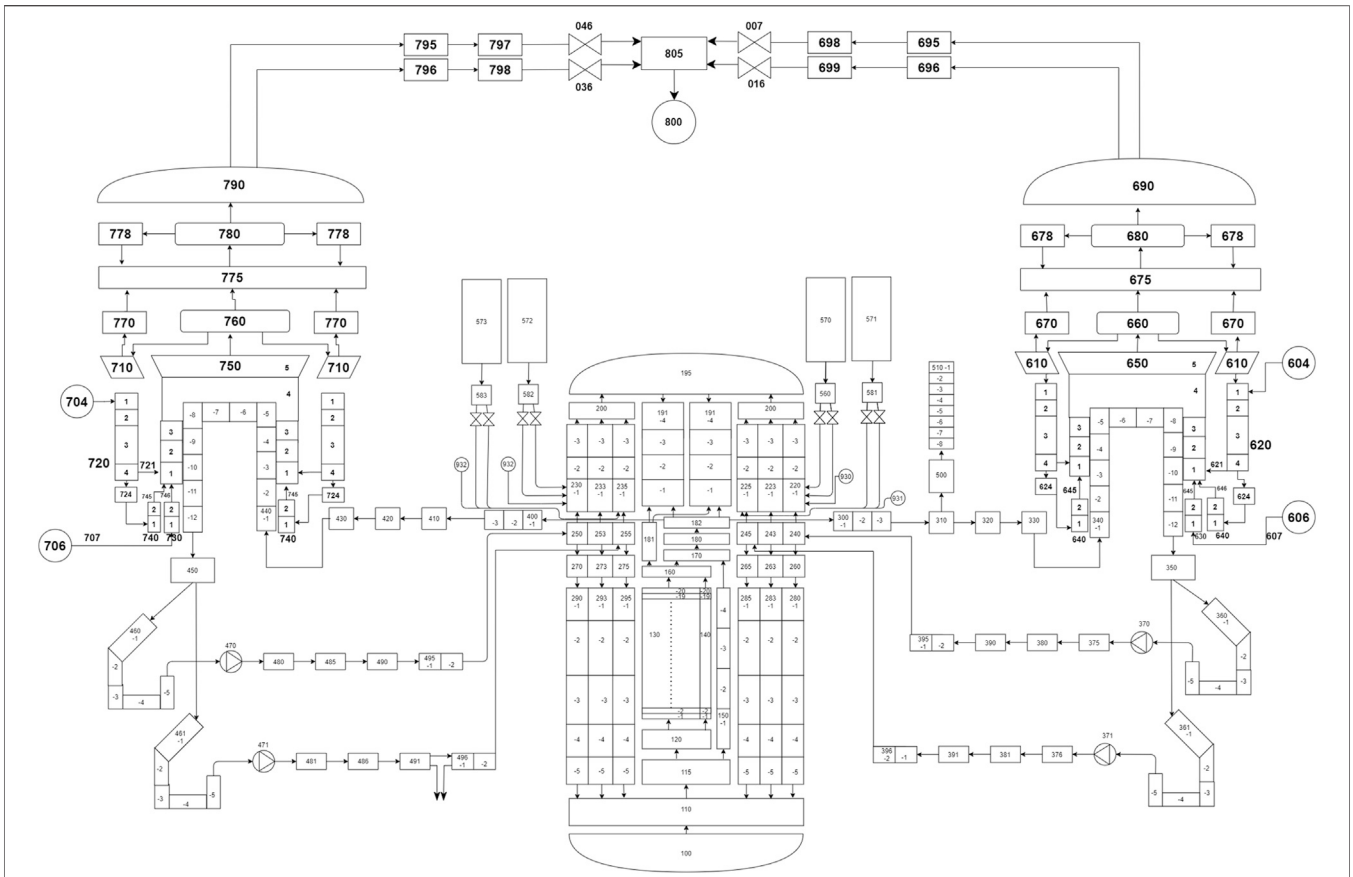


FIGURE 1 | APR1400 nodalization for the LBLOCA thermal-hydraulic model.

fuel alignment plate (FAP), the upper guide structure, the core shroud and the fuel barrel assembly. The crossflow is permitted between the downcomer regions, the upper annulus, the downcomer upstream, between the average channel and the hot channel and between the loop elevation reference regions. The cross flow allows the inventory to move sideways between the hydrodynamic components to represent the secondary cross flow. At the middle part of the reactor pressure vessel (RPV), a bypass is used to connect the bottom region of the core towards the fuel alignment plate without passing through the core site. The bypass allows liquid to move vertically from the lower plenum to the fuel alignment plate (FAP). To simulate the safety injection system (SIS), both the safety injection tanks (SITs) and the safety injection pumps (SIPs) are modelled in the nodalization. There are four units of SITs that are modelled using the accumulator hydrodynamic component. The SIT are connected to the upper annulus and is controlled using a combination of logical and variable trips. A set point that refers to the low pressurizer set point pressure is used to navigate the turning on of the SITs. The SIT valve is divided into two part for each train to represent the low-flow and high-flow conditions of the actual APR1400 fluidic device. The fluidic device act as a flow regulator in the actual APR1400 plant to optimize the usage of the emergency inventory during the loss of coolant scenario. Meanwhile, the SIPs are comprised of 4 units however, only two units of SIP are available in accordance to the conservative assumption adopted by the APR1400 Design Control Document (DCD) for LBLOCA evaluation. The SIPs will be available only for the break side and the opposite direction across from the break, i.e. located at 60° and 240°, respectively. The SIPs are modeled using a time dependent volume and a time dependent junction instead of pump hydrodynamic components. This configuration allows the user to impose flow boundary conditions and control the velocity of the coolant. Each SIP and SIT is connected to the upper annulus model at different circumferential angles based on APR1400 description.

Consistent with APR1400 DCD the decay heat model is based on the ANS-1973 model and a reactivity table is also provided in the code to account for the negative reactivity insertion due to control rod insertion. However, based on the APR1400 DCD, the negative reactivity contribution from the control rod is discredited for conservatism when conducting the LBLOCA analysis. This will allow the SIS capability in managing the accident and maintaining the core integrity to be fully tested during the LBLOCA accident. Regarding the reactor internals, the heat structure components are used and attached to the related

hydrodynamic volumes to reflect the heat transfer boundary conditions and architecture of the APR1400 design. Meanwhile to describe LBLOCA scenario, a double ended guillotine break is placed on the cold leg after the pump discharge. This is achieved by incorporating two trip valves to divert the coolant from the vessel to the time dependent volume attached to each of the trip valves when the accident is initiated.

LBLOCA Model Assumptions

Following the DCD recommendation, a double ended guillotine break (DEGB) equivalent to double the area of the pipe with the largest cross section of the RCS, i.e. the cold leg piping is used in this work. The standard internal diameter of the connecting pipes between the pump discharge and the reactor pressure vessel inlet nozzle is 762 mm (30 inches) which corresponds to the break area of 0.456 m^2. The thermal hydraulic model development is based on several assumptions similar to those reported in the DCD document: 1) LBLOCA occurs at loop B1 near the pump discharge site. 2) Break type is double ended guillotine break (DEGB). 3) Loss of offsite power (LOOP) for the RCPs. 4) No negative reactivity insertion from the control rods. 5) Single emergency diesel generator (EDG) is not functioning causing two out of four safety injection pumps (SIPs) to be non-operable. 6) All safety injection tanks (SIT) are in operation.

Uncertainty Quantification Framework Development

The statistical tool, Dakota (Adams et al., 2020), is used in this work to propagate the uncertainty parameters into the thermal hydraulic model. Dakota is an open source statistical software tool developed by Sandia National Laboratory. It can be used for optimization, sensitivity analysis and uncertainty quantification. The uncertainty propagation process is achieved by developing the uncertainty quantification framework by loosely coupling the best estimate system code, MARS-KS, and the statistical tool, Dakota, via a python script to manage the data exchange process. Several important files such as, the Dakota input file, the python interface script, the MARS steady state file and the MARS transient file are necessary for the uncertainty quantification framework to run smoothly and propagate the uncertainty parameters.

Uncertain Parameters Identification

As indicated in the introduction, the current work explores the possibility of using ML to predict the PCT under the conditions of a LBLOCA, being an important bounding accident scenario. LBLOCA was used in nuclear safety as a design basis for the emergency core cooling system, ECCS, to provide assurance that the ECCS would not violate any of the safety limits and hence preserve the fuel integrity during a loss of coolant accident (LOCA). For LBLOCA, the key performance measure of the ECCS, as defined by the 10CFR50.46 Appendix-K guideline is that the PCT does not exceed the safety limit of 1477 K (2,200°F) to ensure the integrity of the fuel under the accident conditions (Martin and O'Dell).

Now for the ML algorithm to be developed and trained, it is necessary to use a database of the most important system parameters (features) that impact the safety parameter of interest, in this case the PCT. Generally speaking, the model can use a database originating from the plant historic data, which is not possible under DBA conditions. Alternatively, it can be generated using simulation results produced by system codes. In this work, the database is created, using the latter approach.

A BEPU analysis is undertaken to generate a database of the system response under LBLOCA. In general, uncertainty quantification can be achieved using either the input uncertainty propagation approach or the output uncertainty propagation approach (Martin and O'Dell, 2008). The former approach will be followed in this work.

To conduct the BEPU analysis, the uncertainty quantification process requires the identification of the uncertain parameters that can impact the PCT. Those uncertainty parameters can be derived from the phenomena identification and ranking table (PIRT). The PIRT describes the key phenomena and processes relevant to the plant's thermal-hydraulic response for a specific accident condition. Most of the PIRT developed throughout the years centered on the LBLOCA cases due to its importance to nuclear safety as a bounding DBA scenario. Several PIRTs have been developed for LBLOCA scenario such as: Westinghouse PIRT (USNRC, 1988), AP600 PIRT (LA-UR-95-2718, 1995), KREM PIRT (KHNP, 2014), KNGR PIRT (KINS, 2001), APR1400 PIRT (KEPCO, 2014). For the current project, the investigation will focus on the APR1400 PIRT which is based on the KNGR PIRT, which in turn is derived from the Westinghouse PIRT.

Based on the work of (Lee et al., 2014; Kang, 2016), eight key phenomena were considered in this study. The key phenomena underlying the LBLOCA scenario are gap conductance, energy stored in the fuel, decay heat, rewetting process, reflooding heat transfer, critical flow, pump performance and core reflooding as shown in **Table 2**. A total of 19 uncertainty parameters have been derived from these key phenomena. These key uncertain parameters and the statistical information associated with each (range and distribution) are listed in **Table 2**.

Data Pre-Processing

Before propagating the uncertainty parameters into the best estimate thermal-hydraulic system model, it is essential for the uncertainty parameters to undergo a normalization process. The normalization is done with respect to the statistical information available for each uncertain parameter derived from the PIRT. First, the mean value is calculated using the following expression:

$$\mu = \sum_{i=1}^{n} \frac{x_i}{n} \tag{1}$$

Next, the upper and lower limit can be scaled using this mean value as follows:

$$L_{high} = \frac{x_{high}}{\mu} \tag{2}$$

TABLE 2 | Key uncertain phenomena and associated uncertain parameters.

Phenomenon	Uncertainty parameter	Distribution	Range
Heat conductance	Gap conductance	Normal	0.4–1.5
Energy storage	Fuel thermal conductivity	Uniform	0.847–1.153
	Core power	Normal	0.98–1.02
Decay heat	Decay heat	Normal	0.934–1.066
Rewetting process	Groeneveld critical heat flux	Normal	0.17–1.8
	Transition boiling HTC	Normal	0.54–1.46
Reflooding heat transfer	Chen nucleate boiling HTC	Normal	0.53–1.46
	Dittus-Boelter liquid HTC	Normal	0.606–1.39
	Dittus-Boelter vapor HTC	Normal	0.606–1.39
	Film boiling HTC	Normal	0.428–1.58
Critical flow	Break discharge coefficient	Normal	0.729–1.165
Pump performance	two phase head multiplier	Uniform	0.0–1.0
	two phase torque multiplier	Uniform	0.0–1.0
Reflooding	SIT actuation pressure (MPa)	Normal	4.03–4.46
	SIT water inventory (m^3)	Uniform	45.31–54.57
	SIT water temperature (K)	Normal	294.11–321.89
	SIT loss coefficient	Uniform	10.8–25.2
	SIP water temperature (K)		283.0–321.89
System pressure	Pressurizer pressure (MPa)	Normal	11.94–19.08

TABLE 3 | Normalized uncertain parameters.

UP	Parameter description	Mean, μ	Standard deviation, σ	Range, L_{high}-L_{low}
1	Core power	1.0	0.01	0.98–1.02
2	Groeneveld-CHF	1.0	0.414	0.173–1.827
3	Chen nucleate boiling HTC	1.0	0.234	0.553–1.467
4	Transition boiling HTC	1.0	0.230	0.54–1.46
5	Dittus-Boelter liquid HTC	1.0	0.196	0.607–1.393
6	Dittus-Boelter vapor HTC	1.0	0.196	0.607–1.393
7	Film boiling HTC	1.0	0.287	0.426–1.574
8	Break discharge coefficient	1.0	0.115	0.77–1.23
9	Decay heat	1.0	0.033	0.934–1.066
10	Gap conductance	1.0	0.289	0.421–1.579
11	SIT actuation pressure (MPa)	1.0	0.025	0.949–1.051
12	SIT water inventory (m^3)	1.0	0.046	0.907–1.093
13	SIT loss coefficient	1.0	0.20	0.6–1.4
14	Pressurizer pressure (MPa)	1.0	0.113	0.77–1.23
15	Fuel thermal conductivity	-	-	0.847–1.153
16	Pump two phase head multiplier	-	-	0.0–1.0
17	Pump two phase head multiplier	-	-	0.0–1.0
18	SIT water temperature (K)	-	-	0.955–1.045
19	SIP (IRWST) water temperature (K)	-	-	0.936–1.064

$$L_{low} = \frac{x_{low}}{\mu} \tag{3}$$

And finally for the standard deviation can be calculated as follows:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \mu\right)^2} \tag{4}$$

where x_i represents value of uncertain parameter in the sample, and x_{high} and x_{low} are the upper and lower values of the uncertain parameter, respectively. For a normal distribution function, the mean, standard deviation and upper and lower limits are required; while, for a uniform distribution, only the upper and lower limits are required. **Table 3** shows the uncertain parameters after scaling.

Uncertainty Propagation and Database Generation

With the key uncertain parameters identified and scaled appropriately, they are randomly propagated into the thermal hydraulic system code, MARS-KS using DAKOTA. The goal is to generate a large enough sample that can be representative of the realistic system performance in accordance with the USNRC requirement specified in 10CFR50.46 Appendix-K, i.e. the safety criteria should be satisfied with a probability of 95% and a confidence level of 95%. The 95/95 rule has been recognized by the USNRC to have sufficient conservatism for LBLOCA analyses. Usually a large number of samples are required which can be achieved using the Monte-Carlo random sampling technique. To determine the minimum number of samples required for Monte Carlo method to

TABLE 4 | List of considered hyper-parameters.

Hyper parameters	Dictionary
Number of neurons in 1st layer	10, 20, 50, 100
Number of neurons in 2nd layer	10, 20, 50, 100
Number of hidden layers	1, 2, 3, 4, 5
Optimizers	Adam, Adadelta RMSprop
Kernel (weight) initializers	He (normal, uniform), Xavier (uniform, normal)
Activation functions	ReLU, Tanh, Sigmoid
Learning rate	10.0, 1.0, 0.1, 0.01, 0.003, 0.001
Weight L2 regularizer	0.1, 0.001, 0.0001, 0.00001
Bias L2 regularizer	0.1, 0.001, 0.0001, 0.00001
Activity L2 regularizer	0.1, 0.001, 0.0001, 0.00001

achieve the safety criteria according to the 95/95 rule, it is essential to ensure convergence when the average output stabilizes over the number of samples.

ANN Model Development

In this investigation, six ANN model development steps are applied. The six steps are 1) input selection, 2) data splitting, 3) architecture selection, 4) structure selection, 5) model optimization and 6) model validation.

The input data for the ANN model will be a selected set of the uncertain parameters identified earlier and derived using the PIRT for LBLOCA. The feature selection process is important for ANN model development since too many variables will slow down the optimization process and may prevent the model from finding a good solution (Geron, 2019); whereas a few features may not be sufficient for the model to properly learn the system characteristics embedded in the data. A correlation matrix based on Spearman's method is therefore used to identify the key features from the 19 uncertain parameters that impact the PCT the most.

Next, the random sampling technique is applied to split the database into three main categories: one for training (3,022 samples), validating and testing the model (202 and 332 samples, respectively). In order to improve the ANN performance metrics during training, the input and output parameters should have the same scale. Before the propagation of uncertainty, all input parameters have been scaled; hence only the output parameter (peak cladding temperature, *PCT*) needs to be normalized using the min-max scaling function:

$$y_{scaled} = \left(\frac{y_i - y_{min}}{y_{max} - y_{min}} \right) \tag{5}$$

where y_{max} and y_{min} are the maximum and minimum values of *PCT* in the dataset, respectively; while y_i represents the temperature to be scaled and y_{scaled} represents the scaled temperature.

Architecture selection refers to the choice of ANN hyper-parameters. Since there are a lot of hyper-parameters that can be tuned, choosing the suitable set is a delicate task given the large number of degrees of freedom that the user can manipulate during the ML tuning process. The tuning process of the ML model is therefore more an art than a science and depends on the problem at hand as well as the characteristics of the data. Hence, finding a set of optimal hyper-parameters that provide the best model performance without compromising either its predictive accuracy or generalization capability can be computationally challenging. This is particularly true for hyper-parameters as opposed to other model parameters since the former are not learnt by the model during the training process but must be set manually. Various techniques can be employed to search for the most appropriate hyper-parameters: grid search, random search and Bayesian optimization. In this work, the random search method is used to expedite the convergence. **Table 4** shows all the hyper-parameters that are tuned in this study.

ANN tuning is an important step to enhance the model predictability by converging on the most optimum combination of hyper-parameters. In this study, an automatized optimization tool, the Talos (Autonomio Talos, 2019) software, is used. Talos is an open sources software written in Python language. It is compatible with Keras (Chollet, 2015) application programming interface (API) that is suited for the development of artificial neural network (ANN) models. Currently, Talos does not support any other machine learning model other than the ANN architecture and it only supports Keras backend machine learning algorithms.

Initially, the user needs to define the Keras for the ANN algorithm development. Then the user needs to define the search space boundary in the format of key-value pair python dictionary. Afterwards, the scan function is used to run the Talos experiment. The arguments of the scan function include the type of search method (grid or probabilistic), model's name, number of epochs, batch size and search constraints. Talos will generate a list of possible hyper-parameters combinations along with their corresponding values that can be analyzed using the built in command such as report and predict functions. The size of the results list depends on the number of parameters defined in the search space boundary dictionary. The analysis process can be done by analyzing the whole or the specific model combinations.

If the user is satisfied with the value of the performance metric generated from the results list, then the deploy function is used to save and call the model from the defined python dictionary path and hence the Talos experiment is complete.

TABLE 5 | Variables for reducing computation time.

Variables	Types
Search method	Stratified sampling method
Random method	Quantum method
Fraction limit	0.1
Reduction method	Spearman
Reduction interval	1
Reduction window	1
Reduction threshold	0.3
Reduction metric	Mean squared error
Minimized loss	True
Fraction limit	0.2

As mentioned earlier, the random quantum search method is used to optimize the model. To further reduce the computing time to find the optimized model, both early stopping and window reduction strategies are used. Early stopping prevents the Talos tool from evaluating models that shows unproductive permutation when its metrics are no longer improving; whereas, the window reduction strategy allows the Talos tool to compare the upcoming model with the previously evaluated model based on the specific metrics. Once any of the two criteria is satisfied first, the computation will stop and the results list is generated. **Table 5** shows all the parameters used to reduce computation time.

It should be noted that the Talos ability to find the optimized architecture based on hyper-parameters combinations relies heavily on the defined search space boundary. As such, if the final results do not provide reasonable performance metrics, it is essential to redefine the search space dictionary by adding new hyper-parameters or retuning their corresponding values in order to improve the model accuracy. Even though the Talos is an automated optimization tool, it is still based on a trial and error process that required extensive knowledge in regards to the behavior of each hyper-parameter towards the ANN model. However, the Talos can expedite the optimization process.

It is worth noting that neither under-fitting nor overfitting is desirable in machine learning. To prevent under-fitting, a large database has been generated to train the model, as many input parameters are used to develop the model and the training time was increased until the cost function is minimized to an acceptable value. To mitigate over-fitting, number of techniques can be used for example: cross validation, regularization, dropout, and early stopping. An overfitting model tends to have good learning metrics during training but performs poorly during the validation process. To avoid this, the data is split into "training", "testing" and "validation" data subsets. The model uses the "training" subset during learning and used the unseen subsets for validation and prediction, respectively. Validation metrics were therefore generated using a subset of the database unseen during the training process to ensure that the ANN model is not overfitting the data. In addition to cross-validation, regularization and early stopping were also used to make sure the model does not memorize the data. Further a dropout layer was placed between the input and the hidden layer, whereby the drop rate is determined solely based on random search algorithm.

The regression type ANN was evaluated using the mean squared error (MSE) that represents the squared difference between the predicted and actual value as shown in **Eq. 6**:

$$MSE = \frac{\sum_{i=1}^{n}\left(y_{pred} - y_{act}\right)^2}{n} \tag{6}$$

where y_{pred} and y_{act} are the predicted and actual or known value of the dependent variables, respectively; while, n is the number of samples in the dataset. Another important performance metric is the determination coefficient or the R^2 value that measure how well the model predict under sporadic unseen data. The highest value is 1.0 and indicates that the model has strong generalization capability. The R^2 metric is expressed as **Eqs 7**, **8**:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i,act} - y_{i,pred}\right)^2}{\sum_{i=1}^{n}\left(y_{i,act} - \overline{y}\right)^2} \tag{7}$$

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_{i,act} \tag{8}$$

RESULTS AND DISCUSSION

This section is dedicated to the obtained results. The first subsection focuses on the results of the thermal hydraulic model. This is followed by the results of the uncertainty quantification and post-processing of the generated database. Finally, the results of the ML model are presented.

Thermal Hydraulic Model Results

The thermal hydraulic system response is validated against values reported in the APR1400 DCD (KHNP, 2014) for both steady state and transient simulations. The comparison between the MARS and the APR1400 steady state response are shown in **Table 6**.

Based on **Table 6**, the steady state simulation agrees reasonably well with the plant reference data and the calculated variables are considered to be within the acceptable

TABLE 6 | Validation of steady state analysis.

Parameters	MARS	DCD	Error (%)
Power (MWt)	4062.66	4062.0	0.0
RCP flowrate (kg/s)	5272.0	5250.0	0.4
Core flowrate (kg/s)	20367.0	20361.0	0.03
Primary pressure (MPa)	15.52	15.51	0.01
Core inlet temperature (K)	564.3	563.8	0.12
Core outlet temperature (K)	598.4	597.1	0.16
Upper head temperature (K)	563.9	584.5	3.53
Pressurizer level (m)	8.22	8.18	0.5
Secondary pressure (MPa)	6.90	6.86	0.58
Hot rod fuel temperature (K)	1988.7	1985.2	0.18

TABLE 7 | Validation of the timing of key LBLOCA events.

Key events	Event time (s)	
	MARS	DCD
Cold leg break	0.0	0.0
RCP trip and SG secondary isolations	0.0	-
SIP signal reached	6.0	9.5
Maximum PCT during blowdown (1165.1 K/891.95 C)	10.0	-
High flow SIT signal	15.4	-
High flow SIT injection initiated	15.5	14.8
Core reflood begins	31.0	32.5
SIP injection initiated (42.0 s delay)	48.0	48.3
Low flow SIT injection initiated	57.0	-
Maximum PCT during reflood (1141.7 K/868.55 C)	65.0	-
PCT for average core	143.0	-
PCT for the hot core	151.0	-
SIT depletion time	201.0	171.4

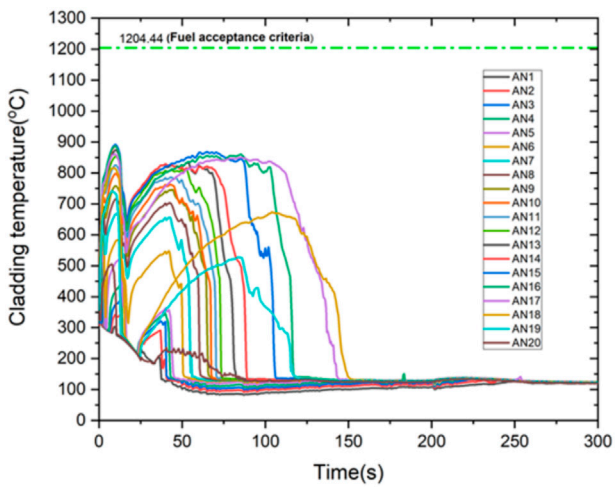


FIGURE 2 | Cladding axial temperature for the hot core.

error limit of less than 5% when compared to the corresponding values reported in the DCD for APR1400. For the transient simulation, key events for the LBLOCA scenario are listed and compared in **Table 7**.

The progression of the LBLOCA is characterized by three different phases; that is, the blowdown, refill, and reflood. At time $t = 0$ where blowdown phase is taking place, the RCP discharge piping starts to break with double ended guillotine break (DEGB) condition. Instantaneously, the loss of offsite power occurred causing all RCPs to coast down. For this scenario the loss of a single emergency diesel generator (EDG) unit is assumed. This leads to the loss of a single safety system train and hence the loss of two SIPs.

During the blowdown phase, the core uncovers and as a result the heat transfer coefficient drops significantly which causes the fuel cladding temperature to increase reaching a maximum of 1165.1 K (891.95°C) at ~10 s which is still below the acceptance criterion of 1447.6 K (1204.44°C) as illustrated in **Figure 2**. Later on, as the decay heat drops the cladding temperature starts to decrease. Further reduction in the cladding temperature is observed due to the large condensation that occurs at the upper guide structure (UGS) and reactor vessel upper head. The condensate passes through the reactor core in reverse direction (from top to bottom) into the lower plenum and headed towards the downcomer region.

The cool down effect continues until the coolant inventory from the upper head of the reactor vessel is depleted. Once the top quenching is over, the fuel-cladding starts to reheat again approximately ~17 s after the break due to the accumulation of decay heat. After this time, there is no longer any cooling mechanism for the core except the one initiated by the safety injection system (SIS) during the reflooding phase.

The refill period starts when the emergency coolant reaches the lower plenum of the vessel and stabilizes till is completely filled and ends when the water level in the lower plenum vessel reaches the core inlet. The SIT high-flow injection starts when the core pressure reaches the set point pressure, approximately 4.2 MPa (~43 kg/cm^2) at ~ 17 s after the pipe break.

The emergency coolant will flow from the downcomer, towards the lower plenum and up into the core. However, even though the SIP is initiated earlier compared to the SIT because of the higher set point pressure approximately, 12.5 MPa (128 kg/cm^2), the SIP starts to inject the emergency coolant at 48.0 s with a 42.0 s delay time. As such after the blowdown phase, the core is cooled initially by the UGS inventory before being assisted by the injection from the SIT followed by the SIP. This delay accounts for the time required for signal actuation time as well as the time needed to start the SIP.

The reflood phase is further subdivided into two phases: the early reflood and late reflood phases. During early reflood phase, sufficient injection from the SIS helps the downcomer to be filled up relatively quickly. However, due to the limited space available in the downcomer combined with the excess amount of emergency coolant, some of the inventory is bypassed through the pipe break regions causing inventory loss. Nonetheless, the amount of coolant available is still sufficient to maintain the core integrity.

The downcomer is nearly filled at around 50 s after the break. At this time, the water level in the downcomer region starts to stabilize as the fluidic device shifts from the high-flow injection to the low-flow injection. The low flow injection stabilizes the downcomer level as the water rises up into the core. The quench front moves vertically upwards to quench the whole core during this times. The maximum flow rate is reached approximately 30 s after the blowdown phase; while the low-flow injection will continue until 200 s. The SIP assist the low flow injection to cover and quench the core. This phase ends when the entire core is quenched from the bottom up gradually and the fuel rod temperature is slightly above the coolant saturation temperature.

During the early refill phase, the steam binding phenomenon may occur which may slow down the process. However, this effect diminishes once the vapor from the upper section of the core no longer received the de-entrainment liquid from the bottom part of the vessel at the surface of the quench front. Afterwards, the steam binding effect starts to disappear after some time which allows the reflooding phase to resume and the reactor core to be filled with water again.

The late reflood phase is marked by the SIT depletion as the SIT low-flow injection comes to an end. During this time, the task of replenishing and providing the emergency inventory for core cooling and core coverage process is achieved solely by the SIPs. The downcomer water level is maintained at a relatively constant value. Both the hot core and average core are finally quenched around 150 s without violating the fuel acceptance criterion of 1477.0 K.

Uncertainty Quantification Results and Database Post Processing

Using the uncertainty quantification framework, the databases for the machine learning model was generated. The Monte Carlo random sampling technique is used to generate 5,000 runs in order to acquire a large data base for the machine learning model. The simulation is conducted using a single PC platform with 3 GHz Intel® Xeon® Gold CPU processor, with 64.0 GB random access memory (RAM), 24 parallel processors and a Windows 10 platform. The time taken to complete the simulation is approximately 3 days.

It is worth noting that, out of the 5,000 samples, only 3,556 samples were successful and used to train the machine learning model. To ensure the number of datasets is enough to represent the 95% probability and 95% confidence criteria, the mean value for the PCT is averaged over the number of samples. As seen in **Figure 3**, the Monte Carlo simulation stabilizes after approximately 2000 runs, hence, a sample size of the 3556 is adequate for to meet the criteria. **Figure 3** shows the spread of PCT which follows a normal distribution with a mean value of 1169 K. The majority of the data are well below the PCT safety criterion of 1477 K. However, there are two data points that lie very close to the safety limit with values of 1462.3 K and 1451.8 K.

As discussed earlier, 19 uncertain parameters are propagated into the thermal hydraulic model using the Dakota uncertainty framework in order to generate the PCT response under LBLOCA scenario. The independent variables are the uncertain parameters (UPs); while the dependent variable is the peak cladding temperature (PCT). However, each UP has different degree of influence towards the PCT. Hence, a sensitivity study is conducted to assess the correlation between the uncertain parameters and the PCT. Spearman's rank correlation, which is a non—parametric measure of the statistical dependency between the variables, is used for the sensitivity analysis. Using the Spearman's correlation coefficients, the strength and the direction of the relationship between the independent and the dependent variables can be evaluated using the following expression:

$$\rho = \frac{\Sigma_i (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\Sigma_i (x_i - \overline{x})^2 \Sigma_i (y_i - \overline{y})^2}} \tag{9}$$

where x_i refers to the input variable, i, $\overline{x}$ is the input variable's mean, y_i is the output variable,i, and $\overline{y}$ is the output variable's

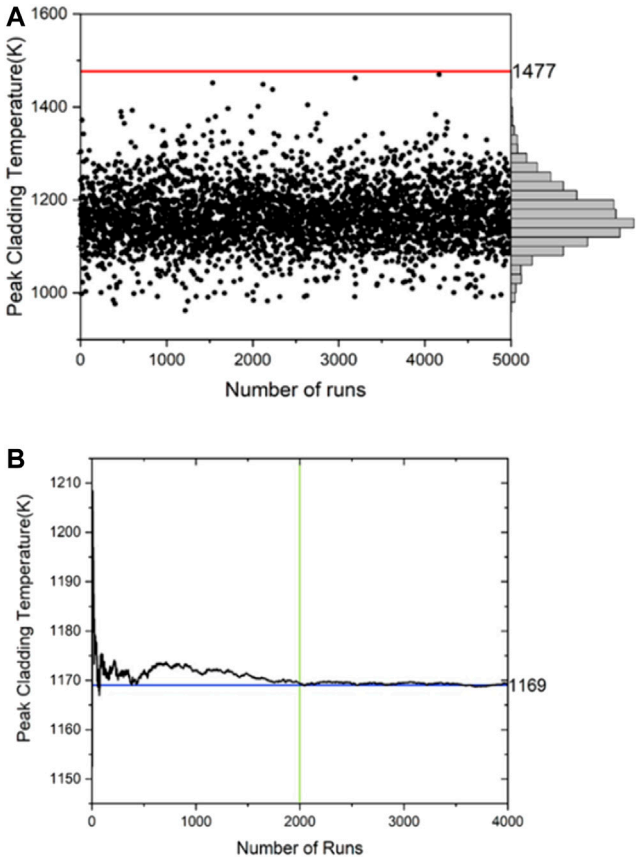


FIGURE 3 | PCT convergence **(A)** and scatter plot **(B)** vs. number of runs.

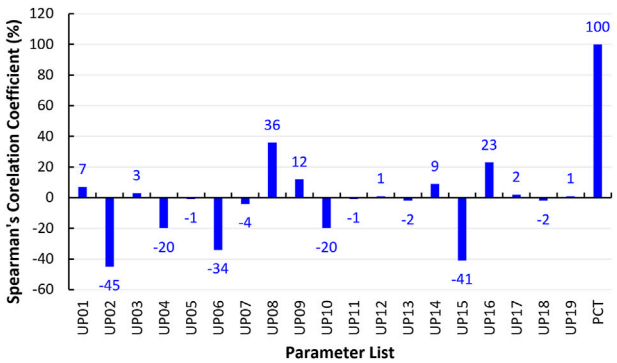


FIGURE 4 | Spearman's correlation matrix.

mean. **Figure 4** shows the Spearman's correlation matrix generated from the database. A positive correlation coefficient means when the independent variable is increased, the dependent variable also increases. While, a negative correlation coefficient means that when the independent variable is increased, the dependent variable decreases. The sensitivity study proves that the uncertainty parameters are independent from each other.

From the sensitivity analysis, a threshold value needs to be defined in order to choose the most significant parameters to reduce the number of inputs for the machine learning model. By selecting a threshold of ±20%, any correlation coefficient higher than the threshold value is deemed to be strongly correlated with the PCT either proportionally or inversely. However, selecting the threshold value is subjective as such it should be tested to find the best value possible.

TABLE 8 | Selected ANN model hyper-parameters.

Hyper-parameters	Values
Number of hidden layers	3
Number of neurons in hidden layer	10
Activation function	Tanh, ReLU
Epoch size	300
Optimizer	Adam
Learning rate	0.001
Regularizes	0.00001
Objective function	MSE
Dropout rate	0.0

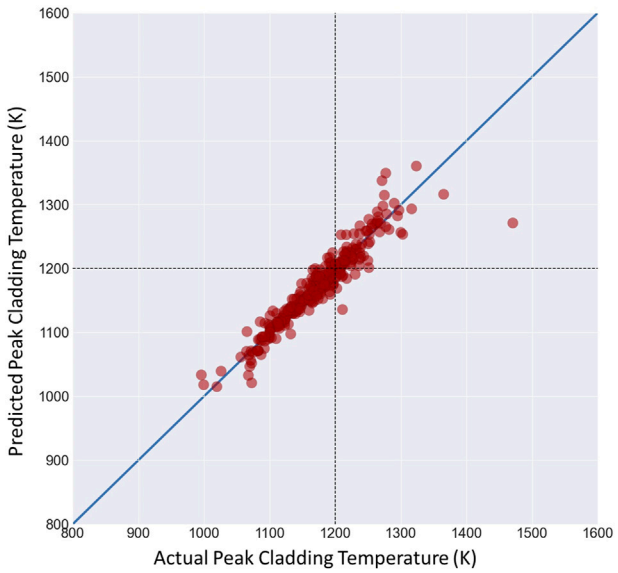


FIGURE 5 | Scatter plot for PCT using 19 uncertain parameters.

TABLE 9 | ML model accuracy corresponding to number of input parameters.

Number of inputs	Spearman's coefficient threshold (%)	R^2	MSE	MAE	MLSE
5	23	0.73319	0.00699	0.0686	0.00336
7	20	0.85655	0.00410	0.04818	0.00189
9	9	0.93484	0.00186	0.03083	0.00084
11	4	0.89104	0.00185	0.03273	0.00091
17	1	0.74024	0.00406	0.04875	0.00198
19	1	0.90242	0.00390	0.0343	0.00117

ANN Model Results

The ANN algorithm has been successfully developed and trained using the database created *via* the uncertainty quantification framework. After tuning, the model was deployed using the Talos optimization tool. 20 models that differ in architectures and hyper-parameters are generated. Among those 20 models, the best model is selected based on the lowest validation metric, i.e. MSE. The final ANN structure recommended by the Talos optimization tool is composed of an input layer, three hidden layers and an output layer **Table 8** shows the hyper-parameters for the selected ANN model achieving the best performance.

The model is trained using the training subset of the available *PCT* database. Next, the validation dataset is used to measure the model ability to learn the salient characteristic of the data. Once the meta-model has been trained, its performance is tested using an unseen subset of data to make the predictions. **Figure 5** shows a comparison between the model predictions of the PCT and the known values produced using the uncertainty quantification

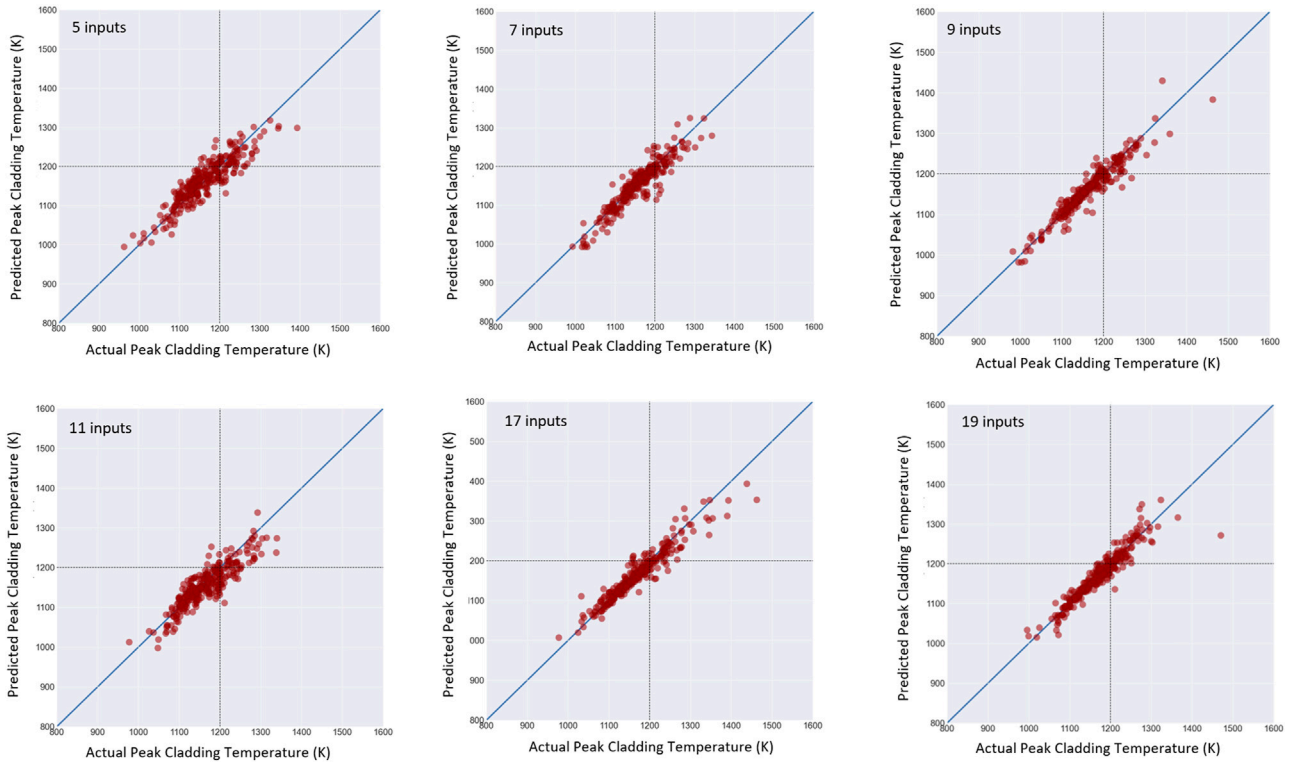


FIGURE 6 | Scatter plot for PCT using different number of uncertain parameters as inputs.

framework for all 19 key uncertain parameters. The ML meta-model predicts the *PCT* with reasonable accuracy (*MSE* = 0.0039); however, it tends to underestimate the high temperatures which is problematic from a safety point of view. This may be due to the fact that unnecessary data from other uncertainty parameters may confuse the ML algorithm and hence impact the model accuracy.

However, the obtained MSE depends on the chosen input parameters. To assess the model sensitivity to the number of input parameters, the model is tested with multiple sets of inputs (5, 7, 9, 11, 17, and 19 UPs) that correspond to different threshold values for the Spearman's correlation coefficient (i.e. level of importance to PCT). **Table 9**, summarizes the performance metrics of the model with different number of uncertain parameters used as inputs to the meta-model. After dimension reduction, the various cases investigated are compared to each other using a number of performance metrics: the determination coefficient (R^2), the mean square error (MSE), the mean arithmetic error (MAE) and the mean logarithmic squared error (MLSE).

Judging by both R^2 and MSE, the model with 9 to 11 parameters achieves reasonable performance. Considering the results presented in **Table 9**, the lowest possible MSE is approximately 0.00185 which is obtained using 11 inputs with an R^2 value of ~0.89. When, the machine learning model is tested with nine inputs, approximately similar results are obtained with a loss function, MSE, of ~0.00186 at an R^2 value of ~0.93. Outside this range (9 to 11 inputs) the ML model performance deteriorates. Given, the aleatory nature of the ANN model optimization which is based on random optimization, the optimum number of input parameters is expected to be within the range 9 to 11 variables with slight variation in performance metrics results (MSE ~0.002, R^2 ~ 0.9). It is worth noting that from a safety perspective, it is better to tune the model for high temperatures.

Figure 6 shows the prediction results in comparison to the actual known PCT values with different number of input parameters. Clearly, the lower number of input parameters does not capture fully the relationship between inputs and outputs embedded in the database. On the other hand, the higher number of input parameters may include unnecessary details that may confuse the model. One would suspect an optimum number of input parameters may exist for better prediction capability as evidenced by the results shown in **Table 9**.

CONCLUSION

The aim of this work is to develop a machine learning (ML) algorithm that is capable of accurately predicting the key safety parameter, PCT, under LBLOCA scenario. The algorithm was trained using a database created using the best estimate code MARS-KS V1.4, with uncertainty quantification using the statistical tool, Dakota to propagate the uncertainty parameters

through the thermal hydraulic model. A Monte Carlo sampling method was used to generate 3,556 successful samples to train the ML model using 19 key uncertain parameters. The Monte Carlo simulation converged after 2,000 samples yielding the required average PCT in consistency with the USNRC requirement of 95% probability and 95% confidence interval.

An ANN model was developed, trained and optimized using the provided database. The ANN model was successfully tuned using the Talos optimization tool to predict the PCT with high accuracy. The optimum model is chosen based on the desirable objective function and the validations metric, MSE. A model with 9 to 11 inputs best represents the data and can be used to predict PCT accurately with a MSE of ~0.002 with R^2 value of ~0.9.

This study successfully shows that ANN can be used as a surrogate to the thermal hydraulics MARS-KS model to predict the PCT value for the LBLOCA scenario using only the key uncertain input parameters with reasonable accuracy. For future work, the framework developed for this project can be used for uncertainty quantification of other key safety parameters such as Departure from Nucleate Boiling Ratio (DNBR) under LBLOCA or other critical scenarios. This is a preliminary step towards developing an expert support system that can be used to guide the operator actions under the stressful accident conditions. As a bounding accident scenario, the analysis of the LBLOCA case paves the way to using machine learning as a decision making tool for design extension conditions as well as severe accidents.

DATA AVAILABILITY STATEMENT

The datasets presented in this article may be available based on personal communication with the corresponding author. Requests to access the datasets should be directed to aya.diab@kings.ac.kr.

AUTHOR CONTRIBUTIONS

CRediT authorship contribution statement: WS: Methodology, Software, Writing—original draft. AD: Conceptualization, Supervision, Analysis, Writing—review and editing.

FUNDING

This research was supported by the 2020 research fund of KEPCO International Nuclear Graduate School (KINGS), Republic of Korea. (Corresponding author: AD).

REFERENCES

Abdel-Aal, R. E., and Al-Haddad, M. N. (1997). Determination of Radioisotopes in Gamma-ray Spectroscopy Using Abductive Machine Learning. *Nucl. Instr. Methods Phys. Res. Section A: Acc. Spectrometers, Detectors Associated Equipment.* 391 (2), 275–288. doi:10.1016/s0168-9002(97)00391-4

Adams, B. M., Bohnhoff, W. J., Dalbey, K. R., Ebeida, M. S., Eddy, J. P., Eldred, M. S., et al. (2020). *Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.12 User's Manual.* Albuquerque, New Mexico, USA: Sandia Technical Report; Sandia National Laboratories.

Ahsan, S. N., and Hassan, S. A. (2013). "Machine Learning Based Fault Prediction System for the Primary Heat Transport System of CANDU Type Pressurized Heavy Water Reactor," in 2013 International Conference on Open Source Systems and Technologies Lahore. doi:10.1109/icosst.2013.6720608

Alketbi, S., and Diab, A. (2021). Using Artificial Intelligence to Identify the Success Window of FLEX Strategy Under an Extended Station Blackout. *Nucl. Eng. Des.* 382 (2021), 111368. doi:10.1016/j.nucengdes.2021.111368

Autonomio Talos [Computer software] (2019). Retrieved from http://github.com/autonomio/talos.

Bae, I. H., Na, M. G., Lee, Y. J., and Park, G. C. (2008). Calculation of the Power Peaking Factor in a Nuclear Reactor Using Support Vector Regression Models. *Ann. Nucl. Energ.* 35, 2200–2205. doi:10.1016/j.anucene.2008.09.004

Chai, J., Sisk, D. R., Bond, I. J., Jarrel, D. B., HatelyMeador, D. D. R. J., and Koehler Watkins Kim, T. M. K. S. W. (2003). *On-line Intelligent Self-Diagnostic Monitoring System for Next Gernaeration Nuclear Power Plants*. Southwest Washington, D.C., USA: United States Department of Energy.

Chang, Y., Wang, M., Zhang, J., Tian, W., Qiu, S., and Su, G. H. (2020). Best Estimate Plus Uncertainty Analysis of the China Advanced Large-Scale PWR during LBLOCA Scenarios. *Int. J. Adv. Nucl. Reactor Des. Technology.* 2, 34–42. doi:10.1016/j.jandt.2020.07.002

Chen, F.-C., and Jahanshahi, M. R. (2018). NB-CNN: Deep Learning-Based Crack Detection Using Convolutional Neural Network and Naïve Bayes Data Fusion. *IEEE Trans. Ind. Electron.* 65 (5), 4392–4400. doi:10.1109/tie.2017.2764844

Chen, L., and Wei, Y.-X. (2009). Nuclide Identification Algorithm Based on K-L Transform and Neural Networks. *Nucl. Instr. Methods Phys. Res. Section A: Acc. Spectrometers, Detectors Associated Equipment.* 598 (2), 450–453. doi:10.1016/j.nima.2008.09.035

Chollet, F. (2015). Keras. GitHub. Retrieved from https://github.com/fchollet/keras.

Dawani, J. (2020). *Hands on Mathematics for Deep Learning*. Birmingham: Packt Publisher.

Devereux, M., Murray, P., and West, G. (2020). A New Approach for Crack Detection and Sizing in Nuclear Reactor Cores. *Nucl. Eng. Des.* 359. doi:10.1016/j.nucengdes.2019.110464

Erdogan, A., and Gekinli, M. (2003). A PWR Reload Optimization Code (XCORE) Using Artificial Neural Networks and Genetic Algorithms. *Ann. Nucl. Energ.* 30 (1), 35–53. doi:10.1016/S0306-4549(02)00041-5

Farber, J. A., and Cole, D. G. (2020). Detecting Loss-Of-Coolant Accidents Without Accident-Specific Data. *Prog. Nucl. Energ.* 128, 103469. doi:10.1016/j.pnucene.2020.103469

Faria, E. F., and Pereira, C. (2003). Nuclear Fuel Loading Pattern Optimisation Using a Neural Network. *Ann. Nucl. Energ.* 30 (5), 603–613. doi:10.1016/s0306-4549(02)00092-0

Farshad Faghihi, K. H., and Seyed, M. (2011). *A Literature Survey of Neutronics and Thermal-Hydraulics Codes for Investigating Reactor Core Parameters: Artificial Neural Networks as the VVER-1000 Core Predictor, Chapter 6*. Nuclear Power Systems Simulations and Operations. Shanghai, China: InTech, 103–122.

Geron (2019). *Hands on Machine Learning with Scikit Learn, Keras & Tensorflow*. Seebastopol: O'Reilly Publisher.

Gomez Fernandez, M., Tokuhiro, A., Welter, K., and Wu, Q. (2017). Nuclear Energy System's Behavior and Decision Making Using Machine Learning. *Nucl. Eng. Des.* 324, 27–34. doi:10.1016/j.nucengdes.2017.08.020

Gomez-Fernandez, M., Higley, K., Tokuhiro, A., Welter, K., Wong, W.-K., and Yang, H. (2020). Status of Research and Development of Learning-Based Approaches in Nuclear Science and Engineering: A Review. *Nucl. Eng. Des.* 359, 110479. doi:10.1016/j.nucengdes.2019.110479

IAEA (2020). *Emerging Technologies Workshop Insight and Actionable Ideas for Key Safeguards Challenge*. Vienna: International Atomic Energy Agency.

KAERI (2004). *MARS Code Manual Volume 1 : Code Structure, System Models and Solution Methods*. Daejeon: Korea Atomic Energy Research Institute.

KAERI (2009). *"MARS Code Manual," KAERI/TR2812/2004 Korea*. Daejon: Atomic Energy Research Institute.

Kamuda, M., and Sullivan, C. J. (2019). An Automated Isotope Identification and Quantification Algorithm for Isotope Mixtures in Low-Resolution Gamma-ray Spectra. *Radiat. Phys. Chem.* 155, 281–286. doi:10.1016/j.radphyschem.2018.06.017

Kang, D. G. (2016). Analysis of LBLOCA Using Best Estimate Plus Uncertainties for Three-Loop Nuclear Power Plant Power Uprate. *Ann. Nucl. Energ.* 90, 318–330. doi:10.1016/j.anucene.2015.12.017

Keller, P. E., and Kouzes, R. T. (1994). Gamma Spectral Analysis via Neural Networks. *Proc. 1994 IEEE Nucl. Sci. Symp.* 1 1, 341–345. doi:10.1109/NSSMIC.1994.474365

KEPCO (2014). *Design Control Document Chapter 15: Transient and Accident Analysis*. Daejeon: Korea Electric Power Corporation.

KHNP (2014). *Design Control Document Chapter 15 Transient and Accident AnalysesKorea Hydro and Nuclear Power*.

Park, H. M., Lee, J. H., and Kim, K. D. (2020). Wall Temperature Prediction at Critical Heat Flux Using a Machine Learning Model. *Ann. Nucl. Energ.* 141, 2020. doi:10.1016/j.anucene.2020.107334

LA-UR-95-2718 (1995). *AP600 Large Break LOCA Phenomena Identification and Ranking Tabulation*.

Lee, C.-J., and Lee, K. J. (2006). Application of Bayesian Network to the Probabilistic Risk Assessment of Nuclear Waste Disposal. *Reliability Eng. Syst. Saf.* 91, 515–532. doi:10.1016/j.ress.2005.03.011

Lee, S. W., Chung, B. D., Bang, Y.-S., and Bae, S. W. (2014). Analysis of Uncertainty Quantification Method by Comparing Monte-Carlo Method and Wilks' Formula. *Nucl. Eng. Technology*. 46 (4), 481–488. doi:10.5516/net.02.2013.047

Lin, Z., Liu, X., Lao, L., and Liu, H. (2020). Prediction of Two-Phase Flow Patterns in Upward Inclined Pipes via Deep Learning. *Energy* 210, 118541. doi:10.1016/j.energy.2020.118541

Ma, J., and Jiang, J. (2011). Applications of Fault Detection and Diagnosis Methods in Nuclear Power Plants: A Review. *Prog. Nucl. Energ.* 53 (3), 255–266. doi:10.1016/j.pnucene.2010.12.001

Martin, R. P., and O'Dell, L. (2008). Development Considerations of AREVA NP Inc.'s Realistic LBLOCA Analysis Methodology. *Sci. Technology Nucl. Installations.* 13. doi:10.1155/2008/239718

Montes, J. L., François, J. L., Ortiz, J. J., Martín-del-Campo, C., and Perusquía, R. (2009). Local Power Peaking Factor Estimation in Nuclear Fuel by Artificial Neural Networks. *Ann. Nucl. Energ.* 36 (1), 121–130. doi:10.1016/j.anucene.2008.09.011

Musoiu, R., Mihaela, R., Prisecaru, L., and Allison, C. (2019). "BEPU Approach in the CANDU 6 Severe Accident Analysis," 2019 International Conference on ENERGY and ENVIRONMENT (CIEM), 73–77. doi:10.1109/ciem46456.2019.8937571

Na, M., Ho Shin, S., Mi Lee, S., Jung, D. W., Pyung Kim, S., Jeong, J., et al. (2004). Prediction of Major Transient Scenarios for Severe Accidents of Nuclear Power Plants. *Nucl. Sci. IEEE Trans.* 51 (5), 313–321. doi:10.1109/tns.2004.825090

Nissan, E. (2019). An Overview of AI Methods for In-Core Fuel Management: Tools for the Automatic Design of Nuclear Reactor Core Configurations for Fuel Reload, Arranging and Partly Spent Fuel. *Designs* 3 (3), 37. doi:10.3390/designs3030037

OECD (1998). *Report on the Uncertainty Method Study*. Paris, France: Organisation for Economic Co-operation and Development.

OECD (2007a). *BEMUSE Phase III Report Uncertainty and Sensitivity Analysis of the LOFT L2-5 Test*. Paris, France: Organisation for Economic Co-operation and Development.

OECD (2007b). *Task Group on Safety Margin Action Plan (SMAP) Safety Margin Action Plan - Final Report*. Paris, France: Organisation for Economic Co-operation and Development.

Petruzzi, M., Cherubini, M., D'Auria, F., Lanfredini, M., and Mazzantini, O. (2016). The BEPU Evaluation Model with RELAP5-3D for the Licensing of the Atucha-II NPP. *Nucl. Technology.* 193, 51. doi:10.13182/nt14-145

Pinheiro, V., Santos, M., Desterro, F., Schirru, R., Márcio, C., and Pereira, N. A. (2020). "Nuclear Power Plant Accident Identification System With "Don't Know" Response Capability: Novel Deep Learning-Based Approaches". *Ann. Nucl. Energ.* 137, 107111. doi:10.1016/j.anucene.2019.107111

Queral, C., Montero-Mayorga, J., Gonzalez-Cadelo, J., and Jimenez, G. (2015). AP1000 Large-Break LOCA BEPU Analysis With TRACE Code. *Ann. Nucl. Energ.* 85, 576–589. doi:10.1016/j.anucene.2015.06.011

Rani Patra, S., Jehadeesan, R., Rajeswari, S., Banerjee, I., A. V Satya Murty, S., Padmakumar, G., et al. (2012). Neural Network Modeling for Evaluating Sodium Temperature of Intermediate Heat Exchanger of Fast Breeder Reactor. *Ac.* 2, 16–22. doi:10.5923/j.ac.20120202.03

Ridlluan, A., Manic, M., and Tokuuhiro, A. (2009). EBaLM-HTP- A Neural Network Thermohydraulic Prediction Model of Advanced Nuclear System Components. *Nucl. Eng. Des.* 239 (2), 308–319. doi:10.1016/j.nucengdes.2008.10.027

Santosh, G., Babar, A. K., Kushwaha, H. S., and Venkat Raj, V. (2003). Symptom Based Diagnostic System for Nuclear Power Plant Operations Using Artificial Neural Networks. *Reliability Eng. Syst. Saf.* 82 (1), 33–40. doi:10.1016/s0951-8320(03)00120-0

Siegelmann, H. T., Nissan, E., Nissan, E., and Galperin, A. (1997). A Novel Neural/symbolic Hybrid Approach to Heuristically Optimized Fuel Allocation and Automated Revision of Heuristics in Nuclear Engineering. *Adv. Eng. Softw.* 28 (9), 581–592. doi:10.1016/s0965-9978(97)00040-9

USNRC (1988). *Development of a Phenomena Identification and Ranking Table for Thermal Hydraulic Phenomena during PWR LBLOCA*. Washington, DC, USA: United State Nuclear Regulatory Comission.

USNRC (1989). *USNRC Regulatory Guide 1.157 Best Estimate Calculation of Emergency Core Cooling System Performance*. Washington, DC, USA: U.S Office of Nuclear Regulatory Research.

Young, D. K., Ye, J. A., Chang-Hwoi, K., and Man, G. N. (2019). Nuclear Reactor Vessel Water Level Prediction During Severe Accidents Using Deep Neural Networks. *Nucl. Eng. Technology.* 51 (3), 723–730. doi:10.1016/j.net.2018.12.019

Zamer, A., Mirza, S. M., and Mirza, N. M. (2014). Core Loading Pattern Optimization of a Typical Two-Loop 300 MWe PWR Using Simulated Annealing (SA) Novel Crossover Genetic Algorithms (GA) and Hybrid GA/SA Schemes. *Ann. Nucl. Energy* 65, 122–131. doi:10.1016/j.anucene.2013.10.024

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Energy Research

ORIGINAL RESEARCH
published: 02 March 2022
doi: 10.3389/fenrg.2022.823395

Model-Based Deep Transfer Learning Method to Fault Detection and Diagnosis in Nuclear Power Plants

Yuantao Yao[1], *Daochuan Ge*[1]*, *Jie Yu*[1] *and Min Xie*[2]

[1]Institute of Nuclear Energy Safety Technology, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, China, [2]Department of Advanced Design and System Engineering, City University of Hong Kong, Kowloon, Hong Kong SAR, China

Deep learning–based nuclear intelligent fault detection and diagnosis (FDD) methods have been widely developed and have achieved very competitive results with the progress of artificial intelligence technology. However, the pretrained model for diagnosis tasks is hard in achieving good performance when the reactor operation conditions are updated. On the other hand, retraining the model for a new data set will waste computing resources. This article proposes an FDD method for cross-condition and cross-facility tasks based on the optimized transferable convolutional neural network (CNN) model. First, by using the pretrained model's prior knowledge, the model's diagnosis performance to be transferred for source domain data sets is improved. Second, a model-based transfer learning strategy is adopted to freeze the feature extraction layer in a part of the training model. Third, the training data in target domain data sets are used to optimize the model layer by layer to find the optimization model with the transferred layer. Finally, the proposed comprehensive simulation platform provides source and target cross-condition and cross-facility data sets to support case studies. The designed model utilizes the strong nonlinear feature extraction performance of a deep network and applies the prior knowledge of pretrained models to improve the accuracy and timeliness of training. The results show that the proposed method is superior to achieving good generalization performance at less training epoch than the retraining benchmark deep CNN model.

Keywords: fault detection and diagnosis, deep learning, transfer learning, freezing and fine-tuning strategy, nuclear power plants

OPEN ACCESS

Edited by:
Wenxi Tian,
Xi'an Jiaotong University, China

Reviewed by:
Xingang Zhao,
Oak Ridge National Laboratory (DOE), United States
Shripad T. Revankar,
Purdue University, United States

***Correspondence:**
Daochuan Ge
daochuan.ge@inest.cas.cn

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *27 November 2021*
Accepted: *07 February 2022*
Published: *02 March 2022*

Citation:
Yao Y, Ge D, Yu J and Xie M (2022) Model-Based Deep Transfer Learning Method to Fault Detection and Diagnosis in Nuclear Power Plants. Front. Energy Res. 10:823395. doi: 10.3389/fenrg.2022.823395

INTRODUCTION

No matter how advanced the energy systems have progressed with state-of-the-art techniques, operation safety and reliability will be a central research topic all the time. Especially for nuclear systems, safeguards are even more critical and cannot be ignored (Perrault, 2019; Matteo et al., 2021; Yao et al., 2021). Most of the severe nuclear leakage events throughout the history of humankind have been caused by operators' inappropriate responses and solutions. Therefore, it is critical to provide administrators with auxiliary information under different nuclear system operation conditions before an accident worsens (Wahlström, 2018; Yoo et al., 2018).

One predictive maintenance approach that has become increasingly valued is fault detection and diagnosis (FDD), to judge (detection) and identify (diagnosis) the type of fault (Yangping et al., 2000; Ma and Jiang, 2011). According to the review work from Zhao et al. (2021), the development of fault diagnosis in nuclear power plants (NPPs) mainly goes through three essential stages: the model-

based method, knowledge-based rule method, and currently popular data-driven method. Model-based approaches fall into two main categories. One is through the statistical anomaly (fault) and average state residuals, such as parity check, wavelet transform, and time–frequency analysis of quantitative models (Zhong et al., 2018). The other is a qualitative analysis based on the physical or graph structure models. Besides, the rule-based approaches are an essential branch of fault diagnosis research, which by triggering specific "if–then" rules to determine results related to measured/detected fault symptoms. The rule libraries are developed using expert judgment and prior knowledge of systems; the most famous rule of which is the fuzzy rule (Xu et al., 2019).

However, modern industrial systems have a nonlinear, considerable time delay, uncertainty factors, which makes it challenging to build precise mathematical models. Therefore, the application of model-based and rule-based methods is limited.

The data-driven method does not require prior knowledge of the object system (mathematical model or expert experience). It takes the monitoring data as the research object to estimate the state of the target system, avoiding the shortcomings of the physical model–based method. Feature extraction and classifier design are two main parts of this method. The popular method in feature extraction is principal component analysis, which reduces the dimension of the data and extracts critical information (Peng and Wang, 2018). The famous classifiers mainly include Support Vector Machine (Yao et al., 2020a), Extreme Learning Machine (Zheng et al., 2019), and artificial neural network (Xin et al., 2019).

With the advances of the Internet of Things, big data, and the continuous improvement of equipment scale (Lee et al., 2014; Wang et al., 2015; Seabra et al., 2016), deep learning–based fault diagnosis methods such as the automatic deep encoder–decoder (Wang and Zhang, 2018), deep belief networks (DBNs) (Tang et al., 2018), and deep convolutional neural networks (CNNs) (Wen et al., 2017) are gradually coming into view. Compared with traditional data-driven methods, high integrated and end-to-end deep networks with multiple hidden layers can learn and fit any nonlinear relationship under sufficient training data, widely applied in different energy systems.

Correa-Jullian et al. (2020) discussed several deep networks–based methods. They applied them to prognose the performance of solar hot water systems under different meteorological conditions. Xu et al. (2020) combined the CNNs with the variational mode decomposition algorithms to accomplish the fault diagnosis of the rolling bearing of wind turbines. Guo et al. (2018) proposed a fault diagnosis approach using a DBN with a model optimization strategy for building energy saving. In nuclear systems, Peng et al. (2018) utilized the feature selection capability of correlation analysis for dimensionality reduction and DBNs for fault identification. Saeed et al. (2020) proposed a fault diagnosis model based on the deep hybrid networks to achieve FDD with different levels. Mandal et al. (2017) introduced a DBN-based detection and diagnosis method for the thermocouple sensor fault. Yao et al. (2022) presented a residual CNN with an adaptive noise elimination procedure for the FDD in small modular reactors.

The designed deep network–based model can learn features from the original data and have overwhelming advantages in solving various fault classification problems. On the other hand, the end-to-end deep learning model has integration advantages compared with traditional machine learning–based manual feature engineering selection. When training diagnostic models based on data-driven methods, we usually default to the same training and test data distribution. Suppose that there are enough training samples for the fault diagnosis task of a given scene. In that case, the model nonlinear relationship can be fitted through parameter optimization to achieve a high-precision diagnosis.

However, in the actual system operation process, the nuclear systems are in a stable operation state most of the time. The difficulty of obtaining fault data leads to a small number of samples, and most faults are in an unmarked condition. Traditional deep learning models such as CNNs and DBNs will overfit when training on small sample data sets and significantly reduce diagnostic performance. On the other hand, the historical training model fails to identify new data once the system runs under different conditions or upgraded or updated environment.

The transfer learning (TL) method is proposed to solve the above problems. It transfers the knowledge learned from the neighboring domain to improve learning performance under insufficient target task training data. In recent years, it has been developed and applied in natural language processing, computer vision, and autonomous driving (Ruder et al., 2019; Zhuang et al., 2020). Furthermore, to apply TL in the FDDs is to relax the constraint that targets domain data, and the source domain data must obey the same distribution. It will reduce the urgency of collecting massive data combined with TL. At present, relevant research in energy is scarce. For nuclear systems, TL-based fault diagnosis, the initial exploration, has not been involved.

To address the above problems, we propose a diagnosis framework based on transferable CNN models to make full use of the prior knowledge of the pretraining model. Compared with the traditional deep network–based diagnosis framework, the proposed method has the following advantages:

1) A novel freezing and tuning transfer strategy based on a pretraining model can be applied in the nuclear systems' fault diagnosis under different operating conditions and equipment.
2) The proposed method does not need to train and optimize the parameters of each layer of the model one by one. Still, it makes full use of the high-dimensional feature extraction capability of the pretraining model for source domain data.
3) The proposed method significantly reduces the model retraining time. Under insufficient data, avoiding data expansion technology minimizes the probability of model overfitting and improves the training performance.
4) The proposed method has good portability. After simple optimization for different target transfer environments, future research can achieve ideal results.

The remainder of the article is organized as follows: *Vanilla CNN Structure* briefly introduces the vanilla CNN structure. *Proposed Method* proposes a TL-based fault diagnosis

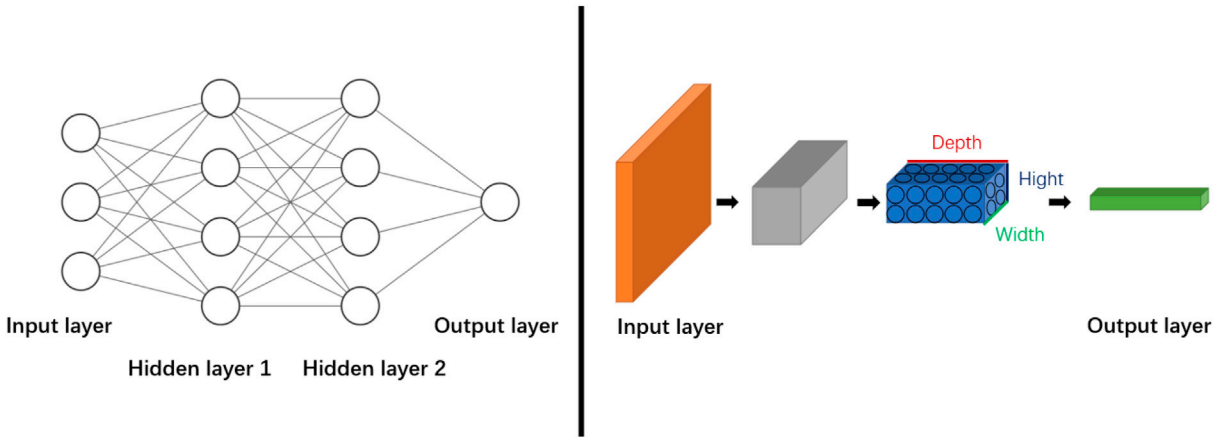


FIGURE 1 | The comparison between fully connected neural networks and the convolutional neural network.

procedure. The case study is presented in *Case Study*. Results and discussions are shown in *Results and Discussion*. *Conclusion* concludes the article and makes an outlook for the future work.

VANILLA CONVOLUTIONAL NEURAL NETWORK STRUCTURE

CNNs (LeCun et al., 2015) are traditional deep networks commonly used in classification research. They are mainly based on the feedforward networks to add the corresponding convolution operation process to extract the high-dimensional characteristics of the input data to be analyzed. From the initial application in speech and image recognition to the current machinery, energy, aerospace, and other fields of abnormal detection, fault diagnosis, time-series prediction, and other applications have a wide range of prospects. Compared with traditional neural networks, neurons in the CNN structure are arranged in three dimensions, as is shown in **Figure 1**.

The neurons in the layers are not fully connected traditionally but are only related to a small area of the previous layer. The CNN structure is mainly composed of three essential parts: the convolutional layer, pooling layer, and fully connected layer, which are executed to make feature extraction, filtering, and output with nonlinear combination, respectively.

PROPOSED METHOD

Brief Structure

Figure 2 shows the scheme of the proposed TL-based diagnosis framework. It mainly contains four critical steps: data set construction, model presetting, transferable model construction, and model optimization and testing.

First, the source and target domains of the research data sets are constructed through the existing comprehensive experimental platform. The source domain data comes from the previous data, and the target domain data comes from different working

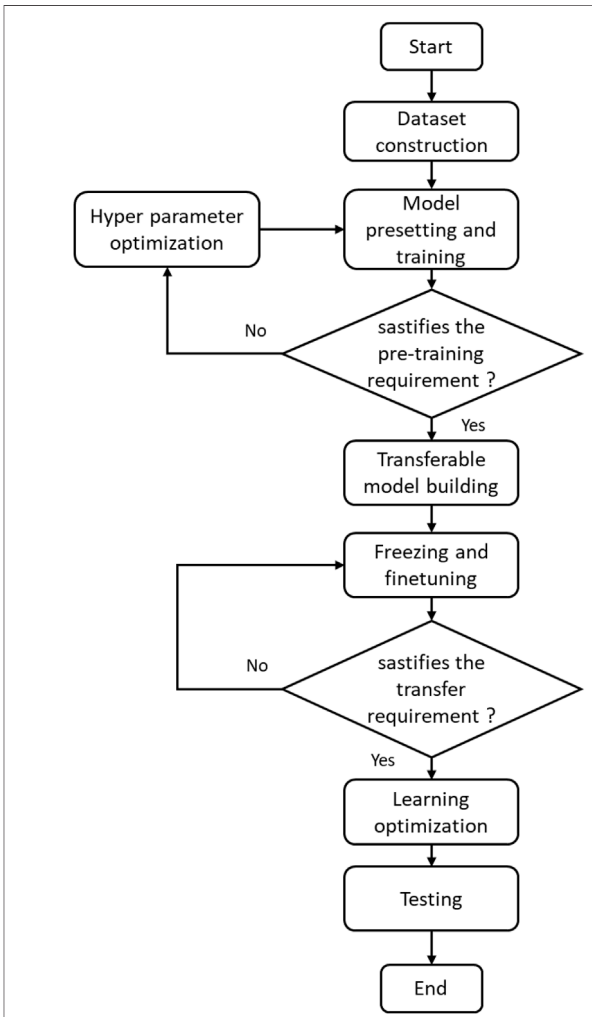


FIGURE 2 | Variations of training accuracy and loss in benchmark convolutional neural network and the proposed transfer diagnosis model aiming at different target domain data sets.

TABLE 1 | Pretrained convolutional neural network model structure.

Layer no	Module no	Layer type	Kernel size	Output size	Stride	Padding
1		Input	—	200 ∗ 316 ∗ 3	—	—
2	M 4	Convolution	128	100 ∗ 158 ∗,128	2	Yes
3		BN	—	100 ∗ 158 ∗,128	—	—
4		ReLU	—	100 ∗ 158 ∗,128	—	—
5		Max pooling	128	50 ∗ 79 ∗,128		
6	M 3	Convolution	32	25 ∗ 40 ∗ 32	2	Yes
7		BN	—	25 ∗ 40 ∗ 32	—	—
8		ReLU	—	25 ∗ 40 ∗ 32	—	—
9		Max pooling	32	13 ∗ 20 ∗ 32		
10	M 2	Convolution	64	7 ∗ 10 ∗ 64	2	Yes
11		BN	—	7 ∗ 10 ∗ 64		
12		ReLU	—	7 ∗ 10 ∗ 64		
13		Max pooling	64	4 ∗ 5 ∗ 64		
14	M 1	Dense	—	256		
15		Dense	—	128		
16		Softmax		11		

conditions or new reactor types. Second, the CNN-based model designed initially is used to complete sufficient training and hyperparameter optimization for the source domain data. Third, a part of the data set in the target domain data set is selected for the transferred model training, which is completed by freezing and tuning some layer parameters. Finally, the hyperparameter tuning is performed on the transferred model, and the target domain data sets are adopted to test the diagnostic performance of the proposed method. The detailed procedure of model presetting and the transferable model building will be introduced.

Model Presetting

The basic structure of the transfer model refers to the previous design, which is modified compared with the vanilla CNN structure (Yao et al., 2020b). It has one input layer, three convolutional layers, three max-pooling layers, and two fully connected layers. Meanwhile, the layers in the network are grouped and divided into specific modules to discuss the influence of different model parts on transfer performance during freezing and tuning. The detailed introduction of the model structure is shown in **Table 1**. On this basis, combined with the state information imaging method, the network input is a two-dimensional matrix image. Furthermore, the number of the convolution kernel is set as a more considerable number (128) to enhance the network feature learning ability. The number of the following convolution kernels is settled as 32 and 64, respectively.

The criterion sets the number of convolution kernels from the least to the most. It gradually increases in multiple relationships to obtain more discriminative features at the higher level of the network. In addition, zero padding is used to make the feature output before and after the convolution operation maintain the same size. Moreover, batch normalization is adopted to avoid network overfitting, thereby maximizing saving the original input information. For the pooling layer parameter setting, the number of feature maps in each component block is the same as that in the convolution layer, verified in the previous work. The Softmax function is selected as the classification function. The Adam optimization method (Kingma and Ba, 2014) is used to make gradient updating, introduced in reference Kingma and Ba (2014) in detail.

Transferable Model Building

The transferable model is based on a fully pretrained CNN–based diagnosis model for the source domain data set. The output of the Softmax function should be replaced by the number of samples in the target domain when constructing the forward transfer model for samples in the target domain. For the hyperparameters of other network layers, layer-by-layer freezing and optimization are adopted, as is shown in **Figure 3**. The detailed procedure is as follows:

Step 1: Replace the output fault categories in the Softmax function according to the target task category.

Step 2: Adopt the target domain data training in a small sample environment, freeze the parameters of the pre-sequence network layer.

Step 3: Adopt a small learning rate design to tune the subsequent connection layer to realize the generalization transfer of the network.

Step 4: Reduce the number of frozen, fixed layers and move them to the transfer connection layer. Similarly, a small learning rate is adopted to tune the transfer connection layer, and the whole process is shown in **Figure 3**.

Step 5: The test data set samples are substituted into the model to obtain its discriminant classes, and the diagnostic performance of the model is tested.

In the *Proposed Method*, Part A, we have selected the basic structural hyperparameters of the CNN model, which includes the step size, number of kernels, batch size, etc. However, the influence of the learning rate on model training is not to be further considered. When using the gradient descent algorithm, if the learning rate is too low, the convergence of the model will be slow. Furthermore, more epochs are needed to complete the training, thus wasting computational power.

On the other hand, if the setting is too large, the model will not converge, reducing the model's diagnostic performance. The

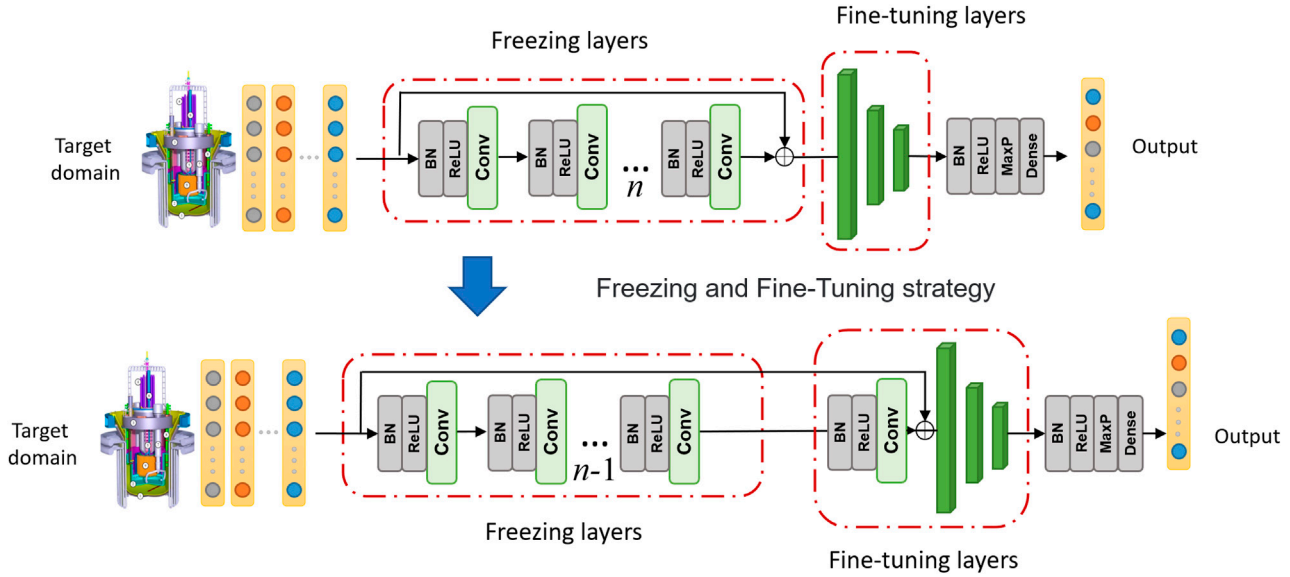


FIGURE 3 | The process of adopted freezing and fine-tuning transfer strategy.

learning rate decay (LDR) is a commonly used method in deep network training. Although the adaptive gradient descent algorithm (Adam) is adopted in the pretraining model to optimize the updating strategy of the gradient, the LDR is still applicable to the training of the model.

Therefore, to improve the efficiency of freezing and fine-tuning the procedure, we divided the training epoch into several parts and adopted an exponential LDR strategy in each part to enhance the training effect of the model, which is as follows:

$$\alpha = D_p^{epoch} * \alpha_0 \tag{1}$$

where D_p represents the LDR in each epoch. P represents the number of division parts. α_0 represents the initial value of the learning rate.

CASE STUDY

Research Platform

Since 2011, the project to build the accelerator-driven systems for nuclear waste transmutation has been developed and researched by the Chinese Academy of Science, including three key stages. Gen-IV China LEAd-based Reactor (CLEAR) was proposed as the reference for the above project (Wu, 2016a).

And then, to test the 1:1 prototype key component and verify the thermal hydraulic performance of the designed CLEAR-I, the integrated nonnuclear test facility CLEAR-S was built commissioning at the end of 2017 (Wu, 2016b). In 2018, the basis experimental hardware system CLEAR assistant simulator was made for the neutron transportation simulation, structure engineering design, and accident security analysis. It utilizes a computerized man–machine interface and digitalized instrumentation and control system. We are currently conducting research and analysis based on nonnuclear test devices and a hardware-in-the-loop simulation experimental platform, which provides data support for further study.

Data Set Description

The data set in this study comes from the CLEAR-I and CLEAR-S accident simulation data stored in the previous research on the simulation experiment platform with the RELAP5 simulation calculation program–based core, which is shown in **Figures 4**, **5**.

The RELAP program is a thermo-hydraulic program developed by Idaho National Laboratory to simulate a transient accident in a light water reactor (Li et al., 2014). It is a one-dimensional transient, two-phase fluid, six hydrodynamic equation and one-dimensional heat conduction. Moreover, point reactor dynamics models are widely used in NPP's accident safety analysis, accident evaluation, experimental analysis, and other fields. RELAP5-HD is a new version developed based on RELAP5. Its most significant feature is that it highly integrates the functions of RELAP5-HD and its three-dimensional (3D) thermo-hydraulic and neutron dynamics modeling capabilities, which can achieve more accurate 3D reactor construction. It can meet the real-time simulation requirements of the simulator. It can be adopted as a simulation program for the thermal-hydraulic system of the reactor simulator.

Source Domain Data Set

The source domain data set D_s is acquired from the CLEAR-I operation data. In the study, 10 typical operation scenarios are included which contain one scenario under standard steady scenario [100% rate full power (RFP)], two power step scenarios (from 100 to 120% and 150% RFP), a severe accident scenario (loss of coolant with a small break), and six scenarios of rotating machinery of component faults in the fan or

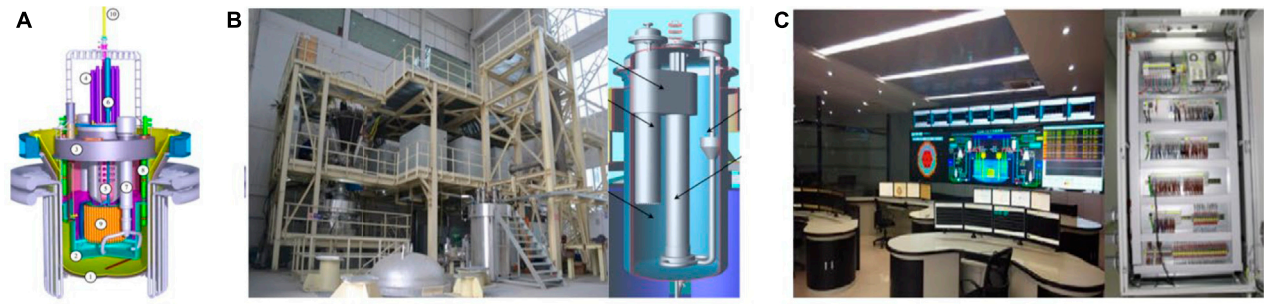


FIGURE 4 | Research platform. **(A)** The designed schematic diagram of China LEAd-based Reactor (CLEAR)-I. **(B)** The experiment and designed schematic of CLEAR-S. **(C)** Control room.

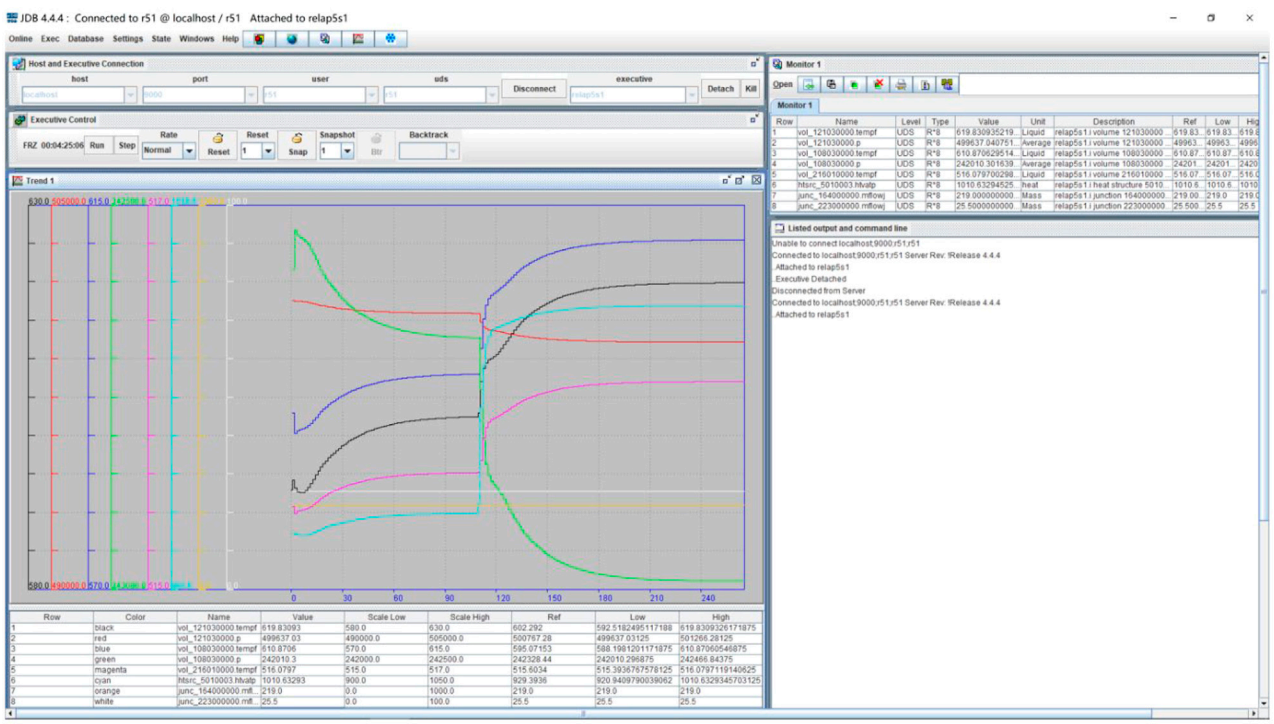


FIGURE 5 | RELAP5–HD communication interface in research platform.

a pump. **Table 2** shows the detailed information about each operation scenario in D_s.

Each scenario starts instantaneously except for scenario 1 (steady-state). The corresponding transient operation or fault is introduced from t = 0 to track the development trend. The sampling time set in the program is 0.25 s, and the sample length is 200, i.e., a total of 50 s. Each scenario contains 100 data samples, 50% of which are obtained through data enhancement. The method adopted is the sliding window method introduced in a previous work (Yao et al., 2020b). Each sample consists of 316 monitoring points from different components. The 316 monitoring points are distributed in 17 key node parameters, which are detailed as follows:

Core (5): Reactivity, power, control rod position, core temperature, and flow rate.

Steam Generator (6): SG primary side temperature, flow rate, and pressure; SG secondary side temperature, flow rate, and pressure.

Pump (2): Main pump flow rate and secondary pump flow rate.

Fan (4): Primary temperature, flow rate, and secondary temperature and flow rate.

TABLE 2 | China LEAd-based Reactor-I source domain data set.

Data set type	Scenario no	Scenario description	Data size
D_s	1	Steady-state with 100% rated full power	200 * 316 *100
	2	20% increasing power steps from steady-state	200 * 316 *100
	3	50% increasing power steps from steady-state	200 * 316 *100
	4	LOCA with the small break	200 * 316 *100
	5	One of the main pump rotor seizures	200 * 316 *100
	6	Two main pump rotor seizure	200 * 316 *100
	7	One of the feed pump rotor seizures	200 * 316 *100
	8	Two feed pump rotor seizure	200 * 316 *100
	9	Fan rotor seizures	200 * 316 *100
	10	Fan speed decreases by 10%	200 * 316 *100
	11	Fan speed decreases by 30%	200 * 316 *100

LOCA, loss of coolant accident.

TABLE 3 | China LEAd-based Reactor (CLEAR)-I source domain data set.

Data set type	Reactor type	Scenario label	Scenario description	Data size
D_{t1}	CLEAR-I	1	Steady state with 120% rated full power	200 * 316 * 80
		2	20% increasing power steps from steady state	200 * 316 * 80
		3	LOCA with the small break	200 * 316 * 80
		4	One primary pump rotor seizure	200 * 316 * 80
		5	Double primary pump rotor seizures	200 * 316 * 80
		6	One secondary pump rotor seizure	200 * 316 * 80
		7	Double secondary pump rotor seizures	200 * 316 * 80
		8	Fan rotor seizures	200 * 316 * 80
		9	Fan speed decreases by 30%	200 * 316 * 80
D_{t2}	CLEAR-I	1	Steady state with 80% or 120% rated full power	200 * 316 * 40&40
		2	20% increasing power steps from steady state	200 * 316 * 40&40
		3	LOCA with the small break	200 * 316 * 40&40
		4	One primary pump rotor seizure	200 * 316 * 40&40
		5	Double primary pump rotor seizures	200 * 316 * 40&40
		6	One secondary pump rotor seizure	200 * 316 * 40&40
		7	Double secondary pump rotor seizures	200 * 316 * 40&40
		8	Fan rotor seizures	200 * 316 * 40&40
		9	Fan speed decreases by 30%	200 * 316 * 40&40

LOCA, loss of coolant accident.

Target Domain Data Sets

The target domain data sets in this study are constructed according to specific transfer tasks. Two different target data sets for the cross-condition and cross-facility mission are built in the case study to verify the validity of the proposed model-based transfer method.

The target domain data sets D_{t1} and D_{t2} are settled to evaluate the ability of the network's transfer and generalization capability under different steady states. D_{t1} shares the same reactor type (CLEAR-I) as D_s but operates at a much higher steady-state power (120% RFP). The sample size in each failure case is 30% of the source domain data set, but the data sample's length is the same. D_{t2}, based on D_{t1}, improves the difficulty of cross-operating conditions transfer, including 80 and 120% RFP steady-state operating conditions of the reactor. Because the monitored parameters will be changed with different operating conditions under accident, the diagnosis task for D_{t2} will be more challenging than that of D_{t1}. To ensure the comparability of experimental results, the total number of samples between two target data sets is unified. **Table 3** shows details of the cross-condition target domain data set.

D_{t3} and D_{t4} are constructed for cross-facility transfer tasks built from a completely different reactor (CLEAR-S), as is shown in **Table 4**. The steady-state and power step operation data were derived from the experimental data on the existing facility. Compared with cross-condition data sets, cross-facility data sets differ significantly from source domain data in distribution characteristics, improving transfer complexity. The purpose of building this type of target domain data set is to explore the ability of the proposed transfer strategy between different facilities. It is worth noting that there are differences in structure between CLEAR-S and CLEAR-I nodes (CLEAR-I nodes are more precisely divided). To ensure that the cross-facility data set is dimensionally the same as the source data, we use zero padding for the default nodes to unify the data dimension, avoiding data heterogeneity.

To be similar to the procedure in the cross-condition transfer task, when compared with D_{t3}, D_{t4} is replaced with the data in

Table 4 | Cross-facility target domain data set.

Data set type	Reactor type	Scenario label	Scenario description	Data size
D_{t3}	CLEAR-S	1	Steady state with 100% rated full power	200 * 316 * 40
		2	20% increasing power steps from steady state	200 * 316 * 40
		3	50% increasing power steps from steady state	200 * 316 * 40
		4	LOCA with the small break	200 * 316 * 40
		5	Main pump rotor seizures	200 * 316 * 40
		6	Secondary pump rotor seizures	200 * 316 * 40
		7	Fan rotor seizures	200 * 316 * 40
		8	Fan speed decreases by 10%	200 * 316 * 40
		9	Fan speed decreases by 30%	200 * 316 * 40
D_{t4}	CLEAR-S	1	Steady state with 100% or 120% rated full power	200 * 316 * 20&20
		2	20% increasing power steps from steady state	200 * 316 * 20&20
		3	50% increasing power steps from steady state	200 * 316 * 20&20
		4	LOCA with the small break	200 * 316 * 20&20
		5	Main pump rotor seizures	200 * 316 * 20&20
		6	Secondary pump rotor seizures	200 * 316 * 20&20
		7	Fan rotor seizures	200 * 316 * 20&20
		8	Fan speed decreases by 10%	200 * 316 * 20&20
		9	Fan speed decreases by 30%	200 * 316 * 20&20

LOCA, loss of coolant accident.

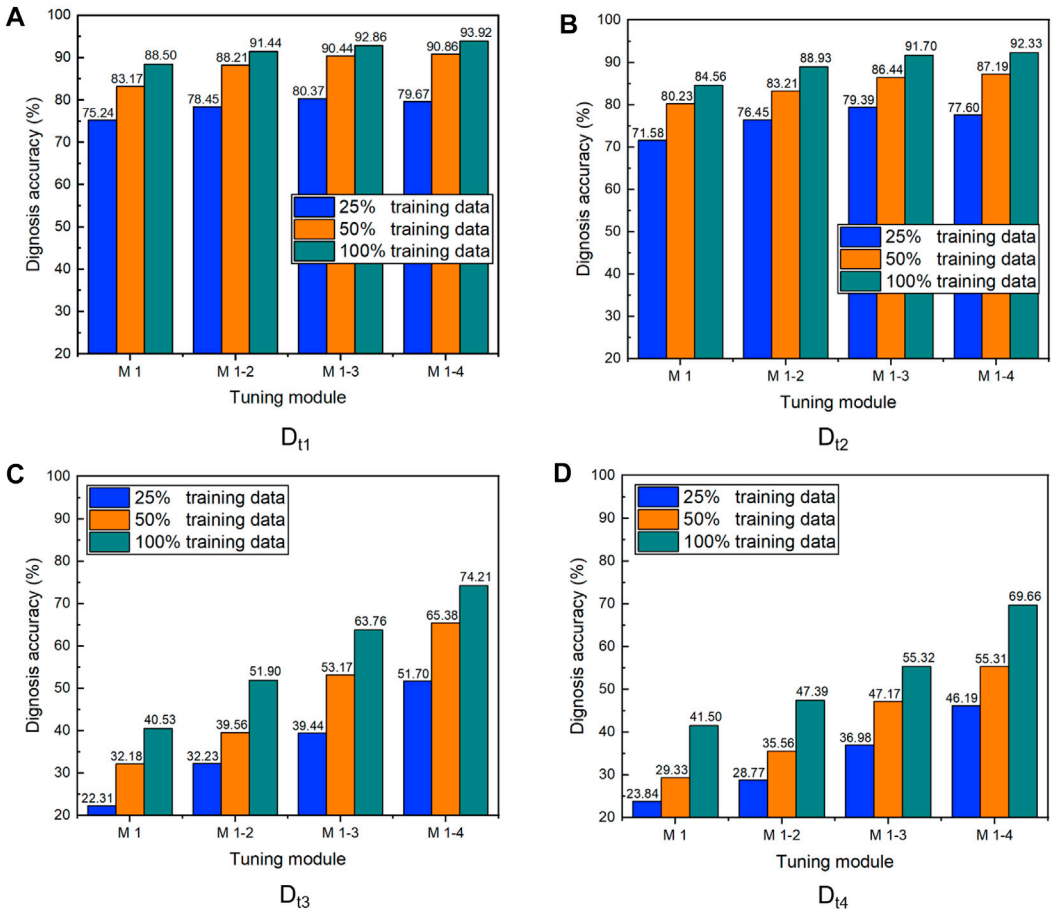


FIGURE 6 | Variations of diagnostic accuracy using the freezing-and-tuning strategy with different training data ratios and tuning modules (**(A)** Dt1; **(B)** Dt2; **(C)** Dt3; **(D)** Dt4).

120% operating conditions to increase the transfer complexity. It is more challenging to transfer the model under different steady-state operating conditions simultaneously under cross-facility conditions.

RESULTS AND DISCUSSION

To highlight the effectiveness of the proposed transfer model, we discussed and analyzed the screening results of model layers. The comparison results between transfer and non-transfer models. We also considered the final training and optimization results for cross-condition and cross-facility transfer tasks. The training procedure of the proposed transfer model was implemented in Keras v2.2.4 on Microsoft Windows 10 operating system based on an Intel Core i7-10750 2.6 GHz CPU with 64 GB RAM and accelerated by the Cuda v11.1 environment with NVIDIA RTX 3070 GPU.

Freezing and Fine-Tuning Result

To discuss the influence of the ratio of training samples on the transfer model's freezing-and-tuning process, we further divide D_{t1} to D_{t4} into three parts, which include 25, 50, and 100% training samples. Meanwhile, the training and test samples' ratio is the same as the pretraining model in the source domain, set as 4:1. All training and testing sites were cross-validated by a percentage of five folds to ensure accuracy. Meanwhile, all the test results are averaged 10 times.

Figures 6A,B show the influence of different training data ratios and tuning modules on the diagnosis accuracy in cross-working condition data sets (D_{t1} to D_{t2}). The horizontal axis shows the included tuning modules. For example, M 1 represents that only the last fully connected layer is tuned. At the same time, M 1–4 illustrate that all modules from 1 to 4 are adjusted layer by layer. The vertical axis shows the diagnostic accuracy of the target domain test data set.

It can be concluded that the diagnostic accuracy of all transfer models aiming at D_{t1} and D_{t2} can reach more than 70% under the most extreme condition of the 25% training samples when only the latter fully connected layers in the model are fine-tuned. These results indicate that the designed model transfer scheme is feasible for cross-condition target domain data sets. On the other hand, the diagnostic accuracy is also improved in the increasing proportion of fine-tuning modules after transfer. Significant improvement has been made in learning the lowest level convolutional pooling module with the M 1–2 strategy. However, the effect of this promotion is gradually decreasing, indicating that the low-level features of the training process of the transfer model have universal value. By contrast, the transfer of high-level features is more complex and abstract. It is challenging to obtain ideal training effects.

When all modules are fine-tuned, the diagnosis accuracy only improves 0.66% compared to that of the M 1–3 strategy under the condition of 100% training samples. To make matters worse, the diagnostic performance of the model decreased by 1.79% under the condition of 25% training samples. According to the structure of the model in the *Proposed Method*, Part B, it can be seen that the large size of the convolution and pooling layer makes it impossible to train a large number of neurons for parameter optimization and update when the number of samples is small. Therefore, over-fitting problems occurred in model training, introducing the decrease in diagnostic performance.

However, we do not need to be pessimistic because this situation will be improved as the number of training samples increases. Therefore, it is often necessary to fix the weight of the underlying parameters rather than fine-tune all model parameters in the whole transfer procedure.

Figures 6C,D show the results of variations of diagnostic accuracy in cross-facility data sets (D_{t3} to D_{t4}). It can be found that when all the training data are used for training and the number of tuning layers is more remarkable than three, the diagnostic accuracy of the target domain sample can reach an ideal result, which is more than 75%. By comparing the data of different transfer modules, it can be found that the freezing and tuning transfer strategies significantly improve the diagnosis accuracy under the cross-facility transfer task.

It shows that the high-dimensional features of the source domain in the transferred layer are beneficial to the generalization of the model in the target domain. In addition, the improvement of diagnostic accuracy of data samples is more significant, indicating that for cross-facility condition data with apparent differences in distribution, sufficient trainable samples are more important for advancing model performance.

Meanwhile, **Table 4** shows the influence of the amount of training data and the number of transferred layers on the training time of the model. It can be found that the increase in the number of tuning layers will significantly prolong the model training time. Compared with M 1–3, tuning all layers in the model (M 1–4) improved 97.36, 86.06, and 77.25% in three different training data levels. The above results mean that tuning a high level with high dimensional characteristics increases model complexity. However, using more training data does not significantly increase the training time of the model compared with tuning more layers. Combined with the above discussion results related to **Figure 6**, we finally selected M 1–3 as the transferred tuning model structure compared with the benchmark CNN in the following discussions.

Compared With the Benchmark Convolutional Neural Network Model

In this section, we verify the effectiveness of the proposed method by comparing the accuracy and loss changes of different models. The benchmark CNN model and the pretraining model are identical in structure to ensure the fairness of comparison results, as is shown in **Table 1**. The former directly uses the insufficient data of the target domain for direct training. By contrast, the latter uses the M1-3 structure for different transfers to perform full pretraining on the CLEAR-I source domain data and then migrates the source domain knowledge to the target domain network.

Figure 7 shows the variations of accuracy and loss of the benchmark CNN model and the proposed transfer model in the training process of 100 epochs for different transfer target data sets. According to the cross-condition target data set results,

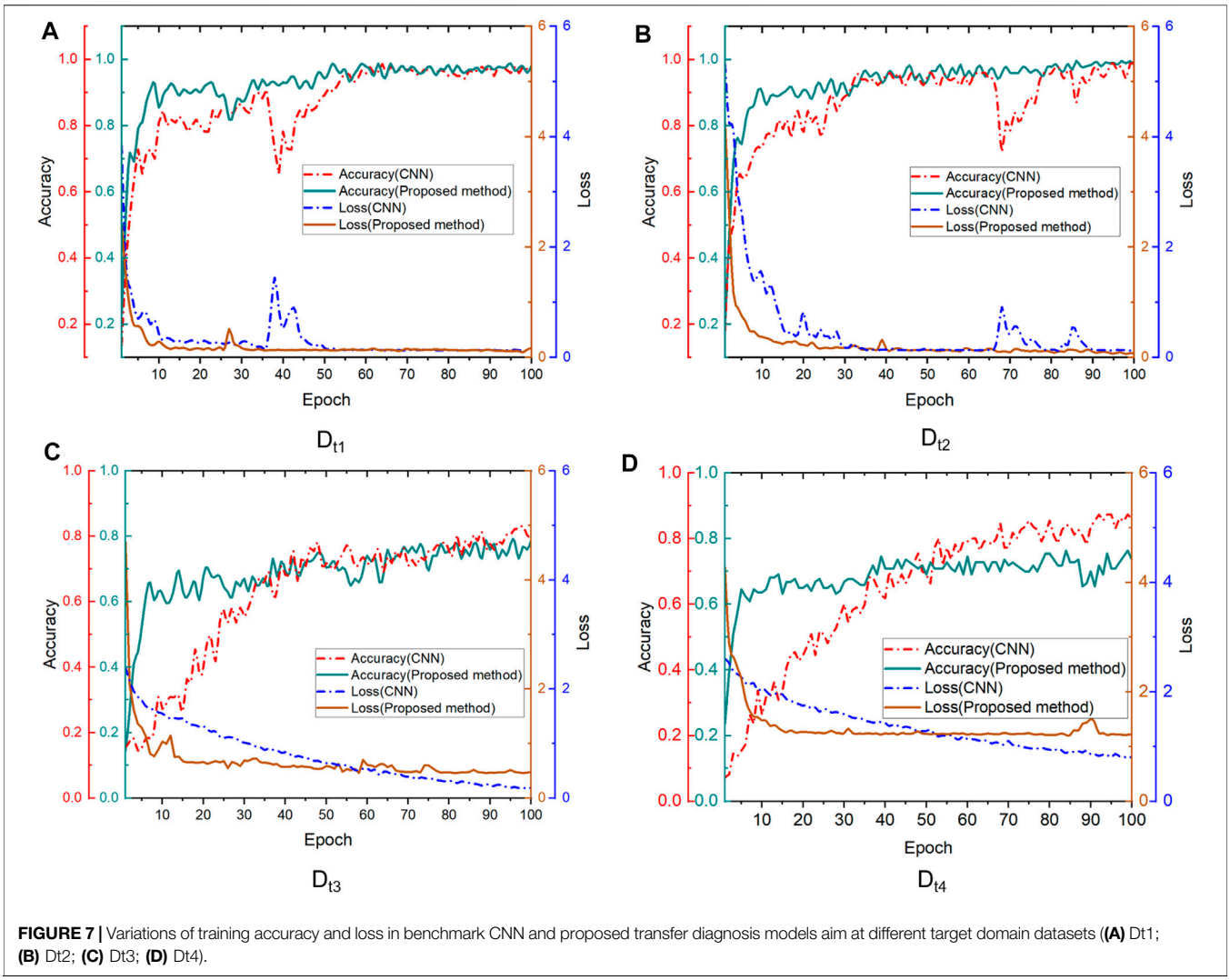


FIGURE 7 | Variations of training accuracy and loss in benchmark CNN and proposed transfer diagnosis models aim at different target domain datasets (**(A)** Dt1; **(B)** Dt2; **(C)** Dt3; **(D)** Dt4).

which are shown in **Figures 7A,B**, both models get 98.5% accuracy after 100 epochs training. The training loss in the proposed transfer model gradually approaches a fixed value. It remains stable after 50 epochs, but the similar target losses of the benchmark CNN model for D_{t1} and D_{t2} gradually stabilized after 60 and 90 epochs, respectively.

The results indicate that the CNN model is more prone to over-fitting for sample data in target domains, thus falling into local optimal solutions. The proposed method adopts the transfer strategy based on the pretraining model, effectively reducing the network's dependence on training parameters and sample numbers. On the other hand, it makes the network parameters establish a better initial value in the searched parameter space, conducive to faster training and convergence of the model.

Figures 7C,D show the training results of the transfer model for cross-facility data sets. It can be found that compared with the benchmark CNN model, the training loss of the transfer model decreases faster in the initial stage, indicating that source domain knowledge plays a specific role in the transfer process. However, with the epoch increases in the later period, training loss was not further reduced, mainly because of the significant difference in sample distribution among different devices. The prior knowledge obtained from the source domain data could not be further generalized to target domain data set to improve diagnostic accuracy. As the CNN model is retrained, its accuracy could be further improved with the epoch increase, which is better than the transfer model after 100 epochs. However, the ideal training effect cannot be achieved due to insufficient samples in the cross-facility target domain.

TABLE 5 | Cross-facility target domain data set.

Tuning module	Training time cost/epoch (s)		
	25% training data	50% training data	100% training data
M 1	0.22	0.44	0.53
M 1–2	0.69	0.83	1.21
M 1–3	1.14	1.65	2.11
M 1–4	2.25	3.07	3.74

TABLE 6 | Cross-facility target domain data set.

Data set	Total training time (s)			Diagnosis accuracy (%)		
	CNN	Proposed method	Proposed method-50	CNN	Proposed method	Proposed method-50
D_{t1}	314.6	147.2	80.3	93.9	94.1	88.6
D_{t2}	319.7	145.6	74.6	89.5	93.2	90.8
D_{t3}	201.7	109.7	59.0	71.2	63.7	59.2
D_{t4}	203.6	107.5	58.2	73.5	57.3	52.2

CNN, convolutional neural network.

Since the proposed strategy can achieve good generalization performance at the initial training stage, we discuss the proposed method with the benchmark CNN's training time and diagnostic accuracy after only 50 epochs. At the same time, we also give the results after 100 epochs, as is shown in **Table 5**.

It can be found that compared with the benchmark CNN model, after training with 100-epoch, the proposed transfer strategy saves 53.21% on average in the overall training time for the target domain data sets. Similar training accuracy is achieved on the cross-condition data sets (D_{t1} and D_{t2}). What is more noteworthy is that after 50 training sessions, the test and diagnosis accuracy of the proposed method for D_{t1} and D_{t2} is close to 90%, which achieves a high model training effect in less time, proving the effectiveness of the proposed transfer strategy. However, the transfer strategy has not achieved satisfactory results in cross-facility data sets (D_{t3} and D_{t4}). Although the accuracy was significantly improved initially, it could not be further improved due to the difference in data distribution, which proved that the existing high-dimensional feature extractor did not realize its function in the target domain samples.

CONCLUSION

In this article, we proposed an FDD method based on the optimized transferable CNN model. The priority knowledge and proposed fine-tuning strategy improved the diagnosis performance of the pretrained model aiming at a new target domain data set. It saved 53.21% of the training time compared with the benchmark CNN model. In addition, acceptable training accuracy could be achieved no more than 50-epoch training, proving that the proposed method has good generalization performance and timeliness.

On the other hand, although the proposed transfer strategy could not achieve ideal diagnostic accuracy for the cross-facility diagnosis task, the model performance increased at the initial stage of training. It indicates that the training model with characteristic information in the source domain data set provided a specific help. However, the data distribution difference is too big between the two data sets. The available training data are limited, leading to worse diagnosis results. Similarly, the CNN model could not obtain ideal training results when the data were missing.

In general, the proposed method ideally solved the problem of cross-condition transfer. Besides, collecting fault data of different domains at the initial stage is time-consuming and essential, which is also to prepare for future TL-related tasks. We will optimize the diagnostic performance of the deep TL model to resolve data distribution differences. In addition, the transferred non–deep learning method can be equally valuable, which will be further discussed and attempted.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

CRediT authorship contribution statement: YY: methodology, analysis, software, writing—original draft. DG: conceptualization, writing—review and editing, funding acquisition. JY: methodology, supervision, project administration. MX: methodology, supervision, funding acquisition.

FUNDING

This work is supported by the Anhui Foreign Science and Technology Cooperation Project- Intelligent Fault Diagnosis in Nuclear Power Plants (No. 201904b11020046) and China's National Key R&D Program (No.2018YFB1900301) and the National Natural Science Foundation of China (No.71901203, 71971181). This work is also funded by Research Grant Council of Hong Kong (11203519 and 11200621), Hong Kong ITC (InnoHK Project CIMDA) and HKIDS (Project 9360163).

REFERENCES

Correa-Jullian, C., Cardemil, J. M., López Droguett, E., and Behzad, M. (2020). Assessment of Deep Learning Techniques for Prognosis of Solar thermal Systems. *Renew. Energ.* 145, 2178–2191. doi:10.1016/j.renene.2019.07.100

Guo, Y., Tan, Z., Chen, H., Li, G., Wang, J., Huang, R., et al. (2018). Deep Learning-Based Fault Diagnosis of Variable Refrigerant Flow Air-Conditioning System for Building Energy Saving. *Appl. Energ.* 225, 732–745. doi:10.1016/j.apenergy.2018.05.075

Kingma, D. P., and Ba, J. (2014). *Adam: A Method for Stochastic Optimization*. arXiv preprint arXiv:1412.6980.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *nature* 521 (7553), 436–444. doi:10.1038/nature14539

Lee, J., Kao, H.-A., and Yang, S. (2014). Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. *Proced. Cirp* 16, 3–8. doi:10.1016/j.procir.2014.02.001

Li, W., Wu, X., Zhang, D., Su, G., Tian, W., and Qiu, S. (2014). Preliminary Study of Coupling CFD Code FLUENT and System Code RELAP5. *Ann. Nucl. Energ.* 73, 96–107. doi:10.1016/j.anucene.2014.06.042

Ma, J., and Jiang, J. (2011). Applications of Fault Detection and Diagnosis Methods in Nuclear Power Plants: A Review. *Prog. Nucl. Energ.* 53 (3), 255–266. doi:10.1016/j.pnucene.2010.12.001

Mandal, S., Santhi, B., Sridhar, S., Vinolia, K., and Swaminathan, P. (2017). Nuclear Power Plant Thermocouple Sensor-Fault Detection and Classification Using Deep Learning and Generalized Likelihood Ratio Test. *IEEE Trans. Nucl. Sci.* 64 (6), 1526–1534. doi:10.1109/tns.2017.2697919

Matteo, F., Carlo, G., Federico, P., and Enrico, Z. (2021). Time-dependent Reliability Analysis of the Reactor Building of a Nuclear Power Plant for Accounting of its Aging and Degradation. *Reliability Eng. Syst. Saf.* 205, 107173. doi:10.1016/j.ress.2020.107173

Peng, B.-S., Xia, H., Liu, Y.-K., Yang, B., Guo, D., and Zhu, S.-M. (2018). Research on Intelligent Fault Diagnosis Method for Nuclear Power Plant Based on Correlation Analysis and Deep Belief Network. *Prog. Nucl. Energ.* 108, 419–427. doi:10.1016/j.pnucene.2018.06.003

Peng, M., and Wang, Q. (2018). False Alarm Reducing in PCA Method for Sensor Fault Detection in a Nuclear Power Plant. *Ann. Nucl. Energ.* 118, 131–139.

Perrault, D. (2019). Nuclear Safety Aspects on the Road towards Fusion Energy. *Fusion Eng. Des.* 146, 130–134. doi:10.1016/j.fusengdes.2018.11.053

Ruder, S., Peters, M. E., Swayamdipta, S., and Wolf, T. (2019). "Transfer Learning in Natural Language Processing," in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorials, 15–18. June.

Saeed, H. A., Peng, M.-j., Wang, H., and Zhang, B.-w. (2020). Novel Fault Diagnosis Scheme Utilizing Deep Learning Networks. *Prog. Nucl. Energ.* 118, 103066. doi:10.1016/j.pnucene.2019.103066

Seabra, J. C., Costa, M. A., and Lucena, M. M. (2016). "September). IoT Based Intelligent System for Fault Detection and Diagnosis in Domestic Appliances," in 2016 IEEE 6th International Conference on Consumer Electronics-Berlin (ICCE-Berlin) (IEEE), 205–208.

Tang, Q., Chai, Y., Qu, J., and Ren, H. (2018). Fisher Discriminative Sparse Representation Based on DBN for Fault Diagnosis of Complex System. *Appl. Sci.* 8 (5), 795. doi:10.3390/app8050795

Wahlström, B. (2018). Systemic Thinking in Support of Safety Management in Nuclear Power Plants. *Saf. Sci.* 109, 201–218. doi:10.1016/j.ssci.2018.06.001

Wang, C., Vo, H. T., and Ni, P. (2015). "December)An IoT Application for Fault Diagnosis and Prediction," in 2015 IEEE International Conference on Data Science and Data Intensive Systems (IEEE), 726–731.

Wang, J., and Zhang, C. (2018). Software Reliability Prediction Using a Deep Learning Model Based on the RNN Encoder-Decoder. *Reliability Eng. Syst. Saf.* 170, 73–82. doi:10.1016/j.ress.2017.10.019

Wen, L., Li, X., Gao, L., and Zhang, Y. (2017). A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method. *IEEE Trans. Ind. Elect.* 65 (7), 5990–5998.

Wu, Y. (2016). CLEAR-S: an Integrated Non-nuclear Test Facility for China lead-based Research Reactor. *Int. J. Energ. Res.* 40 (14), 1951–1956. doi:10.1002/er.3569

Wu, Y. (2016). Design and R&D Progress of China Lead-Based Reactor for ADS Research Facility. *Engineering* 2 (1), 124–131. doi:10.1016/j.eng.2016.01.023

Xin, M., Jiao, W., and Da-zi, L. (2019). "Fault Diagnosis of Nuclear Power Plant Based on Simplified Signed Directed Graph with Principal Component Analysis and Support Vector Machine," in *Proc. 2019 Chin* (Hangzhou, China: Autom. Congr.), 3082–3087. doi:10.1109/CAC48633.2019.8997001

Xu, B., Li, H., Pang, W., Chen, D., Tian, Y., Lei, X., et al. (2019). Bayesian Network Approach to Fault Diagnosis of a Hydroelectric Generation System. *Energy Sci Eng* 7, 1669–1677. doi:10.1002/ese3.383

Xu, Z., Li, C., and Yang, Y. (2020). Fault Diagnosis of Rolling Bearing of Wind Turbines Based on the Variational Mode Decomposition and Deep Convolutional Neural Networks. *Appl. Soft Comput.* 95, 106515. doi:10.1016/j.asoc.2020.106515

Yangping, Z., Bingquan, Z., and DongXin, W. (2000). Application of Genetic Algorithms to Fault Diagnosis in Nuclear Power Plants. *Reliability Eng. Syst. Saf.* 67 (2), 153–160. doi:10.1016/s0951-8320(99)00061-7

Yao, Y., Wang, J., Ge, D., and Xie, M. (2021). "October)Intelligent PHM Based Auxiliary Decision Framework of Advanced Modular Nuclear," in *2021 Global Reliability and Prognostics and Health Management (PHM-Nanjing)* (IEEE), 1–5.

Yao, Y., Wang, J., Long, P., Xie, M., and Wang, J. (2020). Small-batch-size Convolutional Neural Network Based Fault Diagnosis System for Nuclear Energy Production Safety with Big-data Environment. *Int. J. Energ. Res* 44 (7), 5841–5855. doi:10.1002/er.5348

Yao, Y., Wang, J., and Xie, M. (2022). Adaptive Residual CNN-Based Fault Detection and Diagnosis System of Small Modular Reactors. *Appl. Soft Comput.* 114, 108064. doi:10.1016/j.asoc.2021.108064

Yao, Y., Wang, J., Xie, M., Hu, L., and Wang, J. (2020). A New Approach for Fault Diagnosis with Full-Scope Simulator Based on State Information Imaging in Nuclear Power Plant. *Ann. Nucl. Energ.* 141, 107274. doi:10.1016/j.anucene.2019.107274

Yoo, K. H., Back, J. H., Na, M. G., Hur, S., and Kim, H. (2018). Smart Support System for Diagnosing Severe Accidents in Nuclear Power Plants. *Nucl. Eng. Tech.* 50 (4), 562–569. doi:10.1016/j.net.2018.03.007

Zhao, X., Kim, J., Warns, K., Wang, X., Ramuhalli, P., Cetiner, S., and Golay, M. (2021). Prognostics and Health Management in Nuclear Power Plants: An Updated Method-Centric Review with Special Focus on Data-Driven Methods. *Front. Energ. Res.* 9, 294. doi:10.3389/fenrg.2021.696785

Zheng, J., Dong, Z., Pan, H., Ni, Q., Liu, T., and Zhang, J. (2019). Composite Multi-Scale Weighted Permutation Entropy and Extreme Learning Machine Based Intelligent Fault Diagnosis for Rolling Bearing. *Measurement* 143, 69–80. doi:10.1016/j.measurement.2019.05.002

Zhong, M., Xue, T., and Ding, S. X. (2018). A Survey on Model-Based Fault Diagnosis for Linear Discrete Time-Varying Systems. *Neurocomputing* 306, 51–60. doi:10.1016/j.neucom.2018.04.037

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., et al. (2020). A Comprehensive Survey on Transfer Learning. *Proc. IEEE* 109 (1), 43–76.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Energy Research

ORIGINAL RESEARCH
published: 21 March 2022
doi: 10.3389/fenrg.2022.851848

An Improved Method for PWR Fuel Failure Detection Using Cascade-forward Neural Network With Decision Tree

Bing Dong[1], *Kang Yang*[2], *Wei Zhang*[2], *Junlian Yin*[1] *and Dezhong Wang*[1]*

[1]School of Nuclear Science and Engineering, Shanghai Jiao Tong University, Shanghai, China, [2]China Nuclear Power Engineering Company, Beijing, China

When a fuel rod is damaged, determining the degree of fuel failure makes sense. The operators can decide whether to continue operating the reactor or shut it down based on the severity of the fuel failure. The isotopic ratio of two radioactive fission products (FPs) is a typical technique for evaluating the degree of fuel failure, although this is not applicable in the case of little fuel failure but large tramp uranium mass. The feedforward neural network (FFNN) has been used to identify fuel failures in order to overcome the shortcomings of the isotopic ratio method, although there is still inadequacy in the ability to distinguish between an intact fuel rod and a defective fuel rod with a small defect. In this study, we propose a cascade-forward neural network with a decision tree for fuel failure detection that performs well at classifying the degree of fuel failure and, in particular, at differentiating between an intact fuel rod and a defective fuel rod with a small size defect. The input of the neural network is the specific activity of FPs measured in the coolant. The degree of fuel failure is determined by the neural network's output, which is labeled using one-hot encoding. The training set is constructed using the Booth-type diffusion model and the first-order kinetic model. The performance of the improved neural network is demonstrated. It is shown that the improved method is more accurate and responsive than the previous neural network when recognizing the onset of fuel failure. Finally, the most important nuclides are determined through the sensitivity analysis, and the neural network is simplified according to the importance of nuclides and the limitation of the radioactive detector in practical application.

Keywords: nuclear fuel, cladding defect, fuel failure detection, neural network, decision tree, fission product

OPEN ACCESS

Edited by:
Xingang Zhao,
Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Wenzhong Zhou,
Sun Yat-sen University, China
Jiankai Yu,
Massachusetts Institute of Technology, United States
A. Abdelghafar Galahom,
Higher Technological Institute, Egypt
Serkan Akkoyun,
Cumhuriyet University, Turkey

***Correspondence:**
Dezhong Wang
dzwang@sjtu.edu.cn

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *10 January 2022*
Accepted: *17 February 2022*
Published: *21 March 2022*

Citation:
Dong B, Yang K, Zhang W, Yin J and Wang D (2022) An Improved Method for PWR Fuel Failure Detection Using Cascade-forward Neural Network With Decision Tree. Front. Energy Res. 10:851848. doi: 10.3389/fenrg.2022.851848

1 INTRODUCTION

The reactor generates a substantial amount of fission products (FPs) during operation. Under normal operating conditions, the fission products are contained within the fuel cladding which prevents them from escaping into the primary coolant and maintains the coolant's specific activity within the management limit. While the performance of the fuel rods keeps improving, the fuel cladding is inevitably defective during operation for a variety of reasons (Qin et al., 2020), including the following: 1) power ramp defects caused by stress corrosion cracking (SCC) or pellet-cladding interaction (PCI); 2) circumferential cracking caused by hydrogen embrittlement; 3) fabrication defects; and 4) fretting defects caused by interaction with the grid spacer or debris in the primary

coolant (Lewis et al., 2017). When the fuel cladding is defective, FPs can migrate from the fuel cladding gap to the primary coolant, considerably increasing the specific activity of the coolant. If the specific activities or defect sizes exceed the critical values, the reactor must shut down according to the regulation (Likhanskii et al., 2006; Qin et al., 2019). Otherwise, it may increase the risks of occupational exposure and harm the safety of reactor operation (Iqbal et al., 2007; Qin et al., 2020).

In order to determine whether the fuel cladding is defective, either a chemical sampling approach from the let-down flow or the real-time online monitoring system is used to collect specific activities in the primary coolant. By analyzing the specific activities of FPs derived by sampling or online monitoring, fuel rod failures are discovered by various methodologies, such as the release-to-birth ratio method (Zanker, 1989), fitted escape rate coefficient method (Yamamoto and Morishita, 2015), and isotopic ratio method (Kalinichev et al., 2018; Li et al., 2017) (Qin et al., 2016). Generally, these strategies are developed in consideration of the FP release model or reactor operation experience. Among these strategies, the isotopic ratio method is the one that is most frequently used to detect fuel failures during reactor operation. It is based on the ratios of specific activities of two isotopic FPs, which can be used to determine whether the fuel cladding is defective (Menéndez, 2009). In addition, it can identify the degree of fuel failure if the fuel cladding is defective (Li et al., 2017). The advantage of the isotopic ratio method is its simplicity, as it allows for straightforward determination of the specific activities of FPs by analyzing the spectrum obtained from sampling or online monitoring for the primary coolant. The feature of the isotopic ratio method is that it does not require the historical operation data of sampling or online monitoring. The status of the fuel rod can be assessed by the isotopic ratio at a certain time.

The isotopic ratio method has distinct disadvantages. When the defect size is small and the mass of tramp uranium is large, it has been demonstrated that the isotopic ratio increases slowly before reaching the threshold (Dong et al., 2019). Thus, the response of the isotopic ratio method is insensitive for detecting fuel failures. Even in some circumstances, the isotopic ratio method may fail under the influence of tramp uranium. In addition, there is no single standard for the isotopes and the threshold of the isotopic ratio method (Li and Yang, 2008; Menéndez, 2009; Li et al., 2017; Lyu and Xiong, 2019).

To address the limitations of the isotopic ratio method, the neural network–based method for fuel failure detection is developed (Likhanskii et al., 2006). Artificial neural networks (ANNs) are a highly effective technique for establishing a connection between input data and output data in multidimensional space. ANNs have been extensively used in the field of nuclear engineering, particularly in fault diagnosis. Andrews et al. (1999) use ANN models to predict the cesium release fraction from a water reactor under severe accident conditions and demonstrate that the models are capable of reproducing the relationships between the release fraction and time. Guo et al. (2019) apply the deep neural network to detect the damaged fuel assembly by processing the real-time image frame and point out that the deep neural network has more advantages over the traditional computer vision method. Wang et al. (2022) use the deep convolutional neural network (CNN) to detect the system-level fault in the nuclear power plant (NPP) and demonstrate that it is able to improve the NPP fault diagnosis. Ebrahimzadeh et al. (2022) use the feedforward neural network in detection and estimation of fault sensors in the NPP and prove that it has advantages over traditional methods. Zhang et al., (2021) develop a surface crack detection method for nuclear fuel pellets based on the CNN and show that the new method improves the performance of traditional machine vision inspection systems.

In this study, a cascade-forward neural network with a decision tree is proposed for fuel failure detection, which has a good performance to classify the degree of fuel failure, particularly in distinguishing the intact fuel rod and small-size defective fuel rod. The inputs of the neural network are the radioactivity of fission products measured in the coolant. The output of the neural network is the degree of fuel failure, which is labeled using one-hot encoding. The dataset is generated by the Booth-type diffusion model and the first-order kinetic model. The performance of the improved neural network is presented in the study.

The remaining sections of this manuscript are organized as follows. **Section 2** is dedicated to the improved method for fuel failure detection. **Section 3** shows the performance of the improved method and results of sensitivity analysis. Finally, **Section 4** concludes the study.

2 METHODOLOGY

2.1 Conceptual Framework

During normal reactor operation, the major source of specific activity in the coolant is tramp uranium, which has a relatively low level of specific activity. When the fuel cladding is defective, the FPs released from the gap in the fuel cladding dominate the specific activity, that is, in a rather high level. It is obvious that a large defect size will result in a high level of specific activity. As a result, the coolant's specific activity serves as a fingerprint for predicting the status of fuel failure. While it should be noted that if the defect size is very small and the tramp uranium mass is considerable, the tramp uranium contribution may overwhelm the defective fuel rod contribution.

The framework for the detection approach is depicted in **Figure 1**, which is divided into three parts. The blue parts illustrate how the training, validation, and test datasets are generated. Due to the extremely low probability of fuel failure and pursuit of zero fuel rod failure, there is no appropriate fuel failure dataset for training the network. Thus, the dataset is generated using the FP release model, which includes the Booth-type diffusion model and first-order kinetic model, as detailed in the previous study (Dong et al., 2019). The Booth-type diffusion and first-order kinetic models are proven to be appropriate for calculating fission gas release fraction (Lewis et al., 2017). The green parts denote the establishment of a system for detecting fuel failures. The system is trained and

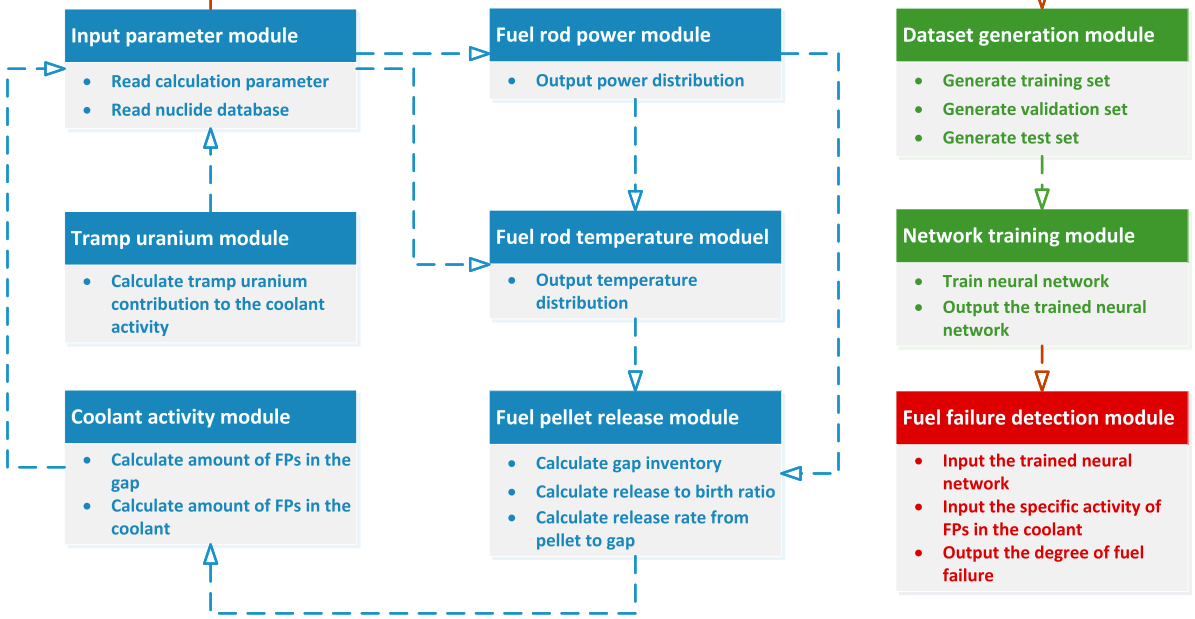


FIGURE 1 | Overview of the fuel failure detection method.

validated on the dataset generated by the blue-outlined technique. The red parts demonstrate how the system is being used to predict the status of a fuel failure.

2.2 Fission Product Release Model

2.2.1 Booth-type Diffusion Model

The FP release from the pellet to the gap is dominated by the diffusion process. Following the assumption from Booth, a one-dimensional diffusion equation is created considering the FP generation in the uranium grain due to fission, eliminated by the decay:

$$\frac{\partial C(r,t)}{\partial t} = \frac{D}{r}\frac{\partial^2 (rC(r,t))}{\partial r^2} - \lambda C(r,t) + B, \tag{1}$$

where C is the concentration of fission products in the uranium grain (m^{-3}); D is the diffusion coefficient of fission products (m^2s^{-1}); λ is the decay constant of the fission products (s^{-1}); and B is the production rate of the fission products $m^{-3}s^{-1}$.

Solving **Eq. 1**, the release-to-birth ratio of FP release can be derived:

$$\frac{R}{B} = \frac{3}{a}\sqrt{\frac{D}{\lambda}}, \tag{2}$$

where a is the radius of the uranium grain.

2.2.2 First-Order Kinetic Model

The FPs in the gap and coolant can be treated as a first-order rate process in which the degree of fuel failure is characterized by the escape rate coefficient ν_i. Then, the equilibrium equations of fission products can be established:

$$\begin{aligned}
\frac{dN_{gi}}{dt} &= R_i + \sigma_j \phi N_{gj} + \sum_k f_{ik}\lambda_k N_{gk} - (\nu_i + \lambda_i + \sigma_i \phi) N_{gi} \\
\frac{dN_{ci}}{dt} &= \nu_i N_{gi} + \tau\sigma_j \phi N_{cj} + \sum_k f_{ik}\lambda_k N_{ck} - \left(\lambda_i + \frac{Q}{W}\eta_i + \beta + \tau\sigma_i\phi + \frac{L}{W}\right) N_{ci}, \\
\frac{dN_{ci}^{\text{tramp}}}{dt} &= R_i^{\text{tramp}} - \left(\lambda_i + \frac{Q}{W}\eta_i + \beta + \tau\sigma_i\phi + \frac{L}{W}\right) N_{ci}^{\text{tramp}}
\end{aligned} \tag{3}$$

where i, j, and k denote the isotopes; N_g is the number of FPs in the gap; N_c is the number of FPs in the coolant; R_i is the FP release rate from the pellet to gap, which can be derived from **Eq. 2**; $\sigma_j \phi N$ is the FP transmutation; $f_{ik}\lambda_k N$ is the FP decay and f_{ik} is the branching ratio; ν_i is the escape rate coefficient from the gap to coolant; $Q/W\eta_i$ is the FP elimination due to purification; β is the FP elimination due to boron control; $\tau\sigma_i\phi$ is the FP elimination due to neutron absorption; and L/W is the FP elimination due to coolant leakage; Among them, Q is the let-down flow rate (kg/s), L is the leakage flow rate (kg/s), η_i is the purification efficiency, and W is the total coolant mass (kg); tramp denotes the contribution of tramp uranium.

2.3 Structure of the Network

In this study, an improved method for fuel failure detection is proposed based on the cascade-forward neural network (CFNN). The CFNN is a type of feedforward neural network, in which there is a direct connection between every two layers of the network. It means that there is an additional connection between the input layer and output layer for a regular three-layer neural network. The advantage of the CFNN is that it accommodates the nonlinear relationship between the input and output by not eliminating the linear relationship between the two (Warsito et al., 2018). Thus, it is suitable for establishing a relationship

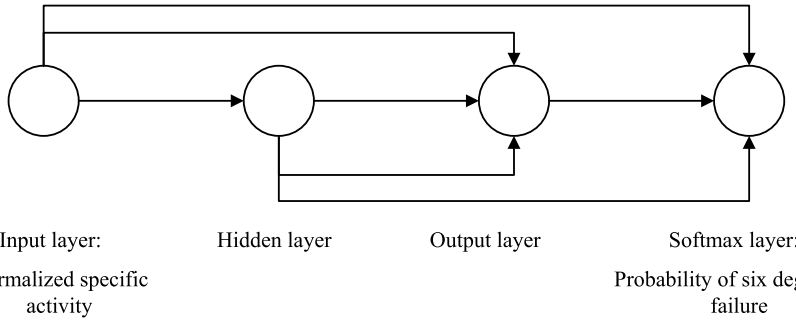


FIGURE 2 | Structure of the CFNN used for fuel failure detection.

TABLE 1 | Typical samples of the input and output of the neural network.

Specific activity (Bq/g)	Normalized specific activity (input)	Probability vector (output)	Degree of failure
\{1.4 357e+01, 4.2 612e-02, 4.4 207e+01, 8.7 714e+01, 1.1 916e+02, 7.0 106e-03, 8.0 954e+01, 3.8 272e+00, 1.1 828e-09, 4.7 286e+00, 6.7 657e+01, 2.4 087e+02, 1.2 158e+02, 3.0 619e+02, 7.1 566e+00, 1.0 031e+02, 4.1 319e+01, 2.1 286e+02, 1.6 813e-07, 8.2 949e-05, 7.0 140e-03, 2.1 082e+02, 5.2 939e+00\}	\{0.089 0, 0.060 3, 0.077 9, 0.088 3, 0.083 5, 0.047 4, 0.098 3, 0.063 0, 0.053 1, 0.006 4, 0.043 5, 0.091 8, 0.076 1, 0.045 4, 0.057 6, 0.044 9, 0.088 1, 0.096 5, 0.038 2, 0.087 1, 0.053 1, 0.098 3, 0.098 0\}	\{4.9 257e-02, 8.9 995e-01, 5.0 790e-02, 0, 0, 0\}	2
\{1.4 453e+02, 2.1 697e-01, 4.4 500e+02, 8.8 301e+02, 1.1 996e+03, 6.9 479e-02, 8.1 495e+02, 3.8 526e+01, 1.1 871e-08, 4.5 788e+01, 6.8 098e+02, 2.4 248e+03, 1.2 239e+03, 1.9 201e+03, 6.1 294e+01, 1.0 097e+03, 4.1 595e+02, 2.1 428e+03, 1.6 831e-06, 8.3 504e-04, 7.0 395e-02, 2.1 224e+03, 5.3 292e+01\}	\{0.896 2, 0.306 9, 0.783 8, 0.888 5, 0.840 5, 0.469 3, 0.989 6, 0.634 0, 0.533 3, 0.061 7, 0.438 2, 0.923 8, 0.766 4, 0.284 6, 0.493 5, 0.452 4, 0.886 9, 0.971 1, 0.382 5, 0.877 2, 0.533 3, 0.989 4, 0.986 5\}	\{9.4 024e-02, 1.0 606e-01, 7.9 991e-01, 0, 0, 0\}	3

between the specific activity in the coolant and degree of fuel failure.

Figure 2 illustrates the structure of the CFNN utilized to identify fuel failures. The network's input is a vector containing the normalized specific activity of FPs in the coolant, and the network's output is the degree of fuel failure, which is encoded in one-hot form. The degree is divided into six categories: degree 1 denotes an unbroken fuel cladding, whereas degrees 2–6 denote a defective fuel cladding. Correspondingly, a high degree indicates a big defect size. After going through a Softmax layer, the value of the neuron represents the probability of that degree of failure occurring.

Perez and Hamawi (2017) show that typical elements in design basis source term determinations can be represented by Kr, Xe, Br, I, Rb, Cs, Sr, Ba, Mo, and Te. Among these elements, Kr, Xe, I, and Cs are most important FPs in fuel failure detection. The nuclides mainly considered in the study are Kr, Xe, I, and Cs, which have a relatively large fission yield and a relatively long half-life, including Kr-85, I-131, I-133, Xe-133, Xe-135, and Cs-137. Simultaneously, considering the decay chain and the representativeness of nuclides in each group, the generated dataset has specific activities of 23 nuclides: Br-83, Kr-85, Kr-85m, Kr-87, Kr-88, Sr-90, Te-131, Te-131m, I-129, I-131, I-133, I-134, I-135, Xe-133, Xe-133m, Xe-135, Xe-135m, Xe-138, Cs-134, Cs-134m, Cs-137, Cs-139, and Pr-143, which are used as the input of the neural network. Then, the neural networks using 23 nuclides as input are trained in this study. **Table 1** shows typical samples of the input and output of the neural network.

2.4 Development of the Neural Network

Prior to training the neural network, various parameters must be adjusted to make the neural network work well. The number of neurons in the hidden layer is obtained empirically using the following formula (Dong et al., 2019):

$$l = \sqrt{0.43nm + 0.12\ m^2 + 2.54n + 0.77\ m + 0.35} + 0.51, \quad (4)$$

where m, n, and l are the corresponding number of neurons in the output, input, and hidden layer, respectively.

Eq. 4 is proven to be an efficient formula to determine the number of neurons in the hidden layer. Since it is shown that the neural networks with the number of hidden layer neurons around the formula all have good performance, the number of hidden layer neurons is determined directly by the formula. The number of hidden layer neurons is 11 for an input of 23 nuclides.

As **Figure 2** shows, neurons pass their value to the next layers. Besides the final layer, an activation function, rectified linear unit (ReLu) function, is applied between the two layers to make the neural network present the nonlinear relationship. The ReLu function helps solve the gradient vanishing problem and increases training efficiency (Choi et al., 2021). The Softmax function is applied as the activation function in the final layer:

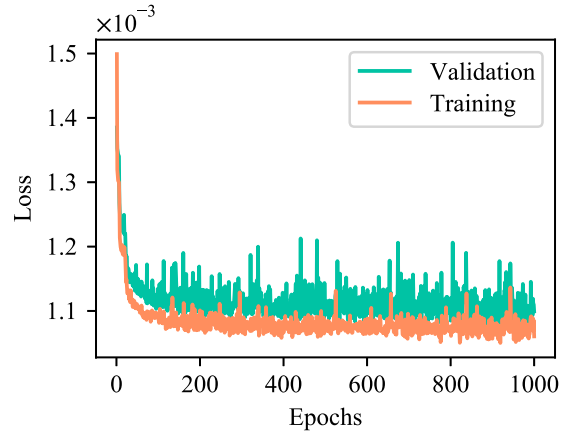


FIGURE 3 | Training and validation loss over epochs.

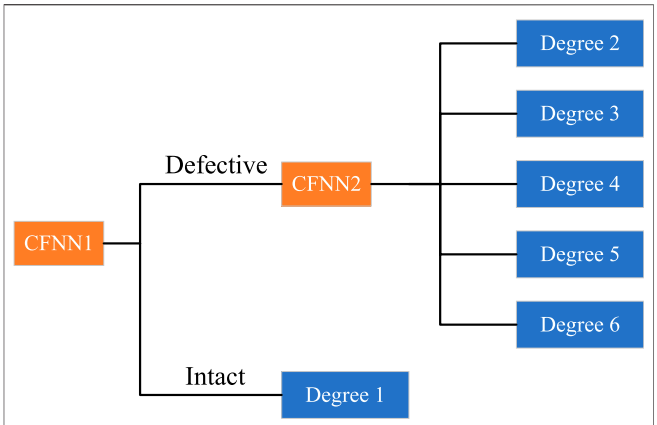


FIGURE 4 | Diagram of the CFNN with the decision tree designed for fuel failure detection.

$$y_i = \frac{e^{x_i}}{\sum_j e^{x_j}}, \tag{5}$$

where x_i and x_j indicate neurons in the previous layer.

After applying a Softmax function, the values of the neurons are normalized to [0,1], and the summation of each value of neurons in the final layer is 1. Then, the value of neurons can represent the probability of the corresponding degree of failure. The advantage of Softmax function is that it can improve the discrimination between the output neurons and hence increase learning efficiency.

Accompanying the Softmax function, cross-entropy is selected as the loss function (de Boer et al., 2005):

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j} y_j^{(i)} \log\left(\hat{y}_j^{(i)}\right), \tag{6}$$

where N is the total number of samples, i denotes the sample number, j denotes the neurons in the final layer, $y_j^{(i)}$ is the real value in the training sample, and $\hat{y}_j^{(i)}$ is the output value of the neural network. Using Softmax function and cross-entropy together will make the training process faster and more stable.

Adam is selected as the optimizer (Kingma and Ba, 2017), and the neural network is trained with a learning rate of 0.01 balancing the convergence and training speed. The loss of training and validation is shown in **Figure 3**. It can be seen that the loss of training and validation almost does not decrease after 1,000 epochs. In addition, aiming to avoid the discrepancy of single training, the neural network is trained 10 times individually and averaged to inference the degree of fuel failure.

2.5 Optimization for Small Defects

The dataset utilized to train the ANN is imbalanced, with the intact condition prevailing. As a consequence, the trained neural networks have a tendency to categorize degree 2 or 3 (small defects) as degree 1 (intact), which significantly underestimates the degree of fuel failure. To overcome this issue, a decision tree (DT) such as the classifier is developed, as shown in **Figure 4**.

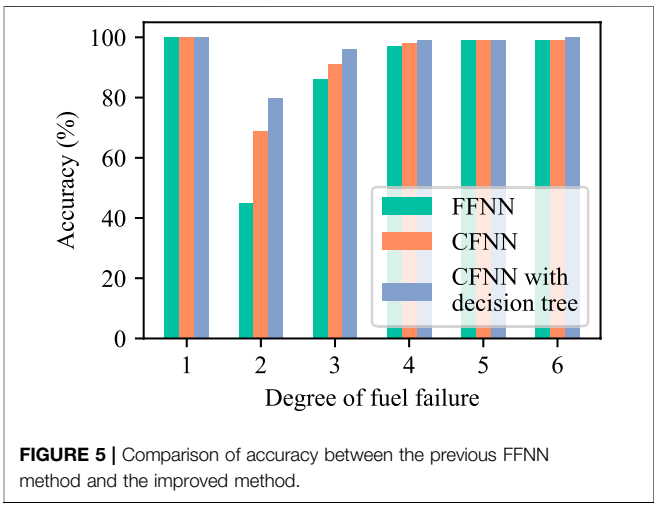


FIGURE 5 | Comparison of accuracy between the previous FFNN method and the improved method.

Two distinct forms of the CFNN are developed in the improved method. The CFNN1 is used for predicting whether the fuel cladding is defective. Its structure is the same as shown in **Figure 2**, except that the output layer contains just two neurons. If the output of the CFNN1 indicates that the fuel cladding is intact, the degree of fuel failure can be assessed to 1. Otherwise, the CFNN2 is used to predict the degree of fuel failure when the fuel cladding is defective. The output of the CFNN2 is the probability of the corresponding degree under defective condition.

2.6 Relative Importance of the Input of the Neural Network

Due to the fact that the neural network is a black-box used to make a connection between the input and output, it is incapable of explaining the mechanisms behind physical

TABLE 2 | Confusion matrix of neural network results for small defects in percentage.

(A) FFNN			
	Actual degree 1	Actual degree 2	Actual degree 3 (%)
Predicted degree 1	100%	55.0%	8.3
Predicted degree 2	0	45.0%	5.7
Predicted degree 3	0	0	86.0
(B) CFNN			
	Actual degree 1	Actual degree 2	Actual degree 3 (%)
Predicted degree 1	100%	30.6%	7.9
Predicted degree 2	0	69.4%	0.9
Predicted degree 3	0	0	91.2
(C) CFNN with DT			
	Actual degree 1	Actual degree 2	Actual degree 3 (%)
Predicted degree 1	100%	20.0%	3.7
Predicted degree 2	0	80.0%	3.1
Predicted degree 3	0	0	93.2

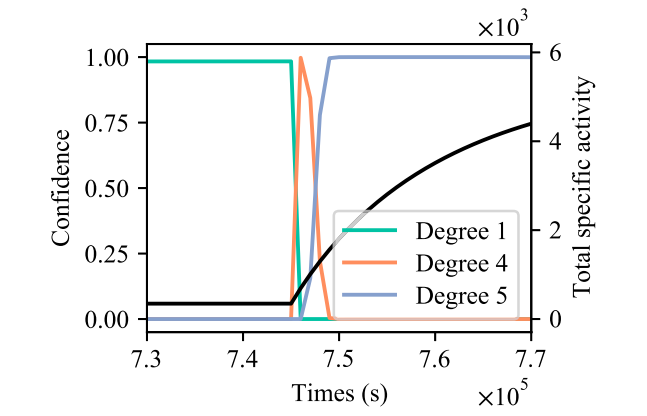


FIGURE 6 | Variation of total specific activity and neural network output after fuel failure of degree 5.

phenomena. While the output of the neural network depends on the magnitude of the weights between connections, the contribution of input variables can be determined by interpreting the weights of the trained neural network. This is meaningful information for reactor operators to select the appropriate FPs predicting the status of fuel failure.

A method to determine the relative importance of the input of the neural network is proposed first by Garson, (1991) and simplified by Gevrey et al., (2003). The method essentially involves partitioning the hidden-output connection weights of each hidden neuron into components associated with each input neuron.

For each hidden neuron, the absolute value of the input hidden layer connection weight is divided by the sum of the absolute value of the input hidden layer connection weight of all input neurons:

$$Q_{nl} = \frac{|W_{nl}|}{\sum_{n}|W_{nl}|}, \tag{7}$$

where W denotes the weight between neuron connections, n denotes the input layer, and l denotes the hidden layer.

Once the parameter Q_{nl} is determined, the sum of Q_{nl} for each hidden neuron is divided by the sum of each hidden neuron of the sum for each input neuron of Q_{nl}. The relative importance of all output weights attributable to the given input variable is then obtained. Then, the relative importance of all output weights for a given input variable can be obtained:

$$RI_n = \frac{\sum_{l} Q_{nl}}{\sum_{l}\sum_{n} Q_{nl}}. \tag{8}$$

Apart from the connection between the input and hidden layer, there are also connections between the input and other layers. Then, the relative importance should also include other connections and be averaged:

$$RI_n^{ave} = \frac{1}{3}\left(RI_n^{1\to 2} + RI_n^{1\to 3} + RI_n^{1\to 4}\right). \tag{9}$$

3 RESULTS AND DISCUSSION

3.1 Performance of the Method

3.1.1 Accuracy of the Method

The accuracy is calculated as the ratio between the number of correct predictions to the total number of predictions. For example, when a fuel cladding shows degree 1 failure, the prediction is correct if the output of the neural network indicates a degree 1 failure. Otherwise, the prediction is wrong. The predictions are made for all samples in the test set using the trained neural network.

It has been proven that the general FFNN predicts the degree of fuel failure and performs well when the defect size is relatively large, while there will be deviations for small defects, particularly under the impact of large mass of tramp uranium. **Figure 5** shows the comparison among the FFNN in the previous study, the CFNN,

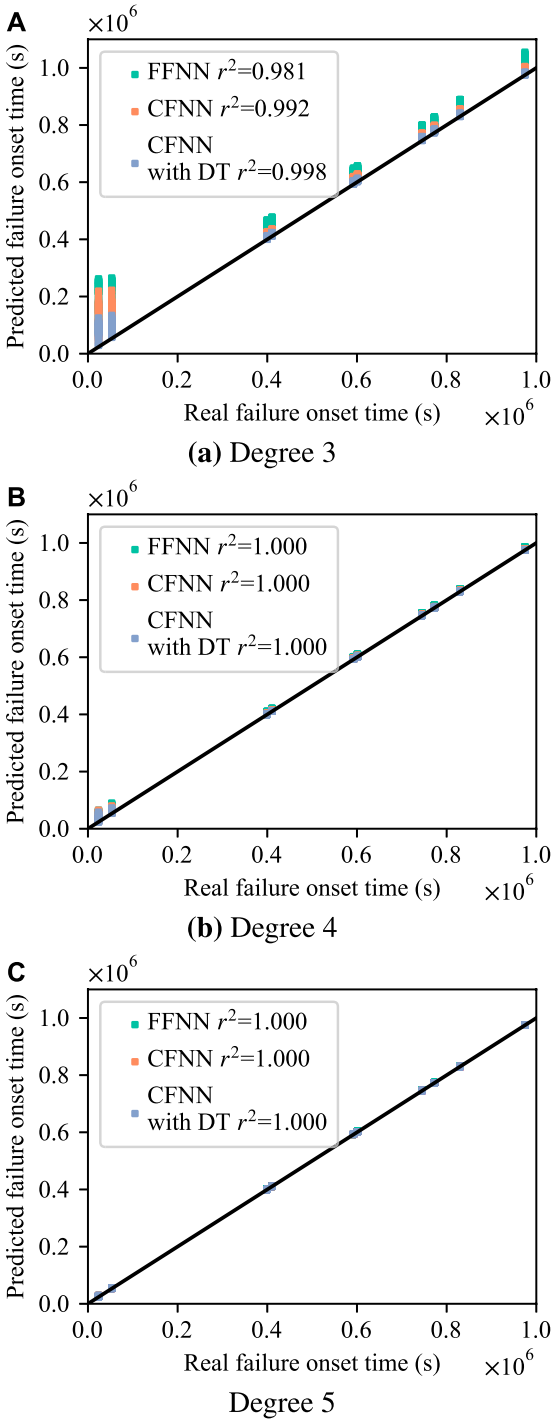


FIGURE 7 | Comparison of response time between the previous FFNN method and the improved method. **(A)** Degree 3, **(B)** Degree 4, and **(C)** Degree 5

and the CFNN with the DT. It can be seen that the CFNN increases the accuracy from 45 to 69% for degree 2, while the CFNN with the DT further increases the accuracy to 80% for degree 2. In addition, there is a minor increase in accuracy for degrees 3–6.

The detailed prediction results for small defects are listed in **Table 2**. As can be observed, both the CFNN and CFNN with DT methods continue to underestimate the degree of fuel failure for degree 2. However, the improved method significantly lowers the underestimation rate. The underestimation phenomenon in the previous study is attributed to an imbalanced dataset, where the sample number of degree 1 is larger than that of other degrees. Supposing the neural network is trivial, it is obvious that classifying all samples into degree 1 has a higher accuracy than classifying all samples into any other degree. Then, the trained neural network has a tendency to identify degree 2 and degree 3 as degree 1, increasing the probability of correct prediction.

Using a DT-like technique can address this issue. As the first step is to decide whether the fuel cladding is defective using a CFNN, the dataset needs to be divided into two categories: intact and defective. Under such condition, the dataset is balanced. After passing through the CFNN1, the samples with intact fuel cladding are classified into degree 1, and other samples with defective fuel cladding go to CFNN2. The dataset of defective fuel cladding of degrees 2–6 is also balanced. Theoretically, the CFNN1 and CFNN2 in the DT are both trained by the balanced dataset, and the degree of fuel failure should not be underestimated. From **Table 2**, it can be seen that some samples are still underestimated. The major reason is that the escape rate coefficient of degree 2 is extremely small ($< 1 \times 10^{-7} s^{-1}$). In addition, some samples of degree 2 are in the cases of short-time after defective fuel cladding. The specific activity in the coolant is not significantly different from the intact condition. This indicates that the feature is not sufficiently noticeable for samples of small defect size. As a result, the CFNN with the DT method still underestimates 20% degree 2 samples.

3.1.2 Responding Speed

When the fuel cladding is defective, it is necessary to rapidly detect the status of fuel failure. On the contrary, the growth in specific activity in the coolant is a gradual process. At the onset of fuel failure, the neural network lacks the confidence in predicting that the fuel cladding is defective, as illustrated in **Figure 6**.

Figure 7 shows the comparison of response time between the previous FFNN method and improved method. The test set is randomly sampled considering the influence of the failure onset time, the escape rate coefficient, and the tramp uranium mass. The sample space is built on 10 random defect time, 10 random tramp uranium mass, and 15 random escape rate coefficients of degrees 3–5, a total of 1,500 samples. As illustrated in **Figure 7A**, there is significant delay in detecting fuel failures using the previous FFNN when the defect size is small and defect occurrence is early. Under these conditions, the response of the CFNN is 15% faster than that of the FFNN. Furthermore, the response of the CFNN with DT is twice faster than that of the FFNN for degree 3 with early defect occurrence. Both the CFNN and CFNN with DT work well for late defect occurrence, with responses more than twice faster than those of the FFNN.

For degrees 4 and 5, the performance of the FFNN is adequate. The results of the CFNN and CFNN with DT are consistent with those of the FFNN and demonstrate no significant improvement

TABLE 3 | Confusion matrix of neural network results of the test set in percentage.

(A) FFNN				
	Actual degree 3	Actual degree 4	Actual degree 5 (%)	Actual degree 6
Predicted degree 3	100%	2.2%	0	0
Predicted degree 4	0	97.8%	0	0
Predicted degree 5	0	0	96	0
Predicted degree 6	0	0	4	0
(B) CFNN				
	Actual degree 3	Actual degree 4	Actual degree 5 (%)	Actual degree 6
Predicted degree 3	99.8%	0	0	0
Predicted degree 4	0.2%	98.2%	0	0
Predicted degree 5	0	1.8%	96	0
Predicted degree 6	0	0	4	0
(C) CFNN with DT				
	Actual degree 3	Actual degree 4	Actual degree 5 (%)	Actual degree 6
Predicted degree 3	95.6%	0%	0	0
Predicted degree 4	4.4	94.8%	0	0
Predicted degree 5	0	5.2%	94.6	0
Predicted degree 6	0	0	5.4	0

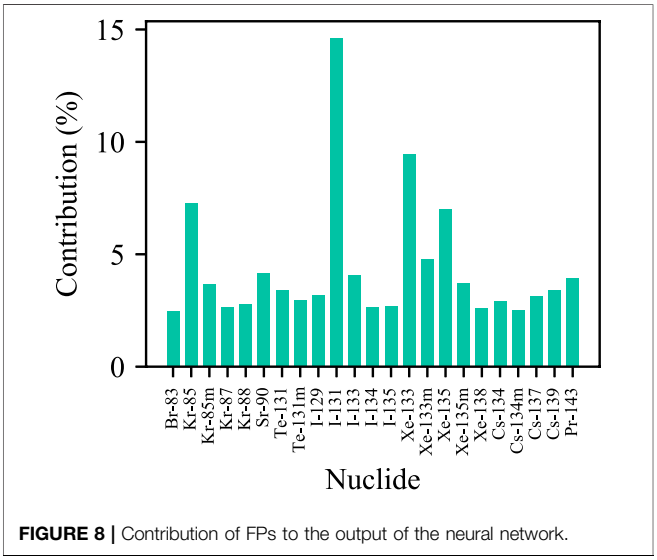


FIGURE 8 | Contribution of FPs to the output of the neural network.

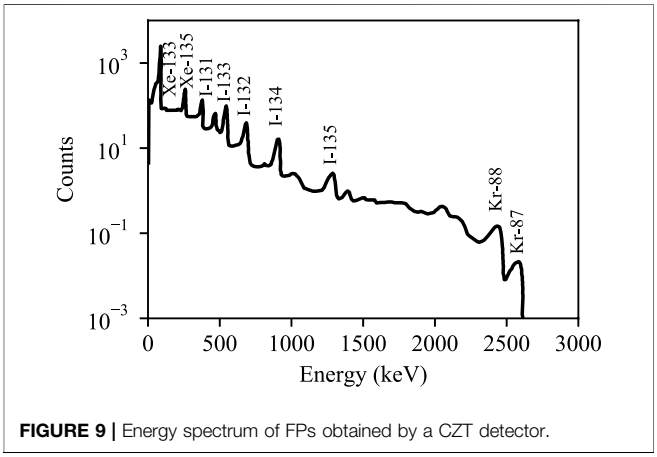


FIGURE 9 | Energy spectrum of FPs obtained by a CZT detector.

in reaction speed, as shown in **Figures 7B,C**. That is because the degrees 4 and 5 correlate to a relatively large defect size, which induces a high increase rate of specific activity in the coolant. The neural network has a high degree of confidence in predicting the occurrence of the defect after the specific activity reaches a critical value.

Table 3 shows the confusion matrix of the neural network results of the test set. The time frame of evaluation is chosen as 48 h from the commencement of fuel failure. Since the structure of the FFNN and CFNN is relatively close, comparing **Table 3A** and **Table 3B**, it can be seen that the accuracy does not decrease, although the CFNN method responds more quickly than the FFNN method. However, the accuracy of the CFNN with DT method is worse than that of the FFNN or CFNN methods, despite its high response speed. This indicates that the CFNN with DT makes a trade-off between accuracy and speed.

The main reason for the fast response of the CFNN with DT method is because the first classifier of the DT is used to detect whether the fuel cladding is defective, which is a binary classification problem. In comparison to the FFNN or CFNN, which are multi-class classification problems, there is a more distinct dividing line between different classes. Since the first classifier of DT responds fast, the second classifier of DT is unable to accurately detect the degree of fuel failure at the moment of the commencement of fuel failure. The feature of the current specific activity may be similar with a minor degree of fuel failure. This phenomenon can also be found in **Figure 6**.

3.2 Contribution of the Nuclide

It is known that hundreds of nuclides are produced during reactor operation, and their fission yields and decay constants are significantly different, which may vary several orders of

magnitude. Because using all nuclides for fuel failure detection is uneconomical and impracticable, it is vital to identify the key nuclides for predicting the status of the fuel cladding.

The neural networks in this study use 23 nuclides as the input, mainly including noble gases, iodine, and cesium. Due to their relatively large fission yield and long half-life, these types of nuclides are already frequently utilized in the isotopic ratio method for fuel failure detection. In order to evaluate the contribution of these nuclides to the determination of the degree of fuel failure and support the further selection of FPs, a sensitivity analysis is performed using the method provided in **Section 2.6**.

After training the neural network, the connection weights are acquired. Then, Q_{nl} in **Eq. 7** can be easily calculated. After obtaining Q_{nl}, the contribution of the input variable can be calculated using **Eq. 8**. Since there are connections between the input and other layers, it should be noted that the contribution derived by **Eq. 8** needs to be averaged.

The contribution of FPs derived by sensitivity analysis of 10 trained neural networks is shown in **Figure 8**. It can be obviously seen that the most important nuclides are I-131, Xe-133, Xe-135, and Kr-85, which is consistent with the prevalent knowledge. These nuclides are already usually used in the isotopic ratio method for fuel failure detection (Menéndez, 2009; Li et al., 2017; Lyu and Xiong, 2019). These four nuclides are critical in the fuel failure detection process. Besides these four nuclides, it can be seen that Xe-133m, Sr-90, and I-133 also have a relatively high contribution.

The previous study has shown that increasing the quantity of input nuclides can improve detection accuracy. While considering the efficiency and limitation of the detection apparatus, some nuclides may be ignored. Although the neural network can be simplified by ignoring some FPs, it is better to not ignore the aforementioned seven nuclides.

3.3 Nuclide Selection in Real Scenario

The ideal neural network should make extensive use of FP-specific activity. While due to the detection limitation of the detector, only a few nuclides can be used as the input of the neural network. Since the specific activity of FPs in the coolant is supposed to be detected by a cadmium zinc telluride (CZT) detector, the nuclides used as the input are selected based on matching the energy spectrum of the CZT detector and generated dataset. A typical spectrum of FPs derived by the CZT detector is shown in **Figure 9**, and the generated dataset has specific activities of 23 nuclides as stated in **Section 2.3**; the nuclides used as the input are selected as I-131, I-133, Xe-133, and Xe-135 combining the sensitivity analysis in **Section 3.2**. Then, the neural networks with an input of four nuclides are trained based on the improved method and used for practical application, which is consistent with the previous study.

4 CONCLUSION

In this study, an improved method based on the CFNN and DT for fuel failure detection is proposed. The method considerably improves the accuracy of fuel failure detection when the defect size of the fuel cladding is small. The CFNN with DT method increases the detection accuracy from 45 to 80% for fuel failure of degree 2. In comparison to the previous FFNN method, the response speed of the CFNN with DT method is more than twice, although at the expense of sacrificing some precision. The sensitivity analysis of the trained neural network indicates that I-131, Xe-133, Xe-135, and Kr-85 are the most important nuclides in fuel failure detection. Besides these four nuclides, Xe-133m, Sr-90, and I-133 also have a relatively high contribution to the neural network for detecting fuel failure.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://data.mendeley.com/datasets/tcscykf5hn/2.

AUTHOR CONTRIBUTIONS

BD: wrote the manuscript and designed the methodology. KY: trained the neural network and revised the manuscript. WZ: provided the detector information. JY: revised the manuscript. DW: supervised and funded the project.

FUNDING

This study is financially supported by the National Natural Science Foundation of China (No. 11675105).

REFERENCES

Andrews, W. S., Lewis, B. J., and Cox, D. S. (1999). Artificial Neural Network Models for Volatile Fission Product Release during Severe Accident Conditions. *J. Nucl. Mater.* 270, 74–86. doi:10.1016/S0022-3115(98)00765-X

Choi, S. H., Choi, H. J., Min, C. H., Chung, Y. H., and Ahn, J. J. (2021). Development of De-noised Image Reconstruction Technique Using Convolutional AutoEncoder for Fast Monitoring of Fuel Assemblies. *Nucl. Eng. Technol.* 53, 888–893. doi:10.1016/j.net.2020.08.020

de Boer, P.-T., Kroese, D. P., Mannor, S., and Rubinstein, R. Y. (2005). A Tutorial on the Cross-Entropy Method. *Ann. Oper. Res.* 134, 19–67. doi:10.1007/s10479-005-5724-z

Dong, B., Xiao, W., Yin, J., and Wang, D. (2020). Detection of Fuel Failure in Pressurized Water Reactor with Artificial Neural Network. *Ann. Nucl. Energ.* 140, 107104. doi:10.1016/j.anucene.2019.107104

Ebrahimzadeh, A., Ghafari, M., and Moshkbar-Bakhshayesh, K. (2022). Detection and Estimation of Faulty Sensors in NPPs Based on thermal-hydraulic Simulation and Feed-Forward Neural Network. *Ann. Nucl. Energ.* 166, 108726. doi:10.1016/j.anucene.2021.108726

Garson, G. D. (1991). Interpreting Neural-Network Connection Weights. *AI Expert* 6, 46–51.

Gevrey, M., Dimopoulos, I., and Lek, S. (2003). Review and Comparison of Methods to Study the Contribution of Variables in Artificial Neural Network Models. *Ecol. Model.* 160, 249–264. doi:10.1016/S0304-3800(02)00257-0

Guo, Z., Wu, Z., Liu, S., Ma, X., Wang, C., Yan, D., et al. (2020). Defect Detection of Nuclear Fuel Assembly Based on Deep Neural Network. *Ann. Nucl. Energ.* 137, 107078. doi:10.1016/j.anucene.2019.107078

Iqbal, M. J., Mirza, N. M., and Mirza, S. M. (2007). Kinetic Simulation of Fission Product Activity in Primary Coolant of Typical PWRs under Power Perturbations. *Nucl. Eng. Des.* 237, 199–205. doi:10.1016/j.nucengdes.2006.06.003

Kalinichev, P., Evdokimov, I., and Likhanskii, V. (2018). A Technique for Detection of WWER Fuel Failures by Activity of Xe Radionuclides during Reactor Operation. *Nucet* 4, 263–270. doi:10.3897/nucet.4.31889

Kingma, D. P., and Ba, J. (2017). Adam: A Method for Stochastic Optimization. *arXiv*. arXiv:1412.6980 [cs]. doi:10.48550/arXiv.1412.6980

Lewis, B. J., Chan, P. K., El-Jaby, A., Iglesias, F. C., and Fitchett, A. (2017). Fission Product Release Modelling for Application of Fuel-Failure Monitoring and Detection - an Overview. *J. Nucl. Mater.* 489, 64–83. doi:10.1016/j.jnucmat.2017.03.037

Li, H., Fu, Y., Mao, L., and Mei, Q. (2017). "Research on Estimating Methods and Application of Fuel Rods Defect," in Proceedings of the 2017 25th International Conference on Nuclear Engineering, Shanghai, China, July 2017 (New York, N.Y.: American Society of Mechanical Engineers (ASME)), 4. doi:10.1115/icone25-66622

Li, L., and Yang, H. (2008). Diagnosis Method for Fuel Failures in Pressurized Water Reactor Nuclear Power Plant. *Nucl. Power Eng.* 29, 135–139.

Likhanskii, V., Afanasieva, E., Sorokin, A., Evdokimov, I., Kanukova, V., and Khromov, A. (2006). "Failed Fuel Diagnosis during WWER Reactor Operation Using the RTOP-CA Code," in Proceedings of the International Conference on WWER Fuel Performance, Modelling and Experimental Support, Bulgaria, 2006 (Vienna, Austria: IAEA), 6.

Lyu, W., and Xiong, J. (2019). Research and Improvement of Diagnosis Method of Fuel Failures for Pressurized Water Reactor Nuclear Power Plant. *Nucl. Power Eng.* 40, 43.

Menéndez, M. A. T. (2009). "Fuel Failure Detection, Characterization and Modelling: Effect on Radionuclide Behaviour in PWR Primary Coolant," (Spain: Polytechnic University of Madrid). Ph.D. thesis.

Perez, P. B., and Hamawi, J. N. (2017). Fission Product Appearance Rate Coefficients in Design Basis Source Term Determinations - Past and Present. *EPJ Web Conf.* 153, 07037. doi:10.1051/epjconf/201715307037

Qin, G., Chen, X., Guo, X., and Ni, N. (2016). On-line Detection of Key Radionuclides for Fuel-Rod Failure in a Pressurized Water Reactor. *Appl. Radiat. Isot.* 114, 71–75. doi:10.1016/j.apradiso.2016.05.017

Qin, G., Wang, Q., Chen, X., Li, F., Li, W., and Guo, X. (2020). Development of Fuel Rod Failure Character Analysis Code for Pressurized Water Reactors. *Nucl. Eng. Des.* 361, 110515. doi:10.1016/j.nucengdes.2020.110515

Qin, G., Wang, Q., Xu, Y., Chen, X., Li, F., and Li, W. (2019). γ-Ray Spectral Analysis Method for Real-Time Detection of Fuel Element Failure. *Ann. Nucl. Energ.* 133, 221–226. doi:10.1016/j.anucene.2019.05.034

Wang, Z., Xia, H., Zhang, J., Annor-Nyarko, M., Zhu, S., Jiang, Y., et al. (2022). A Deep Transfer Learning Method for System-Level Fault Diagnosis of Nuclear Power Plants under Different Power Levels. *Ann. Nucl. Energ.* 166, 108771. doi:10.1016/j.anucene.2021.108771

Warsito, B., Santoso, R., Supartiand Yasin, H. (2018). Cascade Forward Neural Network for Time Series Prediction. *J. Phys. Conf. Ser.* 1025, 012097. doi:10.1088/1742-6596/1025/1/012097

Weifeng, L., and Jun, X. (2019). Research and Improvement of Diagnosis Method of Fuel Failures for Pressurized Water Reactor Nuclear Power Plant. *Nucl. Power Eng.* 40, 43.

Yamamoto, Y., and Morishita, K. (2015). Development of Methodology to Optimize Management of Failed Fuels in Light Water Reactors. *J. Nucl. Sci. Technol.* 52, 709–716. doi:10.1080/00223131.2014.979901

Zanker, H. (1989). Defective Fuel Rod Detection in Operating Pressurized Water Reactors during Periods of Continuously Decreasing Fuel Rod Integrity Levels. *Nucl. Technol.* 86, 239–247. doi:10.13182/NT89-A34292

Zhang, B., Miao, Y., Tian, Y., Zhang, W., Wu, G., Wang, X., et al. (2021). Implementation of Surface Crack Detection Method for Nuclear Fuel Pellets Guided by Convolution Neural Network. *J. Nucl. Sci. Technol.* 58, 787–796. doi:10.1080/00223131.2020.1869622

Conflict of Interest: Author KY and WZ were employed by China Nuclear Power Engineering Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Energy Research

ORIGINAL RESEARCH
published: 21 March 2022
doi: 10.3389/fenrg.2022.849229

An Intelligent Optimization Method for Preliminary Design of Lead-Bismuth Reactor Core Based on Kriging Surrogate Model

Qiong Li[1,2], *Zijing Liu*[1,2]*, *Yingjie Xiao*[1,2], *Pengcheng Zhao*[1,2], *Yanan Zhao*[1,2], *Tao Yang*[1,2] and *Tao Yu*[1,2]*

[1]School of Nuclear Science and Technology, University of South China, Hengyang, China, [2]Hunan Engineering and Technology Research Center for Virtual Nuclear Reactor, University of South China, Hengyang, China

To meet the numerous application demands of lead-bismuth reactors, different design optimization tasks need to be conducted on these reactors based on the existing reactor core solutions. However, the design optimization of lead-bismuth reactors is a challenging task because it is a complex, multi-dimensional, and nonlinear constrained problem. To resolve these issues and improve the efficiency of design optimization, a new method, called the KSM-OLHS-SEUMRE method, based on the Kriging surrogate model (KSM), orthogonal Latin hypercube sampling (OLHS), and space exploration and unimodal region elimination (SEUMRE) algorithm is proposed in this study. Based on this method, a design optimization program of lead-bismuth reactors (DOPPLER-K) is developed, which realizes functions like sample point generation, optimization analysis, pre-post processing of reactor calculation, coupling of the Reactor Monte Carlo (RMC) calculation code and the Steady-state Thermal-hydraulic Analysis Code (STAC). Further, taking lead-bismuth reactors SPALLER-4 and URANUS as prototypes, the proposed intelligent optimization method for preliminary design of lead-bismuth reactor core is verified. The results show that this method can rapidly and accurately find the target scheme satisfying the optimization conditions, and it is three orders of magnitude faster than pure Monte Carlo calculation. Compared with the initial core scheme of URANUS, the optimization rates of fuel loading, total core mass, active zone volume, and total core volume are reduced by 10.8, 11.5, 18.1, and 17.1%, respectively. These results validate the feasibility and efficacy of the proposed method for design optimization of lead-bismuth reactor core.

Keywords: lead-bismuth reactor, intelligent optimization, Kriging surrogate model, SEUMRE algorithm, orthogonal Latin hypercube sampling

OPEN ACCESS

Edited by:
Jun Wang,
University of Wisconsin-Madison, United States

Reviewed by:
Fenglei Niu,
North China Electric Power University, China
Chenglong Wang,
Xi'an Jiaotong University, China

***Correspondence:**
Zijing Liu
liuzijing1123@163.com
Tao Yu
taoyu0626@163.com

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *05 January 2022*
Accepted: *10 February 2022*
Published: *21 March 2022*

Citation:
Li Q, Liu Z, Xiao Y, Zhao P, Zhao Y, Yang T and Yu T (2022) An Intelligent Optimization Method for Preliminary Design of Lead-Bismuth Reactor Core Based on Kriging Surrogate Model. Front. Energy Res. 10:849229. doi: 10.3389/fenrg.2022.849229

1 INTRODUCTION

Over the recent years, lead-bismuth reactors have received considerable attention from the major nuclear energy countries owing to their unique characteristics such as long-life, miniaturization ability, high flux, and natural circulation, which lead to outstanding advantages in the utilization of nuclear energy (Wu, 2018). However, it may be noted that different technical indicators and reactor core design schemes are required in various application scenarios and missions, thus numerous design optimization tasks need to be conducted based on the existing lead-bismuth reactor core

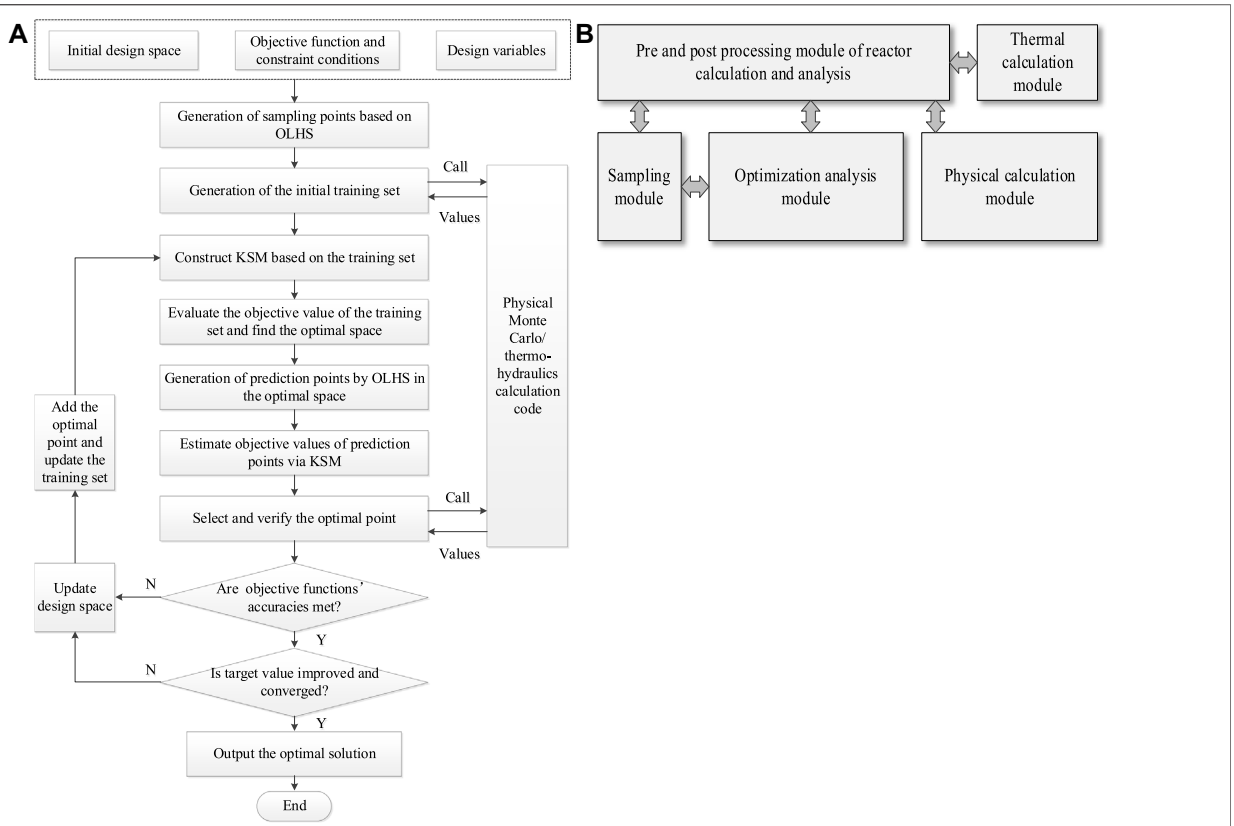


FIGURE 1 | Intelligent optimization method. **(A)** Flow chart of KSM-OLHS-SEUMRE method, **(B)** Functional modules of DOPPLER-K.

schemes. It is well known that since it involves coupled calculation and analysis of physical/thermal-hydraulic characteristics and strength of structural materials, the design optimization of lead-bismuth reactor core is a complex, multi-dimensional, and nonlinear constrained problem, and a large uncertainty exists in the calculation results, which is caused by the nuclear reaction cross section, approximation of calculation model, etc. The existing traditional semi-empirical design methods that rely on the designer experience cannot find the optimal solution efficiently under the influence of multi-factor coupling. On the other hand, the single heuristic optimization algorithms, such as genetic algorithm, simulated annealing algorithm, particle swarm algorithm, etc., need a large number of stochastic samples and Monte Carlo calculations to support the design optimization and suffer from the problems of slow convergence and a trend of falling into local optimum (Meneses et al., 2009; Khoshahval et al., 2011; Zameer et al., 2014). Therefore, it is necessary to develop an intelligent optimization method for preliminary design of lead-bismuth reactor core with random effects, low computational cost, high convergence speed, and strong optimization ability.

To solve these problems such as excessive calculation, slow convergence, and large uncertainty in traditional Monte Carlo computation and single heuristic optimization algorithm, the surrogate modeling technique in which an approximate model is constructed based on the experimental points to replace the complex original model has been considered. The commonly used surrogate models include the response surface method, radial basis function interpolation, Kriging model, and so on (Zhang, 2014). Among them, the Kriging surrogate model (KSM) can not only fit a smooth curve closer to the original objective function but also can be easily adapted to different variation trends, and it covers the inherent uncertainty of objective functions by selecting different regression models and related functions (Kempf et al., 2012). Therefore, although the accuracy of the KSM will decrease when solving the prediction of the objective function with more than 8 design parameters, the KSM is still used in this study due to its overall strong fitting effect and the modeling method that can cover the inherent

TABLE 1 | Commonly used correlation functions and their expressions.

Correlation function	Expression
Exponential function	$R_k(\theta_k, d_k) = \exp(-\theta_k d_k)$
Gaussian function	$R_k(\theta_k, d_k) = \exp(-\theta_k d_k^2)$
Linear function	$R_k(\theta_k, d_k) = \max\{0, 1 - \theta_k d_k\}$
Cubic spline function	$R_k(\theta_k, d_k) = \begin{cases} 1 - 15\zeta_k + 30\zeta_k^3, & 0 \le \zeta_k \le 0.2 \\ 1.25(1 - 15\zeta_k)^3, & 0.2 < \zeta_k < 1 \\ 0 & , \zeta_k \ge 1, \zeta_k = \theta_k d_k \end{cases}$

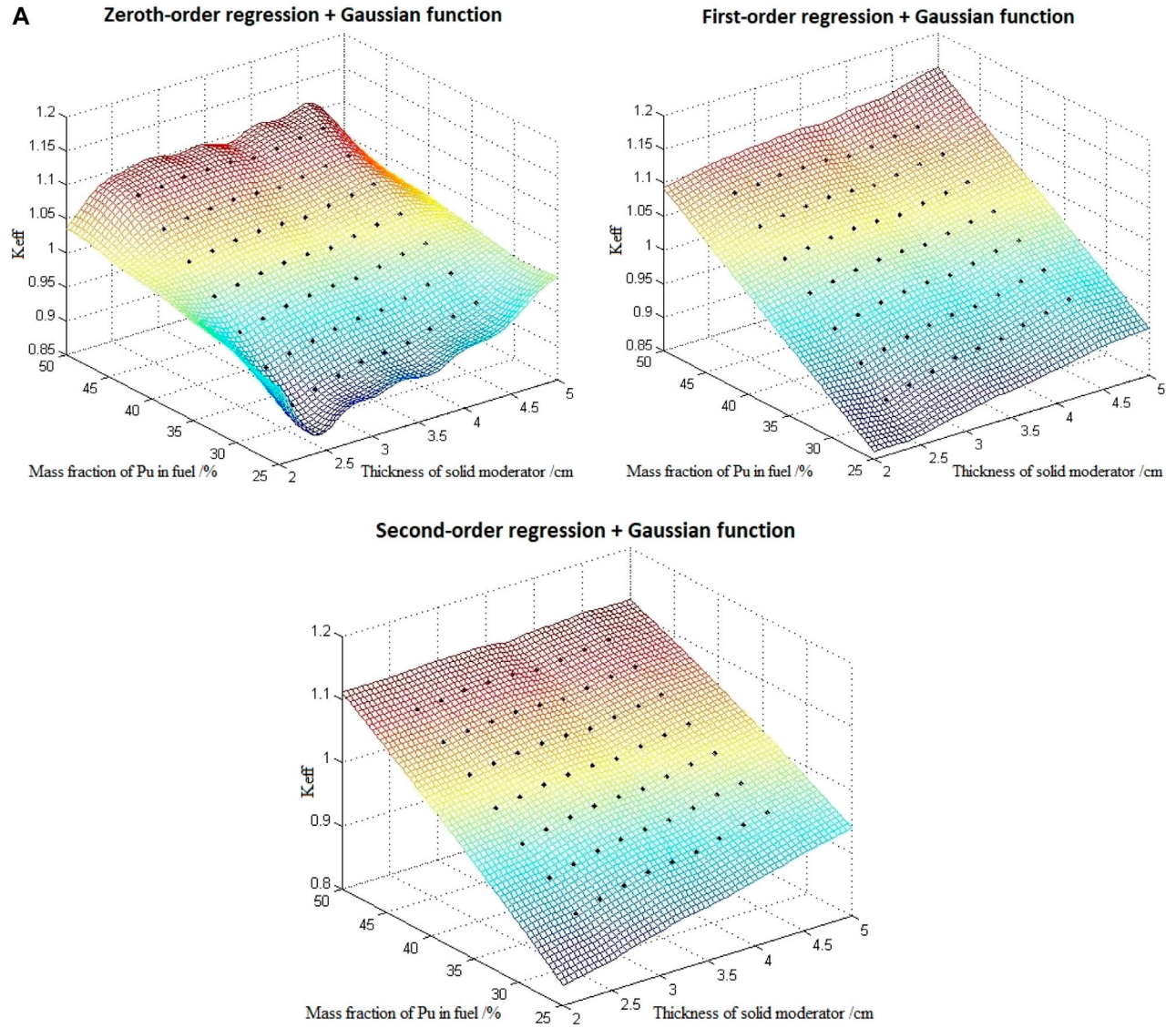


FIGURE 2 | Construction records of KSM, OLHS, and SEUMRE. **(A)** Fitting effect of KSM constructed by different regression models with Gaussian function, **(B)** Fitting effect of KSM constructed by different correlation function with second-order regression model, **(C)** MSE of KSM constructed by exponential/Gaussian correlation function with second-order regression model, **(D)** Comparison between ordinary LHS and OLHS distribution, **(E)** Flow diagram of SEUMRE algorithm.

uncertainty in the reactor calculation. To date, KSM has been widely used in the design optimization of aerospace, machinery, automobiles, and other fields, and it is now being used in the reactor field as well. Zeng et al. (2020) used KSM to provide a generalized framework for the core optimization of sodium-cooled fast reactor, and Kempf et al. (2012) used it to obtain the optimal geometric parameters for a high-flux research reactor. Kim et al. used the KSM for shape optimization of the inlet plenum and rising channels in the pebble-bed modular reactor (Kim and Lee, 2009). Li et al. studied the effects of flow and heat transfer factors in the rod bundle of fast reactor assemblies on the thermal-hydraulic characteristics through the Kriging technology (Song and Yang, 2018). All the above studies verified the good prediction accuracy and high computational efficiency of the KSM. Thus, the optimization method combined with KSM can be a feasible and effective approach for the design optimization of a lead-bismuth reactor core.

In this study, an intelligent optimization method is proposed for preliminary design of lead-bismuth reactor core, which is based on the combination of KSM, orthogonal Latin hypercube sampling (OLHS), and space exploration and unimodal region elimination (SEUMRE) algorithm. This method is called the KSM-OLHS-SEUMRE method. By coupling the reactor Monte Carlo (RMC) code, a Monte Carlo code for reactor core analysis, and the steady-state thermal-hydraulic analysis code (STAC), a

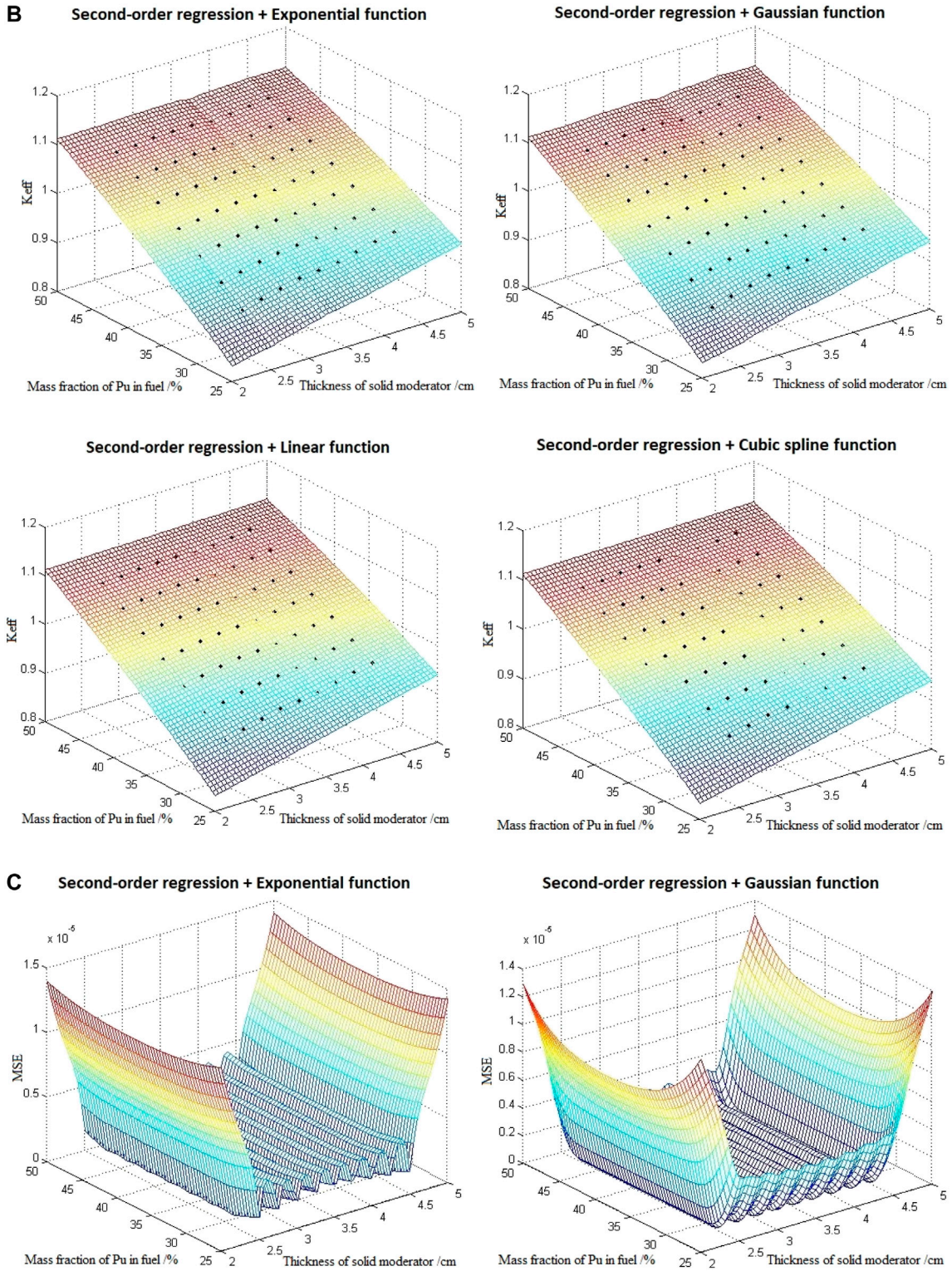


FIGURE 2 | Continued.

design optimization program of lead-bismuth reactors based on KSM-OLHS-SEUMRE (DOPPLER-K) has been developed. Taking lead-bismuth reactors SPALLER-4 and URANUS as prototypes, the proposed optimization method is verified through a rapid search for target design schemes and core parameter optimization.

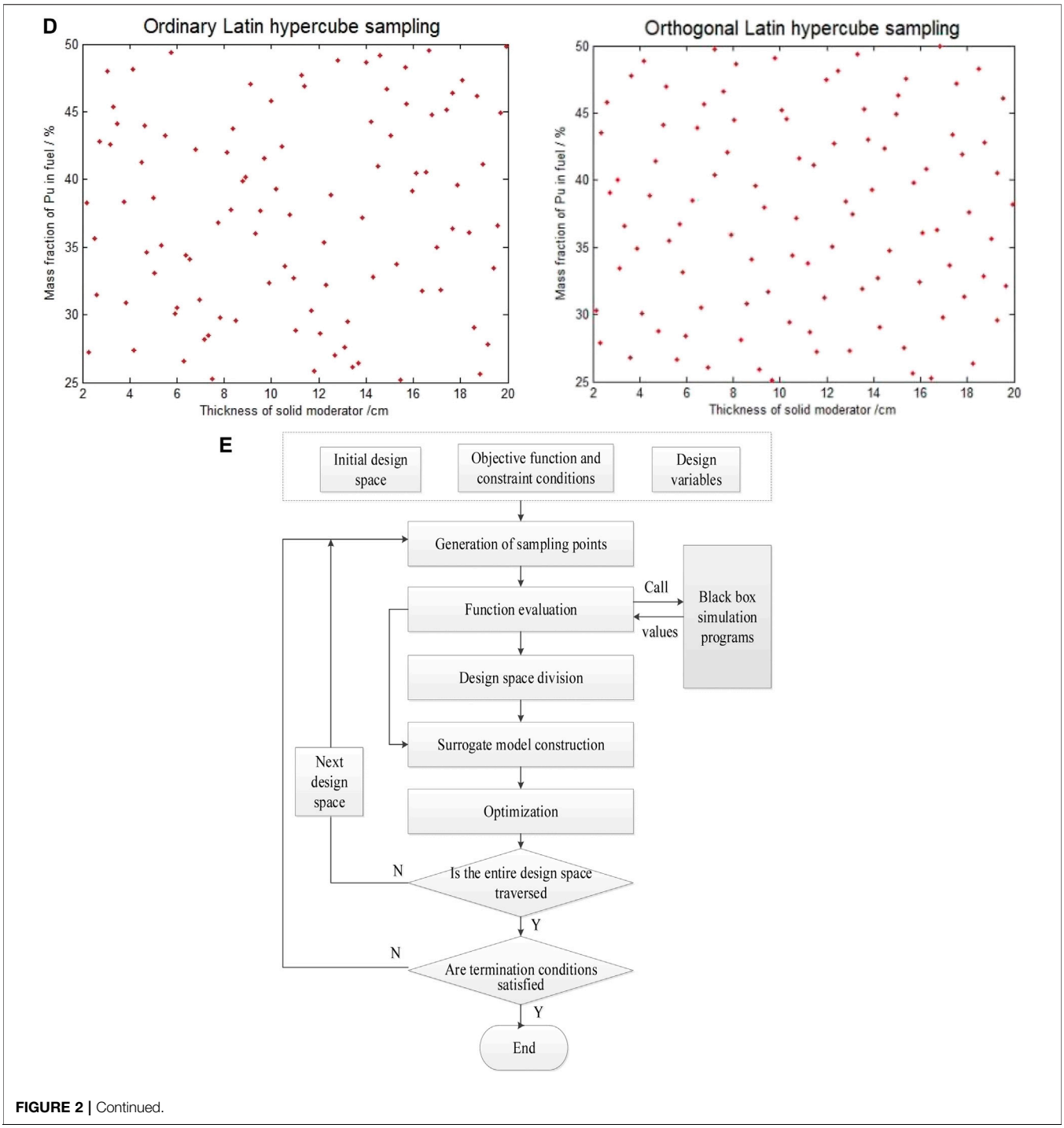


FIGURE 2 | Continued.

The rest of this paper is organized as follows. **Section 2** provides a brief introduction to the KSM-OLHS-SEUMRE method and the DOPPLER-K software. The construction principles of KSM, OLHS, and SEUMRE algorithm are described in **Section 3**. The optimization verification results for SPALLER-4 and URANUS are discussed in **Section 4**. Finally, the study is concluded in **Section 5**.

2 INTELLIGENT OPTIMIZATION METHOD

The intelligent optimization method constructed in this study includes the KSM-OLHS-SEUMRE method (introduced in **Section 2.1**) and DOPPLER-K software (introduced in **Section 2.2**). The construction principles of KSM, OLHS and SEUMRE

TABLE 2 | Design parameters of SPALLER-4 and URANUS.

Design parameter	SPALLER-4	URANUS
Reactor thermal power/MWt	4	100
Refueling cycle/EFPY	10	20
Fuel loading/kg	577.89	17,580
Fuel (Mass fraction of Pu)	PuN-ThN (31/48)	UO_2 (9.55/17.09)
Coolant	^{208}Pb-Bi	^{208}Pb-Bi
Reflector	^{208}Pb-Bi	^{208}Pb-Bi
Shielding	B_4C	B_4C
Moderator	BeO	—
Fuel pin cladding	HT-9	HT-9
Filling gas in the gap of pin	He	He

involved in KSM-OLHS-SEUMRE method are described in **Section 3**.

Different from traditional or single heuristic optimization method, the optimization method adopted in this research firstly uses KSM technology to replace the original Monte Carlo code with slow calculation speed to predict the core characteristic parameters, which not only greatly improves the calculation efficiency, but also realizes the target prediction under the simultaneous coupling of multiple core design parameters. Secondly, when multiple design parameters influence at the same time and the design range is wide, this method makes the optimization avoid the problem of local optimum and difficult convergence in the traditional optimization algorithm as far as possible through sequence iteration, preferential addition point selection and SEUMRE space search technology to improve the optimization search efficiency. Finally, through the developed DOPPLER-K software, the automatic execution of sampling, Monte-Carlo program calling calculation and target optimization analysis can be carried out, which provides great convenience to optimization designers.

2.1 KSM-OLHS-SEUMRE Method

The design optimization of lead-bismuth reactor core generally involves the establishment of a mathematical model, selection of design variables, calculation of physical/thermal characteristic parameters, and determination of the optimal scheme. Based on this strategy and the construction principles in **Section 3**, the optimization method KSM-OLHS-SEUMRE is developed. Its flow chart is shown in **Figure 1A**, and the basic implementation steps are as follows:

Step 1. Establish the mathematical model of the lead-bismuth reactor core to be optimized, including design space and variables, objective functions, and constraint conditions.

Step 2. Generate sample points of the mathematical model by the OLHS technique.

Step 3. Calculate sample points to generate the training set based on the physical Monte-Carlo/thermal-hydraulics calculation code.

Step 4. Construct KSM based on the training set and evaluate the objective values of the training set to find the optimal space.

Step 5. Predict the points of target/constraint functions and assess them to select the optimal point verified by physical Monte Carlo/thermal-hydraulics calculation code.

Step 6. Update the design space and KSM continuously based on the principle of optimal point selection and SEUMRE algorithm to improve the prediction accuracy and overall optimization efficiency of the model.

Step 7. Perform iterative optimization to quickly approximate the objective function and obtain the optimal convergence solution.

2.2 DOPPLER-K Software

Based on the KSM-OLHS-SEUMRE method, the design optimization program of lead-bismuth reactors, called DOPPLER-K software, which couples RMC and STAC and realizes functions, such as sampling, modeling, and optimization, has been developed in MATLAB. The functional modules of DOPPLER-K software are shown in **Figure 1B**.

In the sampling module, all the design variables are sampled and saved in the initial training set using OLHS after the user defines the variables and initial design space.

In the physical calculation module, some characteristic parameters of sample points are calculated by RMC (Kan et al., 2015), a three-dimensional Monte Carlo neutron transport code developed by the Reactor Engineering Analysis Laboratory (REAL), Department of Engineering Physics, Tsinghua University. These characteristic parameters, such as effective multiplication factor (K_{eff}), flux, power, and burnup, can be used as objective functions or constraint conditions for design optimization of the lead-bismuth reactor core.

In the thermal calculation module, thermal-hydraulic characteristic parameters, such as the maximum fuel cladding and pellet temperature, coolant velocity, based on the sample points and reactor core power distribution are used as the constraint conditions for the design optimization of lead-bismuth reactor core. These parameters are calculated by STAC (Zhao et al., 2020), which was developed by one of the authors and includes the physical model of liquid lead/lead-bismuth, calculation model of wall heat transfer and pressure drop, and some models built for lead-bismuth reactor like single channel, closed parallel multichannel, and hottest channel model.

As the coupling interface of sampling, physical/thermal-hydraulic calculation, and optimization analysis, the pre and post processing module automatically matches the reactor core parameter values of sampling points to generate the input file, reads the physical/thermal characteristic parameters of output files after carrying out parallel calculations by calling RMC and STAC automatically, and finally generates the training set or accuracy verification conditions.

In the optimization analysis module, firstly, the KSM is constructed by analyzing the structural characteristics of known objective function values and quantifying their spatial correlation. Secondly, the objective function values of new valuation points are predicted based on the surrogate models after determining the neighborhood range and searching for neighborhood points. Finally, according to the principle of

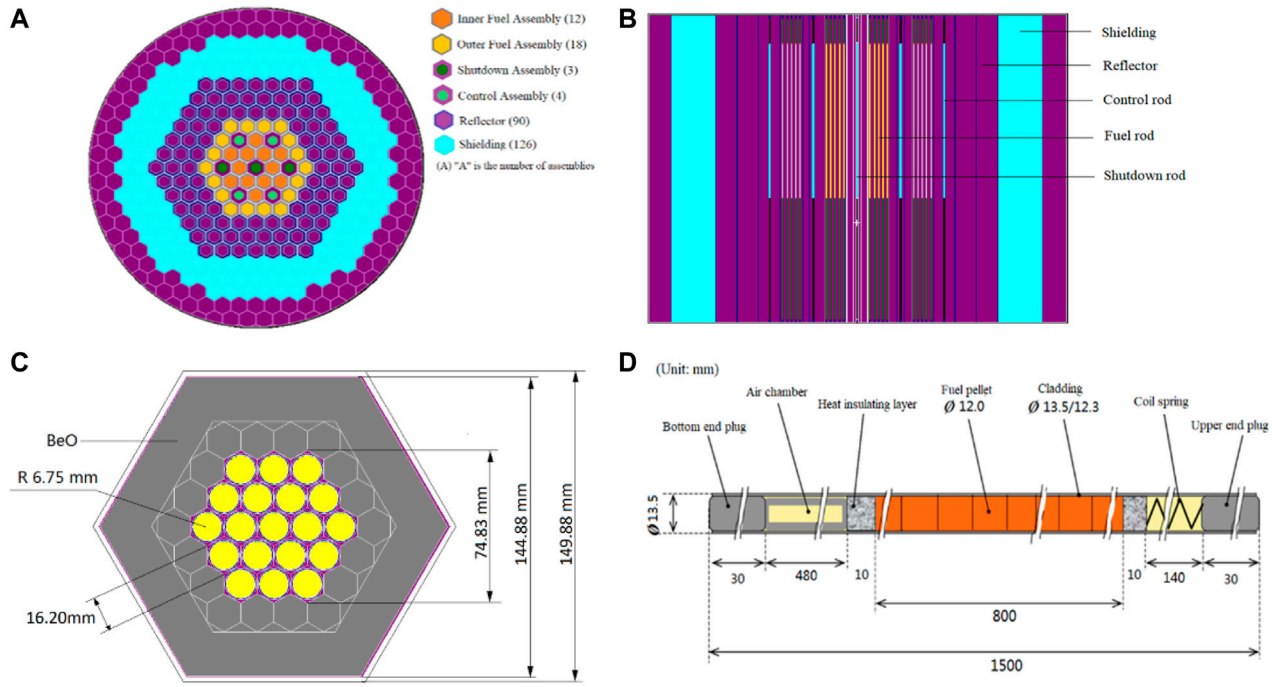


FIGURE 3 | Original core **(A,B)**, fuel assembly **(C)**, and fuel rod **(D)** of SPALLER-4.

optimal point selection and SEUMRE algorithm, the sample points are updated to revise the surrogate model and the design space to approximate the target function iteratively until the optimal solution is obtained.

3 CONSTRUCTION PRINCIPLES OF KRIGING SURROGATE MODEL, ORTHOGONAL LATIN HYPERCUBE SAMPLING, AND SPACE EXPLORATION AND UNIMODAL REGION ELIMINATION

3.1 Construction Principle of Kriging Surrogate Model

Kriging is a regression algorithm for spatial modeling and prediction of random fields based on the theory of correlation functions and structural analysis. In contrast to other interpolation algorithms, the KSM is an optimal unbiased estimation technique, which is composed of a global trend function $F(\beta, x)$ and a random distribution departure $Z(x)$. Its mathematical expression is as follows (Jin et al., 2000):

$$y(x) = F(\beta, x) + Z(x) = \beta * f^{T}(x) + Z(x) \tag{1}$$

Here, $F(\beta, x) = \beta * f^{T}(x)$, where β is the coefficient of the regression function $f^{T}(x)$, and $Z(x)$ indicates that the local error approximation is the key to the accuracy of this model. The characteristics of $Z(x)$ are given as follows:

$$E[Z(x)] = 0 \tag{2}$$

$$Var[Z(x)] = \delta^2 \tag{3}$$

$$Cov\left[Z(x_i), Z\left(x_j\right)\right] = \delta^2 R\left(c, x_i, x_j\right) \tag{4}$$

In other words, the mathematical expectation of random distribution bias is zero, and the variance δ^2 is minimum, which can be calculated by **Eq. 4**, where $Cov[Z(x_i), Z(x_j)]$ is the covariance of the random deviation and $R(c, x_i, x_j)$ is the correlation function, representing the spatial correlation between any two points. $R(c, x_i, x_j)$ is expressed as follows:

$$R\left(c, x_i, x_j\right) = \prod_{k=1}^{n} R_k\left(\theta_k, \left|x_i^k - x_j^k\right|\right) = \prod_{k=1}^{n} R_k\left(\theta_k, d_k\right) \tag{5}$$

Where n is a variable number, θ_k is the proportionality factor, and d_k is the Euclidean distance between two points. It can be seen from these above definitions that the key to the accuracy of KSM is the selection of the correlation function. The commonly used correlation functions include exponential, Gaussian, linear, and cubic spline functions, whose expressions are listed in **Table 1** (Zhang, 2014).

TABLE 3 | Range of design parameters for SPALLER-4.

Design parameter	Range
Thickness of solid moderator/cm	[0, 20]
Mass fraction of Pu in fuel/%	[25, 50]
Fuel pin radius/cm	[0.2, 0.6]
Height of core active zone/cm	[30, 150]
Pitch to diameter ratio (P/D)	[1.01, 1.5]

TABLE 4 | Accuracy verification results of KSM for predicting K_{eff} and burnup.

Contrast group	1	2	3	4	5
Thickness of solid moderator/cm	4.655,531	4.822,177	4.990,846	4.589,912	4.782,832
Mass fraction of Pu in fuel/%	47.202,398	45.410,111	48.931,457	48.822,818	46.664,652
Fuel pin radius/cm	0.291,086	0.277,635	0.260,789	0.211,688	0.217,254
Height of core active zone/cm	112.165,889	115.235,329	118.186,032	103.660,647	116.591,780
Pitch to diameter ratio (P/D)	1.371,007	1.377,327	1.411,726	1.353,392	1.354,788
Third year K_{eff}					
Prediction by KSM	1.050,181	1.035,229	1.032,517	1.016,440	1.024,406
Calculation by RMC	1.050,343	1.035,223	1.033,391	1.017,445	1.023,389
Relative error/%	−0.015,427	0.000,602	−0.084,557	−0.098,739	0.099,373
Burnup/(MW·d·kg^{-1})					
Prediction by KSM	22.947,678	24.660,966	26.864,599	46.352,778	39.158,879
Calculation by RMC	22.796,000	24.446,000	26.894,000	46.544,000	39.396,000
Relative error/%	0.665,370	0.879,352	−0.109,323	−0.410,841	−0.601,890

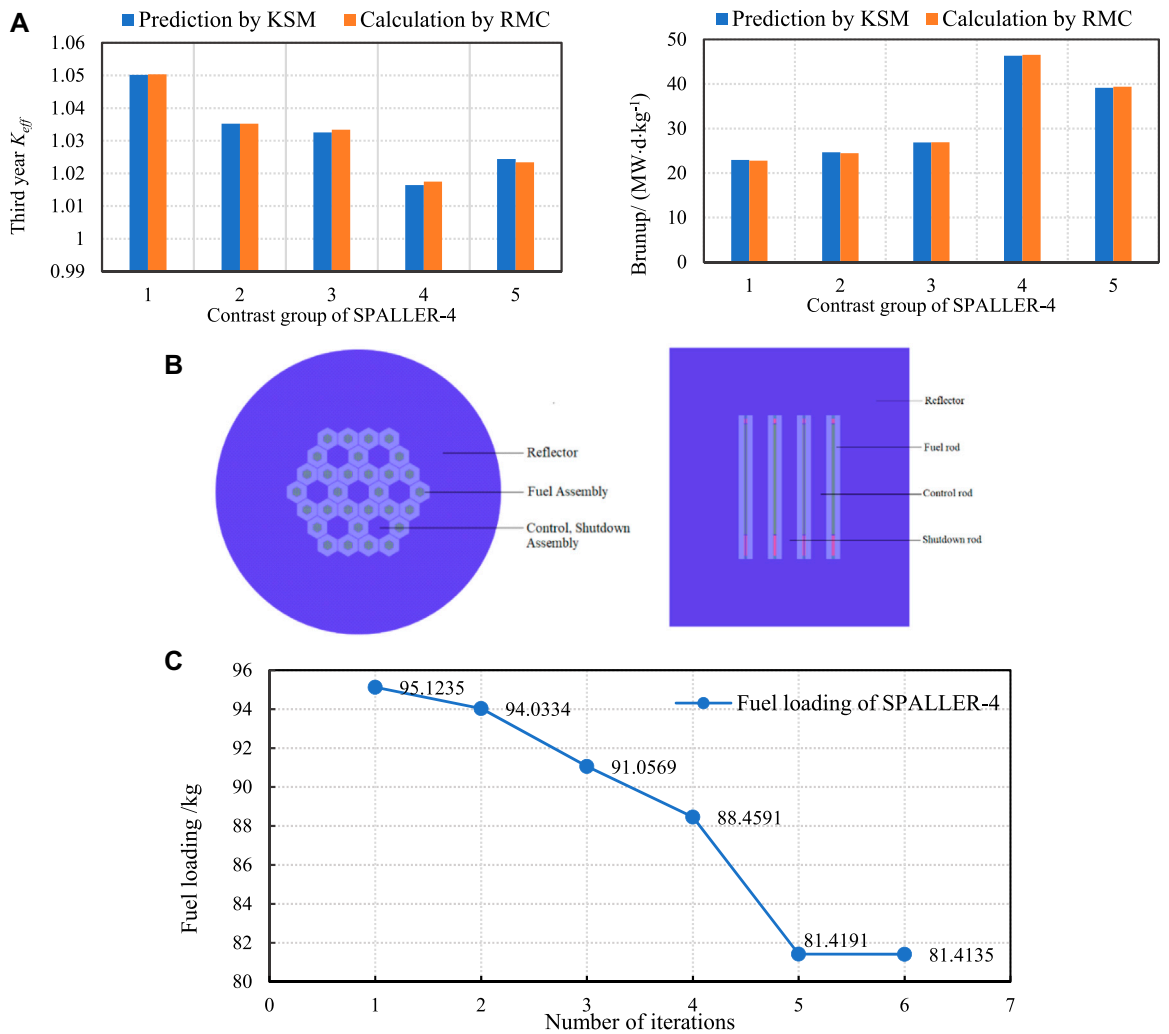


FIGURE 4 | Optimization verification results of SPALLER-4. **(A)**Comparison of K_{eff} and burnup obtained based on KSM and RMC, **(B)** Core structure of SPALLER-4 to be optimized, **(C)** Iterative graph of fuel loading optimization for SPALLER-4.

TABLE 5 | Optimization results of core design scheme for SPALLER-4.

Design parameter	Optimal scheme
Thickness of solid moderator/cm	4.573
Mass fraction of Pu in fuel/%	49.869
Fuel pin radius/cm	0.200
Height of core active zone/cm	100.082
Pitch to diameter ratio (P/D)	1.313
Initial K_{eff}	1.028,058
Third year K_{eff}	
Prediction by KSM	1.005,741
Calculation by RMC	1.005,189
Relative error/%	0.054,959
Burnup/(MW·d·kg^{-1})	
Prediction by KSM	53.702
Calculation by RMC	53.799
Relative error/%	−0.002
Refueling cycle/EFPY	3
Fuel loading/kg	81.414
Maximum temperature of fuel cladding/K	643.162
Maximum temperature of fuel pellet/K	1,133.073

To select the best combination of regression model and correlation function to construct KSM, the selection of regression model was first carried out. As shown in **Figure 2A**, the comparison of the fitting effects of KSM constructed by the combination of zero-order, first-order and second-order regression models with Gaussian correlation functions shows that the fitting effects of KSM constructed by the second-order regression model are more smooth and consistent. Then, the second-order regression model is combined with the exponential, Gaussian, linear and cubic spline correlation functions to construct the model for comparison of fitting effect. As shown in **Figure 2B**, it can be seen that the fitting effect of Gaussian and exponential function is the best. Since the fitting effects of Gaussian and exponential correlation functions are consistent, the mean square error (MSE) is used for experimental comparison. As shown in **Figure 2C**, the MSE of KSM constructed by Gaussian function is slightly smaller than that of exponential function, which is relatively better. Consequently, the second-order regression and Gaussian correlation function are used to construct the KSM in this study.

3.2 Construction Principle of Orthogonal Latin Hypercube Sampling

The selection of sampling points is vital for evaluating the target objects and constructing the surrogate model. Since the sampling points should be selected to represent almost the entire design space and their number should be as less as possible to avoid redundant calculation, it is particularly significant to pick a sampling method with both orthogonality and uniformity.

Based on the ordinary Latin hypercube sampling (LHS) (Pebesma and Heuvelink, 1999), OLHS evenly divides the sampling space into N grids to ensure that there is at least one sample point in each grid, which effectively avoids the uneven distribution of sampling results when sampling fewer points by LHS and facilitates spatially balanced sampling. A comparison between LHS and OLHS distribution is given in **Figure 2D**.

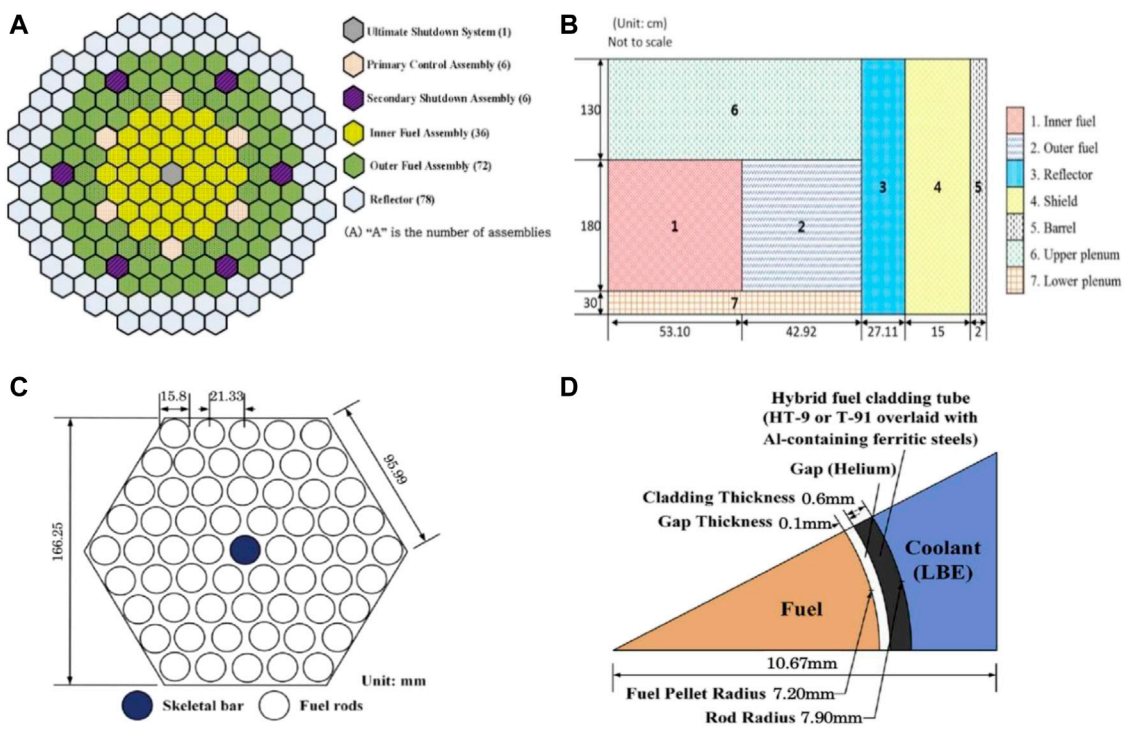


FIGURE 5 | Original core **(A,B)**, fuel assembly **(C)**, and fuel rod **(D)** of URANUS.

TABLE 6 | Original value and initial optimal range of design parameters for URANUS.

Design parameter	Original value	Initial optimal range
Fuel pin radius/cm	0.72	[0.4, 1.0]
Height of core active zone/cm	180	[50, 200]
Pitch to diameter ratio (P/D)	1.35	[1.01, 1.5]

3.3 Construction Principle of Space Exploration and Unimodal Region Elimination

A large number of statistical sampling and calculations are required to ensure the accuracy of KSM when the object to be optimized has a strong degree of non-linearity and a wide range of design variables. Thus, to improve the efficiency of optimized search, the SEUMRE algorithm is introduced to rapidly find the global optimum solution by updating the optimal region repeatedly and iteratively.

The flow diagram of SEUMRE algorithm is shown in **Figure 2E**. First, the design space is divided into different spaces after generating the training data in the initial design space and assessing the objects/constraints, and the most promising space is determined. Next, the surrogate model is constructed to predict the points and select the local optimal point to obtain the next promising space. Finally, the above steps are repeated for iterative search until all the spaces are traversed and the global optimal solution is acquired (Younis and Dong, 2010).

4 OPTIMIZATION VERIFICATION EXAMPLES AND RESULTS

The critical reactor core usually requires thousands of kilograms of fuel because of the small fuel fission cross section under fast-spectrum conditions. The fast-spectrum lead-bismuth reactor is generally heavy and expensive due to a large amount of fuel loading and high-density of lead-bismuth. Therefore, the minimum fuel loading of reactor core is considered as the optimization goal, and lead-bismuth reactors SPALLER-4 (independently developed by one of the authors) and URANUS (Lee, 2017; Kwak and Kim, 2018; designed by Seoul National University, South Korea) are taken as prototypes. The design parameters of the two reactors are listed in **Table 2**. Two optimization models are established after a certain simplification process. Different design variables and constraints are set for the two models to find the optimal reactor core scheme and optimal parameters through the intelligent optimization method (introduced in **Section 2**), and they are verified by the RMC code.

4.1 SPALLER-4 Model

The original structure of SPALLER-4 is shown in **Figure 3**. Considering the limited reactivity adjustment capability of the control rod system and the corrosive effect of lead-bismuth on materials, the SPALLER-4 optimization model can be described as follows:

$$min\, Fs\left(x_s, y_s, l_s, m_s, n_s\right) \text{ s.t. } \begin{cases} 1.005 \le K_{BOC} \le 1.030 \\ 1.0 \le K_{EOC} \\ 3\ EFPY = RC \\ 873\ K \le T_{c,max} \\ 1773\ K \le T_{f,max} \\ 3\ m/s \le V_{c,max} \end{cases} \tag{6}$$

where $min\, Fs(x_s, y_s, l_s, m_s, n_s)$ is the minimum fuel loading under the simultaneous influence of five variable parameters. The constraints of the model include K_{BOC} (K_{eff} at the beginning of the cycle), K_{EOC} (K_{eff} at the end of the cycle), RC (refueling cycle), $T_{c,\max}$ (maximum temperature of cladding), $T_{f,\max}$ (maximum temperature of fuel pellet), and $V_{c,\max}$ (maximum coolant velocity).

4.1.1 Sample Generation

Considering that the increase in the initial reactivity and breeding ability and the decrease in the neutron leakage and parasitic absorption can effectively reduce the fuel loading, and the size of core and lattice affects the fuel loading by changing the neutron energy spectrum and neutron leakage (Michael and Pavel, 2005; Zhang et al., 2020). And basing on the prototype value of the core and the design parameters' values of the classic liquid metal cooled reactor core such as ALFRED (Grasso et al., 2014), ELECTRA (Suvdantsetseg, 2012) and CEFR (Yang and Guo, 2020), the design variables and their ranges are set as shown in **Table 3**. 480 initial sample points are randomly sampled by

TABLE 7 | Accuracy verification results of KSM for predicting K_{eff} and burnup.

Contrast group	1	2	3	4	5
Fuel pin radius/cm	0.728,693	0.737,298	0.738,832	0.740,970	0.737,436
Height of core active zone/cm	164.311,938	157.445,331	156.993,296	153.933,096	157.438,697
Pitch to diameter ratio (P/D)	1.320,736	1.320,794	1.321,136	1.320,508	1.320,326
Twentieth year K_{eff}					
Prediction by KSM	1.000,957	1.000,382	1.000,515	0.999,350	1.000,614
Calculation by RMC	1.001,767	1.000,720	1.000,924	0.999,935	1.000,317
Relative error/%	−0.080,889	−0.033,774	−0.040,893	−0.058,523	0.029,713
Burnup/(MW·d·kg^{-1})					
Prediction by KSM	44.079,677	45.274,617	45.226,628	43.583,013	45.264,459
Calculation by RMC	44.410,000	45.271,000	45.213,000	43.554,000	45.256,000
Relative error/%	−0.743,804	0.007,989	0.030,142	0.066,614	0.018,690

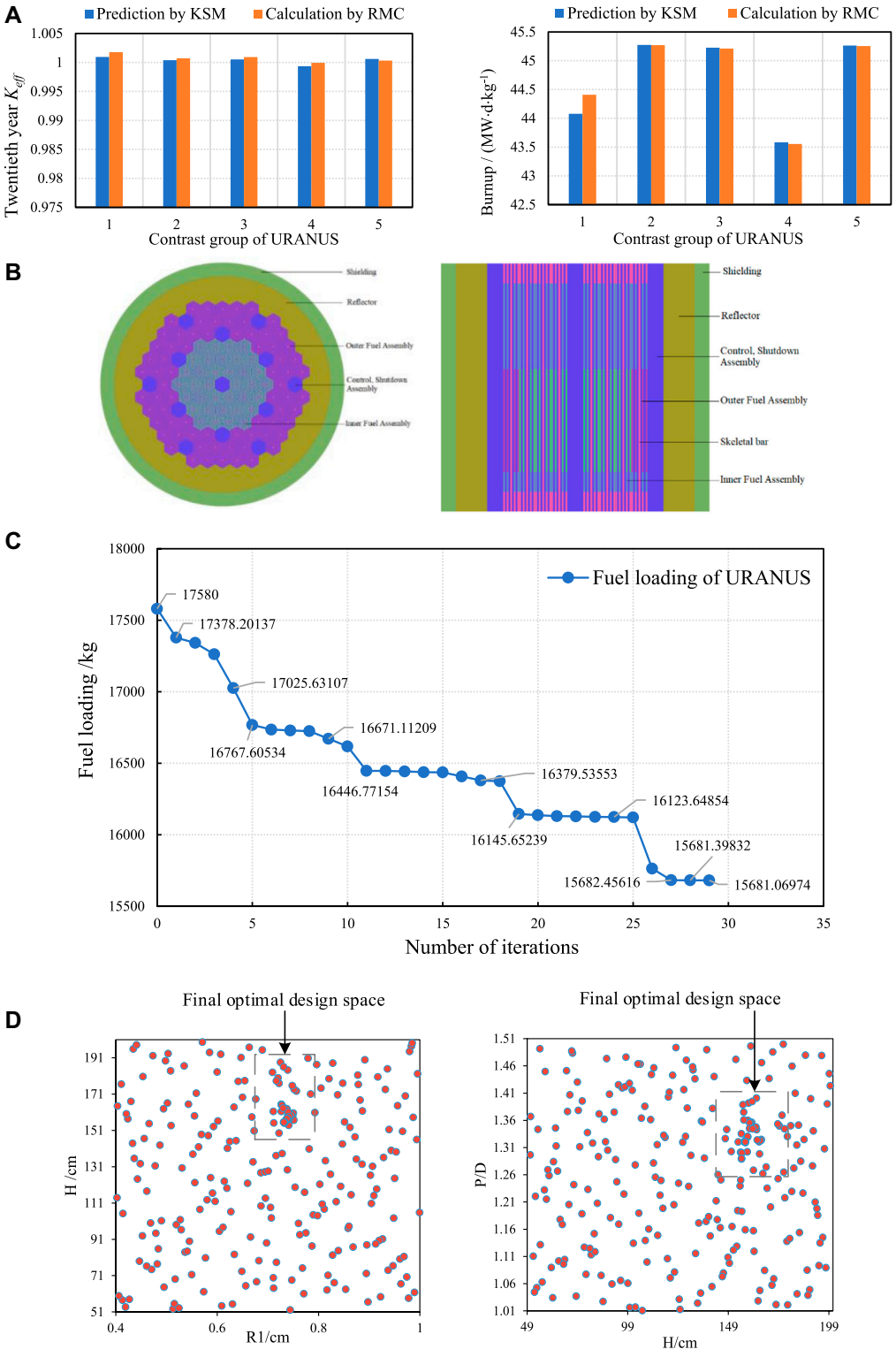


FIGURE 6 | Optimization verification results of URANUS. **(A)** Comparison of K_{eff} and burnup obtained using KSM and RMC code, **(B)** Core structure of URANUS for optimization verification, **(C).** Iterative graph of fuel loading optimization for URANUS, **(D)** Final optimal design space for URANUS.

TABLE 8 | Initial and final optimal range of design parameters for URANUS.

Design parameter	Initial optimal range	Final optimal range
Fuel pin radius/cm	[0.4, 1.0]	[0.7, 0.8]
Height of core active zone/cm	[50, 200]	[150, 185]
Pitch to diameter ratio (P/D)	[1.01, 1.5]	[1.28, 1.41]

OLHS in these ranges and calculated by physical and thermal calculation modules of DOPPLER-K.

4.1.2 Prediction Accuracy of Kriging Surrogate Model

To ensure the prediction accuracy of KSM, the accuracy of surrogate model is verified in each iterative optimization step. Five groups are randomly selected from the training set as the contrast group and are compared with the values calculated by RMC code, and the remaining groups are used as the training set to fit and construct the KSM.

Table 4 and **Figure 4A** show the accuracy verification results of KSM for predicting third year K_{eff} and burnup at full power operating conditions in the optimal design space, which contains 155 groups in the final training set. It can be seen that the relative errors between the predicted K_{eff} and burnup by KSM and the calculated values by RMC are within ±0.1% and ±1%, respectively. This indicates that when a certain amount of training data is ensured, KSM has a high accuracy for linear or nonlinear, directly or indirectly correlated target prediction, which is affected by multiple parameters at the same time. Compared with the RMC code, the calculation time is greatly reduced.

4.1.3 Optimization Results

According to the original design parameters, the core structure of SPALLER-4 (shown in **Figure 4B**) to be optimized is constructed after some simplification. After calculating 480 groups of initial training sets by the RMC code, the intelligent optimization method is used for iterative optimization. 1,060,000 groups of schemes are predicted by KSM in each iteration, and a total of 175 iterations are carried out. After verification by RMC, six groups of solutions meet the optimization constraints, and the final two iterative schemes meet the convergence conditions (shown in **Figure 4C**). In the entire optimization process, 655 groups of training data are calculated.

Table 5 shows the final optimal design scheme, which is verified by RMC. This scheme meets all the constraint limits of initial K_{eff}, refueling interval, steady-state thermal safety conditions, etc. The minimum fuel loading is 81.4135 kg, which is significantly lower than the original fuel loading. At the same time, the relative error between the predicted value by KSM and the calculated value by RMC is within the accuracy range, and the search speed of the target scheme is much higher than that of the traditional Monte Carlo calculation and the single heuristic optimization algorithm.

4.2 URANUS Model

To verify the feasibility of the proposed intelligent optimization method in practical engineering, a core optimization model is established based on the lead-bismuth reactor URANUS (shown in **Figure 5**). In this section, a three-dimensional and six-constraint optimization problem is solved with Ubiquitous, Robust, Accident-forgiving, Nonproliferating and Ultra-lasting Sustainer (URANUS) model for demonstrating the validity of the proposed model in parameter optimization. The optimization model can be described as follows:

$$min\, Fu\left(x_u, y_u, l_u\right) \text{ s.t. } \begin{cases} 1.005 \le K_{BOC} \le 1.035 \\ 1.0 \le K_{EOC} \\ 20\ EFPY = RC \\ 773\ K \le T_{c,max} \\ 1273\ K \le T_{f,max} \\ 3\ m/s \le V_{c,max} \end{cases} \tag{7}$$

TABLE 9 | Optimized design parameters for URANUS.

Design parameter	Initial value	Optimized value
Fuel pin radius/cm	0.72	0.731,415
Height of core active zone/cm	180	155.583,777
Pitch to diameter ratio (P/D)	1.35	1.289,290
Initial K_{eff}	1.028,859	1.030,741
Twentieth year K_{eff}		
Prediction by KSM	—	1.000,729
Calculation by RMC	1.003,136	1.000,958
Relative error/%	—	–0.022,884
Burnup/(MW·d·kg^{-1})		
Prediction by KSM	—	46.577,347
Calculation by RMC	41.524	46.553,000
Relative error/%	—	0.0522,990
Refueling cycle/EFPY	20	20
Fuel loading/kg	17,580.09248	15,681.069,740
Total mass of core (including reflector)/kg	175,459.3633	155,309.949,600
Volume of active zone/m^3	5.213,753	4.269,707
Average volume power density of active zone/(W·cm^{-3})	19.18,004	23.420,813
Total volume of core (including reflector)/m^3	8.573,414	7.105,878
Maximum temperature of fuel cladding/K	600.6,219	604.170,155
Maximum temperature of fuel pellet/K	770.3,892	796.058,922

where $min\,Fu(x_u, y_u, l_u)$ is the minimum fuel loading under simultaneous influence of three variable parameters. The constraints of the model include K_{BOC}, K_{EOC}, RC (refueling cycle), $T_{c,max}$ (maximum temperature of cladding), $T_{f,max}$ (maximum temperature of fuel pellet), and $V_{c,max}$ (maximum coolant velocity).

4.2.1 Sample Generation

Similar to the optimization steps of SPALLER-4 model, the optimization variables of URANUS and their range are listed in **Table 6**. Further, 216 groups of initial sample points are randomly sampled by OLHS in these ranges and calculated by physical and thermal calculation modules of DOPPLER-K.

4.2.2 Prediction Accuracy of Kriging Surrogate Model

The accuracy verification results of KSM for predicting 20th year K_{eff} and burnup at full power operating conditions in the optimal design area, which contains 45 groups of training sets, are shown in **Table 7** and **Figure 6A**. It is clear that the relative errors between the predicted K_{eff} and burnup by KSM and the calculated values by RMC are within ±0.1% and ±1%, respectively.

4.2.3 Optimization Results

According to the original design parameters, the core structure of URANUS (shown in **Figure 6B**) to be optimized is constructed after some simplification. After calculating 216 groups of initial training sets by the RMC code, the intelligent optimization method is used for iterative optimization. Further, 1060000 groups of schemes are predicted by KSM in each iteration, and a total of 37 iterations have been carried out. After verification calculation by RMC, 29 groups of solutions meet the optimization constraints, and the final two iterative schemes meet the convergence conditions (shown in **Figure 6C**). At this time, the final optimal range of design parameters is shown in **Figure 6D** and **Table 8**. Compared to initial design space, the range of optimal space is reduced. In the entire optimization process, 253 groups of training data are calculated.

Table 9 presents a comparison between the initial and optimized schemes. Compared with the initial model under the same constraint limits, the optimized core fuel loading, total mass, volume of active zone, and total volume are reduced by 1,899 kg, 20,149 kg, 0.944,047 m^3, and 1.467,536 m^3, respectively, and the reduction rates are 10.8, 11.5, 18.1, and 17.1%. This verifies the efficacy of the proposed intelligent optimization method for preliminary design and optimization of lead-bismuth reactor core.

5 CONCLUSION

In this study, an intelligent optimization method, which included KSM-OLHS-SEUMRE method and DOPPLER-K software, was developed for addressing the problems of multiple physical parameters, multiple variables, multiple constraints, a large amount of calculation, and low speed in the design optimization of lead-bismuth reactors. Lead-bismuth reactors: SPALLER-4 and URANUS, were used as verification examples to find the optimal scheme and optimal parameters. The entire calculation process could be automatically completed by the developed software. The main results of the study are summarized as follows:

1) KSM has a high prediction accuracy. According to the KSM prediction results for the two optimization examples, it can be concluded that when a certain amount of training data is given, the KSM has a high accuracy for the prediction of multi-constrained, linear or nonlinear, directly or indirectly related objective functions. Meanwhile, compared with the RMC code, which needed approximately 6.5 h to calculate a group of schemes, the KSM only needed more than 2 min to predict 1,060,000 groups of data under the same computer hardware conditions. Therefore, the KSM-based intelligent optimization method can greatly shorten the calculation time and improve the efficiency of preliminary design scheme search and optimization.
2) The proposed method is feasible, efficient, and effective for preliminary design and optimization of lead-bismuth reactor core. The optimal design scheme of SPALLER-4 was obtained using the five-variable and six-constraint model, which indicated that the intelligent optimization method is fast and efficient for the optimal scheme search under the influence of multi-factor coupling. The optimization of URANUS was conducted using a three-variable and six-constraint model, and the optimization rates of fuel loading, total core mass, active zone volume, and total core volume were found to be 10.8, 11.5, 18.1, and 17.1%, respectively, which suggested that the proposed method is effective for parameter optimization of the lead-bismuth reactor core.
3) Comparing the optimization results of SPALLER-4 model and URANUS model, it can be seen that the dimension of design variables will affect the training amount required to ensure the accuracy of Kriging surrogate model. The more design variables there are, the more training volume is required. While the number of objective/constraint functions does not affect the training number. In addition, to a certain extent, the search efficiency and optimization effect of the target scheme are also affected by the initial design variables, the initial optimal range and constraints. Therefore, the intelligent optimization method constructed in this study can better improve the optimization effect after reasonably and comprehensively selecting design variables and setting design scope.

Overall, to achieve different reactor design objectives, the proposed intelligent optimization method can rapidly identify the key design parameters and influence laws. Furthermore, through the rapid estimation of the reactor by this method, the optimal design space and preliminary design scheme of parameters matching the objectives can be obtained to provide reference for the detailed reactor

scheme. Thus, the proposed method opens a new route for the design optimization of miniaturized and lightweight lead-bismuth reactors in remote areas or marine environment in the future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

QL: Methodology, Software, Data curation, Verification, and Writing—original draft and revising. ZL: Conceptualization, Project administration, and Writing—reviewing and editing. YX: Verification. PZ: Writing—reviewing and Supervision. YZ: Writing—reviewing. TY: Resources. TY: Supervision.

FUNDING

This study is supported by the National Natural Science Foundation of China (Grant No. 12005097), Ministry of Equipment Development of Central Military Commission (Grant No. 6142A07190106), Natural Science Fund of Hunan Province (Grant No. 2020JJ5465), Department of Education of Hunan Province (Grant No. 9B494), and Science and Technology Innovation Team Project Fund of Hunan Province (Grant No. 2020RC4053).

ACKNOWLEDGMENTS

The authors would like to thank the Reactor Engineering Calculation and Analysis Laboratory (REAL team) of Tsinghua University which has independently developed the reactor Monte Carlo (RMC) analysis code, all the technology and fund supporters in this study and MJEditor (www.mjeditor.com) for its linguistic assistance during the preparation of this manuscript.

REFERENCES

Grasso, G., Petrovich, C., Mattioli, D., Artioli, C., Sciora, P., Gugiu, D., et al. (2014). The Core Design of ALFRED, a Demonstrator for the European lead-cooled Reactors. *Nucl. Eng. Des.* 278, 287–301. doi:10.1016/j.nucengdes.2014.07.032

Jin, R., Chen, W., and Simpson, T. (2000). Comparative Studies of Metamodeling Techniques under Multiple Modeling Criteria. *Struct. Multidisciplinary Optimization* 23 (1), 1–13. doi:10.2514/6.2000-4801

Kempf, S., Forget, B., and Hu, L.-W. (2012). Kriging-based Algorithm for Nuclear Reactor Neutronic Design Optimization. *Nucl. Eng. Des.* 247 (Jun), 248–253. doi:10.1016/j.nucengdes.2012.03.001

Khoshahval, F., Minuchehr, H., and Zolfaghari, A. (2011). Performance Evaluation of Pso and ga in Pwr Core Loading Pattern Optimization. *Nucl. Eng. Des.* 241 (3), 799–808. doi:10.1016/j.nucengdes.2010.12.023

Kim, K.-Y., and Lee, S.-M. (2009). Shape Optimization of Inlet Plenum in a PBMR-type Gas-Cooled Nuclear Reactor. *J. Nucl. Sci. Technol.* 46 (7), 649–652. doi:10.1080/18811248.2007.9711571

Kwak, J., and Kim, H. R. (2018). Development of Innovative Reactor-Integrated Coolant System Design Concept for a Small Modular lead Fast Reactor. *Int. J. Energ. Res* 42 (13), 4197–4205. doi:10.1002/er.4177

Lee, J. (2017). *Conceptual Neutronic Design of Inverted Core for Lead-Bismuth Cooled Small Modular Reactor*. Graduate School of Seoul National University. Available at: : https://s-space.snu.ac.kr/handle/10371/123541.

Meneses, A. A. d. M., Machado, M. D., and Schirru, R. (20092009). Particle Swarm Optimization Applied to the Nuclear Reload Problem of a Pressurized Water Reactor. *Prog. Nucl. Energ.* 51, 319–326. doi:10.1016/j.pnucene.2008.07.002

Michael, J. D., and Pavel, H. (2005). Reactor Physics Challenges in Gen-Iv Reactor Design. *Nucl. Eng. Techn.* 37 (1), 1–10. Available at: https://www.researchgate.net/publication/229009639.

Pebesma, E. J., and Heuvelink, G. B. M. (1999). Latin Hypercube Sampling of Gaussian Random Fields. *Technometrics* 41 (4), 303–312. doi:10.1080/00401706.1999.10485930

Song, L., Yang, H., and Zhiwei, Z. (2018). Fast Reactor Fuel Assembly Design Based on the Kriging Method. *At. Energ. Sci. Techn.* 052 (007), 1288–1293. doi:10.7538/yzk.2017.youxian.0650

Suvdantsetseg, E. (2012). Design Studies of ELECTRA: European Lead-Cooled Training Reactor. *Other Phys. Top.* 177 (3), 303–313. doi:10.3109/09553002.2012.638357

Wang, K., Li, Z., She, D., Liang, J. g., Xu, Q., Qiu, Y., et al. (2015). RMC - A Monte Carlo Code for Reactor Core Analysis. *Ann. Nucl. Energ.* 82 (10), 121–129. doi:10.1016/j.anucene.2014.08.048

Wu., Y. (2018). Research Progress and Application Prospects of lead-based Reactors. *Mod. Phys. knowledge* 30 (04), 37–41.

Yang, H., and Guo, M. (2020). Design Innovation and Realization of Experimental Fast Reactors in China. *At. Energ. Sci. Techn.* 54 (S01), 199–206. doi:10.7538/yzk.2020.zhuankan.0432

Younis, A., and Dong, Z. (2010). Metamodelling and Search Using Space Exploration and Unimodal Region Elimination for Design Optimization. *Eng. Optimization* 42 (6), 517–533. doi:10.1080/03052150903325540

Zameer, A., Mirza, S. M., and Mirza, N. M. (2014). Core Loading Pattern Optimization of a Typical Two-Loop 300MWe PWR Using Simulated Annealing (SA), Novel Crossover Genetic Algorithms (GA) and Hybrid GA(SA) Schemes. *Ann. Nucl. Energ.* 65 (mar), 122–131. doi:10.1016/j.anucene.2013.10.024

Zeng, K., Stauff, N. E., Hou, J., and Kim, T. K. (2020). Development of Multi-Objective Core Optimization Framework and Application to Sodium-Cooled Fast Test Reactors. *Prog. Nucl. Energ.* 120 (2020), 103184–184. doi:10.1016/j.pnucene.2019.103184

Zhang, Y. (2014). *Global Sensitivity Analysis of Multi-Parameter Nonlinear System and Research on Dynamic Surrogate Model*. Hunan University.

Zhang, Y., Wang, C., Lan, Z., Wei, S., Chen, R., Tian, W., et al. (2020). Review of Thermal-Hydraulic Issues and Studies of Lead-based Fast Reactors. *Renew. Sust. Energ. Rev.* 120 (C), 109625. doi:10.1016/j.rser.2019.109625

Zhao, P., Liu, Z., Tao, Y., Jinsen, X., Zhenping, C., and Shen, C. (2020). Code Development on Steady-State thermal-hydraulic for Small Modular Natural Circulation lead-based Fast Reactor. *Nucl. Eng. Techn.* 52 (12), 2789–2802. doi:10.1016/j.net.2020.05.023

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 06 April 2022
doi: 10.3389/fenrg.2022.837823

Denoising Method of Nuclear Signal Based on Sparse Representation

*San-Jun He, Na Sun, Ling-Ling Su, Bin Chen and Xiu-Liang Zhao**

School of Nuclear Science and Technology, University of South China, Hengyang, China

Nuclear signals are sensitive to noise which may affect final monitoring results significantly. In order to suppress the nuclear signal noise, a sparse representation method, which is based on the sparse representation of signals and a matching pursuit algorithm, has been proposed for denoising. Time–frequency matching "atoms" have been selected for building an over-complete library by training atoms matching with the characteristics of nuclear signals regardless of the noise. The best time–frequency matching atoms have been extracted by sparsely representing the noisy signals with an Orthogonal Matching Pursuit (OMP) algorithm and the library. The residual ratio threshold has been chosen as a stopping criterion in the OMP algorithm for avoiding the influence of improper selection of iterations on denoising results. At the end, the pulse matching the atom extracted by each iteration has been optimized by performing effective sparse representation on the original noiseless nuclear signal component in noisy nuclear signals. The proposed method has been used to denoise the simulated and measured signals and has been compared with the nuclear denoising result using traditional wavelet theory. The results show that the proposed method can accurately suppress the noise interference of nuclear signals, and the denoising effect is better than that of the traditional wavelet method.

Keywords: sparse representation, nuclear signal, signal processing, denoising method, noise reduction

OPEN ACCESS

Edited by:
Xingang Zhao,
Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Zhiyuan Zha,
Nanyang Technological University, Singapore
Paolo Mercorelli,
Leuphana University, Germany
Yong Lv,
Wuhan University of Science and Technology, China

****Correspondence:***
Xiu-Liang Zhao
zhaoxiul@163.com

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *17 December 2021*
Accepted: *04 March 2022*
Published: *06 April 2022*

Citation:
He S-J, Sun N, Su L-L, Chen B and Zhao X-L (2022) Denoising Method of Nuclear Signal Based on Sparse Representation. Front. Energy Res. 10:837823. doi: 10.3389/fenrg.2022.837823

INTRODUCTION

Nuclear radiation detection is one of the key technologies in nuclear analysis and also the fundamental of nuclear science and technology and has been widely applied in the operation of nuclear power and its safety, uranium mining and metallurgy, environmental monitoring and radiation protection, homeland security and nuclear non-proliferation, industrial buildings and radiation imaging, and other fields. The nuclear information, such as energy, time, and position, is usually converted into voltage in visual and thus can be expediently analyzed and used for obtaining valuable information (KNOLL, 2000). As the nuclear signals are usually very weak and there is large electromagnetic noise in a practical environment, therefore, noise suppression has become a problem that must be solved in the analysis and processing of nuclear signals in nuclear measurement systems (Williams, 2005; Hashemian and Bean, 2011; To-Po Wang and Zong-Wei Li, 2014; Min et al., 2015).

Sparse representation of signals is a method to represent the original signal as accurately as possible by less specific information, first proposed by Mallet and Zhang in 1993 (Mallat and Zhifeng Zhang, 1993), in which the signal can be represented by a linear superposition of different basis signals, and the set of these basis signals is called a dictionary, and the basis signals are called atoms. The sparse representation of signals is based on the adaptive selection of a small number of atoms to represent the signal with full consideration of the signal characteristics, and because the method does not require the atoms to have orthogonality, the selection of atoms in the dictionary is flexible and

can better characterize the signal and retain the frequency- and time-domain information of the original signal to the maximum extent, and it has many advantages such as a wide range of signal representation, concise representation method, and strong representation capability (Mallat and Zhifeng Zhang, 1993; Zhang et al., 2017; Othmen et al., 2021; Shi et al., 2021), which has been more often applied in the fields of image restoration, image denoising, and signal recognition (Deeba et al., 2020; Maqsood and Javed, 2020; Alotaibi, 2021; Balnarsaiah and Rajitha, 2021). Sparse representation methods include both sparse decomposition algorithms and construction of over-complete atomic dictionaries, and the main sparse decomposition algorithms are convex optimization algorithms and greedy algorithms. Typical convex optimization algorithms include Basis Pursuit (BP) (Ekanadham et al., 2011), and in 2015, Selesnick et al. constructed convex sparse representation models based on parametric non-convex functions and proposed corresponding sparse decomposition algorithms, which achieved superior performance in signal denoising and other aspects (Selesnick et al., 2014; Parekh and Selesnick, 2015). Convex optimization algorithms have a more rigorous mathematical optimization solution process, and compared with greedy algorithms, convex optimization algorithms can find sparser or more accurate solutions, but the computational complexity is high and will take a lot of time. Typical greedy algorithms include Matching Pursuit (MP) (Mallat and Zhifeng Zhang, 1993), Orthogonal Matching Pursuit (OMP) (Yi and Song, 2015), which is developed on the basis of MP, Regularized Orthogonal Matching Pursuit (ROMP) (Sajjad et al., 2015), Sparsity Adaptive Matching Pursuit (SAMP) (Wang et al., 2020), Compressive Sampling Matching Pursuit (CoSaMP) (Huang et al., 2017), Subspace Pursuit (SP) (Li et al., 2015), and other methods, and all of them can achieve sparse signal reconstruction very well. Matching Pursuit class algorithms are commonly used for image sparse representation, and Rubinstein's team (Rubinstein et al., 2008) used Batch Orthogonal Matching Pursuit (Batch-OMP) to achieve fast noise reduction and sparse representation processing of image signals. Greedy class algorithms with mature theory, low complexity, and fast running speed are widely used for signal sparse decomposition. The construction of over-complete dictionaries mainly includes conformal dictionaries and learning dictionaries. Constructed dictionaries are constructed by analyzing the signal feature structure, using parametric wavelets as atoms, and obtaining a large number of different atoms by changing parameters, such as the Gabor time–frequency atom dictionary (Mallat and Zhifeng Zhang, 1993) and the chirplet time–frequency atom dictionary (Mann and Haykin, 1995). The learned dictionaries are mainly learned from training samples and have good adaptability, but in application scenarios where the signal interference noise is relatively strong, the learned dictionaries may not be optimal and do not perform well for sparse representations of other signals of the same type; moreover, dictionary learning algorithms are generally high in complexity and are not suitable for dealing with large-scale datasets. The current typical dictionary learning methods include the method of optimal directions (MOD) (Engan et al., 2000), K-SVD (Aharon et al., 2005), and online dictionary learning (ODL) (Celik and Bilge, 2017). At the moment, sparse decomposition's application and research in nuclear signal processing is still in its infancy. In 2011, Trigano T et al. (Trigano et al., 2011) conducted a study on activity estimation of radioactive source based on the sparse representation of signals method and investigated the efficiency of this approach on simulation and real datasets. And also in 2018, Zhang (Zhang et al., 2018) investigated the rapid and effective extraction method of nuclear pulse signals based on the sparse representation method.

In this paper, a sparse representation method has been applied for denoising nuclear signals. As nuclear signal matching atoms, Gabor time–frequency atoms and chirplet time–frequency atoms, which can accurately correlate with the characteristics of the original nuclear signal, were first produced. Gabor and chirplet atoms exhibit good time–frequency aggregation, according to the uncertainty principle, and the nuclear signal is a type of uncertain signal with unpredictable time and amplitude. The time–frequency features of the nuclear signal can be completely revealed utilizing the sparse representation of the signal generated using the Gabor dictionary and chirplet dictionary. Then, the Orthogonal Matching Pursuit (OMP) algorithm was applied for searching the best matched atom in the noisy nuclear signals from the over-complete library composed of time–frequency atoms, and the threshold of residual ratio was taken as the stopping criterion of OMP algorithm. Because the matched atoms obtained from each iteration can only effectively sparse represent the original nuclear signals without noise components, the aim of nuclear signal denoising can be achieved. In this work, the above methods have been used to denoise the simulated and measured nuclear signals, respectively (Chen et al., 2009; Zhou et al., 2011). The results prove that the method proposed in this paper is more effective and superior compared with the traditional wavelet denoising method.

ORIGINAL RESEARCH ARTICLE

Sparse Representation Theory of Signals

Any signal $f \in R^N$ can be represented as a linear combination of atoms $(\phi_\gamma(t))_{\gamma\in\Gamma}$ in the dictionary D:

$$f = \sum_{\gamma\in\Gamma} \alpha_\gamma \phi_\gamma \ , \tag{1}$$

where α_γ is the expansion coefficient. Since the dictionary D is non-orthogonal and over-complete, a signal f has various possible representations in the dictionary D. Solving the sparse signal representation coefficient in a certain atomic dictionary D is equivalent to solving the following optimization problem:

$$(P_0): \ \min \|x_0\| \ \ s.t. \ y = Dx, \tag{2}$$

where $\|x\|_0 = |\{i: x(i) \neq 0\}|$ is the number of non-zero terms in the coefficient vector x.

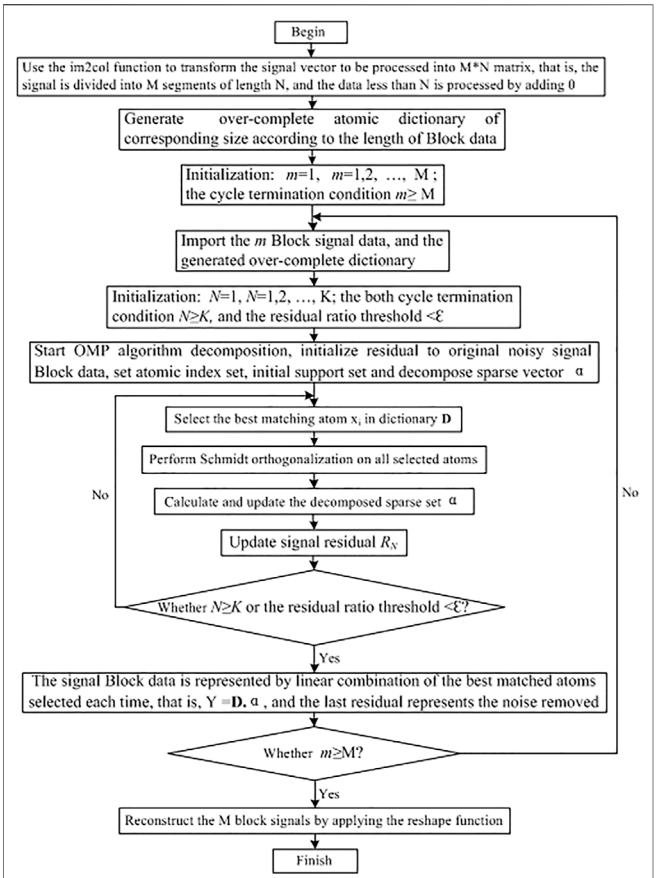


FIGURE 1 | Flow chart of the sparse decomposition algorithm based on block processing.

The denoising of the nuclear signals infected by noise can be composed of two parts, namely, the original nuclear signals without noise and the noise signal, and its mathematical model can be expressed as (SARKAR et al., 2012)

$$f = f_p + f_n, \tag{3}$$

where f is the nuclear signal infected by noise, f_p is the original nuclear signal without noise, and f_n is the noise signal.

Basically, f_p has a particular structure, but f_n does not, or f_p and f_n have different structures. If there is some kind of atom P_γ, and its atomic structure is related to f_p and unrelated or has weak correlation to f_n, when sparse representation based on the decomposition algorithm is done to f_p in an over-complete dictionary of atoms composed of P_γ atoms, the inner product of atoms and f_p must be greater than that of atoms and f_n. Based on the rational of MP algorithm, the first extracted nuclear signal must be the original nuclear signal without noise and can be formulized as

$$f = \sum_{k=1}^{k} < R^k f, p_\gamma^k > p_\gamma^k + R^{k+1} f, \tag{4}$$

where $\sum_{k=1}^{k} < R^k f, p_\gamma^k > p_\gamma^k$ is the original nuclear signal without noise, $R^{k+1} f$ represents the noise signal, and p_γ^k denotes matched atoms for nuclear signals.

Sparse Representation and Decomposition Algorithm

Orthogonal Matching Pursuit Signal Sparse Decomposition Algorithm

The process of Matching Pursuit algorithm (Mallat and Zhifeng Zhang, 1993) is presented as follows: Firstly, the atom x_{r_0} that best matches the signal y to be decomposed is selected from the over-perfect dictionary to satisfy the following conditions:

$$\left| < y , x_{r_0} > \right| = sup \left| < y , x_{r_i} > \right|. \tag{5}$$

The signal can be decomposed into components and residuals on the optimal atom:

$$y = < y , x_{r_0} > x_{r_0} + R_1, \tag{6}$$

where R_1 is the residue after the optimal matching of the original signal with the optimal atom. The same decomposition process above can be carried out for the residue after the best matching:

$$R_t = < R_t , x_{r_t} > x_{r_t} + R_{t+1}. \tag{7}$$

After the T step decomposition, the signal is decomposed into

$$y = \sum_{t=0}^{T-1} < R_t , g_{r_t} > g_{r_t} + R_T. \tag{8}$$

And a small number of atoms can represent the main component of the signal, namely,

$$y \approx \sum_{t=0}^{T-1} < R_t , g_{r_t} > g_{r_t}. \tag{9}$$

From **Equation 6**, it can be seen that the sparse decomposition of signals by the matching tracking algorithm is a continuous iterative process. Without limiting the residual energy threshold and decomposition iteration times, signals can be decomposed indefinitely on a fixed atomic dictionary.

The OMP algorithm uses the Gram–Schmidt orthogonalization method to normalize the matched atoms x_{r_t} at each step of MP decomposition (Yi and Song, 2015), which can not only accelerate the convergence rate but also avoid introducing unnecessary components when residual errors are projected on the atoms x_{r_t}. The specific process assumes $u_0 = x_{r_0}$, the most matching atom x_{r_t} is selected according to **Equation 7**, and then x_{r_t} is normalized:

$$u_t = x_{rt} - \sum_{t=0}^{N-1} \frac{< x_{rt}, u_{t-1} >}{u_{t-1}^2} u_{t-1}. \tag{10}$$

After N iterations, the system output signal y is decomposed:

$$y = \sum_{t=0}^{N-1} \frac{< x_{rt}, u_{t-1} >}{u_{t-1}^2} u_{t-1} + R_N. \tag{11}$$

The Nuclear Signal Over-Complete Atomic Dictionary

The Gabor Atomic Over-Complete Dictionary

According to the principle of sparse representation, the sparse representation of a signal can be achieved in any over-complete atomic dictionary (Mallat and Zhifeng Zhang, 1993). From the point of view of obtaining a better sparse representation of the signal, the over-complete dictionary of atoms chosen or

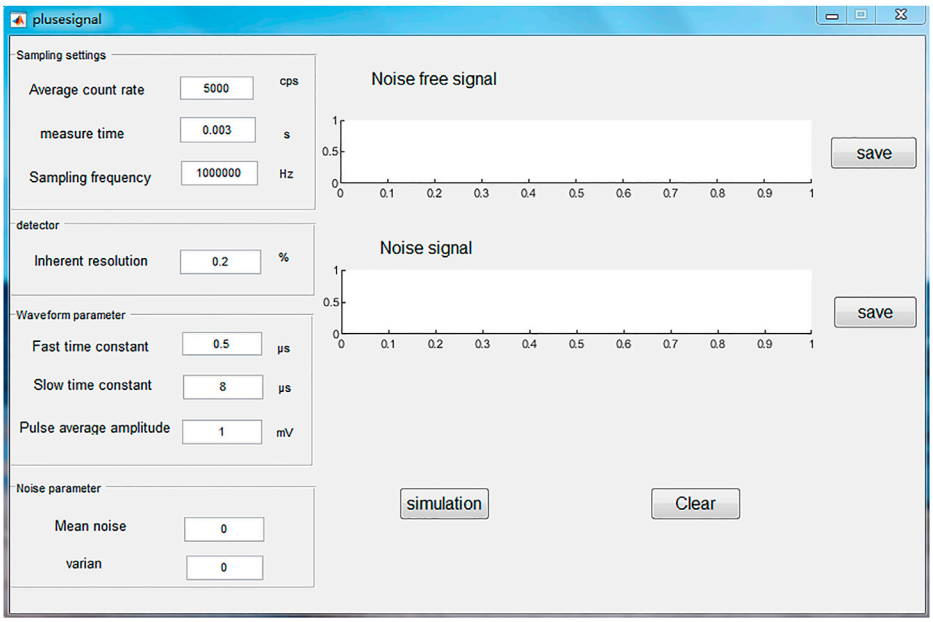


FIGURE 2 | Simulation GUI interface of the nuclear signal.

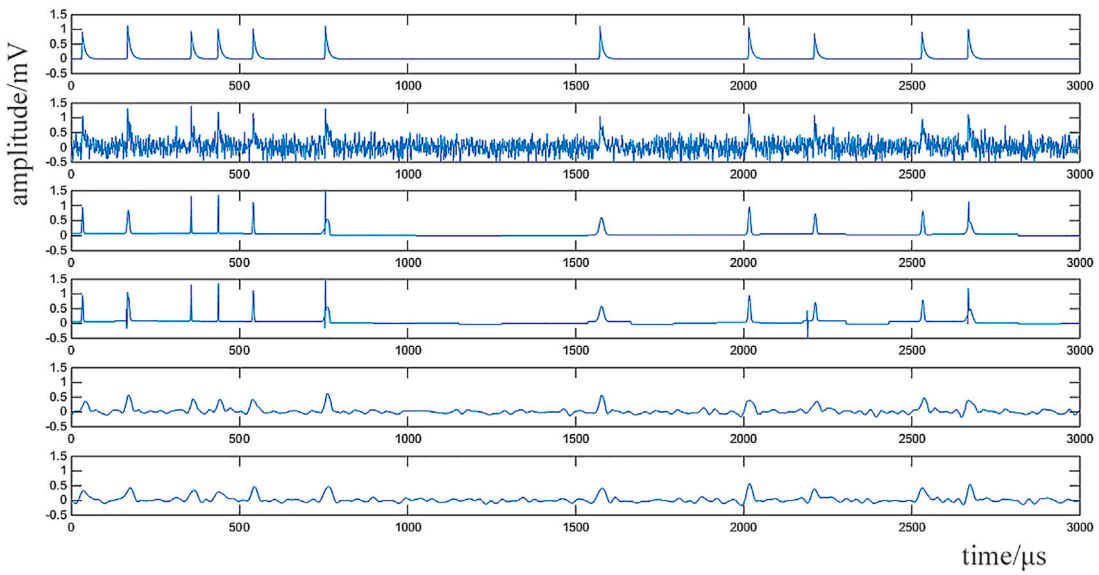


FIGURE 3 | Denoising results of the nuclear pulse simulation signal.

constructed should match as closely as possible the intrinsic structure and properties of the signal, so that as few atoms as possible can be used for the representation, and the representation results will be more sparse. In addition, in order to better describe the time-varying characteristics of non-stationary nuclear signals from the perspective of time–frequency analysis, the atoms in the dictionary should have good resolution in both the time domain and the frequency domain. In this study, a Gabor atom is firstly used to construct an over-complete dictionary due to the best time–frequency aggregation. It is expressed as follows:

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) cos(vt + w), \tag{12}$$

TABLE 1 | Nuclear signal sparse decomposition algorithm based on block processing.

Input: Raw nuclear signal vector y
Processing: Use the im2col function to transform the signal vector to be processed into an M*N matrix, that is, the signal is divided into M segments of length N, and the data less than N are processed by adding 0
Task: Select the corresponding time–frequency atoms according to nuclear signal characteristic construction and generate an over-complete atomic dictionary of corresponding size according to the length of Block data
Initialization: Set the current cycle number m = 1, the maximum cycle number m = M, m = 1,2, …, M, and the cycle termination condition $m \geq$ M
Repeat steps 1 to 6
1. Input: Import the m Block signal data and the generated over-complete dictionary
2. Initialization: Set the cumulative number of stopping updates of the current iteration N = 1 and the maximum number of stopping iterations N=K. And set both $N \geq$ K and the residual ratio threshold
3. $q(R_{N-1}) = \frac{R_N - \xi R_{N-1}}{\xi R_{N-1}} < \varepsilon$, N = 1,2, …, K, as the calculation iteration termination condition.
4. Start OMP algorithm decomposition, initialize residual to original noisy signal Block data, set the atomic index set, initialize support set, and decompose sparse vector $\boldsymbol{\alpha}$
5. Repeat steps 1 to 5
1) The atom x_{r_0} that best matches the signal Block data y is selected from the over-complete dictionary to satisfy the following conditions: $\lvert < y, x_{r0} > \rvert = \sup \lvert < y, x_{ri} > \rvert$
2) Schmidt orthogonal processing for all selected atoms: set $u_0 = x_{ro}$ $u_t = x_{rt} - \sum_{t=0}^{N-1} \frac{< x_{rt}, u_{t-1} >}{u_{t-1}^2} u_{t-1}$
3) Sparse is decomposed by the following calculation: $y = \sum_{t=0}^{N-1} \frac{< x_{rt}, u_{t-1} >}{u_{t-1}^2} u_{t-1} + R_N$ and the sparse vector $\boldsymbol{\alpha}$ is updated
4) Update signal residual R_N $R_N = R_{N-1} - \frac{< x_{rN-1}, u_{N-2} >}{u_{N-2}^2} u_{N-2}$, N = 2,3. . .,K
5) Determine if termination conditions are met: $N \geq K$ or $q(R_{N-1}) = \frac{R_N - \xi R_{N-1}}{\xi R_{N-1}} < \varepsilon$, $\xi = \sqrt{E[(R_N)^2]/E[(R_{N-1})^2]}$ If one of the above conditions is satisfied, stop the iteration; if not, set N=N+1 and return to step 1
6. Processing: Signal Block data are represented by a linear combination of the best matched atoms selected each time, that is, Y = **D.α**, and the last residual represents the noise removed
7. Determine whether to meet the loop termination condition: $m \geq$ M; if satisfied, stop the loop and execute the next step; if not, set m = m+1 and return to step 1
Processing: Use the reshape function to reorganize the M-segment Block data to reconstruct the complete signal

TABLE 2 | Calculation results of denoising effect evaluation indexes.

Index	Method	1	2	3	4	5	6	7	8	9	10	Overall
RMSE	Signal with noise	0.0373	0.0369	0.0397	0.0419	0.0386	0.0382	0.0408	0.0397	0.041	0.0412	0.0395
	Chirplet	0.0068	0.0059	0.0062	0.006	0.0059	0.0073	0.0071	0.0057	0.0066	0.0067	0.0064
	Gabor	0.0094	0.007	0.0069	0.0055	0.0075	0.0112	0.0076	0.0073	0.0088	0.0068	0.0078
	Db4	0.0091	0.0091	0.0103	0.0082	0.0091	0.0099	0.0095	0.0116	0.0102	0.0124	0.0099
	Db8	0.0089	0.0105	0.0095	0.0084	0.0098	0.0117	0.0106	0.0108	0.0109	0.0112	0.0102
NCC	Signal with noise	0.5901	0.6131	0.5848	0.5405	0.5977	0.6523	0.6073	0.6137	0.6009	0.6325	0.6033
	Chirplet	0.8243	0.8772	0.843	0.8246	0.8609	0.8677	0.8538	0.8808	0.8524	0.8712	0.8556
	Gabor	0.7361	0.8494	0.8157	0.8351	0.8101	0.7766	0.8329	0.8431	0.7914	0.8695	0.8160
	Db4	0.7467	0.7977	0.7174	0.7393	0.7686	0.8062	0.7878	0.7306	0.7555	0.7363	0.7586
	Db8	0.7511	0.7603	0.7423	0.7288	0.7473	0.7681	0.7597	0.7509	0.737	0.7669	0.7512
SNR	Signal with noise	14.29	14.02	14.02	13.78	14.14	14.18	13.90	14.01	13.88	13.85	14.01
	Chirplet	21.66	22.27	22.09	22.19	22.30	21.38	21.49	22.47	21.82	21.71	21.94
	Gabor	20.29	21.55	21.58	22.16	21.25	19.53	21.12	21.36	20.53	21.67	21.16
	Db4	20.40	20.42	19.87	20.88	20.39	20.05	20.22	19.34	19.93	19.05	20.05
	Db8	20.51	19.79	20.22	20.75	20.10	19.33	1976	19.67	19.61	19.50	19.92

where $g(t) = e^{-\pi t^2}$ represents a Gaussian window function, $\gamma = (s, u, v, w)$ is the atomic time–frequency parameter, s is the scaling factor, u is the shift factor, v is the frequency factor, and w is the phase factor. Signal sparse representation requires high redundancy and enough diversity of atomic dictionary. To meet the design requirements, we discretize the atomic time–frequency parameters into

$$\gamma = (a^j,\ pa^j \Delta u,\ ka^{-j} \Delta v,\ i\Delta w). \tag{13}$$

Here, $a = 2$, $\Delta u = 1/2$, $\Delta v = \pi$, $\Delta w = \pi/6$, $0 < j \ll log_2 N$, $0 \ll P \ll N2^{-j+1}$, $0 \ll k \ll 2^{j+1}$, $0 \ll i \ll 12$, and N represents the number of sampling points of a frame signal processed.

The Chirplet Atomic Over-Complete Dictionary

The chirplet atom (Mann and Haykin, 1995) is the most widely used atom after the Gabor atom. On the principle of Gabor atom, and then the chirplet atom, a one-dimensional frequency modulation parameter is added, which makes the chirplet

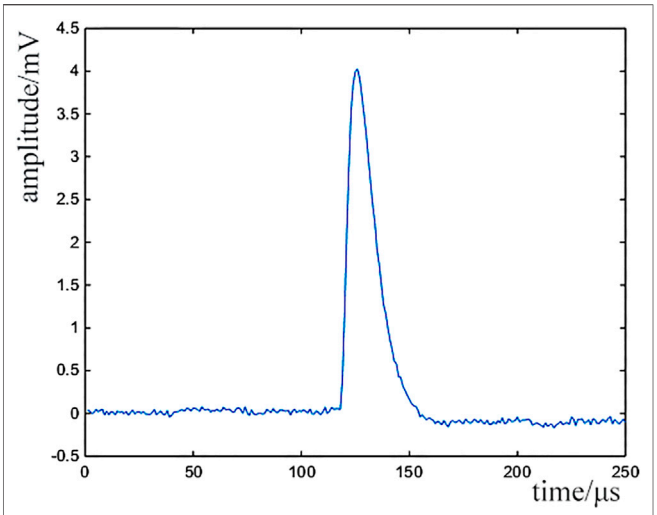


FIGURE 4 | Measured nuclear pulse signals in the laboratory.

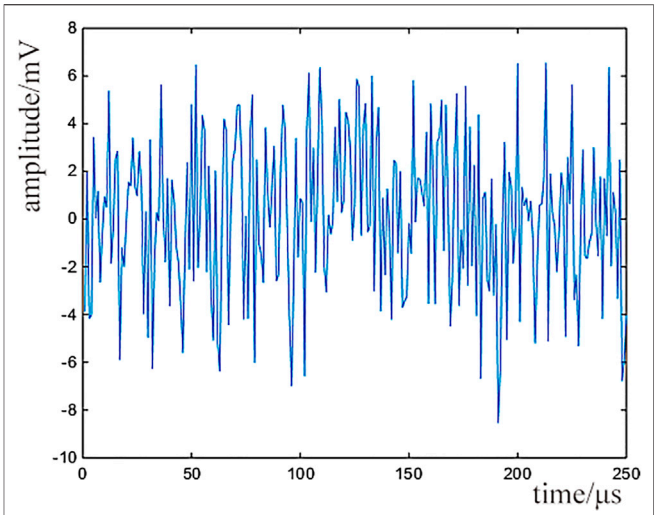


FIGURE 5 | Measured nuclear pulse signals with noise superimposed in the laboratory.

atom have a good matching effect on the linear frequency modulation signal. It can be expressed as follows:

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) exp\left(j\left(\xi(t-u) + \frac{1}{2}c(t-u)^2\right)\right), \quad (14)$$

where $g(t) = e^{-\pi t^2}$ represents a Gaussian window function, $\gamma = (s, u, \xi, c)$ is the atomic time–frequency parameter, s is the telescopic scale, u is the shift factor, ξ is the modulation factor, namely, frequency center, and c is the linear frequency modulation factor responding signal frequency over time. The real part of the time–frequency atom can be expressed as

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) cos\left(\xi(t-u) + \frac{1}{2}c(t-u)^2\right). \quad (15)$$

According to the optimal discretization method, the atomic parameter set γ is discretized:

$$\begin{aligned} \gamma &== (s, u, \xi, c, \omega) \\ &= (a^j, pa^j \Delta u, ka^{-j} \Delta\xi, la^{-2j} \Delta c, i\Delta w), \end{aligned} \quad (16)$$

in which $a = 2$, $\Delta u = 1/2$, $\Delta\xi = \pi$, $\Delta w = \pi/6$, $0 < j \ll log_2 N$, $0 \ll P \ll N2^{-j+1}$, $0 \ll k < 2^{j+1}$, $0 \ll l < 2^{j+1}$, $0 \ll i \ll 12$, and N represents the number of sampling points of a frame signal processed.

The Termination Conditions of Residual Threshold

The iterative termination conditions of the OMP algorithm are mainly composed by the hard and soft threshold methods. The former refers to the fixed iteration termination number K, and the original signal was replaced with the linear combination of K original signals. This method is simple but has the flaw that the K value is difficult to determine accurately. When K is too small, the original noiseless signal component will be lost, while the noise component will be introduced in reverse with a very large K. On the contrary, the soft threshold method holds that the iteration is terminated when the residual signal is less than a certain threshold. Its denoising effect is fine when the signal-to-noise ratio is high, whereas at low signal-to-noise ratios, a larger noise component will impact on the judgment of the residual error threshold. Thus, no matter how many times iteration was done, the residual all cannot reach the specified threshold. In addition, when the number of iterations is too much, a noise component will further be introduced, which also will influence the denoising effect. A termination condition of the residual ratio threshold of the signal denoising was introduced in the study of Liang and Que (2010) based on the MP theory, which avoids the influence on the judgment of the threshold of the residual ratio when the noise energy is large; as a consequence, the noise disturbance was reduced, and the robustness of the sparse representation was improved. Taking $R^k f$ and $R^{k+1} f$, respectively, as the k-th and k+1-th residuals, the residual error ratio is

$$q(R^K f) = \frac{\left\| R^{K+1} f - \xi R^K f \right\|}{\xi R^K f}, \quad (17)$$

where $\xi = \sqrt{E[(R^{k+1} f)^2]/E[(R^k f)^2]}$ and E(·) denotes the expectation value.

Steps of Nuclear Signal Denoising Method Based on Sparse Representation

The steps of the nuclear signal denoising method based on sparse representation are presented as follows:

1) The number of dictionary contents will be huge, when the length of the signal to be processed is large enough for the size of the over-complete dictionary used in sparse decomposition which depends on the length of the signal to be processed. To solve this problem, the collected data sequence is divided into blocks. A Block is a segment of an entire data sequence (Rubinstein et al., 2008).

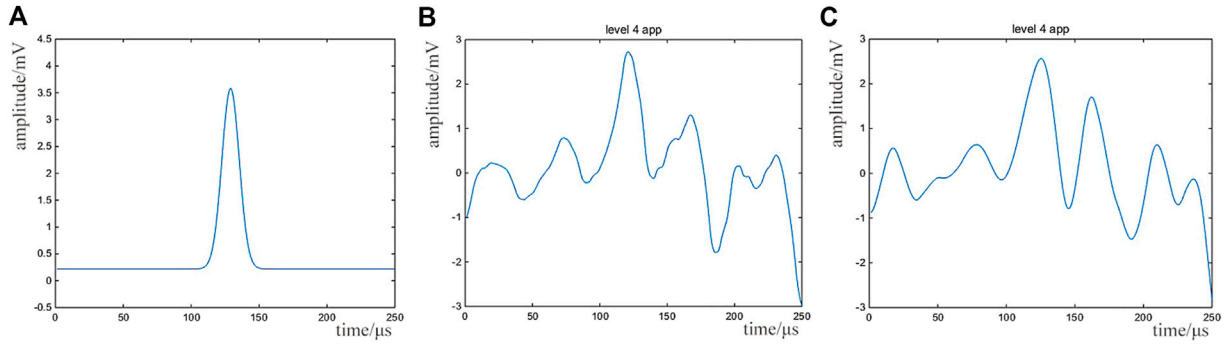


FIGURE 6 | Denoising result of nuclear pulse signals measured in the laboratory: **(A)** denoising result based on sparse representation; **(B)** denoising result by wavelet morphology–wavelet method (Db4); **(C)** denoising result by wavelet morphology–wavelet method (Db8).

TABLE 3 | Evaluation indexes of denoising effect of measured nuclear signals in the laboratory.

Denoising method	RMSE	NCC	SNR
Before denoising	10.4730	0.1760	−10.2007
Method in this paper	0.0971	0.9275	10.1279
Wavelet Db4	0.8023	0.4989	0.9568
Wavelet Db8	0.6742	0.5787	1.7121

2) The corresponding time–frequency atoms are constructed according to the characteristics of the nuclear signal, and the over-complete atomic dictionary is generated according to the length of the Block.
3) For each Block, OMP decomposition is performed separately. Set the cumulative number of stopping iterations $N = 1$, and set the maximum allowed number of stopping iterations to K.
4) According to the threshold condition of residual ratio, do determination when the OMP calculation iteration termination condition is satisfied. If not, return to step 3); if so, a Block denoising ends.
5) Each Block is processed separately and then spliced.

Flow chart of the sparse decomposition algorithm based on block processing is shown in **Figure 1**, and the detailed algorithm is shown in **Table 1**.

Simulation Verification and Analysis

Nuclear Signal Simulation

The simulation experiment is designed to verify the effectiveness of the method based on sparse representation. According to the random statistical law of nuclear event (Bertuccio and Pullia, 1993; Georgiev and Gast, 1993), the waveform shape, amplitude, adjacent pulse time interval, and system interference noise characteristics of the nuclear signal are statistically described, and then the simulated nuclear signal is generated on this basis.

1) The mathematical model of pulse waveform

To select an appropriate signal mathematical model according to the type of preamplifier after the detector, in this paper, a resistance–capacitance feedback preamplifier is employed to simulate the output pulse waveform. It is approximated by a double exponential function:

$$s_n(t) = A_n \times \left(e^{-(t-t_n)/\tau_1} - e^{-(t-t_n)/\tau_2}\right) \times u(t-t_n), \quad (18)$$

where A_n represents the amplitude of the nth pulse waveform, t_n stands for the formation time of the pulse waveform, τ_1 and τ_2 show the corresponding slow time constant and fast time constant, respectively, and the function $u(t-t_n)$ is the first step function..

2) The pulse time interval satisfies the exponential distribution rule

$$dI(t) = me^{-mt}dt, \quad (19)$$

where m represents the average counting rate of pulses.

3) The pulse amplitude is proportional to the energy loss of the incident particle in the detector, which has random fluctuation characteristics. Generally, the pulse amplitude distribution of the nuclear signal meets the conditions of normal distribution:

$$H(A) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(A-\bar{A})^2/2\sigma^2}, \quad (20)$$

where $\bar{A}$ is the average pulse amplitude and σ is the amplitude standard deviation determined by the intrinsic energy R of the detector and the average pulse amplitude, $\sigma = R \times (\bar{A}/2.355)$.

4) The statistical characteristics of noise interference

The interference noise will be introduced in the measurement of the nuclear signal due to the influence of electronic devices and environment. The white noise distribution satisfies the normal distribution rule and is superposed linearly with the nuclear

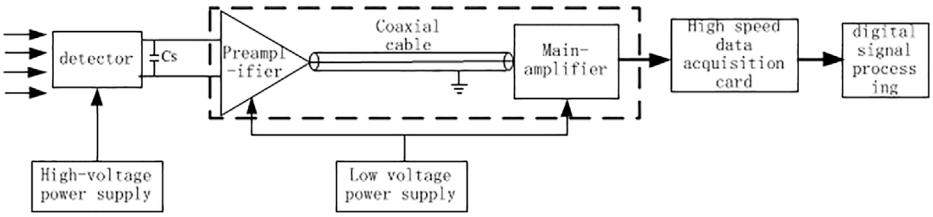


FIGURE 7 | Structural block diagram of the neutron measurement system.

signal to form an observation signal. The slow and fast time constants in the dual exponential function are set to 20 and 0.5 μs, respectively, based on the pulse waveform of the nuclear signal. For the pulse amplitude, the natural resolution of the detector is set to 20%, and the average pulse amplitude is set to 1 V, which obeys the normal distribution. The generation time of the pulse waveform is obtained randomly through exponential distribution according to the pulse time interval and the setting of the average count rate of nuclear signal pulse of 6,000 cps. For the interference noise, the mean value and standard deviation of the noise signal are set as 0 and 0.2 V, respectively, which obeys the additive superposition rule. **Figure 2** shows the original nuclear signal obtained by sampling frequency 1 MHz and sampling time 1 ms.

Analysis of Denoising Effect of Nuclear Signal

The wavelet analysis method is selected to denoise the simulated signal to compare and illustrate the denoising effect of this method. Wavelet packets, respectively, choose wavelet Db4 and wavelet Db8, which are currently widely applied in the nuclear signal denoising area, and the denoising results are shown in **Figure 3**: straight from the top, in turn, plots present the original nuclear signal without noise, the nuclear signal with noise, the result of the sparse decomposition based on the Gabor dictionary, the result of the sparse decomposition based on the chirplet dictionary, the result processed by Db8, and the result processed by Db4.

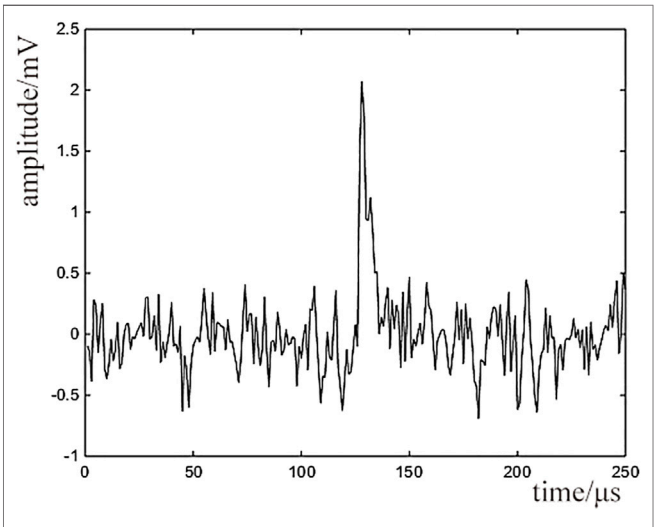


FIGURE 8 | Nuclear pulse signal output by the neutron detection system built in the laboratory.

In this work, three parameters were introduced, which formed the evaluation index of denoising effect, that is, signal-to-noise ratio (SNR), root mean square error (RMSE), and normalized correlation coefficient (NCC). The SNR was used to evaluate the noise energy; the smaller the value of SNR, the lower the noise energy in the signal. The RMSE was used to evaluate the overall error between the recovered signal after denoising and the original noise bureau broadcast signal. The smaller the value of RMSE is, the lower the error is. The NCC reflects the degree of similarity between the recovered signal and the original ideal signal waveform without noise after denoising, and the closer it is to 1, the more similar the two waveforms are.

Based on the results in **Figure 3** and **Table 2**, the following conclusions can be drawn by comparing the denoising results and denoising evaluation indexes of the four methods:

1) The denoising effect of the nuclear signal based on sparse representation introduced in this paper is better than that of wavelet analysis. The waveform of each pulse recovery signal remains consistent, the pulse trend remains the same, and the error based on sparse representation is the smallest compared with the original pulse signal after denoising.
2) The denoising results obtained by denoising the nuclear signal based on sparse representation are related to the selection of over-complete dictionaries, and the results vary obviously by different over-complete dictionaries.
3) The selection of wavelet basis has great influence on the denoising effect of the nuclear signal based on wavelet analysis. Compared with that of the original signal, the amplitude of the denoised signal has big error, and its waveform was distorted.

The Verification of Measured Signals

In the laboratory, a γ pulse nuclear signal was obtained through the nuclear measurement system composed of an NaI detector and γ radiation source ^{60}Co. **Figure 4** shows the waveform measured by the above measurement system in the laboratory by the sampling frequency of 100 M Hz. Due to the weak interference in the laboratory, the nuclear signal can

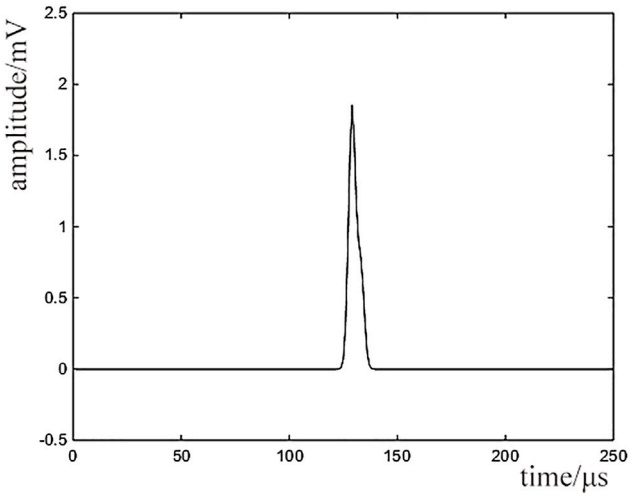


FIGURE 9 | Nuclear pulse signal obtained by sparse decomposing and reconstruction.

be easily differentiated. In this paper, background noise of the laboratory is measured and amplified and then superimposed into the measured signal to simulate the noisy nuclear signal. Its waveform is shown in **Figure 5**.

The method introduced in this paper is adopted to carry out denoising processing for the above nuclear signal with noise, and the result is shown in **Figure 6A**. Meanwhile, the denoising results by wavelet morphology–wavelet method were also measured and are shown in **Figures 6B,C**.

Due to the small noise interference of the original measured nuclear signal, it can be approximately equivalent to the ideal noise-free waveform, and the evaluation index of denoising effect of each method is calculated. The results are shown in **Table 3**.

Based on the results in **Figure 6** and **Table 3**, the method in this paper can still restore the measured nuclear pulse signal in the laboratory under high noise, whereas the denoising effect of wavelet method is poor with large energy loss and large distortion. In general, the method of denoising the nuclear signal introduced in this paper has high accuracy, small waveform distortion, and good retention of the time characteristics and amplitude of the original nuclear pulse signal. So, its denoising effect is obviously better than that of the wavelet method.

To validate the denoising effect of the method in this paper for a weak signal under a long cable transmission nuclear measurement system, a set of neutron detection systems was built in the laboratory, which modeled the measurement system in the core pool of a sodium-cooled fast reactor. In the system, the Am–Be neutron source and LB-125 fission ionization chamber were applied, and the output signal of the fission ionization chamber was transported through a 10 m long cable to the preamplifier for amplification to improve the SNR of the detector and then sent to the linear amplifier through a long shielded cable. The structural block diagram of the neutron measurement system is shown in **Figure 7**.

The nuclear pulse signal shown in **Figure 8** was read out by the neutron detection system built in this paper after the main amplifier. Due to the long distance between the preamplifier and the fission ionization chamber, interference noise can easily be mixed into the measurement process through the transmission cable, so large noise has been superimposed in the nuclear pulse signal measured by the experiment.

The nuclear pulse signal denoising method adopted in this paper was used for sparse decomposition and reconstruction of the nuclear pulse signal output by the neutron detection system built in the laboratory. **Figure 9** shows the result obtained through adjusting the iteration threshold parameters.

RMSE, NCC, SNR, and other parameters cannot be used for evaluation, as there is no way to obtain the nuclear pulse without noise through the neutron detection system. However, from the reconstructed pulse image, we can obviously find the method can effectively extract the nuclear signal from the system with a random noise signal and maintain the time information of the original pulse.

CONCLUSION

In this paper, a sparse representation–based nuclear signal denoising method is proposed for nuclear signal extraction in a strong noise interference environment. Firstly, the Gabor time–frequency atoms and chirplet time–frequency atoms are constructed, and then the sparse decomposition and reconstruction of the signal are performed by the Batch Orthogonal Matching Pursuit (Batch-OMP) algorithm, and the residual ratio threshold is used as the termination condition of the iteration of the algorithm. The simulation results show that this method outperforms the traditional wavelet method in all indexes, with high accuracy and low error, and retains the kernel signal characteristics. The experiments prove that the method can effectively extract the kernel signal in the noisy environment and retain the original pulse information well.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, and further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

S-H and NS completed the algorithm and simulation and wrote the manuscript. L-S, S-H, NS, BC, and X-Z completed the experimental research. L-S, BC, and X-Z revised the article.

FUNDING

This work was supported by the National Natural Science Foundation of China for Youth (12005098), the Research Foundation of Education Bureau of Hunan Province, China (Grant No. 18C0472), and the Open Fund Project of Nuclear Fuel Cycle Technology and Equipment Collaborative Innovation Center of Hunan Province (2019KFY14).

REFERENCES

Aharon, M., Elad, M., and Bruckstein, A. (2005). K-svd:design Dictionaries Sparse Representation. *SPARS* 5, 9–12.

Alotaibi, N. (2021). A Novel Method to Denoise Images Based on a Meta-Heuristic Algorithm and Pre-learned Dictionary. *Ijies* 14 (1), 203–211. doi:10.22266/ijies2021.0228.20

Balnarsaiah, B., and Rajitha, G. (2021). *Denoising and Optical and Sar Image Classifications Based on Feature Extraction and Sparse Representation*. arXiv preprint arXiv:2106.01896.

Bertuccio, G., and Pullia, A. (1993). A Method for the Determination of the Noise Parameters in Preamplifying Systems for Semiconductor Radiation Detectors. *Rev. Scientific Instr.* 64 (11), 3294–3298. doi:10.1063/1.1144293

Celik, C., and Bilge, H. S. (2017). Content Based Image Retrieval with Sparse Representations and Local Feature Descriptors : A Comparative Study. *Pattern Recognition* 68, 1–13. doi:10.1016/j.patcog.2017.03.006

Chen, G. Y., Bui, T. D., and Krzyżak, A. (2009). Invariant Pattern Recognition Using Radon, Dual-Tree Complex Wavelet and Fourier Transforms. *Pattern Recognition* 42 (9), 2013–2019. doi:10.1016/j.patcog.2008.10.008

Deeba, F., Kun, S., Ali Dharejo, F., and Zhou, Y. (2020). Sparse Representation Based Computed Tomography Images Reconstruction by Coupled Dictionary Learning Algorithm. *IET image process* 14 (11), 2365–2375. doi:10.1049/iet-ipr.2019.1312

Ekanadham, C., Tranchina, D., and Simoncelli, E. P. (2011). Recovery of Sparse Translation-Invariant Signals with Continuous Basis Pursuit. *IEEE Trans. Signal. Process.* 59 (10), 4735–4744. doi:10.1109/tsp.2011.2160058

Engan, K., Aase, S. O., and Husøy, J. H. (2000). Multi-frame Compression: Theory and Design. *Signal. Process.* 80 (10), 2121–2140. doi:10.1016/s0165-1684(00)00072-4

Georgiev, A., and Gast, W. (1993). Digital Pulse Processing in High Resolution, High Throughput, Gamma-ray Spectroscopy. *IEEE Trans. Nucl. Sci.* 40 (4), 770–779. doi:10.1109/23.256659

Hashemian, H. M., and Bean, W. C. (2011). Sensors for Next-Generation Nuclear Plants: Fiber-Optic and Wireless. *Nucl. Sci. Eng. Eng. J. Am. Nucl. Soc.* 169, 262–278. doi:10.13182/nse10-48

Huang, F., Tao, J., Xiang, Y., Liu, P., Dong, L., and Wang, L. (2017). Parallel Compressive Sampling Matching Pursuit Algorithm for Compressed Sensing Signal Reconstruction with OpenCL. *J. Syst. Architecture* 72 (C), 51–60. doi:10.1016/j.sysarc.2016.07.002

Knoll, G. F. (2000). *Radiation Detection and Measurement*. 3rd ed. New Jersey: John Wiley and Sons.

Li, H., Wimalajeewa, T., and Varshney, P. K. (2015). *On the Detection of Sparse Signals with Sensor Networks Based on Subspace Pursuit*. IEEE.

Liang, Wei., and Que, Pei-Wen. (2010). Residual Rat Io Iteration Termination Condition for MP Method. *JOURNAL SH ANGHAI JIAOTONG UNIVERSIT Y* 44 (02), 171–176. doi:10.1007/BF00961363

Mallat, S. G., and Zhifeng Zhang, Z. (1993). Matching Pursuits with Time-Frequency Dictionaries. *IEEE Trans. Signal. Process.* 41, 3397–3415. doi:10.1109/78.258082

Mann, S., and Haykin, S. (1995). The Chirplet Transform: Physical Considerations. *IEEE Trans. Signal. Process.* 43 (11), 2745–2761. doi:10.1109/78.482123

Maqsood, S., and Javed, U. (2020). Multi-modal Medical Image Fusion Based on Two-Scale Image Decomposition and Sparse Representation. *Biomed. Signal Process. Control.* 57, 101810. doi:10.1016/j.bspc.2019.101810

Min, M.-G., Lee, J.-K., Ji, Y.-H., Jo, S.-H., and Kim, H.-J. (2015). Evaluation of Electromagnetic Interference Environment of the Instrumentation and Control Systems in Nuclear Power Units. *Nucl. Eng. Des.* 285 (apr), 15–22. doi:10.1016/j.nucengdes.2014.12.038

Othmen, F., Lazzaretti, A., and Baklouti, M. (2021). A Sparse Representation Classification for Noise Robust Wrist-Based Fall Detection. *14th Int. Conf. Health Inform.*, 409–416. doi:10.5220/0010238804090416

Parekh, A., and Selesnick, I. W. (2015). Convex Denoising Using Non-convex Tight Frame Regularization. *IEEE Signal. Process. Lett.* 22 (10), 1786–1790. doi:10.1109/lsp.2015.2432095

Rubinstein, R., Zibulevsky, M., and Elad, M. (2008). *Efficien T Implementation of the K-SVD Algorithm and the Batch-OMP Method*. Technical report. Israel: Department of Computer Science, Technion, Technical CS–08.

Sajjad, M., Mehmood, I., and Ba Ik, S. W. (2015). Sparse Coded Image Super-resolution Using K-Svd Trained Dictionary Based on Regularized Orthogonal Matching Pursuit. *Bio-medical Mater. Eng.* 26 (Suppl. 1s1), S1399. doi:10.3233/bme-151438

Sarkar, P., Chakraborty, C., and Ghosh, M. (2012). Content Based Leukocyte Image Retrieval Ensembling Quaternion Fourier Transform and Gabor-Wavelet Features[C]//2012. *12th Int. Conf. Intell. Syst. Des. Appl.*, 345–350. doi:10.1109/ISDA.2012.6416562

Selesnick, I. W., Parekh, A., and Bayram, I. (2014). Convex 1-d Total Variation Denoising with Non-convex Regularization. *IEEE Signal. Process. Lett.* 22 (2), 141–144. doi:10.1109/LSP.2014.2349356

Shi, M., Zhang, F., Wang, S., Zhang, C., and Li, X. (2021). Detail Preserving Image Denoising with Patch-Based Structure Similarity via Sparse Representation and SVD. *Computer Vis. Image Understanding* 206, 103173. doi:10.1016/j.cviu.2021.103173

To-Po Wang, T. P., and Zong-Wei Li, Z. W. (2014). Significant Reduction of Electromagnetic Interference for Fine-Motion Control Rod Drive in a Nuclear Reactor. *IEEE Trans. Ind. Electron.* 61 (10), 5582–5589. doi:10.1109/tie.2013.2297352

Trigano, T., Sepulcre, Y., and Roitman, M. (2011). *On Nonhomogeneous Activity Estimation in Gamma Spectrometry Using Sparse Signal Representation*. IEEE.

Wang, Y., Peng, Y., Liu, S., Li, J., and Wang, X. (2020). Sparsity Adaptive Matching Pursuit for Face Recognition. *J. Vis. Commun. Image Representation* 67 (7), 102764. doi:10.1016/j.jvcir.2020.102764

Williams, T. (2005). Chapter 8-Electromagnetic Compatibility. *Circuit Designers Companion*. Amsterdam: Elsevier.

Yi, S., and Song, L. (2015). Sparse Signals Recovery from Noisy Measurements by Orthogonal Matching Pursuit. *Inverse Probl. Imaging* 9 (1), 231–238. doi:10.3934/ipi.2015.9.231

Zhang, Jiangmei, Zhu, Qingping, and Ji, Haibo (2018). A Pulse Signal Recovery Method Based on Sparse Representation. *J. Beijing Inst. Tech.* 27 (96), 5–12.

Zhang, Z., Xu, Y., and Yang, J. (2017). A Survey of Sparse Representation: Algorithms and Applications. *IEEE Access* 3, 490–530. doi:10.1109/ACCESS.2015.2430359

Zhou, B., An, Y. L., and Chen, C. Z. (2011). Fault Diagnosis for Low-Speed Rolling Bearing Using Stress Wave and Wavelet Analysis. *Amr* 199-200, 1031–1035. doi:10.4028/www.scientific.net/amr.199-200.1031

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 12 April 2022
doi: 10.3389/fenrg.2022.852146

Research on Thermal-Hydraulic Parameter Prediction Method of the Small Lead–Bismuth Fast Reactor Core Based on Adaptive RBF Neural Network

Hong Wu[1], *Ren Li*[2], *Pengcheng Zhao*[1], *Tao Yu*[1]* *and Yanan Zhao*[1]*

[1]School of Nuclear Science and Technology, University of South China, Hengyang, China, [2]College of Nuclear Science and Technology, Harbin Engineering University, Harbin, China

In this study, a cladding surface temperature prediction method based on an adaptive RBF neural network was proposed. This method can significantly improve the accuracy and efficiency of the thermal safety evaluation of the lead–bismuth fast reactor. First, based on the sub-channel analysis program SUBCHANFLOW, the core sub-channel model of the small lead–bismuth fast reactor SPALLER-100 was established. Second, the calculated 2000 groups of core power distribution and coolant flow distribution data were used as training samples. The adaptive RBF neural network model was trained to predict the surface temperature of fuel elements in the lead–bismuth fast reactor. Finally, by comparison, the effectiveness and superiority of the adaptive RBF neural network method were proved. The results indicate that the relative error of the maximum temperature of the fuel cladding predicted using the adaptive RBF neural network method was less than 0.5%, which can be used for the rapid prediction of the thermal and hydraulic parameters of the lead–bismuth fast reactor.

Keywords: RBF neural network, adaptive algorithm, small lead–bismuth fast reactor, thermal safety, SUBCHANFLOW

OPEN ACCESS

Edited by:
Jun Wang,
University of Wisconsin-Madison, United States

Reviewed by:
Mingjun Wang,
Xi'an Jiaotong University, China
Ataollah Rabiee,
Shiraz University, Iran

***Correspondence:**
Tao Yu
yutao29@sina.com
Yanan Zhao
chinazhaoyanan@hotmail.com

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *10 January 2022*
Accepted: *11 March 2022*
Published: *12 April 2022*

Citation:
Wu H, Li R, Zhao P, Yu T and Zhao Y (2022) Research on Thermal-Hydraulic Parameter Prediction Method of the Small Lead–Bismuth Fast Reactor Core Based on Adaptive RBF Neural Network. Front. Energy Res. 10:852146. doi: 10.3389/fenrg.2022.852146

INTRODUCTION

As one of the six originally selected GEN IV nuclear energy systems of the Generation IV International Forum, the lead-cooled fast reactor (LFR) has attracted continuous and widespread research upsurge worldwide (Pioro 2016; Alemberti 2017; Forum 2014). The distinctive configurations and features offer the LFR distinctive advantages in the aspects of long-term fuel sustainability, safety, economics, proliferation resistance, and physical protection.

So far, major nuclear powerhouses have proposed their own LFR development road map and relevant conceptual designs. In terms of technology maturity, Russia's BREST–300 takes the considerably leading position, which is expected to operate in 2026 (Forum 2014; Zabudko et al., 2021). In parallel, activities are also carried out on SVBR-100, which is based on the previous naval propulsion systems. Meanwhile, Japan has developed a small LFR (LSPR) and a direct-contact PBWFR (Takahashi et al., 2008; Alemberti et al., 2014). Europe proposed the industrial-size plant ELFR design along with its demonstrator called ALFRED (Alemberti et al., 2020). In the United States, only limited development of the SSTAR has been implemented (Smith

et al., 2008). Moreover, a number of innovative LFR conceptual designs that are in various stages have been carried out for different purposes worldwide (Forum 2014). It is worth noting that the research on LFR systems in China has received great emphasis from research institutes to universities (Wu et al., 2016). One of the representative LFR activities in China is the CLEAR series carried out by the Institute of Nuclear Energy Safety Technology (INEST) within the Chinese Academy of Sciences (CAS), which adopts a pool-type configuration and use lead–bismuth eutectic (LBE) as the primary coolant. Other research institutes including the China Institute of Atomic Energy (CIAE) and Nuclear Power Institute of China (NPIC) also carried out their own LFR system concepts (Pioro 2016; Ma et al., 2019).

Actually, in the past decade, a large number of the major LFR design and engineering problems have been tackled, and improvements have been implemented in the practices. Issues such as system integration, component design, performance assessment, lead technology, and safety analysis (including accident mitigation) have got remarkable achievements. However, some technological problems still exist that needed to be resolved, for example, material corrosion, fuel development, and further safety validation. Among these problems, a common problem is the detection and prediction of the cladding maximum temperature, since the cladding maximum temperature is a key parameter of the LFR's thermal safety criteria. It is well known that the high boiling temperature of lead allows the LFR to require neither pressurization nor concerning the overheating of the primary coolant. However, the cladding maximum temperature still needs to be considered in the LFR thermal safety analysis due to its higher coolant operating temperature. The chemical reaction between the LBE and the cladding material, and the failure of the cladding are closely related to the cladding temperature.

In recent years, the neural network has been proven that it is qualified to provide accurate and fast thermal parameter prediction. The representative application is reported in Cong et al. (2013), which uses an artificial neural network and wavelet analysis to carry out the nonlinear research of reactor thermal-hydraulic analysis. Cong's work proves that the neural network method is feasible in thermal-hydraulic analysis. Subsequently, much research has been carried out to verify the feasibility and accuracy of the neural network method in different aspects of reactor thermal-hydraulic analysis. Wang used the BP artificial neural network method to predict the three key parameters of core fuel refueling of Qinshan phase II PWR (Wang et al., 2020). Based on the regularized radial basis function (RBF) neural network model, Peng studied the power distribution of the ACP-100 modular reactor. It is not only concluded that the method can accurately reconstruct the axial power distribution of the reactor core but also proved that the method has good robustness and can overcome the inherent uncertainty in the power distribution reconstruction (Peng et al., 2014). Furthermore, Xia constructed a real-time three-dimensional distribution monitoring system of core power by using the nuclear measurement system and RBF neural network, which improved the accuracy and real-time performance of monitoring

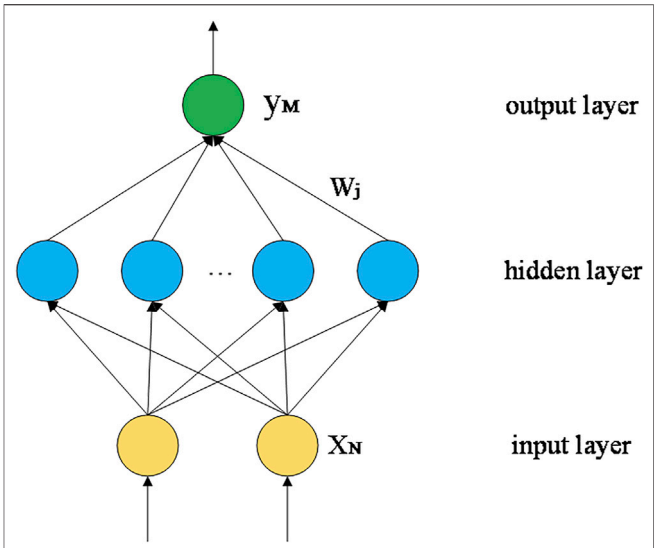


FIGURE 1 | The structure of the RBF neural network.

(Xia et al., 2014). Chen established a feature fusion neural network with seven layers to predict the key safety parameters of the Qinshan reactor. The prediction results show great agreement with the simulation data conducted using the COSMO code (Chen et al., 2022). Although the neural network method has been widely used in the prediction of thermal-hydraulic parameters of reactors and shows great agreement beyond expectation, the relevant research on lead–bismuth fast reactors is still insufficient.

In the present study, the adaptive RBF neural network method is selected to predict the cladding surface temperature of the SPALLER-100 reactor after comparing the performance of several neural network methods. (At present, the BP neural network and RBF neural network are often used to study, so this study takes the BP neural network as a typical comparison.) The training data samples used as a training set and prediction set are obtained by SUBCHANFLOW program. The performance and generalization ability of the adaptive RBF neural network were also verified.

MATHEMATICAL MODEL AND METHOD

A Brief Introduction of the RBF Neural Network

The radial basis function (RBF) neural network is a feedforward neural network with a three-layer structure, namely, the input layer, the output layer, and the hidden layer as is shown in **Figure 1** (Hartman, Keeler, and Kowalski 1990; Park and Sandberg 1991). The basic mathematical model of the RBF neural network is a locally distributed non-negative nonlinear function with central radial symmetric decay. It can approach any nonlinear function with arbitrary precision and has the ability to approximate the error of global, which fundamentally solves the local optimization problem of the BP neural network. Moreover,

it has a compact topology so that the structural parameters can realize separation learning and achieves quick convergence. This characteristic is quite suitable for the real-time control.

The output of hidden layer neurons is as follows:

$$h_j = \exp\left(-\frac{\|x - c_j\|^2}{2b_j^2} \right), \quad (1)$$

where $x = [x_i]^T$ represents the input of the network, the hidden layer output of the network is expressed as $h = [h_j]^T$, h_j is the output of the jth neuron in the hidden layer, $c = [c_{ij}] = \begin{bmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nm} \end{bmatrix}$ is the coordinate vector of the center point of the Gaussian basis function of the jth neuron in the hidden layer, $i = 1, 2, 3, \ldots, n$, $j = 1, 2, 3, \ldots, m$; $b = [b_1, b_2, \ldots, b_m]^T$, and b_j is the width of the Gaussian basis function of the jth neuron in the hidden layer. The implied number of layers in this article is 20, and the transfer function is tanh.RBF network weights are as follows:

$$\omega = [\omega_1, \ldots, \omega_m]^T, \quad (2)$$

The output of RBF network is as follows:

$$y_m t = \omega_1 h_1 + \ldots + \omega_m h_m, \quad (3)$$

The error index of the RBF neural network can be written as follows:

$$E(t) = \frac{1}{2}\left(y(t) - y_m(t)\right)^2, \quad (4)$$

In addition, the RBF neural network has the characteristics of self-learning, self-organizing, and self-adaptive functions. Meanwhile, the RBF neural network has the uniform approximation to nonlinear continuous functions and high learning efficiency. The advantages expressed before offers the RBF neural network the capability of large-scale data fusion and data high-speed parallel processing. Presently, the RBF neural network has been successfully applied to the aspects of nonlinear function approximation, time series analysis, data classification, pattern recognition, information processing, image processing, system modeling, control, fault diagnosis, etc. (Seshagiri and Khalil 2000; Li et al., 2004; Wang and Yu 2008).

Adaptive RBF Neural Network

An adaptive algorithm is a process aimed at approaching the target continuously, which is based on a gradient algorithm. By introducing the adaptive algorithm into the conventional neural network, the "over-fitting" phenomenon can be effectively eliminated. Thus, it can significantly reduce the dependence on the accuracy of the neural network identifier and dramatically improve the weakness of the conventional neural network.

According to mature literature, compared with the adaptive BP neural network, the adaptive RBF neural network can effectively improve the performance of the controller when the system has large uncertainty and has a better prediction effect (Zhu et al., 2008). In view of this, the adaptive gradient descent (Adam) algorithm is adopted to overcome the drawbacks of falling into local minimum and slow convergence that the traditional BP neural network has. The flowchart of the adaptive RBF neural network algorithm is demonstrated in **Figure 2**.

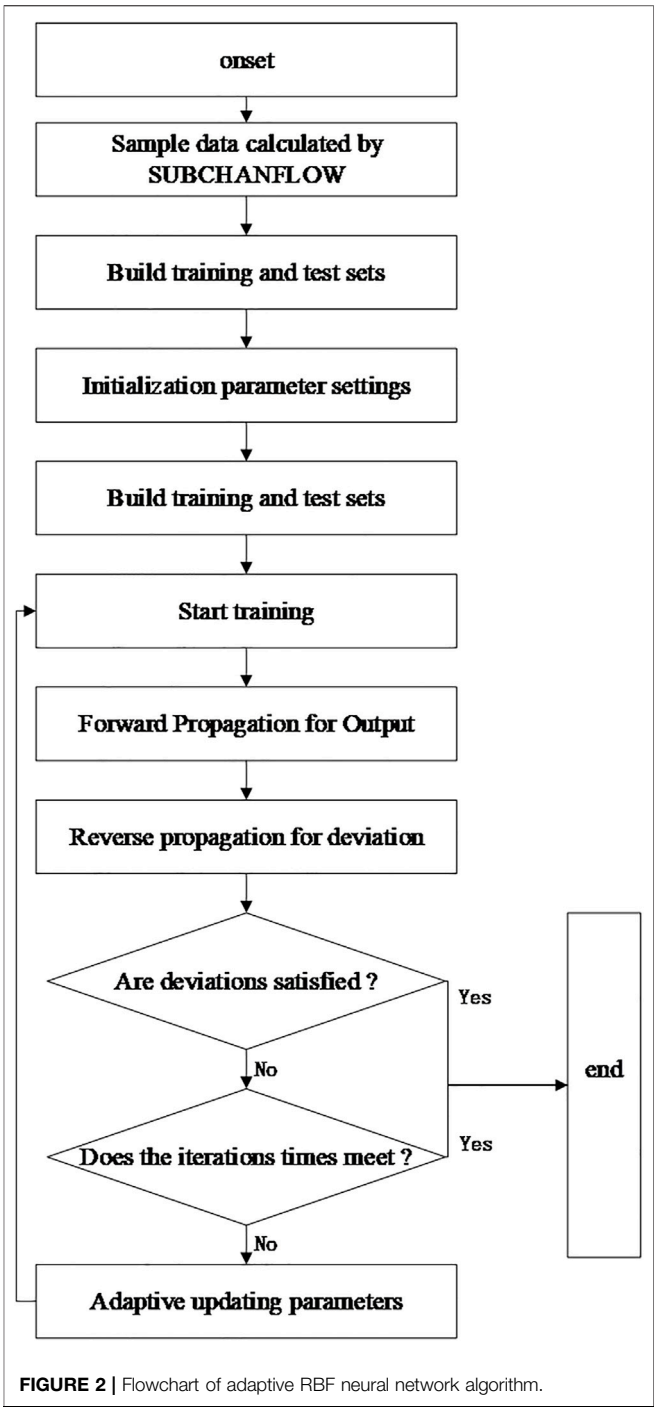


FIGURE 2 | Flowchart of adaptive RBF neural network algorithm.

The Adam algorithm updates the parameters as follows:

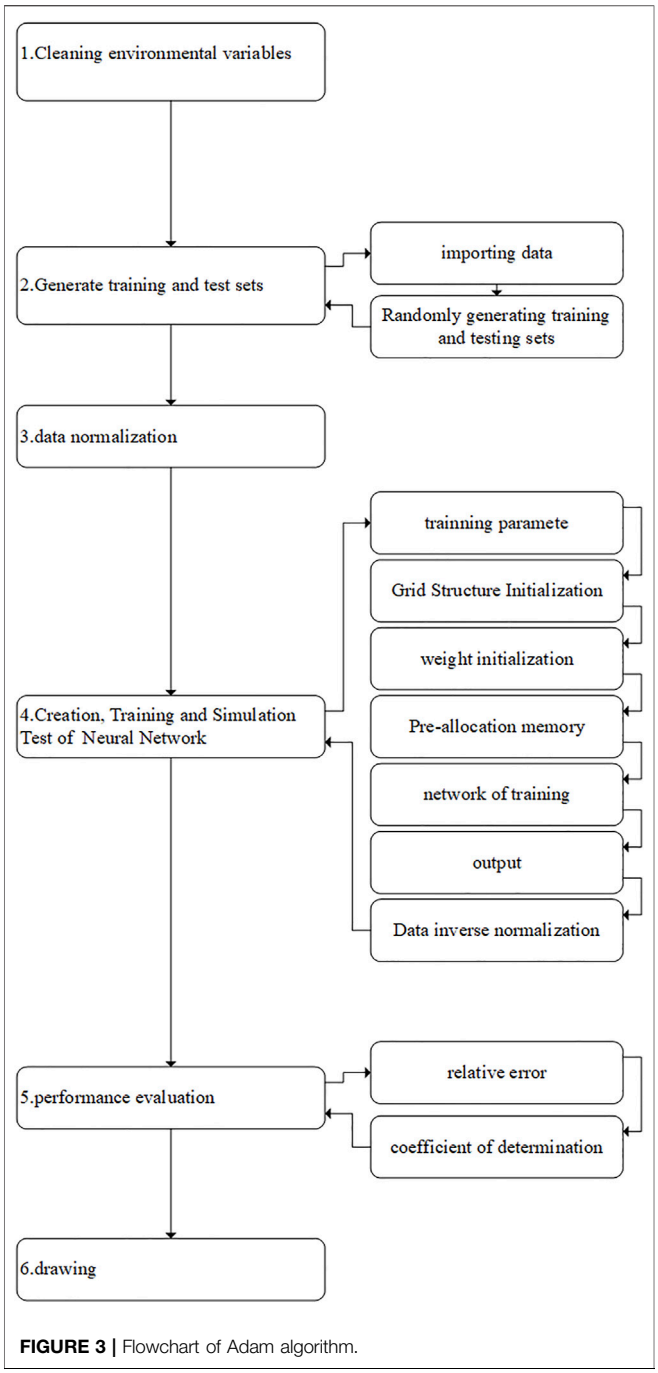


FIGURE 3 | Flowchart of Adam algorithm.

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon}\hat{m}_t, \tag{5}$$

where η is the learning rate, which controls the update ratio of weights and takes a smaller value, which will make the training converge to better performance, t is the iteration time, $\hat{m}_t$ is the weighted average of the gradient, and $\hat{v}_t$ is the weighted deviation. During training, β represents the error signal between the output layer and the hidden layer, β_1 is the exponential decay rate of the first moment estimation, and β_2 is the exponential decay rate of the second moment estimation. In this article, $\eta = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$. The detailed process is shown in **Figure 3**.

SPALLER-100 Introduction

SPALLER-100 is a small lead–bismuth fast reactor with a thermal power of 100 MW (Liu et al., 2020). The schematic diagram of the small lead–bismuth fast reactor SPALLER-100 core is shown in **Figure 4** (cross-section view) and the main parameters of the SPALLER-100 is listed in **Table 1**. The core of the SPALLER-100 is hexagonal and consists of 48 fuel assembly, 13 control rod components, 66 reflector components, and 126 shielding components. The coolant and reflector were 208 Pb–Bi, and the shielding material was B_4C. In this study, the SPALLER-100 is chosen as the research target.

SUBCHANFLOW Code Description

The data used for training the RBF neural network is conducted using the SUBCHANFLOW code. SUBCHANFLOW is a sub-channel flow code to analyze thermal-hydraulic phenomena in the core of pressurized water reactors, boiling water reactors, and innovative reactors operated with gas or liquid metal as coolant, which is developed by the Karlsruhe Institute of Technology (Imke and Sanchez. 2012).

The SUBCHANFLOW code can handle rectangular and hexagonal geometry fuel rod types. The total flow or each channel flow can be selected as the boundary conditions. According to the friction force at the inlet of the tube bundle, the flow can be automatically allocated to the parallel channel. In addition, the given inlet and outlet pressure difference boundary can be used for steady-state calculation. The inlet fluid temperature and outlet pressure are always given as boundary conditions. In this study, the SPALLR-100 core channel is divided and numbered first. The nodalization scheme of the SPALLER-100 core is shown in **Figure 4A**. The nodalization scheme of the SPALLER-100 core is shown in **Figure 4B**. The heat conduction of fuel rod (heating part) in the SUBCHANFLOW is solved using the standard finite volume method. The convective heat transfer coefficient between the fuel rod and coolant is calculated according to the empirical relationship between the heat transfer form and coolant flow pattern. The constitutive relation used in the SUBCHANFLOW code is listed as follows:

(1) Physical properties model: the thermophysical properties data of lead–bismuth alloy are from the HLMC handbook.
(2) Thermal conductivity model: SUBCHANFLOW uses the full implicit finite difference method to calculate the heat conduction process in fuel core and cladding materials.
(3) Heat convection model: the general heat transfer equation of liquid metal heat transfer:

$$\mathrm{Nu} = \mathrm{A} + \mathrm{B} \cdot \mathrm{Pe}^{\mathrm{C}}, \tag{6}$$

(4) Pressure loss model: the Novendstern model and Rehme model are used for the pressure drop calculation.

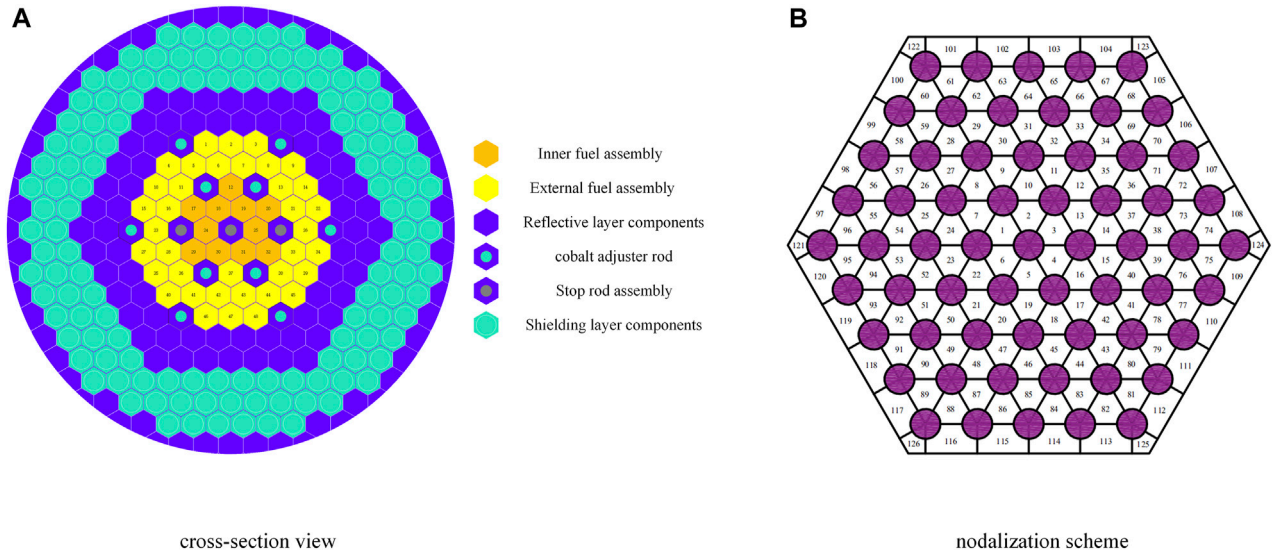


FIGURE 4 | Cross-section view and the nodalizaiton scheme of SPALLER-100 core.

TABLE 1 | Main parameters of the SPALLER-100.

Parameter	Numerical value
Number of fuel rods in the assembly	61
Internal and external diameters of fuel rod (cm)	1.2/1.35
Number of components	48
Rod diameter ratio	1.7
Average linear power density (kW/m)	22.77
Average volume power density (MW/m^3)	29.37
Active zone length (cm)	150
Cladding thickness (mm)	12.7
Fuel rod gap width (mm)	0.15
Cladding thickness (mm)	0.6
Cladding material	Stainless steel
Fuel material	UO_2

PERFORMANCE ANALYSIS OF THE ADAPTIVE RBF NEURAL NETWORK

The feasibility, accuracy, and efficiency of the RBF neural network are verified based on the steady-state data in this section. First, the hottest assembly in the core is found according to the SUBCHANFLOW calculation results. This searching process repeats 100 times to guarantee the result's reliability. Second, focusing on the hottest assembly, several groups of data were randomly selected with the power ranging from 0 to 1,200 kW and the mass flow ranging from 1,200 to 2,200 kg/s. Third, these data are calculated using the SUBCHANFLOW code as an input. Finally, 2000 groups of effective data samples are obtained. Among these 2000 groups of data samples, 1900 groups are selected as the training set, and the remaining 100 groups are selected as the prediction set. Then, the prediction model is evaluated by comparing the error between the prediction results and the calculation result.

Figure 5A and **Figure 6A** demonstrate the error band between the prediction result and the calculation result of the adaptive BP neural network and adaptive RBF neural network separately. It can be observed that the prediction results conducted using the adaptive RBF neural network show a good agreement with the calculation results in the cladding maximum temperature, since the error bound is within 5%. Meanwhile, **Figure 5B** and **Figure 6B** show the comparison results between the predicted and experimental values of the two methods. By comparing the two figures, it can be seen that the fitting results of the two lines in **Figure 6B** are better, that is, the adaptive RBF neural network shows a better performance in predicting cladding maximum temperature than the adaptive BP neural network.

Table 2 illustrates the efficiency and accuracy of different methods after 50 times prediction. The adaptive RBF neural network reaches the average relative error of 0.10 within 6 s and 160 iteration times, which is fully superior to the adaptive BP neural network. Therefore, the adaptive RBF neural network prediction model has better accuracy and feasibility in predicting cladding maximum temperature.

THE TRANSIENT PREDICTION PERFORMANCE ANALYSIS

In the transient response analysis part, the axial and radial power of each fuel rod in the fuel assembly is assumed and uniformly distributed for simplification, since the power distribution has a little influence on the transients. All the coolant channels in a single assembly can be merged into a large channel centered on the fuel rod with the equivalent heating perimeter and wetted perimeter. The initial power is set to 30 MW. **Figure 7** shows the variation of the coolant mass flow and the variation of the

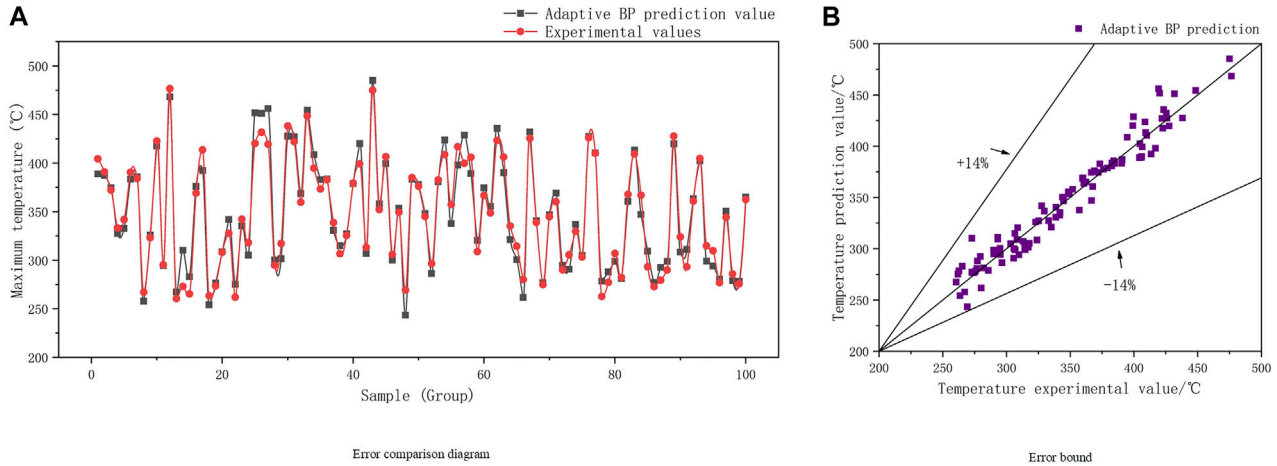


FIGURE 5 | The error comparison diagram and the error bound of adaptive BP neural network.

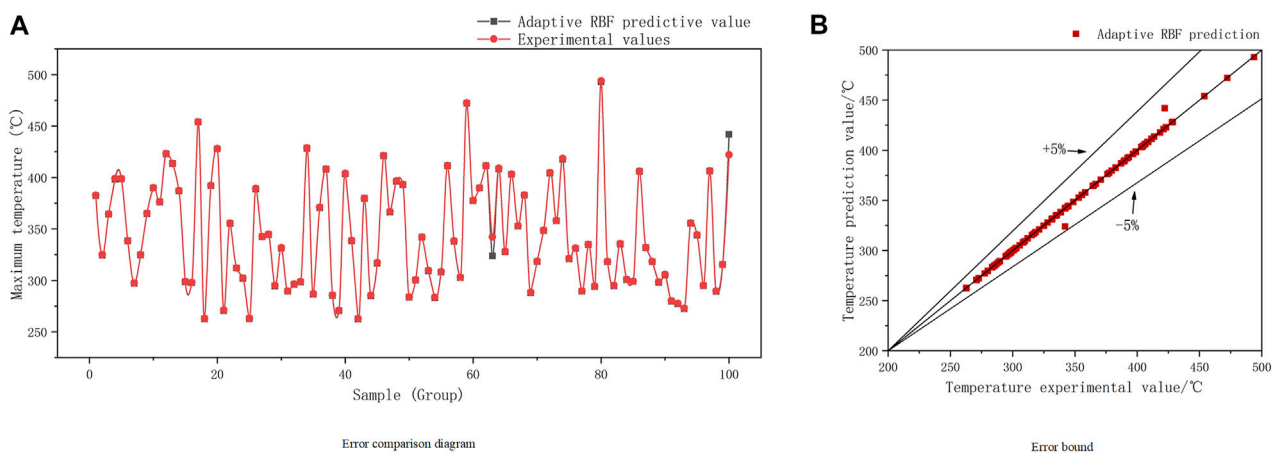


FIGURE 6 | The error comparison diagram and the error bound of adaptive RBF neural network.

TABLE 2 | Comparison of calculation efficiency.

Model	Iteration times	Calculation time/s	Average absolute error (°C)	Average relative error (%)
Adaptive BP neural network	241	8	9.67	0.26
Adaptive RBF neural network	160	6	0.72	0.10

Prediction performance to the mass flow variation.

maximum cladding temperature conducted using the SUBCHANFLOW code.

Among these calculation data, 800 groups were randomly selected in the 80s data sample, 750 groups were used as the training set, and the remaining 50 groups were used as the prediction set. Similar to *Performance Analysis of the Adaptive RBF Neural Network*, the mass flow and the heating power is considered as the input and the cladding maximum temperature is considered as the output. The neural network is operated to predict the cladding maximum temperature in the following 5s, that is, 80–85s. Thus, the error bound between the prediction results and the calculation results is used to evaluate the transient performance of different prediction methods, which is shown in **Figures 8A,B**, **Figures 9A,B**.

Both the adaptive BP neural network and adaptive RBF neural network show remarkable transient prediction ability

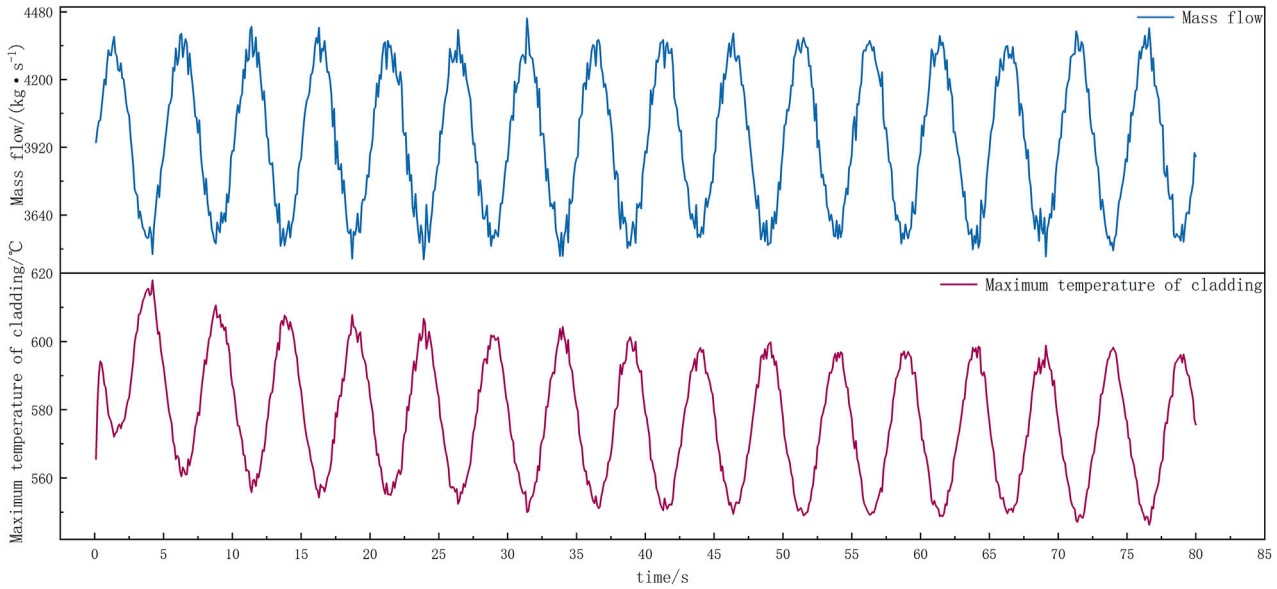


FIGURE 7 | Variation of coolant mass flow and cladding maximum temperature within 80s.

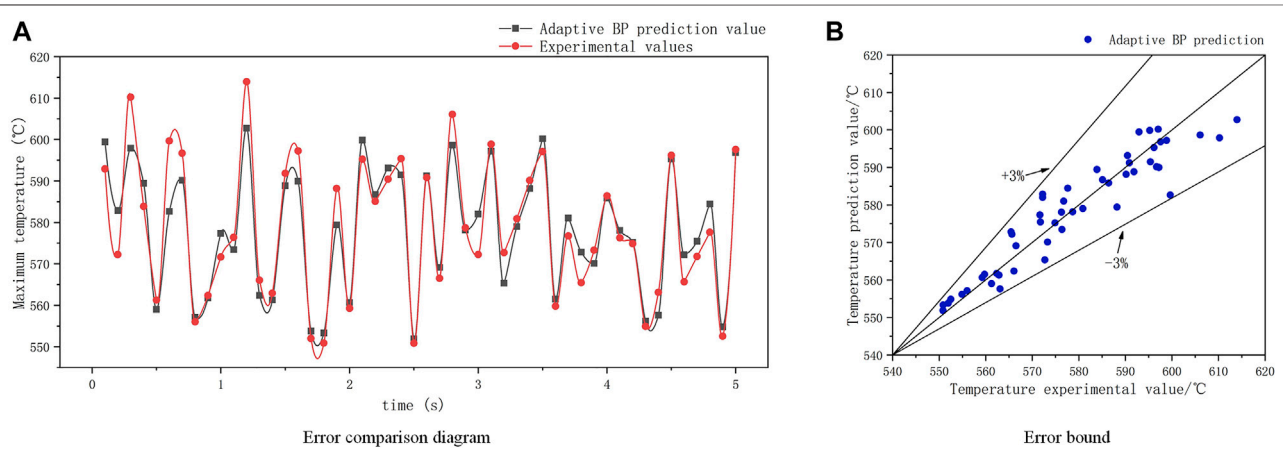


FIGURE 8 | The error comparison diagram and the error bound of adaptive BP neural network under coolant mass flow variation condition.

to the mass flow variation with the error bound being within 3%. It can be concluded from **Figures 8B**, **9B** that the difference between the two prediction results is not large. However, in the error point diagram given by **Figures 8A**, **9A**, it can be clearly seen that the results of adaptive RBF are compared with adaptive BP, and most of the data are close to the center line, indicating that its stability is better. The average errors are shown in **Table 3**. The maximum relative error of adaptive RBF neural network is 2.1%, and the average absolute error is 2.94°C, which proves the adaptive RBF neural network is able to deal with transient conditions as well. Also, it is reasonable to infer that the adaptive RBF neural network will have a better accuracy when extending the prediction time, since the preorder prediction result will influence the following prediction result.

Prediction Performance to the Power Variation

Similarly, transient responses to the power variation are verified by setting the initial coolant mass flow to 4,000 kg/s, and assuming the core power changes. **Figure 10** shows the change of core power and the cladding maximum temperature calculated using the SUBCANFLOW code within 80s. The effectiveness of the adaptive RBF neural network under power variation conditions is analyzed.

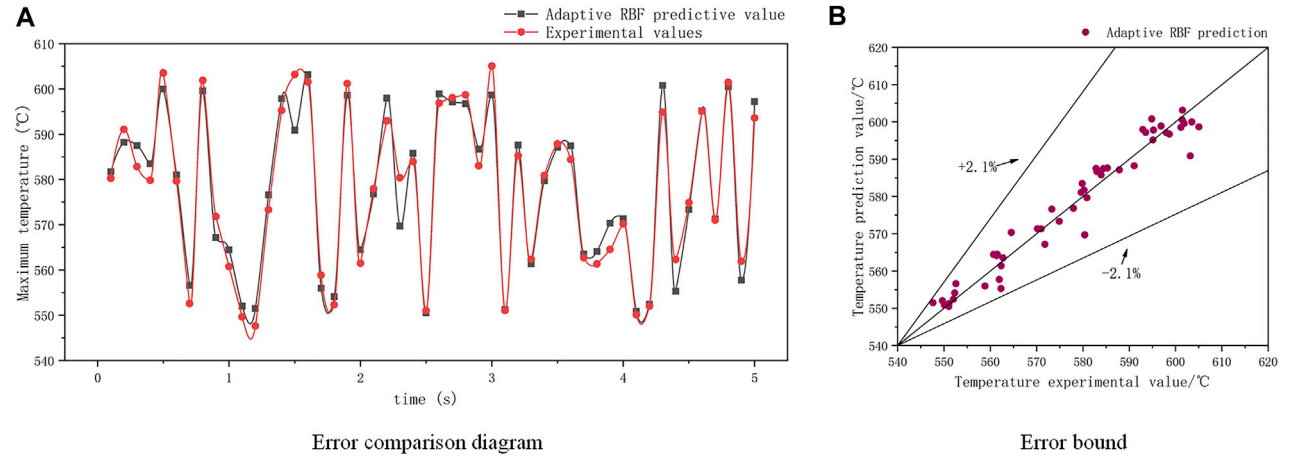


FIGURE 9 | The error comparison diagram and the error bound of adaptive RBF neural network under coolant mass flow variation condition.

TABLE 3 | Comparison of average error under mass flow variation condition.

Model	Average relative error (%)	Average absolute error (°C)
Adaptive BP neural network	−0.04	4.20
Adaptive RBF neural network	0.03	2.92

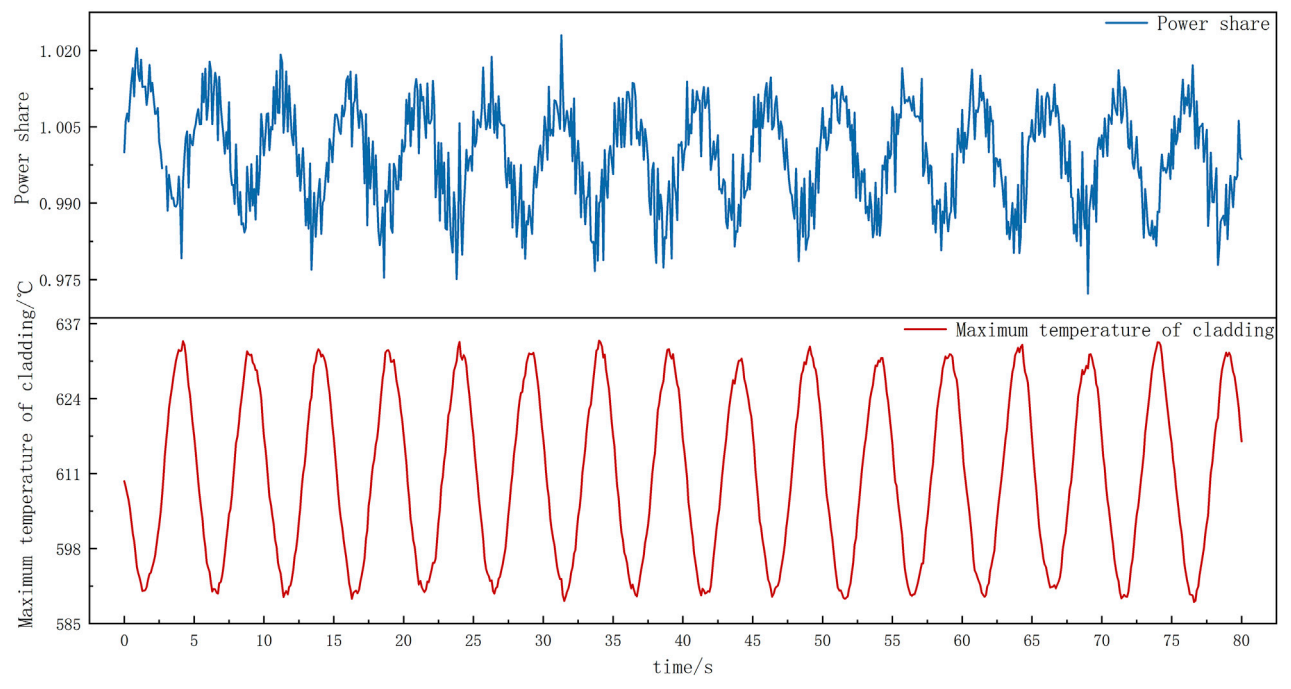


FIGURE 10 | Variation of reactor power and cladding maximum temperature within 80s.

Figure 11 gives the predicting error bound of the adaptive BP neural network and adaptive RBF neural network during power variation condition separately. The average error data are shown in **Table 4**. It can be seen that compared with the adaptive BP results, the adaptive RBF has more error points close to the center line, and its accuracy is better. In addition,

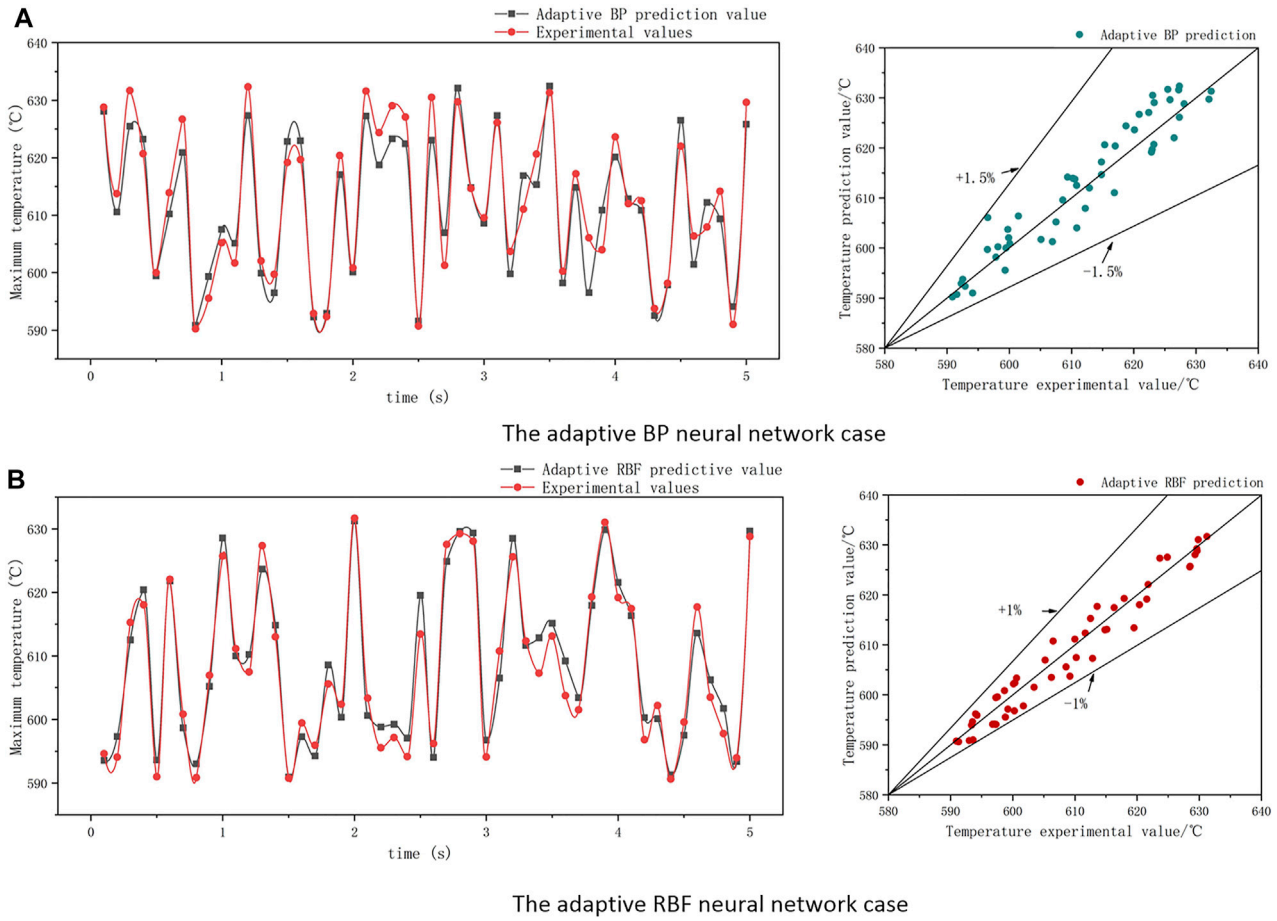


FIGURE 11 | The accuracy comparation between adaptive BP neural network and adaptive RBF neural network.

TABLE 4 | Average error results calculated by different methods.

Model	Average absolute error (°C)	Average relative error (%)
Adaptive BP neural network	3.29	−0.16
Adaptive RBF neural network	2.31	0.09

the correctness of the aforementioned results can be verified from the point-line diagrams of the predicted and experimental values given in **Figure 11A** and **Figure 11B**. As shown in **Table 4**, the error of adaptive RBF neural network is less than 1%, which is slightly better than the adaptive BP neural network prediction method.

Combined with the conclusion conducted in *Prediction Performance to the Power Variation*, it can be concluded that the adaptive RBF neural network shows good effectiveness and superiority in predicting cladding maximum temperature under power variation and coolant mass flow variation conditions. Thus, the adaptive RBF neural network can be used to real-time predict the dynamic value of LFR cladding maximum temperature, which is obviously beneficial for the reactor's safety under both transient conditions and accident.

CONCLUSION

This study analyzes the performance of the adaptive RBF neural network in predicting the cladding maximum temperature for the typical LFR. The feasibility, accuracy, and efficiency of the adaptive RBF neural network under both steady-state and transient conditions are evaluated. The conclusions drawn from the study are summarized as follows:

(1) A cladding maximum temperature prediction method based on the adaptive RBF neural network for the LFR is proposed. The SUBCHANFLOW program is used to generate data for the RBF neural network training.
(2) By comparing the adaptive RBF neural network and the adaptive BP neural network, the adaptive RBF neural network shows full superiority. The adaptive RBF neural network has good feasibility, accuracy, and efficiency in predicting the cladding maximum temperature of the lead–bismuth fast reactor.
(3) The adaptive RBF neural network can accurately predict the trend of the cladding maximum temperature in short time

under the transient conditions of power variation and coolant mass flow variation.

(4) The real-time thermal-hydraulic parameter prediction capability of the adaptive RBF neural network is of great significance for the LFR's thermal safety.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

RL contributes to the nodalizaiton scheme and debugging of the SUBCHANFLOW code.

FUNDING

This study is funded by the Research Foundation of Education Bureau of Hunan Province, China (Grant No. 20B490) and Hunan Science and Technology Innovation Team Project (Grant No. 2020RC4053).

REFERENCES

Alemberti, A., Caramello, M., Frignani, M., Grasso, G., Merli, F., Morresi, G., et al. (2020). ALFRED Reactor Coolant System Design. *Nucl. Eng. Des.* 370, 110884. doi:10.1016/j.nucengdes.2020.110884

Alemberti, A., Smirnov, V., Smith, C. F., and Takahashi, M. (2014). Overview of lead-cooled Fast Reactor Activities. *Prog. Nucl. Energ.* 77, 300–307. doi:10.1016/j.pnucene.2013.11.011

Alemberti, A. (2017). The lead Fast Reactor: an Opportunity for the Future?

Chen, Y., Wang, D., Kai, C., Pan, C., Yu, Y., and Hou, M. (2022). Prediction of Safety Parameters of Pressurized Water Reactor Based on Feature Fusion Neural Network. *Ann. Nucl. Energ.* 166, 108803. doi:10.1016/j.anucene.2021.108803

Cong, T., Su, G., Qiu, S., and Tian, W. (2013). Applications of ANNs in Flow and Heat Transfer Problems in Nuclear Engineering: a Review Work. *Prog. Nucl. Energ.* 62, 54–71. doi:10.1016/j.pnucene.2012.09.003

Houssin, D., Dujardin, T., Cameron, R., Tam, C., Paillere, H., Baroni, M., and Vance, R. (2015). Technology Road-Map-Nuclear Energy (No. NEA-IEA--2015). Organisation for Economic Co-operation and Development. Hartman, Eric J, James D Keeler, and Jacek M %J Neural Computation Kowalski. *Layered neural networks with Gaussian hidden units as universal approximations* 2, 210–215.

Imke, U., and Sanchez, V. H. (2012). *Validation of the Subchannel Code SUBCHANFLOW Using the NUPEC PWR Tests (PSBT)*. Science and Technology of Nuclear Installations.

Li, Y., Qiang, S., Zhuang, X., and Kaynak, O. (2004). Robust and Adaptive Backstepping Control for Nonlinear Systems Using RBF Neural Networks. *IEEE Trans. Neural Netw.* 15 (3), 693–701. doi:10.1109/tnn.2004.826215

Liu, Z., Zhao, P., Zhang, B., Yu, T., Xie, J., Chen, Z., et al. (2020). Study on Core Conceptual Design of Ultra-long Life Small Naturally Circulating Pb - Bi Fast Reactor. *At. Energ. Sci. Tech.* 54 (7), 1254–1265.

Ma, Z., Ma, Z., Wu, Y., Gao, F., Hei, B., and Su, G. H. (2019). Design and R&D Progress of Core Assembly Deformation Test Facility for China Demonstration Fast Reactor. *Nucl. Eng. Des.* 348, 65–77. doi:10.1016/j.nucengdes.2019.04.002

Park, J., and Sandberg, I. W. (1991). Universal Approximation Using Radial-Basis-Function Networks. *Neural Comput.* 3 (2), 246–257. doi:10.1162/neco.1991.3.2.246

Peng, X., Ying, D., Li, Q., and Wang, K. (2014). Application of Regularized Radial Basis Function Neural Network in Axial Power Distribution Reconstruction of Reactor Core. *Nucl. Power Eng.* 35 (S2), 12–15.

Pioro, I. (2016). *Handbook of Generation IV Nuclear Reactors*. Woodhead Publishing.

Seshagiri, S., and Khalil, H. K. (2000). Output Feedback Control of Nonlinear Systems Using RBF Neural Networks. *IEEE Trans. Neural Netw.* 11 (1), 69–79. doi:10.1109/72.822511

Smith, C. F., Halsey, W. G., Brown, N. W., Sienicki, J. J., Moisseytsev, A., and Wade, D. C. (2008). SSTAR: The US lead-cooled Fast Reactor (LFR). *J. Nucl. Mater.* 376 (3), 255–259. doi:10.1016/j.jnucmat.2008.02.049

Takahashi, M., Uchida, S., Yamada, Y., and Koyama, K. (2008). Safety Design of Pb–Bi-Cooled Direct Contact Boiling Water Fast Reactor (PBWFR). *Prog. Nucl. Energ.* 50 (2-6), 269–275. doi:10.1016/j.pnucene.2007.11.082

Wang, Duan., Wang, Weice., and Pan, Cuijie. (2020). Key Parameters of Core Refueling for Pressurized Water Reactors Based on Adaptive BP Neural Network. *At. Energ. Sci. Tech.* 54, 112–118.

Wang, S., and Yu, D. L. (2008). Adaptive RBF Network for Parameter Estimation and Stable Air-Fuel Ratio Control. *Neural Networks* 21 (1), 102–112. doi:10.1016/j.neunet.2007.10.006

Wu, Y., Bai, Y., Song, Y., Huang, Q., Zhao, Z., and Hu, L. (2016). Development Strategy and Conceptual Design of China lead-based Research Reactor. *Ann. Nucl. Energ.* 87, 511–516. doi:10.1016/j.anucene.2015.08.015

Xia, Hong., Li, Bin., and Liu, Jianxin. (2014). Research on Three-Dimensional Power Distribution Method of Pressurized Water Reactor Core Based on RBF Neural Network. *At. Energ. Sci. Tech.* 4, 48.

Zabudko, L. M., Grachev, A. F., Zherebtsov, A. A., Lachkanov, E. V., Mochalov, Y. S., Skupov, M. V., et al. (2021). Status on Performance Study of Mixed Nitride Fuel Pins of BREST Reactor Type. *Nucl. Eng. Des.* 384, 111430. doi:10.1016/j.nucengdes.2021.111430

Zhu, Q., Fei, S., Zhang, T., and Li, T. (2008). Adaptive RBF Neural-Networks Control for a Class of Time-Delay Nonlinear Systems. *Neurocomputing* 71 (16-18), 3617–3624. doi:10.1016/j.neucom.2008.04.012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 12 April 2022
doi: 10.3389/fenrg.2022.852349

Multivariate Time Series Prediction for Loss of Coolant Accidents With a Zigmoid-Based LSTM

Shanshan Gong, Suyuan Yang, Jingke She, Weiqi Li and Shaofei Lu*

College of Computer Science and Electronic Engineering, Hunan University, Changsha, China

Post-LOCA prediction is of safety significance to NPP, but requires a processing coverage of non-linearity, both short and long-term memory, and multiple system parameters. To enable an ability promotion of previous LOCA prediction models, a new gate function called zigmoid is introduced and embedded to the traditional long short-term memory (LSTM) model. The newly constructed zigmoid-based LSTM (zLSTM) amplifies the gradient at the far end of the time series, which enhances the long-term memory without weakening the short-term one. Multiple system parameters are integrated into a 12-dimension input vector to the zLSTM for a comprehensive consideration based on which the LOCA prediction can be accurately generated. Experimental results show both accuracy evaluations and LOCA progression produced by the proposed zLSTM, and two baseline methods demonstrating the superiority of applying zLSTM to LCOA predictions.

Keywords: LOCA, prediction, multivariate time series, zigmoid, LSTM

OPEN ACCESS

Edited by:
Xianping Zhong,
University of Pittsburgh, United States

Reviewed by:
Guang Hu,
Karlsruhe Institute of Technology (KIT), Germany
Sai Zhang,
Idaho National Laboratory (DOE), United States

****Correspondence:***
Jingke She
shejingke@hnu.edu.cn

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *11 January 2022*
Accepted: *28 February 2022*
Published: *12 April 2022*

Citation:
Gong S, Yang S, She J, Li W and Lu S (2022) Multivariate Time Series Prediction for Loss of Coolant Accidents With a Zigmoid-Based LSTM. Front. Energy Res. 10:852349. doi: 10.3389/fenrg.2022.852349

1 INTRODUCTION

Loss of coolant accident (LOCA) is a severe accident that causes safety threat to nuclear power plants (NPPs). Obviously, it is of great importance to systematically analyze, prevent, and predict LOCAs such that effective decision-making support can be offered to the emergency response strategy. The prediction of the LOCA progression trends, as one of the significant emergency measures, provides evaluation of safety threats ahead of their physical occurrence and allows the emergency response strategy to plan accordingly before worse scenarios emerge. However, the non-linearity of LOCAs and associated complex system factors prevent accurate LOCA predictions. As a coupling result influenced by multiple system parameters, the prediction for LOCA progression also faces multivariant processing challenges, which makes the system modeling more complicated.

In the past decades, various attempts have been taken for process predictions in NPPs. A series of assumptions based on statistical methods and mathematical equations are applied for process predictions such as 1) monitoring the real-time condition of LOCA *via* time–frequency domain reflectometry (TFDR) (Lee et al., 2017) and 2) using RELAP5/MOD3.3 code to predict the LOCA of the main stream break on generation III reactor (Yang et al., 2019). The aforementioned research studies rely on effort-consuming system modeling and have made feasible progress on LOCA prediction, but the challenges of multivariate processing/coupling remain for further investigation.

Using data-based artificial intelligence (AI) approaches has become an effective way to solve the non-linearity problem with the progress of machine learning, especially when enormous simulated NPP data from previous research studies have founded a firm database for AI applications.

A variety of traditional machine learning algorithms have been applied to NPPs. An abnormal operation state detection method of NPP based on an unsupervised deep generative model is

established by using variational auto encoders (VAE) and isolation forest (iForest) (Li et al., 2021). Moshkbar-Bakhshayesh and Ghafari (2022) used support vector machine (SVM) as a machine learning–based method to predict the vessel water level. Xiang et al. (2020) proposed a clustering algorithm for the transient detection in NPPs. Furthermore, Wang et al. (2021a) utilized the clustering algorithm together with SVM and principal component analysis (PCA) for the sensor anomalies in NPPs, which is also reviewed in Hu et al. (2021).

By stacking multiple hidden layers, deep neural network (DNN) has stronger non-linear feature extraction ability. It was utilized to predict the vessel water level (Koo et al., 2018) as well as to identify the fault diagnosis scheme (Santos et al., 2019).

Convolutional neural network (CNN) is a variant of DNN and is usually used for image processing. Viewing NPP sensor data as images, CNN was applied to event identification (Lin et al., 2021; Pantera et al., 2021) and break size estimation (Lin et al., 2022). The mentioned traditional machine learning algorithms (SVM) and deep learning methods (DNN and CNN) can deal with non-linearity, while the sequential data-dependency and multiple physical factors are not taken into account.

Recurrent neural network (RNN), as a classical example, has been successfully applied to sequential data modeling in former explorations. Several long short-term memory (LSTM)–based models cover both the non-linearity and time correlation of LOCAs. For example, the LSTM-based expert system was adopted to predict LOCA behaviors (Mira et al., 2020; Santhosh et al., 2010; Chen et al., 2021) and to evaluate abnormal operation conditions in NPPs (She et al., 2020; Wang et al., 2021b). The coolant flowrate variation was analyzed by She et al. (2021) using a combination of CNN and LSTM. PCA and LSTM were used to identify the fault diagnosis scheme Saeed et al. (2020).

It is necessary to consider modeling non-linearity, multivariate processing, and long-term memory for accurate prediction of LOCA. The aforementioned literatures ignored that LSTM cannot model longer time series. To fully cover the non-linearity, time correlation, and multivariate processing for LOCA predictions, this study proposes an improved LSTM model in which a new gate function called 'zigmoid' is constructed. With rigorous experimental verifications conducted on simulated LOCA datasets, the zLSTM is proved to be more accurate and efficient for post-LOCA predictions.

This article starts with **Section 1** as the introduction and illustrates the zigmoid function in **Section 2**. After the presentation of the verification experiments in **Section 3**, this article is then concluded in **Section 4**.

2 ZIGMOID METHOD

2.1 Zigmoid for Better Long-Term Memory

Established for the sequential processing problems, RNN obtained preliminary short-term memory. To enable the long-term memory, Hochreiter and Schmidhuber (1997) made a gate-level innovation on RNN and created LSTM that is capable for both short and long-term processing. Nevertheless, the contribution of x_t in LSTM will decay in k timesteps by f_t^k when f_t is a constant (Tallec and Ollivier, 2018). This gives the unit an effective decay period timescale of $O(\frac{1}{1-f_t})$. In case the LSTM is required to have a memory of 1,000 steps, the forget gate parameter f_t has to reach 0.999 since the effective decay period timescale = $1/(1 - f_t)$. However, it is very challeging to achieve this, according to the following analysis of LSTM.

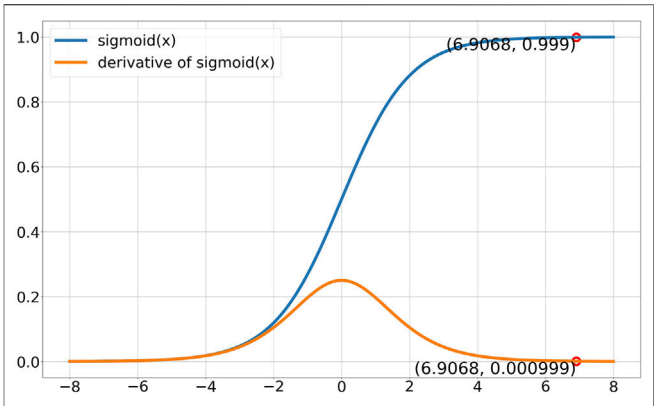


FIGURE 1 | Sigmoid properties.

The standard LSTM process is defined as follows:

$$i_t = sigmoid\,(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \tag{1}$$

$$\boldsymbol{f_t = sigmoid\,(W_{fx}x_t + W_{fh}h_{t-1} + b_f)} \tag{2}$$

$$o_t = sigmoid\,(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{3}$$

$$\hat{c}_t = \tanh\,(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \hat{c}_t \tag{5}$$

$$h_t = o_t \odot \tanh\,(c_t) \tag{6}$$

where W_{ix}, W_{ih}, b_i, W_{fx}, W_{fh}, b_f, W_{ox}, W_{oh}, b_o, W_{cx}, W_{ch}, and b_c are trainable parameters.

However, the derivative of sigmoid is 0.000999 when $f_t = 0.999$, as shown in **Figure 1**, which causes LSTM to be untrainable at this stage. In other words, the LSTM's long-term memory ability is weakened at the far end of the time series and cannot guarantee accurate prediction in LOCAs.

Since the sigmoid function in the forget gate determines the long-term memory of LSTM, a natural idea is to amplify the derivative of the sigmoid function such that model training is still feasible for LSTM even when f_t reaches 0.999. For this purpose, zigmoid is constructed by embedding a transfer function within the original sigmoid.

$$zigmoid\,(x) = sigmoid\,(trans\,(x)) \tag{7}$$

$$trans\,(x) = \begin{cases} e^{\beta * x} - 1 & x \geq 0 \\ 1 - e^{-\beta * x} & x < 0 \end{cases} \tag{8}$$

where β is a hyper-parameter.

The derivative of zigmoid is

$$\frac{dzigmoid\,(x)}{dx} = \frac{dsigmoid\,(x)}{dtrans\,(x)} * \frac{dtrans\,(x)}{dx} \tag{9}$$

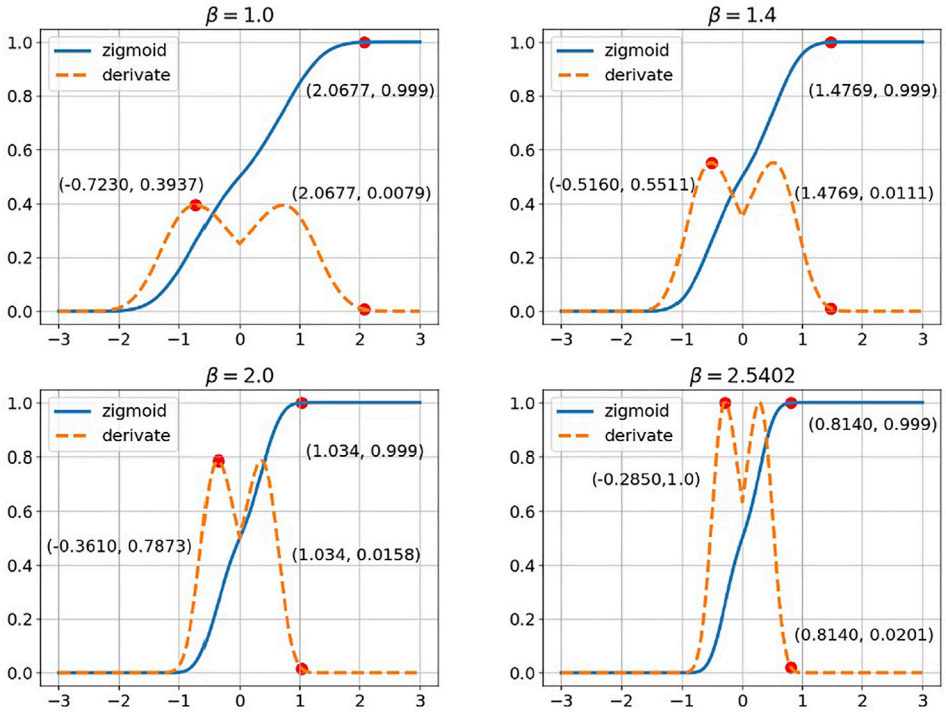


FIGURE 2 | Zigmoid properties.

$$\frac{dtrans(x)}{dx} = \begin{cases} \beta * e^{\beta * x} & x \geq 0 \\ \beta * e^{-\beta * x} & x < 0 \end{cases} \tag{10}$$

When β is a large value, that is, $\max(\frac{dzigmoid(x)}{dx}) \geq 1.0$, this may lead to gradient explosion. The numerical experiments in PyTorch show that such a case can be avoided when $\beta \leq 2.5402$. On the other hand, the lower bound of β is set to 1.0 because a small value of β cannot achieve effective gradient amplification. Therefore, β is ranged to [1.0, 2.5402] to ensure the system stability.

As shown in the comparison of **Figures 1**, **2**, zigmoid amplifies the derivate at the far end of the time axis where sigmoid failed to do so.

2.2 zLSTM

The aforementioned defined zigmoid with ranged β is expected to alleviate the gradient problem for long time series. The following attempts are then conducted to build a new variant of LSTM using the proposed zigmoid (zLSTM) as shown in **Figure 3**:

1. Replace sigmoid function in the forget gate with zigmoid such that the gradient can be effectively amplified (**Figure 3**).
2. Replace i_t with $(1 - f_t)$ (Cho et al., 2014) in order to reduce the trainable parameters.

$$f_t = \boldsymbol{zigmoid}\left(W_{fx}x_t + W_{fh}h_{t-1} + b_f\right) \tag{11}$$

$$o_t = sigmoid\left(W_{ox}x_t + W_{oh}h_{t-1} + b_o\right) \tag{12}$$

$$\hat{c}_t = \tanh\left(W_{cx}x_t + W_{ch}h_{t-1} + b_c\right) \tag{13}$$

$$c_t = f_t \odot c_{t-1} + \left(1 - f_t\right) \odot \hat{c}_t \tag{14}$$

$$h_t = o_t \odot \tanh\left(c_t\right) \tag{15}$$

where W_{ix}, W_{ih}, b_i, W_{fx}, W_{fh}, b_f, W_{ox}, W_{oh}, b_o, W_{cx}, W_{ch}, and b_c are trainable parameters.

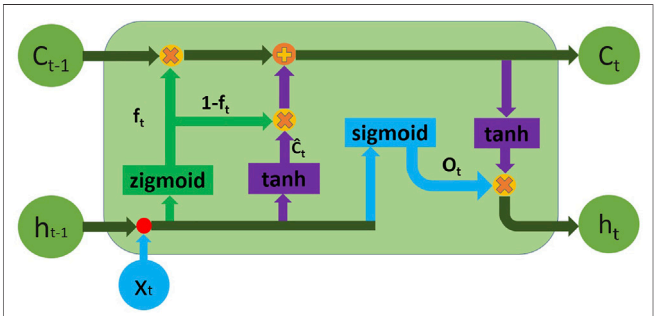


FIGURE 3 | Structure of zLSTM.

2.3 Hyper-Parameter β

An appropriate hyper-parameter β is of great significance for controlling the intensity of the derivative amplification. β should be a value that amplifies the gradient enough for the network to learn long-term information.

Further information to be learned is that the smaller gradient is in zigmoid. For an input sequence with length L:

$$\frac{1}{1 - f_t} = L \tag{16}$$

$$f_t = zigmoid\left(x\right) \tag{17}$$

Therefore,

$$\frac{dzigmoid(x)}{dx} = \frac{L-1}{L^2}\beta\left(1 + log(L-1)\right) \tag{18}$$

The gradient has to be greater than a certain value g_{min} such that the networks can continue the learning. The g_{min} is observed and suggested in this study as 0.01.

$$\frac{L-1}{L^2}\beta\left(1 + log(L-1)\right) \geq g_{min} \tag{19}$$

yields

$$\beta \geq \frac{g_{min}L^2}{(L-1)\left(1 + log(L-1)\right)} \tag{20}$$

Thus, the β value can be calculated as

$$\beta = \begin{cases} 1.0 & \beta_{lim} < 1.0 \\ \beta_{min} & 1.0 \leq \beta_{lim} \leq 2.5402 \\ 2.5402 & \beta_{lim} > 2.5402 \end{cases} \tag{21}$$

where $\beta_{lim} = \frac{g_{min}L}{(L-1)(1+log(L-1))}$.

2.4 Capabilities of zLSTM

zLSTM obtained by the aforementioned procedure has the following capabilities:

1. zLSTM contains basic properties of LSTM:
   a. Non-linearity: As a variant of RNN, zLSTM inherits the non-linearity processing ability.
   b. Short-term memory: Zigmoid maintains LSTM's short-term memory ability.
2. As an improved version of LSTM, zLSTM has advantages such as
   a. Long-term memory: Compared with sigmoid, zigmoid has a greater gradient with the same output. This enhances the long-term performance by allowing the model to conduct learning over the full length of the time series. Longer inputs are then allowed to be fed into the model, enriching the information used for predictions.
   b. Reduced trainable parameters: Given that the sum of forget gate and input gate is 1, the input gate can be omitted to deduct parameters. Gate reduction brings reduced trainable parameters such that the computing time and resource are less than traditional models.

With the properties and advantages mentioned previously, this work proposes zLSTM as a better solution for LOCA prediction due to the following considerations:

1. Compared to traditional machine learning algorithms, such as the ones mentioned in Moshkbar-Bakhshayesh and Ghafari (2022), and feed-forward neural network (Santos et al., 2019), zLSTM can achieve better feature extraction with its LSTM kernel that performs the calculation along timesteps. It is then suggested for the non-linear LOCA process, whose variation features are hard to capture.
2. Long term features can be captured by a combination of existing models, such as CNN + LSTM (Wang et al., 2021c), with a sacrifice of more hyper-parameters and more tuning tricks, which burdens the model hyper-parameter processing and deployment. zLSTM, on the contrary, avoids such process by using an improved structure without additional hyper-parameters.
3. The training process of zLSTM is more executable due to reduced gate and parameters, allowing it to generate training/predicting results with less time and efforts.
4. Compared to those baseline models, the LOCA prediction from zLSTM has better credibility and enhanced generalization performance due to zLSTM's lower overfitting probability and fewer trainable parameters.

3 EXPERIMENTS

3.1 Datasets

The datasets are obtained from LOCA simulations using an industry-grade NPP simulation platform (Sun et al., 2017). The simulations are carried out at 100% reactor power for LOCA cases, that is, break sizes of 0.9, 0.95, 1.0, 1.5, and 2.0 cm^2.

There are a total of twelve crucial system parameters selected as the modeling features:

1. pressurizer water level;
2. coolant average temperature;
3. steam generator No. 1 water level;
4. steam generator No. 2 water level;
5. loop 1 coolant flowrate;
6. loop 2 coolant flowrate;
7. pressurizer pressure;
8. stream generator No. 1 output pressure;
9. stream generator No. 2 output pressure;
10. reactor power;
11. cold leg temperature;
12. hot leg temperature.

3.2 Data Preprocessing

To reduce the influence of multiple dimensions, the dataset is preprocessed using the z-score method such that fast convergency can be achieved during the model training process.

$$x_{scaled} = \frac{x - \bar{x}}{\sigma} \tag{22}$$

where $\bar{x}$ and σ denote mean and variance of x, respectively.

3.3 Metrics

As common metrics for regression task evaluation, mean squared error (MSE) and mean absolute error (MAE) are chosen as the performance judgment for the proposed zLSTM model.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{23}$$

TABLE 1 | Hyper parameters of zLSTM model.

Item	Value
Units of input layer	1[a] or 12[b]
Units of zLSTM of 1st hidden layer	128
Units of zLSTM of 2nd hidden layer	64
Units of output layer	1

[a]*1 for zLSTM univariate (zLSTM-U).*
[b]*12 for zLSTM multivariate (zLSTM-M).*

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{24}$$

where y_i and $\hat{y}_i$ are the original value and predicted value, respectively.

3.4 Hyper-Parameter Setting

The zLSTM structure consists of one input layer, two hidden layers, and one output layer. Details are provided in **Table 1**.

3.5 Baseline Methods

There have been two similar investigations performed by She et al. (2020) and She et al. (2021). However, their major purpose was to verify the feasibility and effects of applying deep learning methods to the LOCA predictions. Neither of them covers the multivariate processing performance that requires theoretical innovation on the NN itself, such as defining a new zigmoid function for the LSTM model. To demonstrate the superiority of using the zigmoid method for multivariate processing, these two previous cases are selected as the baseline, and the prediction accuracy represented by MSE and MAE is compared among all the three methods. The mentioned two previous works, LSTM and CNN-LSTM, are compared with zLSTM to demonstrate its superiority on post-LOCA predictions.

3.6 Model Training

All datasets are randomly split into three subsets, that is, a training set (60%), a validation set (20%), and a test set (20%). Multivariate time series data needed for model training are derived by applying the rolling update method. Following the previous work, the window sizes for LSTM and CNN-LSTM remain 5 and 50. zLSTM uses the same window size as the compared baseline, which means that it uses window size 5 when comparing with LSTM and 50 for CNN-LSTM. The training parameters are optimized using Adam algorithm (Kingma and Adam, 2014) with a learning rate 10^{-3} for all models. When the training starts, a sliding window moves from the first row of the training dataset and provides a series of training input data describing the parameter variation during the period limited by the window size. The model learns and memorizes the variations such that it can reproduce similar ones once the test data is fed to it. The model is trained during such iterations until desired loss value is reached. More training process details are provided in She et al. (2020).

3.7 Model Verification Experiments

The performance of the proposed zLSTM is verified through experiments designed to predict crucial parameters of LOCA, in which both univariate scenario and multivariate scenario are tested using zLSTM and the two baseline methods.

The first crucial parameter chosen as the prediction feature is loop 1 flowrate since it is the most impacted parameter during a LOCA. Flowrate data from the test dataset are the so called "single input" for the univariate scenario. As for the multivariate scenario, all the twelve system parameters are integrated into a vector x_t and fed into the zLSTM for a coupled prediction processing. A single output (loop 1 flowrate prediction) is generated by zLSTM's single-cell output layer that merges the processing results of the 12-dimension vector. The multivariate experiment is only for zLSTM since both baseline methods are originally single-input models without parameter-coupling capability. With a diversity consideration, similar univariate and multivariate experiments are conducted to predict the pressurizer water level as well.

The univariate test is necessary since the two baseline methods are oriented to only one system parameter prediction. During this experiment, the memorizing performance of the models for long- and short-term information is tested, allowing the zLSTM to present its long-term memory advantage with the amplified gradient. For a fair play, zLSTM used for this experiment takes the same single input as the baseline methods. Such univariate-input zLSTM is named zLSTM-univariate (zLSTM-U).

The multivariate test, on the other hand, is to confirm a lower loss value when the prediction is generated with an algorithm (zLSTM) that takes all associated parameters into account. In this case, system parameters associated to the predicted feature are fed to zLSTM as multivariate inputs, naming it zLSTM-multivariate (zLSTM-M).

3.7.1 Prediction of Loop 1 Coolant Flowrate

As mentioned in the verification process introduction, the univariate experiment uses the single input value for LSTM, CNN-LSTM, and zLSTM-U. The multivariate experiment, which is for zLSTM-M only, yields a single predicted feature (flowrate or water level) using a 12-dimension vector containing all the key system parameters. The experiments in this subsection focus on the variation of the loop 1 coolant flowrate during LOCAs of five different break sizes. Regarding the discussion in **Section 3.6**, both zLSTM-U and zLSTM-M select window size 5 to run the univariate and multivariate tests against LSTM, which forms a test group {LSTM, zLSTM-U5, zLSTM-M5}. When comparing to CNN-LSTM that has window size 50, the test group becomes {CNN-LSTM, zLSTM-U50, zLSTM-M50}. Model performance is evaluated using MSE and MAE for each of the six models, providing twelve accuracy evaluation results for each of the five LOCA cases. **Table 2** shows all these 60 results for the loop 1 flowrate predictions. The predicted LOCA trends are plotted in **Figures 4**, **5**.

TABLE 2 | Prediction accuracy evaluations for loop 1 flowrate.

	Metric	Model	0.9	0.95	1.0	1.5	2.0
LSTM vs. zLSTM	MAE	LSTM	1.751×10^{0}	1.841×10^{0}	3.023×10^{-3}	1.114×10^{0}	4.800×10^{-3}
		zLSTM-U5	1.598×10^{0}	1.605×10^{0}	2.732×10^{-3}	1.040×10^{0}	3.928×10^{-3}
		zLSTM-M5	***9.235 × 10⁻¹***	***9.243 × 10⁻¹***	***2.008 × 10⁻³***	***9.603 × 10⁻¹***	***3.443 × 10⁻³***
	MSE	LSTM	1.039×10^{1}	1.048×10^{1}	2.700×10^{-5}	7.702×10^{0}	8.000×10^{-5}
		zLSTM-U5	9.835×10^{0}	9.850×10^{0}	2.300×10^{-5}	4.659×10^{0}	5.700×10^{-5}
		zLSTM-M5	***3.458 × 10⁰***	***3.444 × 10⁰***	***1.800 × 10⁻⁵***	***3.984 × 10⁰***	***5.000 × 10⁻⁵***
CNN-LSTM vs. zLSTM	MAE	CNN-LSTM	2.353×10^{0}	2.270×10^{0}	2.615×10^{-3}	1.223×10^{0}	3.848×10^{-3}
		zLSTM-U50	9.006×10^{-1}	8.550×10^{-1}	2.003×10^{-3}	1.030×10^{0}	3.187×10^{-3}
		zLSTM-M50	***7.217 × 10⁻¹***	***7.241 × 10⁻¹***	***1.630 × 10⁻³***	***1.025 × 10⁰***	***2.330 × 10⁻³***
	MSE	CNN-LSTM	2.140×10^{1}	2.322×10^{1}	2.700×10^{-5}	6.878×10^{0}	5.700×10^{-5}
		zLSTM-U50	4.054×10^{0}	4.030×10^{0}	1.500×10^{-5}	5.231×10^{0}	5.300×10^{-5}
		zLSTM-M50	***2.126 × 10⁰***	***2.121 × 10⁰***	***1.300 × 10⁻⁵***	***4.141 × 10⁰***	***4.400 × 10⁻⁵***

Underline data, the best result of univariate experiments; bold data, the best result of multivariate and univariate experiments.

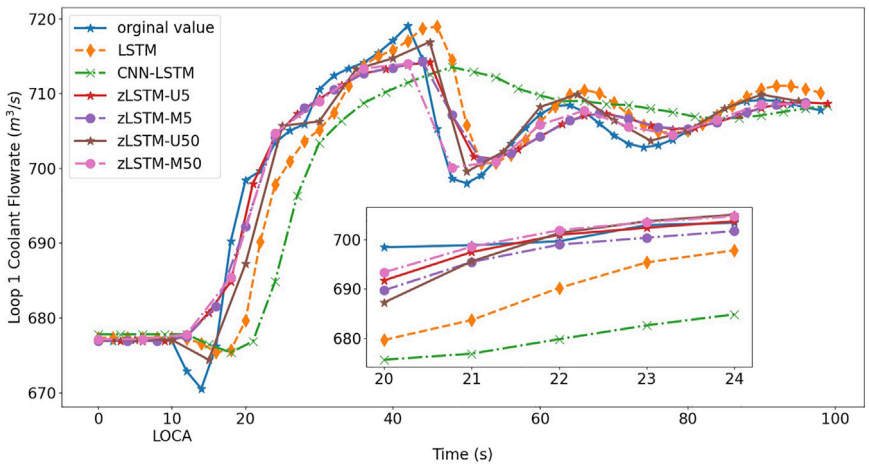


FIGURE 4 | Prediction of loop 1 coolant flowrate on break size 0.9 cm^2.

3.7.2 Prediction of Pressurizer Water Level

As another crucial system parameter describing the LOCA behavior, the pressurizer water level is predicted in this subsection by experiments same as in **Section 3.7.1**. **Table 3** presents the 60 MSE/MAE data as the accuracy evaluations of the two test groups. Meanwhile, the water level variation illustrated by all the models is presented in **Figures 6**, **7**.

3.7.3 Result Analysis

The prediction accuracy metrics listed in **Tables 2**, **3** describe the prediction performance of the tested models. The upper half of each table presents the prediction errors from zLSTM (both U and M) against those from LSTM. It can be seen that zLSTM achieves lower errors than LSTM on both MAE and MSE. For instance, the average MAE difference between LSTM and zLSTM-U5 for flowrate prediction is 0.26 in **Table 2**, giving a 28.7% improvement on prediction accuracy. Additionally, zLSTM working on multivariate mode also beats its univariate mode due to the advantages brought by the 12-dimension parameter vector. The average MAE difference between zLSTM-U5 and zLSTM-M5 is 0.24, giving an accuracy improvement of 28.8%. Similar comparison is reflected by the lower half of each table, where zLSTM once again proves its superiority over CNN-LSTM with accuracy improvements such as 29.52% for pressurizer water level prediction (MSE in **Table 3**, CNN-LSTM vs. zLSTM-U50).

Figure 4 to **Figure 7** visually illustrate the predicted LOCA trends in different LOCA cases. The mini graphs within the figures amplify chosen segments of the trends, offering a better view to the model performance. After the LOCA occurs at $t = 10s$, the loop 1 flowrate and the pressurizer water level experience dramatical variations, and then approach a stable state with help from the emergency response system. During the entire process, the zLSTM group {U5, M5, U50, M50} represents a more precise prediction performance. At the beginning of the LOCA, it is the zLSTM that grasps the suddenly inserted non-linear variation using its efficient short-term memory, producing a prediction close

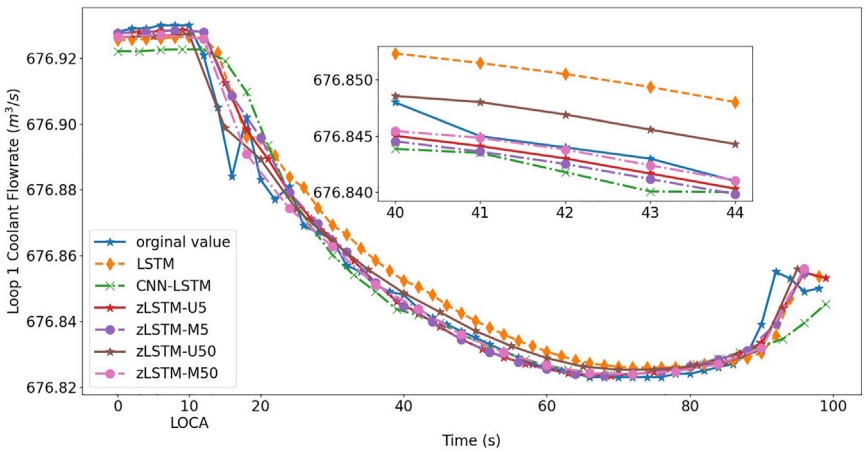


FIGURE 5 | Prediction of loop 1 coolant flowrate on break size 1.0 cm^2.

TABLE 3 | Prediction accuracy evaluations for pressurizer water level.

	Metric	Model	0.9	0.95	1.0	1.5	2.0
LSTM	MAE	LSTM	1.505×10^{-1}	1.519×10^{-1}	4.167×10^{-2}	4.791×10^{-1}	8.124×10^{-2}
vs.		zLSTM-U5	1.402×10^{-2}	1.393×10^{-2}	3.855×10^{-2}	4.698×10^{-1}	7.919×10^{-2}
zLSTM		zLSTM-M5	***8.800×10^{-4}***	***8.990×10^{-4}***	***1.214×10^{-6}***	***1.346×10^{-3}***	***5.000×10^{-6}***
	MSE	LSTM	7.775×10^{-2}	7.898×10^{-2}	2.423×10^{-3}	3.794×10^{-1}	9.233×10^{-3}
		zLSTM-U5	7.723×10^{-2}	7.723×10^{-2}	2.126×10^{-3}	3.453×10^{-1}	8.841×10^{-3}
		zLSTM-M5	***1.618×10^{-2}***	***1.894×10^{-2}***	***9.000×10^{-4}***	***2.939×10^{-2}***	***1.761×10^{-3}***
CNN-LSTM	MAE	CNN-LSTM	9.192×10^{-2}	9.238×10^{-2}	3.150×10^{-2}	4.221×10^{-1}	3.595×10^{-2}
vs.		zLSTM-U50	5.561×10^{-2}	4.991×10^{-2}	2.770×10^{-2}	3.777×10^{-1}	3.373×10^{-2}
zLSTM		zLSTM-M50	***1.264×10^{-2}***	***1.410×10^{-2}***	***5.940×10^{-4}***	***2.448×10^{-2}***	***9.870×10^{-4}***
	MSE	CNN-LSTM	3.231×10^{-2}	3.234×10^{-2}	1.359×10^{-3}	3.216×10^{-1}	2.872×10^{-3}
		zLSTM-U50	1.227×10^{-2}	1.131×10^{-2}	1.196×10^{-3}	2.416×10^{-1}	2.029×10^{-3}
		zLSTM-M50	***4.600×10^{-4}***	***4.860×10^{-4}***	***6.000×10^{-6}***	***8.740×10^{-4}***	***2.00×10^{-5}***

Underline data, the best result of univariate experiments; bold data, the best result of multivariate and univariate experiments.

to the actual trend. How the zigmoid function can enhance the model's long-term memory is well verified when the zLSTM models generate better predictions at the far end of the time axis. zLSTM-M50 is the one that grasps the progression trends most accurately and persistently, which demonstrate the importance and effect of using multivariate processing (12-dimension vector) and wide data window (size 50).

The analysis conducted to investigate further explanations is presented as follows:

1. The multivariate mode of zLSTM (zLSTM-M) allows the prediction to be generated based on the coupling of system parameters, that is, the prediction comprehensively considers all the 12 critical system parameters relevant to LOCA progression. Sufficient information provided by such a 12-variable input vector guarantees improved prediction accuracy.

2. LOCA predictions for small breaks received higher loss values than those for big ones. They confirm the difficulties of learning and simulating a process with dramatical variations, for example, a small LOCA. Inflect points shorten the time period necessary for information gathering, preventing the model from sufficient evaluation of the progression process. The loss values rise along with the number of inflect points, showing that more inflect points cause more missing information during learning and prediction.

3. It is observed that a larger input window size gives the zLSTM model a better performance since the window size determines the coverage of critical information. More accurate results are generated when the model is capable of learning comprehensively by capturing more useful information from the predicted process.

To summarize, the application of zigmoid function to LSTM has enhanced the short and long-term memory of the model. With the input vector integrated using 12 system parameters, the zLSTM-M model can be even more comprehensive to the

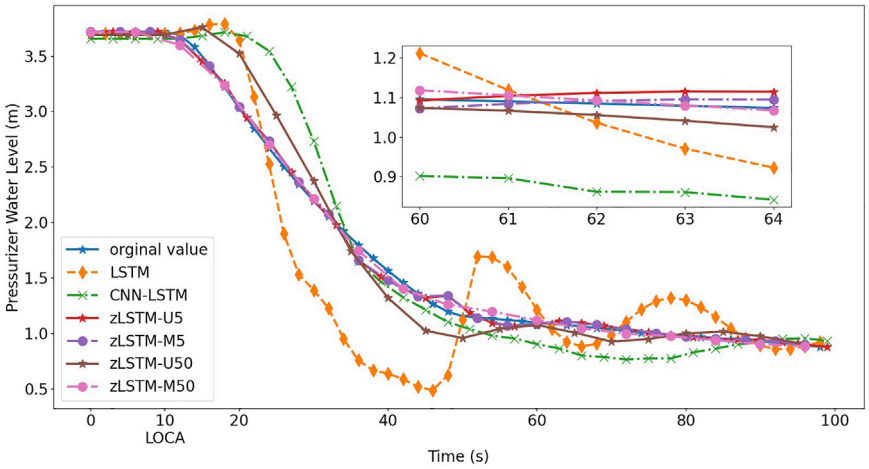


FIGURE 6 | Prediction of pressurizer water level on break size 0.9 cm^2.

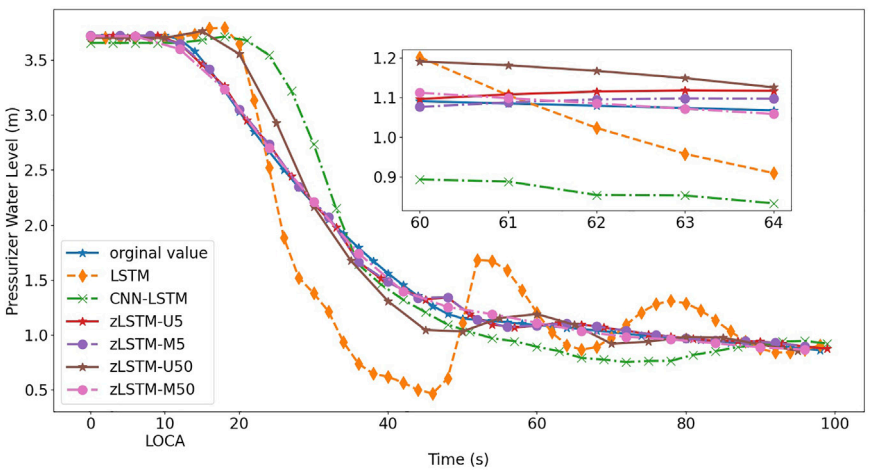


FIGURE 7 | Prediction of pressurizer water level on break size 0.95 $cm.^2$

multivariate environment of LOCA, allowing the predicted feature to be more accurate.

4 CONCLUSION

A new gate function zigmoid is raised as a solution to the far-end gradient problem of RNN class models, which is proposed to cover the non-linearity, time correlation, and multivariate processing for LOCA predictions. Proved through theoretical analysis, the zigmoid function is embedded into traditional LSTM to form zLSTM that is capable of effectively memorizing both short- and long-term information. Its multivariate processing is enabled by using a 12-dimension input vector that integrates 12 system parameters. The multivariate mode gathers all-sided system information that eliminates blind spots during the prediction process. The verification experiments successfully demonstrate the aforementioned advantages of the zLSTM model. The accuracy metrics (MAE/MSE) of zLSTM is kept lower than traditional models for both univariate and multivariate scenarios. During the LOCA progression, the parameter trends are followed by zLSTM's prediction, with the smallest deviation according to the experiment figures. All these findings confirm zLSTM to be a better method for LOCA predictions.

In addition to the achievements, there are a few issues remaining for future investigation. First, zLSTM is constructed by replacing only the forget gate in LSTM. Possible further enhancement could be obtained with more applications of the zigmoid function. The next is the model training process that may be improved using more actual NPP data. Last but not least, inflect points in the LOCA trend cannot be well followed by the

prediction curve generated from the deep learning models, which implies alternative solutions in future explorations.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SG proposed the idea of using zLSTM methods for LOCA prediction and established the main structure of the deep learning models used in this work. SY performed the data preprocessing and model training task. JS drafted most parts of the article and coordinated the cooperation of all the co-authors. WL was responsible for the experiment results analysis as well as preparing the tables and figures. SL provided key instructions to the group members to ensure accurate and efficient research methodologies. All authors contributed to the article and approved the submitted version.

FUNDING

The authors would like to acknowledge the financial and technical support received from the following research projects and institutions, including but not limited to, National Key Research and Development Project (2020YFB1713400), and The Industrial Internet Innovation and Development Project of China (TC19084DY).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenrg.2022.852349/full#supplementary-material

REFERENCES

Chen, Y., Lin, M., Ren, Y., and Wang, T. (2021). Research on Simulation and State Prediction of Nuclear Power System Based on Lstm Neural Network. *Sci. Technology Nucl. Installations* 2021, 8839867. doi:10.1155/2021/8839867

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., et al. (2014). "Learning Phrase Representations Using RNN Encoder–Decoder for Statistical Machine Translation," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, October 25 and 29, 2014 (Doha, Qatar: Association for Computational Linguistics), 1724–1734.

Hochreiter, S., and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Comput.* 9 (8), 1735–1780. doi:10.1162/neco.1997.9.8.1735

Hu, G., Zhou, T., and Liu, Q. (2021). Data-driven Machine Learning for Fault Detection and Diagnosis in Nuclear Power Plants: A Review. *Front. Energ. Res.* 9, 663296.

Kingma, D., and Adam, J. B. (2014). *A Method for Stochastic Optimization*. San Diego, CA USA: Computer Science.

Koo, Y. D., Man, G. N., Kim, K. S., and Kim, C. H. (2018). "Prediction of Nuclear Reactor Vessel Water Level Using Deep Neural Networks," in 2018 International Conference on Electronics, Information, and Communication (ICEIC), Honolulu, HI, USA, 24-27 Jan. 2018 (IEEE). doi:10.23919/elinfocom.2018.8330616

Lee, C.-K., Kwon, G.-Y., Chang, S. J., Jung, M. K., Park, J. B., Kim, H.-S., et al. (2017). Real-Time Condition Monitoring of LOCA via Time-Frequency Domain Reflectometry. *IEEE Trans. Instrum. Meas.* 66, 1864–1873. doi:10.1109/tim.2017.2664578

Li, X., Huang, T., Cheng, K., Qiu, Z., and Tan, S. (2021). Research on Anomaly Detection Method of Nuclear Power Plant Operation State Based on Unsupervised Deep Generative Model. *Ann. Nucl. Energ.* 167, 108785.

Lin, T.-H., Chen, C., Wu, S.-C., Wang, T.-C., and Ferng, Y.-M. (2022). Localization and Size Estimation for Breaks in Nuclear Power Plants. *Nucl. Eng. Technology* 54 (1), 193–206. doi:10.1016/j.net.2021.07.007

Lin, T.-H., Wang, T.-C., and Wu, S.-C. (2021). Deep Learning Schemes for Event Identification and Signal Reconstruction in Nuclear Power Plants with Sensor Faults. *Ann. Nucl. Energ.* 154, 108113. doi:10.1016/j.anucene.2020.108113

Mira, B., Pigg, C., Kozlowski, T., Deng, Y., and Qu, A. (2020). Neural-based Time Series Forecasting of Loss of Coolant Accidents in Nuclear Power Plants - Sciencedirect. *Expert Syst. Appl.* 160.

Moshkbar-Bakhshayesh, K., and Ghafari, M. (2022). Prediction of Steam/water Stratified Flow Characteristics in Npps Transients Using Svm Learning Algorithm with Combination of thermal-hydraulic Model and New Data Mapping Technique. *Ann. Nucl. Energ.* 166, 108699.

Pantera, L., Stulík, P., Vidal-Ferràndiz, A., Carreño, A., Ginestar, D., George, I., et al. (2021). Localizing Perturbations in Pressurized Water Reactors Using One-Dimensional Deep Convolutional Neural Networks. *Sensors* 22 (1), 113.doi:10.3390/s22010113

Saeed, H. A., Peng, M.-j., Wang, H., and Zhang, B.-w. (2020). Novel Fault Diagnosis Scheme Utilizing Deep Learning Networks. *Prog. Nucl. Energ.* 118, 103066. doi:10.1016/j.pnucene.2019.103066

Santhosh, M. K., Thangamani, I., Mukhopadhyay, D., Verma, V., Rao, V. V. S. S., Vaze, K. K., et al. (2010). "Neural Network Based Diagnostic System for Accident Management in Nuclear Power Plants," in 2010 2nd International Conference on Reliability, Safety and Hazard - Risk-Based Technologies and Physics-of-Failure Methods (ICRESH), Mumbai, India, 14-16 Dec. 2010 (IEEE), 572–578. doi:10.1109/icresh.2010.5779613

Santos, M. C. d., Pinheiro, V. H. C., Desterro, F. S. M. d., Avellar, R. K. d., Schirru, R., Santos Nicolau, A. d., et al. (2019). Deep Rectifier Neural Network Applied to the Accident Identification Problem in a Pwr Nuclear Power Plant. *Ann. Nucl. Energ.* 133, 400–408. doi:10.1016/j.anucene.2019.05.039

She, J. K., Xue, S. Y., Sun, P. W., and Cao, H. S. (2020). "The Application of LSTM Model to the Prediction of Abnormal Condition in Nuclear Power Plants," in *Nuclear Power Plants: Innovative Technologies for Instrumentation and Control Systems, the Fourth International Symposium on Software Reliability, Industrial Safety, Cyber Security and Physical Protection of Nuclear Power Plant (ISNPP)*. doi:10.1007/978-981-15-1876-8_46

She, J., Shi, T., Xue, S., Zhu, Y., Lu, S., Sun, P., et al. (2021). Diagnosis and Prediction for Loss of Coolant Accidents in Nuclear Power Plants Using Deep Learning Methods. *Front. Energ. Res.* 9, 665262. doi:10.3389/fenrg.2021.665262

Sun, P., Zhao, H., Liao, L., Zhang, J., and Su, G. (2017). Control System Design and Validation Platform Development for Small Pressurized Water Reactors (Spwr) by Coupling an Engineering Simulator and Matlab/simulink. *Ann. Nucl. Energ.* 102 (APR), 309–316. doi:10.1016/j.anucene.2016.12.034

Tallec, C., and Ollivier, Y. (2018). *Can Recurrent Neural Networks Warp Time?* Arxiv.

Wang, H., Peng, M. J., Ayodeji, A., Xia, H., Wang, X. K., and kang, Z. (2021). Advanced Fault Diagnosis Method for Nuclear Power Plant Based on Convolutional Gated Recurrent Network and Enhanced Particle Swarm

Optimization. *Ann. Nucl. Energ.* 151, 107934. doi:10.1016/j.anucene.2020.107934

Wang, H., Peng, M. J., Yue, Y., Saeed, H., Cheng, M. H., and Liu, Y. K. (2021). Fault Identification and Diagnosis Based on Kpca and Similarity Clustering for Nuclear Power Plants. *Ann. Nucl. Energ.* 150, 107786.

Wang, M-D., Lin, T-H., Jhan, K-C., and Wu, S-C. (2021). Abnormal Event Detection, Identification and Isolation in Nuclear Power Plants Using Lstm Networks. *Prog. Nucl. Energ.* 140, 103928. doi:10.1016/j.pnucene.2021.103928

Xiang, L., Fu, X. M., Xiong, F. R., and Bai, X. M. (2020). Deep Learning-Based Unsupervised Representation Clustering Methodology for Automatic Nuclear Reactor Operating Transient Identification. *Knowledge-Based Syst.* 204, 1061782020.

Yang, J., Yang, Y., Deng, C., and Ishii, M. (2019). Best Estimate Plus Uncertainty Analysis of a Large Break Loca on Generation Iii Reactor with Relap5. *Ann. Nucl. Energ.* 127 (MAY), 326–340. doi:10.1016/j.anucene.2018.12.019

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 13 April 2022
doi: 10.3389/fenrg.2022.849874

Intelligent Optimization Method for Core Flow Zoning of Long-Cycle Lead-Bismuth-Cooled Reactor

Qingyuan Yu[1], Shengqi Dai[1], Pengcheng Zhao[1,2], Yanan Zhao[1], Yingjie Xiao[1], Liangxing Peng[1] and Tao Yu[1]**

[1]School of Nuclear Science and Technology, University of South China, Hengyang, China, [2]Science and Technology on Reactor System Design Technology Laboratory, Nuclear Power Institute of China, Chengdu, China

Flow zoning is an important way to achieve core outlet temperature flattening. Appropriate zoning can improve safety and economy. This study combines an artificial intelligence optimization algorithm with a parallel multi-channel model to develop a model for calculating reactor core flow zoning based on the modern optimization theory, convergence analysis of a genetic algorithm, differential evolution algorithm, and quantum genetic algorithm is carried out for long-life reactor flow partitioning. Using the optimized algorithm, two flow rates are determined using power distribution at the beginning of the core life as the sample data and the maximum power of each fuel assembly during the entire life as the sample data. Comparative analysis of two different flow zoning schemes is implemented on a small long-life natural circulation lead-bismuth fast reactor, SPALLER-100. The findings of this study show that the quantum genetic algorithm has the best convergence for the long-life reactor among the three intelligent optimization algorithms, and it can quickly provide optimal results. In flow zoning scheme calculations based on the core power distribution at the beginning of reactor life, the maximum outlet temperature of the fuel assembly exceeds the thermal safety limit of the reactor, and in the flow zoning scheme calculations based on the average core power distribution during the whole reactor life, the maximum outlet temperature of the fuel assembly is 140 K lower than the maximum outlet temperature obtained in the previous scheme, remaining below the thermal safety limit. The optimal number of partitions for the SPALLER-100 reactor is determined to be 5, and increasing the number of zones only slightly improved the thermal safety performance of the reactor.

Keywords: lead-bismuth reactor, flow zoning, intelligent optimization algorithm, quantum genetic algorithm, safety performance

OPEN ACCESS

Edited by:
Jun Wang,
University of Wisconsin-Madison, United States

Reviewed by:
Mingjun Wang,
Xi'an Jiaotong University, China
Chenglong Wang,
Xi'an Jiaotong University, China
Ronghua Chen,
Xi'an Jiaotong University, China

****Correspondence:***
Pengcheng Zhao
zpc1030@mail.ustc.edu.cn
Tao Yu
taoyu0626@163.com

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *06 January 2022*
Accepted: *22 March 2022*
Published: *13 April 2022*

Citation:
Yu Q, Dai S, Zhao P, Zhao Y, Xiao Y, Peng L and Yu T (2022) Intelligent Optimization Method for Core Flow Zoning of Long-Cycle Lead-Bismuth-Cooled Reactor. Front. Energy Res. 10:849874. doi: 10.3389/fenrg.2022.849874

INTRODUCTION

From the perspective of thermal-hydraulic design, the core flow distribution of nuclear reactors is one of the most important issues. At the core outlet, the mixing of coolant temperatures causes temperature oscillation with a specific frequency and amplitude, resulting in thermal stress in the thermal measurement device and the central measurement column. Core flow zoning is an important means to realize core outlet temperature flattening. Flow distribution between different components is realized by opening specific component pins. High-power components have larger flows than low-power components, because of which, different components have similar outlet temperatures. The

non-uniform distribution of the outlet temperatures of the reactor core will greatly limit the operation life of the reactor. Reasonable zoning can improve the safety and economy of the reactor. Therefore, to improve the operation life of the long-life reactor and realize the rapid optimization of core outlet temperature flattening, it is urgent to develop a new core flow zoning method.

Flow zoning to ensure reactor safety and performance has been investigated by many researchers. Nagy et al. (2012) analyzed the influence of flow zoning on the life of a graphite-moderated of a molten salt reactor and examined the fuel proliferation efficiency. The fuel proliferation efficiency or the life of graphite was increased by separating the core into two radial regions with different amounts of graphite. Basualdo et al. (2020) coupled the sub-channel code SUBCHANFLOW with the reactor dynamics diffusion code PARCS. This coupling aimed to improve the physical model involved in the core analysis by enhancing the prediction accuracy. Moreover, it enabled the detailed prediction of local thermal-hydraulics parameters. Zhao et al. (Chen et al., 2014) developed a non-uniform heating parallel channel flow field calculation code for a small modular natural circulation lead- or lead alloy-cooled fast reactor (LFR) and compared it with the CFD method to verify the accuracy of the code, and then, based on the results of optimization analysis, they proposed a 10 MW natural circulation LFR core flow distribution optimization design. Liu et al. (Yizhe et al., 2012) developed the hydraulic characteristic calculation code DAEMON to calculate the flow distribution of the whole core under different working conditions, according to the design characteristics of the core and the primary circuit of the China Experimental Fast Reactor (CEFR). Zhou et al. (2018) developed and verified a code for the thermal-hydraulic design and optimization of a CFR600 fast reactor; the functions included in the code were fast reactor full-pile graphical modeling, fine sub-channel automatic division, thermal-hydraulic analysis considering heat transfer between components, and automatic flow zoning optimization. Li et al. (2019) divided the reactor core into external, middle, and internal fuel zones, calculated the minimum coolant flow required for each fuel assembly in each fuel zone, and then, divided the flow of the reactor core according to the different values of adjacent minimum coolant flows. Based on intelligent optimization algorithm, Wang et al. (2020) established the optimization model of flow partition design. Considering the maximum fuel temperature limit and the cladding temperature limit as boundary conditions, the total flow of the reactor core was minimized; the average temperature of the reactor core outlet was the highest, providing a novel method to address flow partition in large reactor cores. Zhang et al. (2017) used genetic algorithm (GA) and artificial neural network (ANN) (Cong et al., 2011) to predict CHF and succeeded to correlate the existing CHF data with better accuracy than the existing empirical correlations. Moreover, Chen et al. (2010) proposed mechanism-based correlations for LBB leakage by genetic algorithm. The presented correlations provide higher precision than the existing correlation.

With the increase of the number of fuel assemblies and partitions, resulting in poor convergence of the algorithm, longer time required for calculation, which cannot realize the rapid optimization design of core outlet temperature flattening. The power of each fuel assembly changes with changes in core life. In model establishment and calculation of results, the impact of different life periods on reactor power distribution should be considered. The designed flow zoning method should also ensure that the reactor is below the thermal safety limit during the entire life period. Based on the modern optimization theory, the study about artificial intelligence optimization algorithms with a parallel multi-channel model, and develops a reactor core flow zoning calculation model; the convergence of genetic algorithm, differential evolution algorithm, and quantum genetic algorithm for addressing the flow zoning problem of long-life reactors is analyzed, and the convergence optimal algorithm is obtained. According to the obtained optimal algorithm, for a small long-life natural circulation lead-bismuth fast reactor, SPALLER-100, considering two data samples, power distribution at the beginning of life and the maximum power of each fuel assembly in the whole life, two different flow zoning schemes are compared and analyzed, and the optimal flow zoning scheme is determined. It can be applied to other types of reactors to provide reference help for the design of the reactor, which will be strengthened in future study.

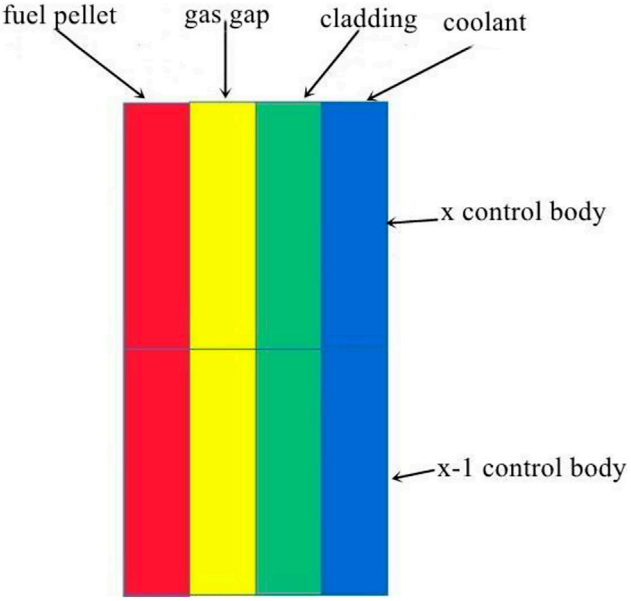


FIGURE 1 | Fuel rod heat conduction model.

MATHEMATICAL PHYSICAL MODEL

Fuel Rod Heat Conduction Model

In this study, the fuel rod heat transfer model is used to describe the heat transfer process of a fuel rod in a reactor core, without considering the exchange of mass and momentum between channels. The heat conduction model is established using the rod bundle fuel rod, and its axial and radial control bodies are divided, as shown in **Figure 1**.

The axial heat conduction of fuel rods and the heat source inside the cladding are ignored, and the heat transfer between control bodies at a certain height in the core active zone is expressed as follows:

(1) Coolant controller:

$$Q(x) = Wc_p\left(t_{f,x} - t_{f,x-1}\right) \tag{1}$$

(2) External surface of the cladding control body:

$$t_{cs}(x) = t_f(x) + \frac{q_l(x)}{\pi d_{cs} h_x} \tag{2}$$

(3) External surface of the cladding control body:

$$t_{ci}(x) = t_{cs}(x) + \ln\left(d_{cs}/d_{ci}\right)\frac{q_l(x)}{2\pi k_{\text{clad}}} \tag{3}$$

(4) External surface of the fuel pellets control body:

$$t_u(x) = t_{ci}(x) + \ln\left(d_{ci}/d_u\right)\frac{q_l(x)}{2\pi k_{\text{gas}}} \tag{4}$$

(5) Fuel pellets control body central temperature:

$$t_o(x) = t_u(x) + \frac{q_l(x)}{4\pi k_u} \tag{5}$$

In the above equations, Q is the heat absorbed by the coolant (W), W is mass flow (kg/s), Cp is the average constant pressure specific heat capacity ($J \cdot kg^{-1} \cdot K^{-1}$), x is the axial control body symbol, $q_l(x)$ is the linear power density of the xth axial controller (W/m), h is the convective heat transfer coefficient ($W \cdot (m^{-2}K^{-1})$), k_{cald} is the thermal conductivity of the cladding ($W/(m \cdot K)$), k_{gas} is the thermal conductivity of interstitial gas ($W/(m \cdot K)$), k_u is the thermal conductivity of the core ($W/(m \cdot K)$), d_{cs} is the largest outer diameter of fuel rods (m), d_{ci} is the inner diameter of the cladding (m), d_u is the core diameter (m), $t_f(x)$ is the temperature of the xth coolant control body (K), $t_{cs}(x)$ is the x cladding control body surface temperature (K); $t_{ci}(x)$ is the x cladding to control the body surface temperature (K), $t_u(x)$ is the x core control body surface temperature (K), $t_o(x)$ is the xth core control center temperature (K).

Parallel Multi-Channel Model

The parallel multi-channel model considers that the required heat pipe is isolated and closed. To facilitate calculation, it is assumed that there is no exchange momentum, energy, and mass of coolant with adjacent channels at any point in the core height. The parallel multi-channel model is most suitable for analyzing closed channels, but it can also be used for open channels to simplify calculation. Due to the existence of transverse mixing, the parallel multi-channel model is rough, and transverse mixing engineering heat pipe factors should be used to control increase in enthalpy. The parallel multi-channel model includes following steps.

(1) Enter data and determine the number of partitions.
(2) Identify the hottest pipe in each partition according to the partition scheme.
(3) Considering the reactor thermal-hydraulic pressure and core power distribution, determine the hottest tube of the cladding.
(4) Calculate the outlet temperature of the hottest pipe coolant, and obtain the hottest pipe flow in each zone.
(5) Calculate the total flow of the whole reactor, and obtain the average core outlet temperature (T_{out}) according to the relationship between the average core outlet temperature and the total reactor flow. The relationship between average core outlet temperature and total reactor flow is given as follows:

$$T_{out} = T_{in} + P\Big/\left(C_p W\right) \tag{6}$$

where T_{in} is the core inlet temperature (K), P is the total core power (W).

Designing the Intelligent Optimization Algorithm for Life Cycle Traffic Zoning

The genetic algorithm (Zhang, 2017), differential evolution algorithm (Yang and Gu, 1997), and simulated annealing algorithm (Gai, 2017) are contemporary commonly used optimization algorithms. The novel quantum genetic algorithm (QGA) (Ying et al., 2018; Liu et al., 2020) provides new ideas for solving optimization problems. In this study, the genetic algorithm, differential evolution algorithm, and QGA are coupled with the parallel multi-channel model for determining the most suitable intelligent optimization algorithm to calculate the flow partition of a long-life miniaturized reactor, and then the convergence of each algorithm is analyzed. The calculation results of these three algorithms are compared, and the optimal algorithm to solve the core flow zoning problem is determined. The main steps of solving flow zoning are as follows.

(6) Import the initial data and design the initial partition scheme. The partition number (K), of this scheme starts from 1.
(7) For each component, the partition number k is randomly allocated between 1 and K.
(8) According to the partition results obtained in the second step, calculate the average core outlet temperature using the parallel multi-channel model.
(9) Obtain the optimal flow zoning scheme under the number of zones and its corresponding total core flow and average core outlet temperature. If the outlet temperature meets the requirements, output the zoning scheme to end the cycle. Otherwise, obtain a new zoning scheme using the intelligent optimization algorithm.
(10) Obtain another new partition scheme using the intelligent optimization algorithm.

In this study, three intelligent optimization algorithms will be used for calculation, namely, genetic algorithm, differential

evolution algorithm, and QGA. The main steps of the genetic and differential evolution algorithms are as follows.

Coding Design.

It is assumed that there are k possibilities in the flow area of each fuel assembly and s fuel assemblies in the whole core. To ensure that the algorithm can determine the optimal zoning scheme, the optimization domain should include all flow zoning schemes. The binary code is used to code the m_i of the fuel assembly flow zone, and the required binary bit number is $\log_2 K$ up integer, denoted as T. For example, when the number of zones is 4, the required binary bit number is 2, and the binary codes of flow areas 1–4 are 00, 01, 10, and 11, respectively.

Individual Fitness Design.

To achieve the optimization goal of minimizing the total core flow in the flow area, the fitness function is defined as:

$$Fit = \sum_{j=1}^{K} \left(n_j \cdot \min q_j \right) \tag{7}$$

where K is the number of zones, j is the number of the flow zone, and n_j is the number of fuel assemblies in the jth flow zone. $\min q_j$ is the minimum flow rate required by the maximum power component in the jth flow zone, which is determined by the fuel rod temperature limit, that is, the minimum flow not exceeding the temperature limit.

Evolution and Mutation.

This is the core step of the three algorithms, and it is also what differentiates them. The genetic algorithm and the differential evolution algorithm use the same evolutionary operation but different mutation operations. The genetic algorithm performs mutation processing according to the previously set mutation operation, while differential genetic algorithm uses mutation vector to realize population evolution:

$$v_{i,G+1} = x_{best,G} + \lambda g \left(x_{r1,G} - x_{r2,G} \right) \tag{8}$$

where λ is a random number from [0,1], and $x_{r1,G}$ and $x_{2r,G}$ are different chromosomes.

The QGA combines two special operation methods, namely, quantum computing and genetic algorithm, and thus, it has the advantages of two operation methods and is a probabilistic optimization algorithm. QGA is based on the state vector representation of quantum, and the probability amplitude of quantum bit is applied to chromosome coding, because of which, a chromosome can express the superposition of multiple states. The quantum logic gate is used to realize the chromosome update operation, thereby realizing the optimal solution of the target.

The main differences between QGA and genetic algorithm is the different coding methods and the population evolution modes. QGA replaces the chromosome coding operator of genetic algorithm with a quantum bit probability symbol. The standard QGA uses a quantum revolving gate to update the population rather than the evolution and crossover operation used in the standard genetic algorithm. The quantum revolving gate and its renewal process are given as follows:

TABLE 1 | Quantum rotation gate updating table.

x_i	b_i	$f(x_i) \geq f(b_i)$	$\Delta\theta i$	$s(\alpha_i, \beta_i)$			
				$\alpha_i\beta_i > 0$	$\alpha_i\beta_i < 0$	$\alpha_i = 0$	$\beta_i = 0$
0	0	False	0	—	—	—	—
0	0	True	0	—	—	—	—
0	1	False	δ	+1	-1	0	±1
0	1	True	δ	−1	+1	±1	0
1	0	False	δ	−1	+1	±1	0
1	0	True	δ	+1	-1	0	±1
1	1	False	0	—	—	—	—
1	1	True	0	—	—	—	—

"—"—"nil";δ—;Δθ$_i$—;b $_i$—;x $_i$—;f—.

$$U(\theta_i) = \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{bmatrix} \tag{9}$$

$$\begin{bmatrix} \alpha_i' \\ \beta_i' \end{bmatrix} = U(\theta_i) \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} \tag{10}$$

$$\theta_i = s(\alpha_i, \beta_i) \Delta\theta_i \tag{11}$$

where $(\alpha_i, \beta_i)^T$ and $(\alpha'_i, \beta'_i)^T$ are the probability amplitudes before and after the update of the ith qubit rotation gate of the chromosome, respectively; U is the quantum rotation gate; θ_i is the rotation angle, calculated from **Equation 11**, and $\Delta\theta_i$ is the magnitude of the rotation angle; $s(\alpha_i, \beta_i)$ is the direction of rotation. The values of δ and $s(\alpha_i, \beta_i)$ are given by 8 cases consisting of x_i, b_i, $f(x_i)$ and $f(b_i)$. The specific value-giving strategy is shown in **Table 1**. Typically, $0.01\pi \leq \delta \leq 0.05\pi$.

Interpret the iterative steps of the algorithm: When the maximum number of iterative steps designed by the algorithm is reached, the original partition number is updated, K = K + 1, and then, return to the first step to continue the calculation. If the maximum number of iteration steps designed by the algorithm is not reached, the new partition scheme generated by the intelligent optimization algorithm is returned to the second step to continue the calculation. The algorithm flow is shown in **Figure 2**.

Algorithm Test

To test the rationality and convergence of the three algorithms in flow zoning, CEFR was selected as the research object to test the adaptability of the three intelligent optimization algorithms. The half core power distribution of CEFR is shown in **Figure 3**. Under the premise of considering the design margin, the limiting conditions were set as follows: the maximum temperature of the cladding is ≤569°C, the temperature of the fuel center is ≤2,100°C, and the number of zones is 4. The three intelligent optimization algorithms were used to calculate the traffic partition separately. According to the partition results, the results obtained by the three intelligent optimization algorithms were consistent. The CEFR traffic is divided into four zones, which is the same as the actual partition results. Therefore, all three intelligent optimization algorithms were

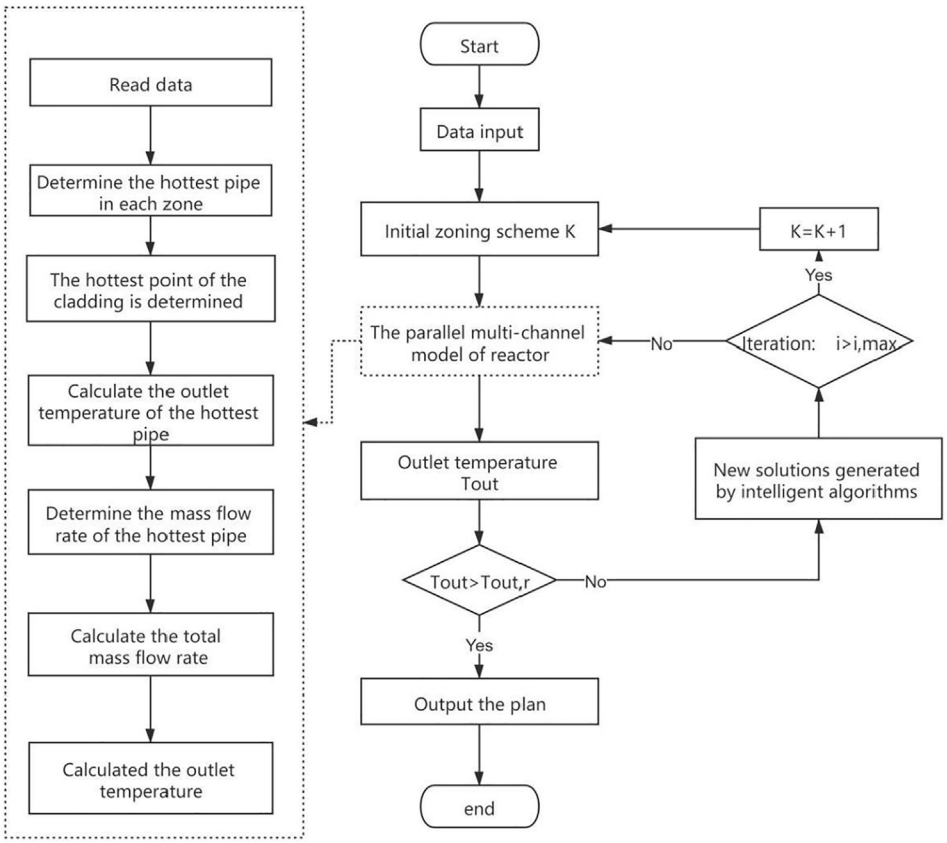


FIGURE 2 | Flow chart of flow partitioning algorithm design.

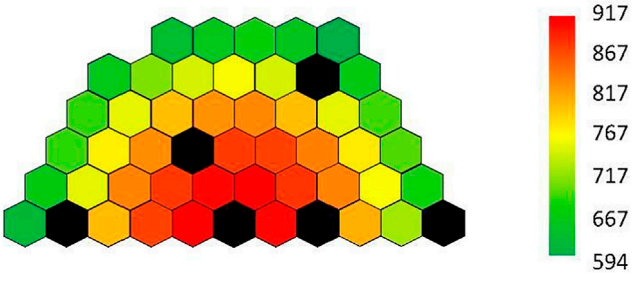


FIGURE 3 | Power distribution of CEFR half core power.

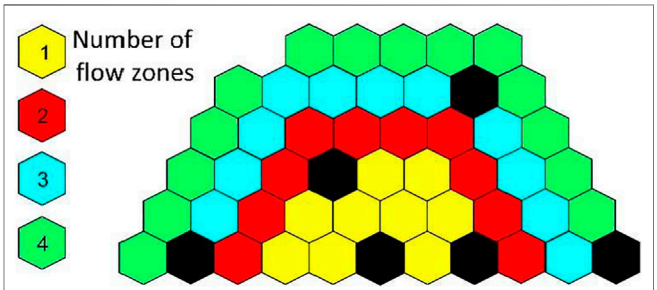


FIGURE 4 | Flow zoning results of half CEFR.

feasible for traffic zoning. **Figure 4** shows the traffic partition schemes calculated using the three optimization algorithms.

Table 2 shows that the total iterative steps of QGA and genetic algorithm are set to 1,000 steps; the QGA tends to converge only after 101 iterations, and the calculation time is 13,368 s. The genetic algorithm converges after 316 iterations, and the calculation time is 15,757 s. Compared with the genetic algorithm, for the QGA, there was 15.16% reduction in time and 68.03% reduction in iteration steps. Therefore, the QGA has obvious advantages over the genetic algorithm in convergence speed; however, the advantage in the speed of iterative calculation is not very obvious. Compared with the differential evolution method, the advantages of QGA are obviously reflected. The differential evolution algorithm has 1,298 convergence steps, and the total time consumption is 63,774 s. Comparatively, for QGA, there is 79.03% reduction in time and 92.21% reduction in iterative steps. Compared with the differential evolution algorithm, the QGA has obvious advantages in terms of total time consumption and convergence speed. Comparing these three algorithms, we know that the QGA has obvious advantages over the other two algorithms in terms of convergence steps and total time required for calculation, and

TABLE 2 | Algorithm convergence analysis results.

Algorithm Type	Total time(s)	Total Iteration Steps	Convergence Steps
Quantum genetic algorithm	13,368	1,000	101
Genetic algorithm	15,757	1,000	136
Differential evolution algorithm	63,774	1,500	1,298

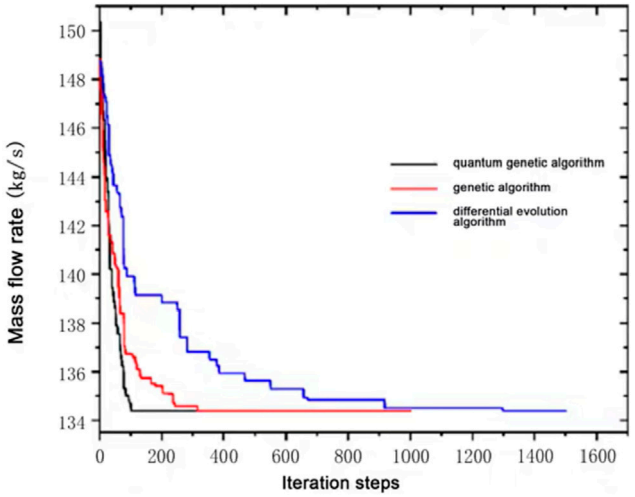


FIGURE 5 | Adaptive curves of the three intelligent algorithms.

it can be used as an intelligent optimization algorithm for the optimization of core flow zoning calculation of a long-life small reactor.

To compare the convergence between algorithms, adaptive curves of QGA intuitively, genetic algorithm, and differential evolution algorithm were extracted, as shown in **Figure 5**. QGA has obvious advantages in convergence compared with the other two algorithms. The final mass flows obtained by the three algorithms have relatively close values.

OPTIMIZATION OF SPALLER-100 FLOW ZONING DESIGN

SPALLER-100 Core Introduction

SPALLER-100 is a small long-life natural circulation lead- or lead-bismuth-based fast reactor that named SPALLER-100 designed at the University of South China. **Figure 6** shows the core section diagram of the SPALLER-100 reactor.

The thermal power of SPALLER-100 reactor core is 100 MW, the refueling cycle is 30 years, the operating temperature is 260–400°C, and the coolant temperature is 200°C under the condition of refueling shutdown. The height of the reactor core active zone is 1.5 m, and the equivalent diameter is 1.7 m. It comprises 48 fuel assemblies, 13 control rod components, 66 reflector components, and 126 shielding components. SPALLER-100 single fuel assembly has 61 fuel rods, and it uses PuN-ThN fuel. The effective density of fuel is 85% TD, and the cladding material is HT-9. He is filled in the gap between the fuel pellet and the cladding. The diameter of the fuel pellet is 12 mm, and the thickness of the cladding is 0.6 mm. The inner diameter of the fuel rod is 12 mm, and the outer diameter is 13.5 mm. The in-pile loading of nuclear fuel is 5,475.88 kg, and the loading of fissile nuclides is 891.02 kg. The core activity is divided into internal and external zones. There are 12 fuel modules in the internal zone, with a Pu mass fraction of 20.5%, and 36 fuel modules in the external zone, with a Pu mass fraction of 30.8%. **Table 3** shows the limit conditions of thermal-hydraulic design for the SPALLER-100 reactor.

Design of Core Flow Zoning Scheme for SPALLER-100 Full Life Cycle

Based on the power distribution at the beginning of SPALLER-100 life, flow partition calculation was carried out using QGA. Flow rates for zones 1–4 were 79.14 kg/s, 72.26 kg/s, 65.05 kg/s and 57.85 kg/s, respectively. When the number of core zones was 4, the outlet temperature and mass flow rate of SPALLER-100 were 522.92°C and 3,143.38 kg/s, respectively. **Figure 7** shows the flow zoning scheme when the number of core zones is 4.

The containment temperature of the reactor was limited to 550°C. The closed channel model does not consider the transverse mixing between the components, thereby retaining the margin of 25°C. For the closed channel, the calculation results showed 24 boxes of fuel assembly with temperature >525°C in the whole life period; the outlet temperature of 12 boxes of fuel assembly was much higher than the limiting temperature. For the sub-channel, the calculation results showed 20 boxes of fuel assembly with temperature >550°C in the whole life period; the outlet temperature of 12 boxes of fuel assembly was much higher than the limiting temperature. It was concluded that the calculation of flow zoning based on the power distribution of a single time point can lead to serious security risks. Therefore, core flow zoning during the whole life period should be considered to ensure the safety of the core during operation.

The maximum power of each fuel assembly during the whole life of the reactor was considered as sample data, and whole life flow partition calculation was carried out for SPALLER-100. On this basis, the relationship between the minimum flow achieved by different partition numbers and the average outlet temperature of the active zone was analyzed, as illustrated in **Figure 8**.

Figure 8 shows that when the number of zones was 5, the average outlet temperature of the reactor core was close to saturation. Increasing the number of zones slightly improved the average outlet temperature of the reactor core, and so, it is reasonable to divide the reactor into five zones. The total core flow was 3,424.20 kg/s, and the average outlet temperature of the core

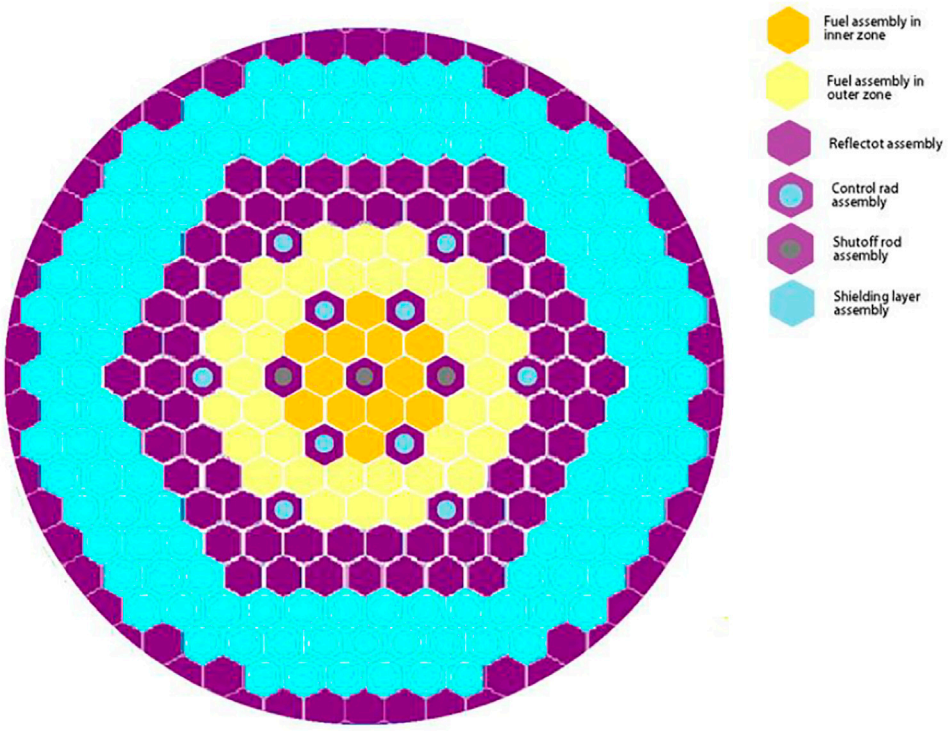


FIGURE 6 | Cross-sectional schematic diagram of SPALLER-100 core.

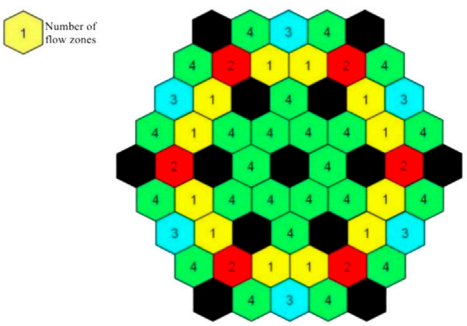
TABLE 3 | Thermal hydraulic design limits of SPALLER-100.

Project	Design Limit	
Maximum Temperature Limit of Fuel/°c	**2,300**	
Maximum temperature limit of cladding/°C	normal condition	550
	accident conditions	650
Maximum flow rate limit of coolant/(m/s)	2	
Minimum temperature of coolant/°C	200	

was 504.38°C. The partition results are shown in **Figure 9**. Flow rates of zones 1-5 are 82.36 kg/s, 79.14 kg/s, 72.29 kg/s, 60.05 kg/s and 58.00 kg/s, respectively.

Verification of SPALLER-100 Full Life Core Flow Zoning Scheme

To verify the accuracy of this procedure, the power of each component at 36 different times during the whole life of the reactor was extracted, and the maximum power of each component was selected as the sample data. The parallel multi-channel model designed in this study was used for calculations, and the calculated outlet temperature distribution is shown in **Figure 12**. The maximum power of each component was selected as the sample data. The fuel assemblies are shown in **Figure 10**, and the axial power distribution of the core is shown in **Figure 11**. These data were input into the card parameters, and the sub-channel code SUBCHANFLOW(Ferraro et al., 2020) was



Number of partitions	Single zone flow
1	79.14 kg/s
2	72.26 kg/s
3	65.05 kg/s
4	57.85 kg/s

FIGURE 7 | Flow partitioning results for SPALLER-100 (beginning of core).

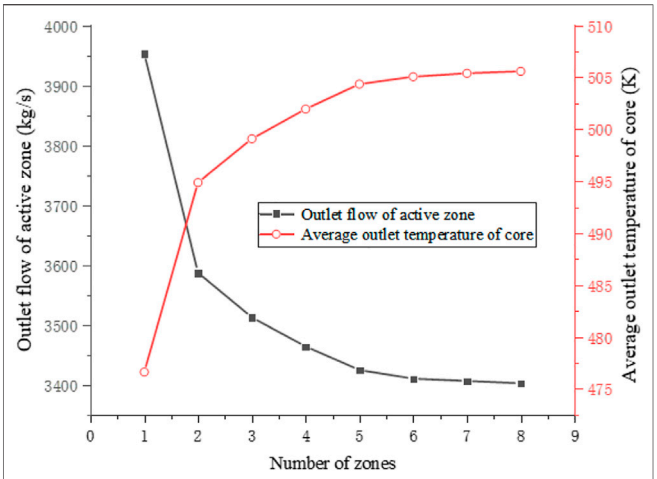


FIGURE 8 | Total active zone flow and active zone exit temperature for different number of zones.

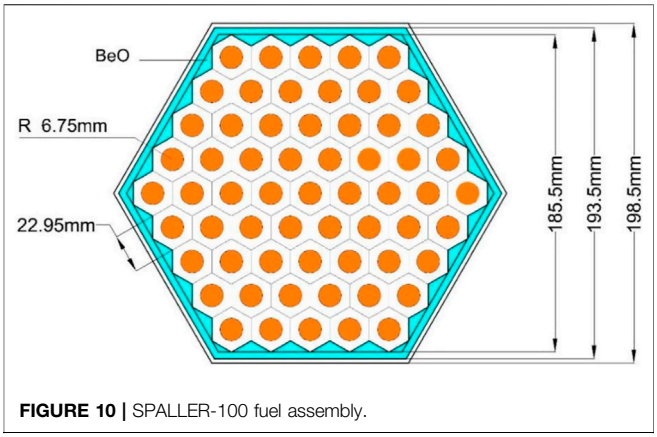


FIGURE 10 | SPALLER-100 fuel assembly.

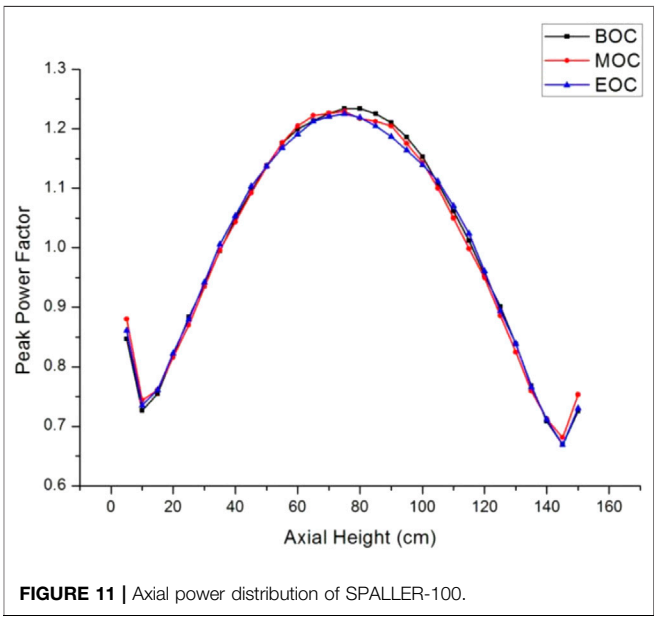


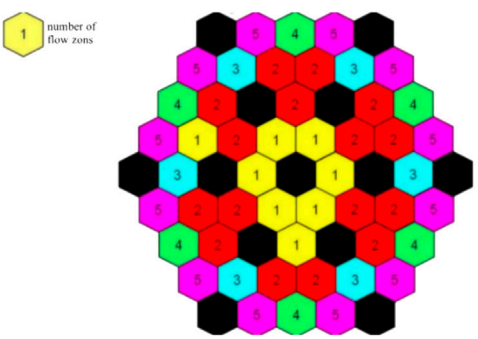
FIGURE 11 | Axial power distribution of SPALLER-100.

used for calculation. The calculated outlet temperatures are presented in **Figure 13**.

Analyses shown in **Figure 12**, **Figure 13** show that under the same flow partition, the calculation results of the proposed model are in good agreement with the calculation results of the sub-channel code SUBCHANFLOW. The minimum relative error, the maximum relative error, and the average relative error were 0.5, 2.9, and 2.3%, respectively. Therefore, it can be considered that the calculation results of the proposed parallel multi-channel model design are sufficiently credible.

To verify the impact of flow zoning on the reactor core outlet temperature, the maximum power of each fuel assembly in the whole life was taken as sample data, and the coolant mass flow distribution of each fuel assembly in the core and the coolant outlet temperature of each fuel assembly without zoning were compared and analyzed. First, after partition treatment, the outlet temperature was calculated to obtain total coolant mass flow of each fuel assembly partition. Then, the coolant mass flow obtained above was averaged for each fuel assembly to obtain the outlet temperature without zoning. The SUBCHANFLOW program was used to calculate the results. According to the calculation results, the following conclusions were obtained.

When the core coolant was treated in zones, and the number of flow zones was 5, the average outlet temperature of the core coolant was 508.50°C; the difference between the maximum and



Number of partitions	Single zone flow
1	82.36 kg/s
2	79.14 kg/s
3	72.29 kg/s
4	60.05 kg/s
5	58.00 kg/s

FIGURE 9 | SPALLER-100 final partition scheme.

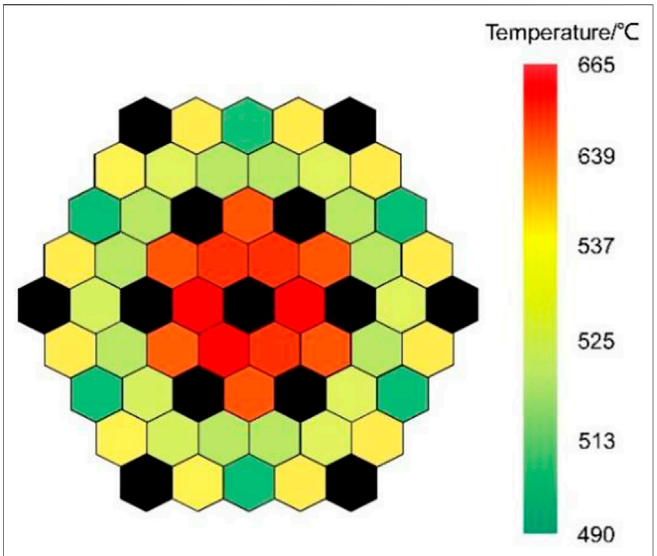


FIGURE 12 | Maximum outlet temperature distribution at 36 groups of power.

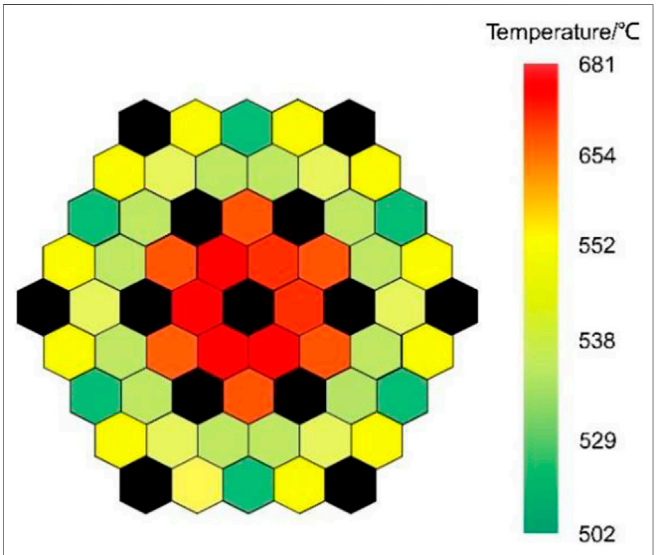


FIGURE 13 | Sub-channel program simulates and calculates the maximum outlet temperature of each component.

minimum temperatures was 22.44 ˚C. When the core coolant was not partitioned, the average coolant outlet temperature of the core was 507.82˚C; the difference between the maximum and minimum temperatures was 76.85˚C. When the total amount of coolant passing through remained unchanged, the average temperature change at the core outlet caused by the zoning and non-zoning treatments was slight, but the temperature fluctuation after zoning treatment is smaller; the difference between the maximum and minimum temperatures was small, and the outlet temperature was better flattened. Therefore, core flow treatment can flatten the core coolant outlet temperature. The detailed outlet temperatures are illustrated in **Figures 14**, **15**.

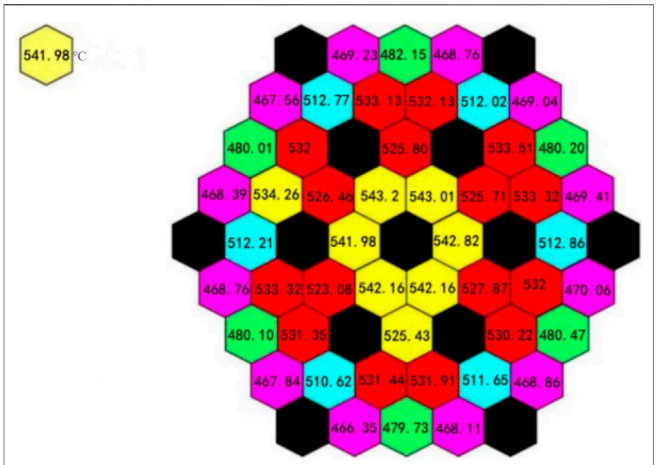


FIGURE 14 | Outlet temperature without zoning.

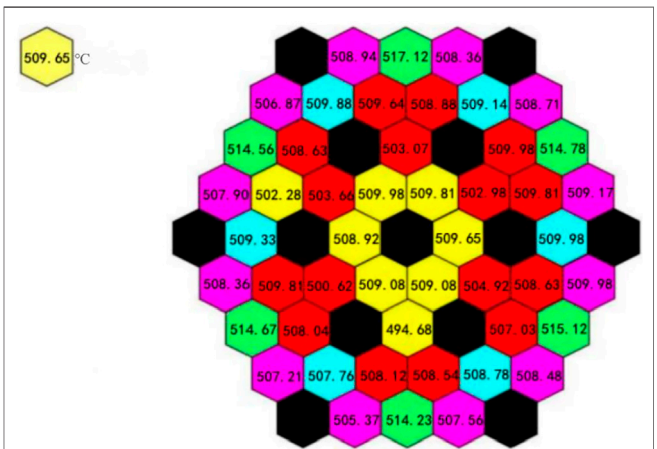


FIGURE 15 | Outlet temperature under zoning.

CONCLUSION

Reasonable core flow zoning of nuclear reactor can improve its safety and economy. Core flow zoning is also an essential means to flatten the core outlet temperature. Based on the modern optimization theory, artificial intelligence optimization algorithms were coupled with a parallel multi-channel model to construct a reactor core flow zoning calculation model. The convergence of different optimization algorithms was analyzed, and the optimal optimization algorithm was determined. This optimal optimization algorithm was used to carry out comparative analysis of two different flow zoning schemes using sample data of power distribution at the beginning of the life and the maximum power in the whole life of SPALLER-100. Following conclusions were obtained.

(1) Considering CEFR as the research object, for three intelligent optimization algorithms, namely, genetic

algorithm, differential evolution algorithm, and QGA, convergence analysis was carried out for reactor flow zoning. The results showed that QGA has the best convergence in long-life miniaturized reactor flow zoning, compared with the genetic algorithm. The total time consumption decreased by 15.16%. Compared with differential evolution algorithm, the total time consumption decreased by 79.03%.

(2) The flow zoning schemes designed based on the sample data of power distribution at the beginning of life and the maximum power of each fuel assembly in the whole life were compared and analyzed. The maximum outlet temperature of the fuel assembly based on the power distribution at the beginning of life exceeded the thermal safety limit of the reactor. Based on the power distribution of the whole reactor life, the maximum outlet temperature of the fuel assembly decreased by 140 K and was maintained below the thermal safety limit. The optimal number of zones of SPALLER-100 reactor at this time was 5. Increasing the number of zones led to challenges in flow distribution and did not significantly improve the thermal safety performance of the reactor.

(3) Flow zoning can bring a good flattening effect to the core outlet temperature. The SUBCHANFLOW program was used to analyze SPALLER-100. Taking the maximum power of the core in the whole life period as the calculation data, the following conclusions were obtained. When the core coolant was treated in zones and the number of flow zones was 5, the average outlet temperature of the core coolant was 508.50°C, and the difference between the maximum and minimum temperatures is 22.44°C. When the core coolant was not partitioned, the average coolant outlet temperature of the core was 507.82°C, and the difference between the maximum and minimum temperatures was 76.85°C. When the total amount of coolant passing through remained unchanged, the average temperature change at the core outlet caused by zoning and non-zoning treatments was only slight; however, the temperature fluctuation after zoning treatment was smaller; the difference between the maximum and minimum temperatures was smaller, and the outlet temperature was better flattened. Therefore, core flow treatment can flatten core coolant outlet temperature.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

QY: Data curation, Writing–original draft. SD: Conceptualization, Methodology, Software. PZ: Supervision. YZ: Visualization, Investigation. YX: Software, Validation. LP: Writing–review; editing. TY: Writing–review; editing.

ACKNOWLEDGMENTS

This work is supported by the "Characteristics Analysis and Influence Research of Core Power Redistribution and Pool Complex Thermal Stratification for Pump-driven Lead-cooled Fast Reactor under Asymmetric Operation Conditions" of Natural Science Foundation of Hunan Province (Grant No. 2019JJ40239).

REFERENCES

Basualdo, J. R., Sánchez-Espinoza, V., Stieglitz, R., and Macián-Juan, R. (2020). Integration of the Subchannel thermal-hydraulic Code SubChanFlow into the Reactor Dynamics Code PARCS: Development and Testing Based on a Computational Benchmark. *Prog. Nucl. Energ.* 119, 103138. doi:10.1016/j.pnucene.2019.103138

Chen, R. H., Su, G. H., Qiu, S. Z., and Fukuda, K. (2010). Prediction of CHF in Concentric-Tube Open Thermosiphon Using Artificial Neural Network and Genetic Algorithm. *Heat Mass. Transfer* 46 (3), 345–353. doi:10.1007/s00231-010-0575-9

Chen, Z., Zhao, P., Zhou, G., and Chen, H. (2014). Study of Core Flow Distribution for Small Modular Natural Circulation lead or lead-alloy Cooled Fast Reactors. *Ann. Nucl. Energ.* 72, 76–83. (in Chinese). doi:10.1016/j.anucene.2014.04.032

Cong, T., Chen, R., Su, G., Qiu, S., and Tian, W. (2011). Analysis of CHF in Saturated Forced Convective Boiling on a Heated Surface with Impinging Jets Using Artificial Neural Network and Genetic Algorithm. *Nucl. Eng. Des.* 241 (9), 3945–3951. doi:10.1016/j.nucengdes.2011.07.029

Ferraro, D., García, M., Valtavirta, V., Imke, U., Tuominen, R., Leppänen, J., et al. (2020). Serpent/SUBCHANFLOW Pin-By-Pin Coupled Transient Calculations for a PWR Minicore. *Ann. Nucl. Energ.* 137, 107090. doi:10.1016/j.anucene.2019.107090

Gai, J. (2017). *Improvement and Research of QuantumGenetic Algorithm[D]*. Jinzhou: Bohai University.

Li, S., Zhou, Z., and Feng, Y. (2019). *Coolant Flow Zoning Method in Fast Reactor core[P]*. BeiJing. CN109615110A-04-12(in Chinese).

Liu, Z., Zhao, P., and Zhang, B. (2020). Research on Core Concept Design of Ultra-long Life Small Natural Circulation Lead-based Fast Reactor[J]. *Nucl. Sci. Eng.* 54 (07), 1254–1265. (in Chinese). doi:10.7538/yzk.2019.youxian.0720

Nagy, K., Kloosterman, J. L., Lathouwers, D., and van der Hagen, T. H. J. J. (2012). The Effects of Core Zoning on the Graphite Lifespan and Breeding Gain of a Moderated Molten Salt Reactor. *Ann. Nucl. Energ.* 43, 19–25. doi:10.1016/j.anucene.2011.12.025

Wang, X., Wang, D., and Qi, S. (2020). Optimized Design Method for Core Flow Zoning of Sodium-Cooled Fast Reactor Based on Genetic Algorithm[J]. *Nucl. Sci. Eng.* 54 (09), 1660–1665. doi:10.7538/yzk.2020.youxian.0093

Yang, R., and Gu, J. (1997). An Efficient Simulated Annealing Algorithm for Global Optimization[J]. *Syst. Eng. —Theory Pract.* 17 (5), 30–33.

Ying, M., Wang, H., and L, H. (2018). Research on Self-Adaptive Quantum Genetic Algorithm. *Computer Eng. Appl.* 54 (20), 99–103.

Yizhe, L., Yijun, X., and Zhifeng, H. (2012). Calculation and Test of Core Flowrate Distribution of CEFR[J]. *Nucl. Sci. Eng.* 33 (S1), 100–103. (in Chinese).

Zhang, J., Chen, R. H., Wang, M. J., Tian, W. X., Su, G. H., and Qiu, S. Z. (2017). Prediction of LBB Leakage for Various Conditions by Genetic Neural Network and Genetic Algorithms. *Nucl. Eng. Des.* 325, 33–43. doi:10.1016/j.nucengdes.2017.09.027

Zhang, Q. (2017). *Research on Particle Swarm Optimization Algorithm and Differential Evolution algorithm[D]*. Jinan: Shandong University.

Zhou, Z., Yang, H., and Li, S. (2018). Preliminary Development of thermal-hydraulic Design Code for CFR600 Core [J]. *Nucl. Sci. Eng.* 52 (01), 56–63. (in Chinese). doi:10.7538/yzk.2020.youxian.0093

Conflict of Interest: We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled " Intelligent Optimization Method for Core Flow Zoning of Long-Cycle Lead-Bismuth-Cooled Reactor ".

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 12 May 2022
doi: 10.3389/fenrg.2022.852325

Measurement and Discrimination of Asymmetric Non-uniform Strain Distribution Based on Spectrum Characterization of FBG Sensors

Zhichun Fan[1,2]*, *He Yan*[2], *Zhiyong Huang*[2] and *Jing Liu*[1]

[1]School of Ocean Information Engineering, Jimei University, Xiamen, China, [2]Institute of Nuclear and New Energy Technology, Key Laboratory of Advanced Reactor Engineering and Safety of Ministry of Education, Collaborative Innovation Center for Advanced Nuclear Energy Technology, Tsinghua University, Beijing, China

The asymmetric deformation of glass-to-metal seals is an important defect that would lead to the failure of the pressure boundary in nuclear reactors. In this research, an efficient approach to measure the asymmetric deformation and prevent the potential failure was proposed based on the spectrum characterization of fiber Bragg grating (FBG) sensors. The asymmetric deformation was induced from small-size large-gradient nonuniform strain in sealing materials; as a result, it could be monitored through the spectra of embedded FBGs affected by strain variations. The theoretical analysis of the FBG spectrum was carried out using the transfer matrix model (TMM) to validate the measuring feasibility and reliability. Based on the theoretical results, the asymmetric deformation was measured by the distributed embedded FBG experimentally. By combining the reconstructed spectrum and the experimental results, the asymmetric deformation of glass-to-metal was proved to be monitored, and the defect was able to be prevented during the manufacturing process effectively via the proposed method.

Keywords: glass-to-metal seal, fiber Bragg grating (FBG), transfer matrix method (TMM), non-uniform strain distribution, nuclear power plant

OPEN ACCESS

Edited by:
Xianping Zhong,
University of Pittsburgh, United States

Reviewed by:
Huaping Wang,
Lanzhou University, China
Xizhen Xu,
Shenzhen University, China

***Correspondence:**
Zhichun Fan
fanzhichun@jmu.edu.cn

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *11 January 2022*
Accepted: *12 April 2022*
Published: *12 May 2022*

Citation:
Fan Z, Yan H, Huang Z and Liu J (2022) Measurement and Discrimination of Asymmetric Non-uniform Strain Distribution Based on Spectrum Characterization of FBG Sensors. Front. Energy Res. 10:852325. doi: 10.3389/fenrg.2022.852325

1 INTRODUCTION

The glass-to-metal seal has played an important role in the hermeticity of pressure boundaries applied in nuclear and renewable energy industries. The compressive strain generated in sealing glass during the manufacturing process with the compaction of the steel shell was the main factor for maintaining good hermeticity in glass-to-metal at harsh environments. Defects such as asymmetric deformation would be induced by the special model design, which would lead to the hermetic failure of glass-to-metal seals. Many advanced techniques [digital image correlation (Van Lancker et al., 2016) and photoluminescence spectroscopy (Li S et al., 2022)] were carried out recently to monitor the strain/stress in real-time; however, it was difficult for these methods to be applied in remote sensing under harsh nuclear environments. Based on our previous research studies (Fan et al., 2019; Fan et al., 2020), the embedded fiber sensing technique is demonstrated in this research. Fiber Bragg grating (FBG) sensors have been developed in the past 40 years, and they have been applied in various sensing schemes (Zaghloul et al., 2018; Fan et al., 2019; Morana et al., 2019). FBG has emerged as a reliable, *in situ*, nondestructive tool for monitoring, diagnosing, and controlling civil structures, and the versatility of FBG sensors

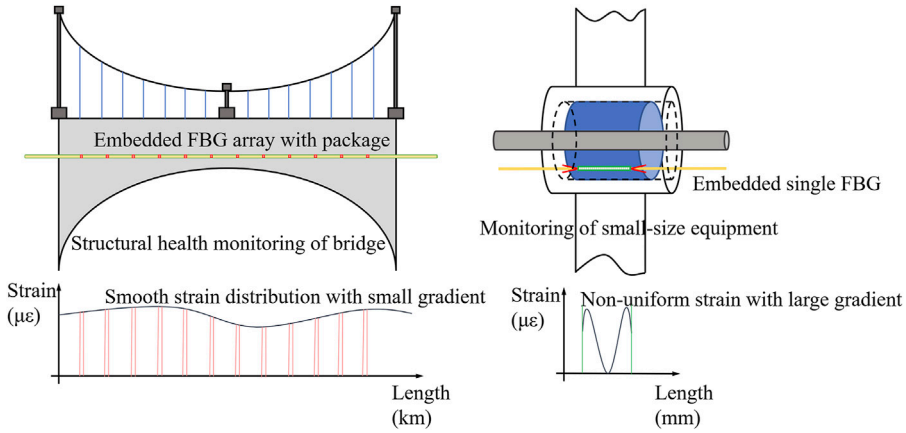


FIGURE 1 | Sensing scheme of sizable and small-size equipment with the related measuring strain distribution.

represents a key advantage over other technologies in the structural health monitoring (SHM) field (Majumder et al., 2008).

Generally, asymmetric deformation would cause the local strain distribution to vary significantly, so the measurement of strain distribution was essential in this research. Strain measurement has attracted research interest for its important role in SHM. The strain distribution of large equipment (bridges, composite structures, etc.) is regarded as being uniform in the measuring region of small sensors, so the strain signals would be stable and clear, as shown in **Figure 1** (Ling et al., 2006; Chen et al., 2017; Goossens et al., 2019; Xiong et al., 2019; Li J et al., 2022). However, for small-size equipment such as glass-to-metal seals, the strain distribution was nonuniform with a large gradient to guarantee hermeticity, so it would be challenging to realize accurate strain measuring because the spectrum would be broadened and distorted by nonuniform strain (Kersey et al., 1997). Previous results showed that it was feasible to characterize the properties of strain distributions with a chirped spectrum (Fan et al., 2019). This phenomenon was widely applied to identify nonuniform strain induced by defects [crack locations in composites (Okabe et al., 2004), crack propagations (Jin et al., 2019), and transverse loads detection (Rajabzadeh et al., 2019)], and the results demonstrated its feasibility.

This research performed a practical monitoring method for the hermetic material of glass-to-metal seals, which could realize the discrimination of nonuniform strain distribution and prevention of asymmetric deformation with a high measuring resolution. The spectrum reconstruction of FBG under nonuniform strain was studied using the combined transfer matrix model (TMM) and finite element method (FEM). The relationship between the gradient parameter of strain distribution and the full width at half maximum (FWHM) of the FBG spectrum was obtained. The characteristic parameters (Bragg wavelength shift and FWHM) of the chirped spectrum were proved to monitor asymmetric deformation efficiently through the *in situ* monitoring experiments, and the accuracy was verified by the numerical simulations.

TABLE 1 | Parameters of the UFBG.

Parameters	Value
Refractive index	1.452
Bragg wavelength (nm)	1550
Length (mm)	12
Elastic modulus (GPa)	73
Poisson's ratio	0.17
Average index change	1×10^{-4}

2 NUMERICAL SIMULATION

2.1 Transfer Matrix Method

According to the index modulation depth distribution of the grating region, FBG can be divided into apodized FBG (AFBG) and uniform FBG (UFBG). The feasibility of AFBG to measure the strain was verified in the previous research. UFBG was also applied to monitor the small-size nonuniform strain distributions because its spectrum was more sensitive and would generate appreciable distortions with the nonuniform strain (Jin et al., 2019). The parameters of applied UFBG in this article are shown in **Table 1**.

FBG is similar to a wavelength-selective reflection filter. A narrow band of the incident optical field is reflected by coherent scattering from the index variations with a wavelength λ_B given by (Hill and Meltz, 1997)

$$\lambda_B = 2n_{eff}\Lambda \tag{1}$$

where Λ is the grating period and n_{eff} is the modulation index. The n_{eff} perturbation in the core for UFBG is described by

$$\Delta n_{eff}(z) = \overline{\Delta n}_{eff}\left\{1 + \upsilon \cos\left[\frac{2\pi}{\Lambda}\right]\right\} \tag{2}$$

where $\overline{\Delta n}_{eff}$ is the average change of the modulation index and υ is the fringe visibility.

The grating forces couple between propagating modes since they impose a dielectric perturbation to the waveguide. The

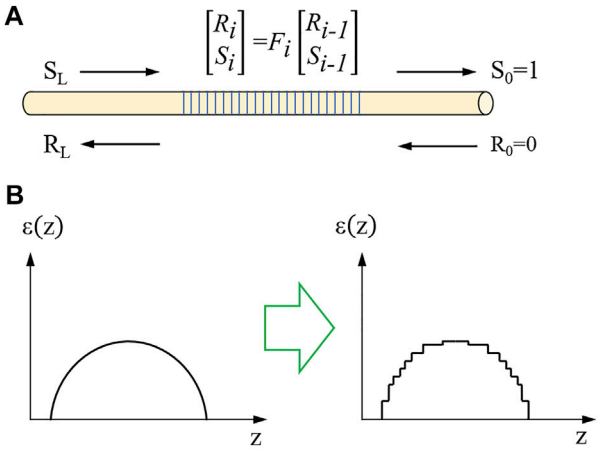


FIGURE 2 | (A) Propagation of a waveguide in an optical fiber. **(B)** Discrete strain distribution by the TMM.

coupled mode theory is effective to describe this behavior. A set of coupled first-order differential equations is used to describe the propagation:

$$\frac{dR}{dz} = j\hat{\sigma}R(z) + j\kappa S(z) \tag{3}$$

$$\frac{dS}{dz} = -j\hat{\sigma}S(z) - j\kappa^* R(z) \tag{4}$$

where $R(z)$ and $S(z)$ represent the amplitude of forward- and backward-propagation, respectively. $\hat{\sigma}$ is the general "dc" self-coupling coefficient, which is a function of the propagating wavelength λ, given by

$$\hat{\sigma} = 2\pi n_{eff}\left(\frac{1}{\lambda} - \frac{1}{\lambda_D}\right) + \frac{2\pi}{\lambda}\Delta n_{eff} - \frac{1}{2}\phi'(z) \tag{5}$$

where $\phi'(z) = d\phi/dz$ and $\lambda_D = 2n_{eff}\Lambda_0$ is the design wavelength for a Bragg scattering within an infinitesimal variation of the effective index ($\Delta n_{eff} \to 0$). κ is the "ac" coupling coefficient defined as

$$\kappa = \frac{\pi}{\lambda}\nu\overline{\Delta n}_{eff} \tag{6}$$

This equation represents the uniform grating with a constant average refractive index change. However, the strain distribution of glass-to-metal was nonuniform with a large gradient, and it would cause a chirp in the grating period. If roughly treated as a uniform strain distribution, the measuring results would generate notable deviations. In this paper, the transmission and reflection spectra from the two-mode coupling can be reconstructed using the transfer matrix method (TMM), whereby the grating is divided into finite discrete uniform sections represented by a 2×2 matrix. The matrix for the whole FBG can be obtained by multiplying all the discrete matrices (**Figure 2**).

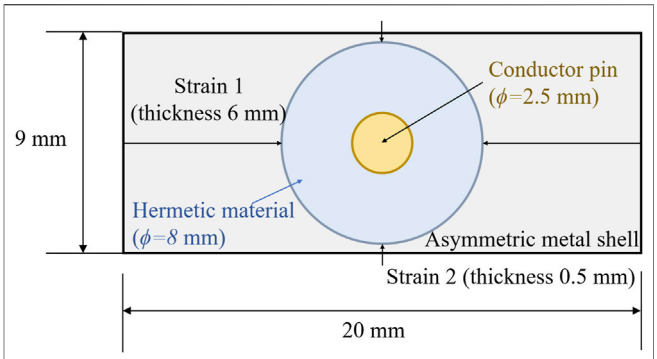


FIGURE 3 | Schematic diagram of a specially designed glass-to-metal sealing model.

The calculation accuracy of the TMM depends on the section number N, and when $N \simeq 100$ is more sufficient to accurately model the chirped gratings, the characteristics of the reconstructed spectrum led to a convergence after N increased above 100. Defining R_i and S_i to be the field amplitudes after traversing the ith section, the propagation through this uniform section is described by

$$\begin{bmatrix} R_i \\ S_i \end{bmatrix} = F_i \begin{bmatrix} R_{i-1} \\ S_{i-1} \end{bmatrix} \tag{7}$$

$$F_i = \begin{bmatrix} \cosh(\gamma_B \Delta z) - i\frac{\hat{\sigma}}{\gamma_B}\sinh(\gamma_B \Delta z) & -i\frac{\kappa}{\gamma_B}\sinh(\gamma_B \Delta z) \\ i\frac{\kappa}{\gamma_B}\sinh(\gamma_B \Delta z) & \cosh(\gamma_B \Delta z) + i\frac{\hat{\sigma}}{\gamma_B}\sinh(\gamma_B \Delta z) \end{bmatrix} \tag{8}$$

The components of the transfer matrix F_i are defined subsequently, where $\gamma_B = \sqrt{\kappa^2 - \hat{\sigma}^2}$ and Δz is the length of each section. The coefficients κ and $\hat{\sigma}$ have local values at the ith section.

The limits of the grating were defined as $0 \le z \le L$. The boundary conditions of the matrix are $R(0) = 1$ for $z \le 0$ and $S(L) = 0$ for $z \ge L$. For the entire grating, the matrix F can be written as

$$\begin{bmatrix} R(0) \\ S(0) \end{bmatrix} = F\begin{bmatrix} R(L) \\ S(L) \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ S(0) \end{bmatrix} = F\begin{bmatrix} R(L) \\ 0 \end{bmatrix} \tag{9}$$

where $F = \prod_{i=1}^{N} F_i$. Fi can be expressed as

$$F_i = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \tag{10}$$

In addition, substituting **Eq. 10** in **Eq. 9** results in

$$\begin{bmatrix} 1 \\ S(0) \end{bmatrix} = \begin{bmatrix} f_{11}R(L) \\ f_{21}R(L) \end{bmatrix} \tag{11}$$

Thus, reflectivity as a function of wavelength can be calculated by

$$r(\lambda) = \left|\frac{S(0)}{R(0)}\right|^2 = \left|\frac{f_{21}}{f_{11}}\right|^2 \tag{12}$$

TABLE 2 | Geometrical parameters of the designed model.

Parameters (mm) component	Section			Height
	Length	Width	Diameter	
Metal shell	20	9	—	30
Hermetic material	—	8	5	—
Conductor pin	—	2.5	30	—

Similarly, the transmitted spectrum can be found as

$$t(\lambda) = \left|\frac{R(L)}{R(0)}\right|^2 = \left|\frac{1}{f_{11}}\right|^2 \tag{13}$$

2.2 Finite Element Simulation

The finite element model of glass-to-metal seals was built to obtain the strain distributions in glass. The thickness of the metal shell was a main factor determining the distribution of compressive strain, so it was designed in a rectangular shape to induce asymmetric deformation defect. The length and width were set to 6 and 0.5 mm (**Figure 3**), respectively, which could generate remarkable asymmetric strain along the circumference of the hermetic material. The geometrical parameters of the model are shown in **Table 2**.

To prove the feasibility and measuring resolution of the FBG spectrum, the TMM and FEM were combined to reconstruct the spectrum response with specific nonuniform strain in glass-to-metal seals. This spectrum reconstruction method combined the advantages of both FEM and TMM, and it has been applied by many researchers to predict spectrum variations under large-gradient strain distribution (Kakei et al., 2018; Rajabzadeh et al., 2019; Fazzi et al., 2019; Stathopoulos et al., 2019). The strain of the fiber core transferred directly to the surface of the sealing glass to simplify the simulation (Wang et al., 2019) because the bare fiber was well-fused with molten glass during the manufacturing process and there was little difference between the mechanical properties of fiber and sealing glass (the main component for both was SiO_2).

3 EXPERIMENTAL SETUP

AFBG and UFBG were embedded in the hermetic material to monitor the nonuniform strain distribution. The experimental setup is shown in **Figure 4**. The FBGs used in this research were type II gratings inscribed by femtolaser provided by Femtofibertech. The previous results showed that FBG had good endurance at high temperature (1000 °C) (Fan et al., 2020). The experimental model was consistent with the finite element model (**Table 1**). The experimental model was manufactured with FBGs by a gasket. To achieve distributed strain and asymmetric deformation monitoring, four sensors (including 3 AFBGs and 1 UFBG) were embedded simultaneously as shown in **Figure 4A**, which were located at the same distance (3 mm) with varied metal thickness. The shell thickness of path 1 to path 4 was 6 mm, 5.5 mm, 1 mm, and 0.5 mm, respectively, which ensured the spectra of FBGs would be affected by different strain distributions.

The model was designed under a specific heating process (20°C to 450°C, 10°C/min) to make the hermetic material fuse well with the metal shell and FBGs (**Figure 4B**). After cooling down to room temperature, the glass-to-metal model with embedded FBGs was obtained, and the hermetic reliability was guaranteed by the compressive strain formed in the sealing glass. The FBGs embedded in four paths were connected to the interrogator to record the real-time spectrum and characteristic parameters. The experimental results were analyzed and compared with the reconstructed spectra obtained by the TMM. The accuracy and feasibility of the proposed method to measure the large-gradient nonuniform strain and prevent the asymmetric deformation defect were demonstrated.

4 RESULTS AND DISCUSSION

4.1 Spectrum Reconstruction to Large-Gradient Nonuniform Strain

The axial strain distributions in the hermetic material were extracted along the measuring path of FBG. A total of 16 groups of strain with different gradients were obtained under

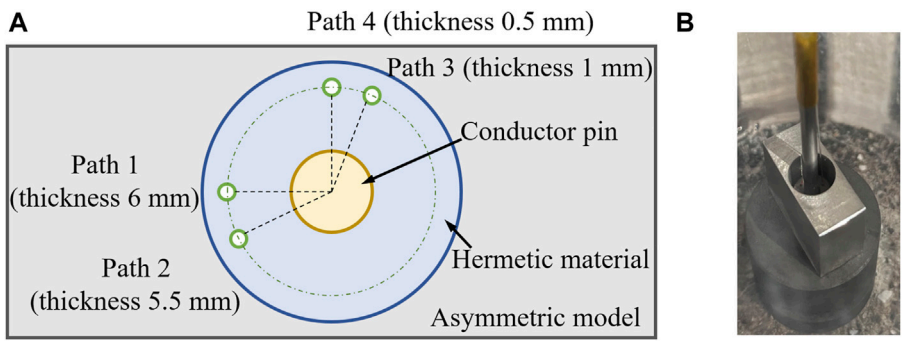


FIGURE 4 | Experimental setup for the asymmetric model with distributed FBGs: **(A)** the schematic diagram and **(B)** the photograph.

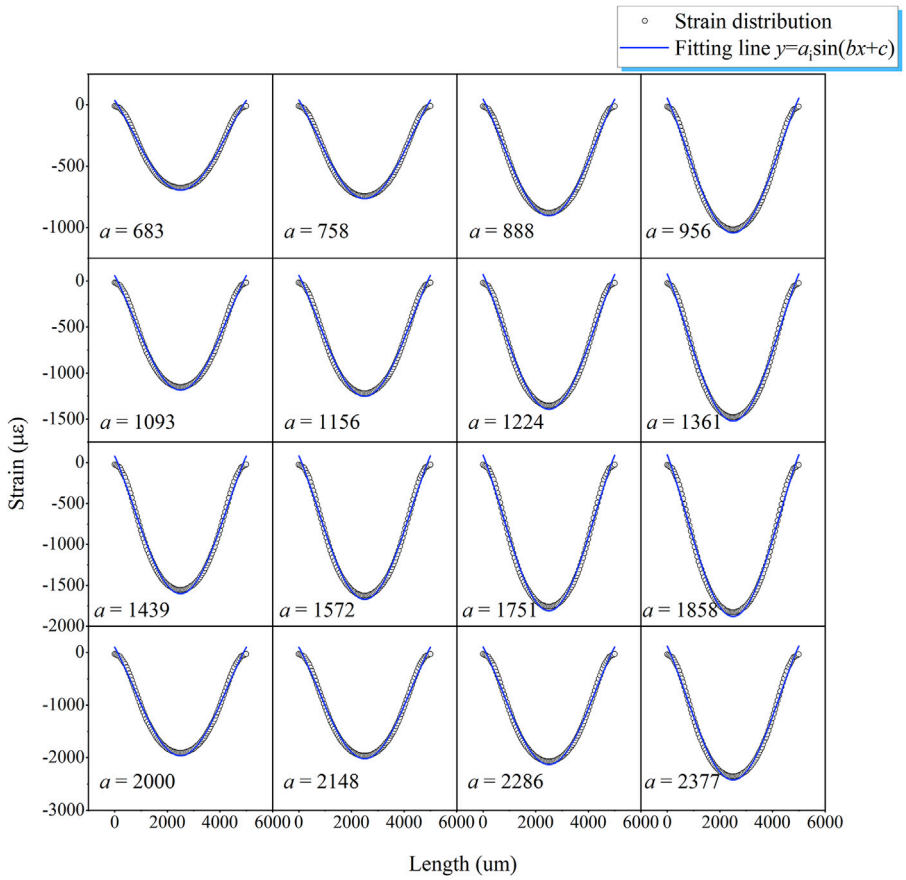


FIGURE 5 | Fitting curve of the large-gradient nonuniform strain distributions in glass-to-metal with gradient parameters *a*.

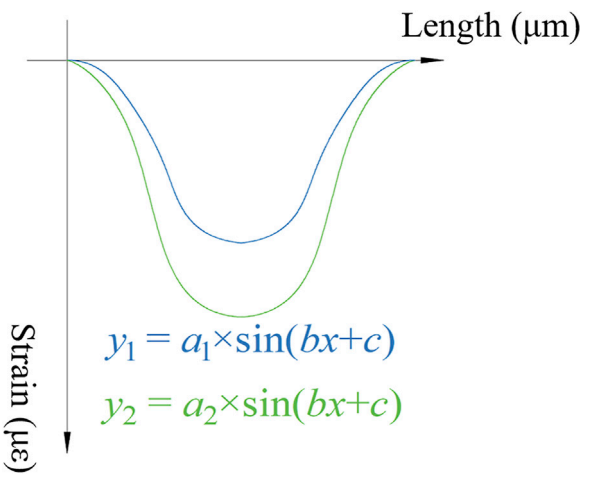


FIGURE 6 | Features of axial strain distribution in hermetic material.

different working conditions from 20°C to 300°C (**Figure 5**). The sine function was applied to fit the strain distributions with different gradient parameters *a*, as shown in **Figure 5**. The gradient parameter *a* changed from 680 to 2400 with the increase in the amplitude. All the fitting functions were close to the original strain distributions with a deviation of less than 5%.

The strain distributions along the measuring paths and their variation range are summarized in **Figure 6**. It was shown that those axial strains could be fitted by the sine curve method, and the following three features could be obtained as a conclusion:

1) When the length z = 0 and z = 5,000, the strain was 0.
2) All the distributions could be summarized by the sine curve y = a^*sin (bx + c).
3) b = 6.7e-4, and c = 3.028 (based on features 1 and 2).

According to the features, all the axial strain distributions in the hermetic material could be described by the sine curve fitting method with an appropriate parameter *a*, which represented the difference between the maximum and the minimum value of the distribution. The strain distributions could be assigned to the grating region by the TMM.

A total of 16 groups of sine curve fitting functions were imposed to FBG through the coupling parameters of the transfer matrix, and the spectra were reconstructed as shown in **Figure 7**. The FWHM tended to broaden (from 0.78 to 2.81 nm) with increasing *a*, and the relationship between these two parameters

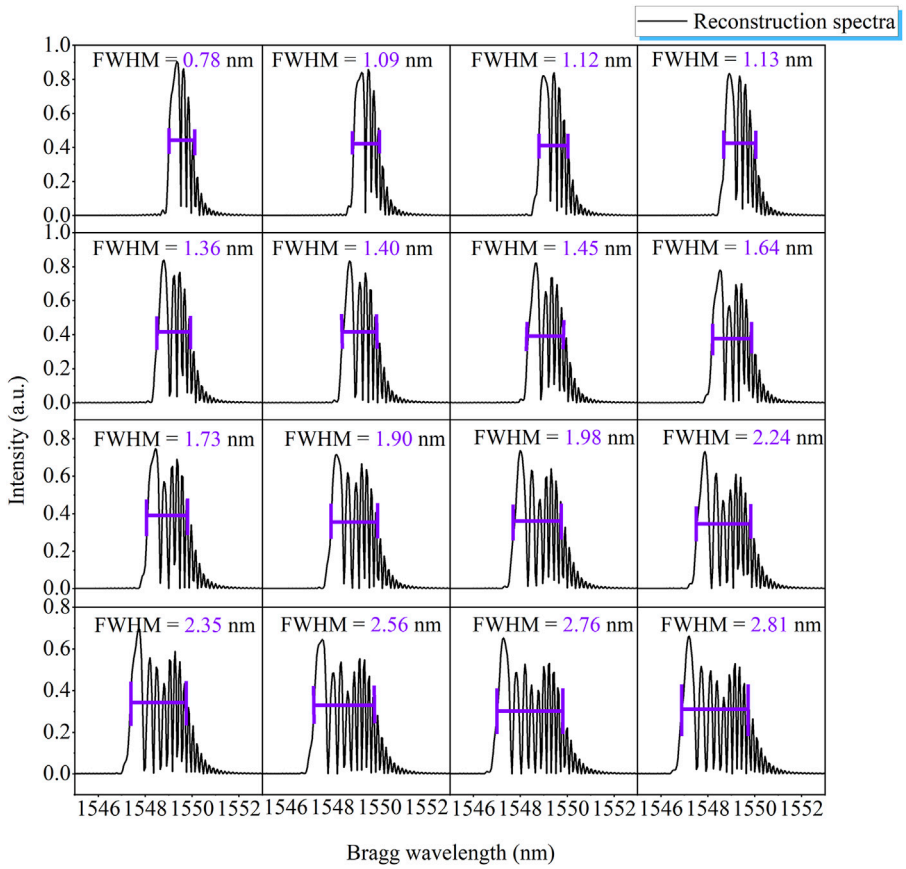


FIGURE 7 | Related spectra variations of FBG induced by different nonuniform strain distributions in glass-to-metal.

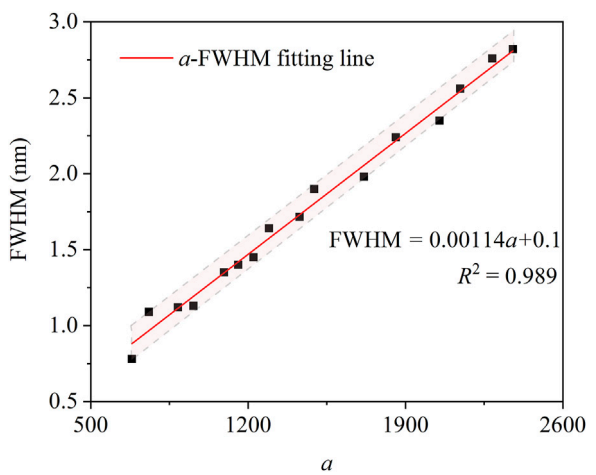


FIGURE 8 | Fitting relationship between the FWHM of FBG and the gradient parameters *a*.

is defined in **Figure 8**. Considering that the resolution of the interrogator was 0.05 nm, consequently, the measuring resolution of the nonuniform strain distribution in the hermetic material was about 70 με. The deviation between the fitting curve and the origin data was around 8%. As a result, the FWHM was proved to characterize the variation in axial strain in the hermetic material. The gradient and distribution of nonuniform strain could be calculated by the spectrum based on the defined relationship. The combined FEM and TMM method was feasible to reconstruct the spectrum response in the hermetic material under different operating conditions.

4.2 The Monitoring Results of Asymmetric Model

The deformation and strain contour of the designed model were obtained by the finite element method. The temperature of the model changed from 380°C to 20°C to simulate the manufacturing process of glass-to-metal. The deformation was extracted along two radial paths, as shown in **Figure 9A**. It was shown that the deformation was extremely asymmetric along radial path 1 and path 2, of which the difference Δl was about 0.008 mm (**Figure 9B**), which was about 10% compared with the origin radius. The asymmetric deformation would be induced in the hermetic material. Then the strain was extracted along axial path 1 and 2 (**Figure 10A**), which were related to experimental path 1

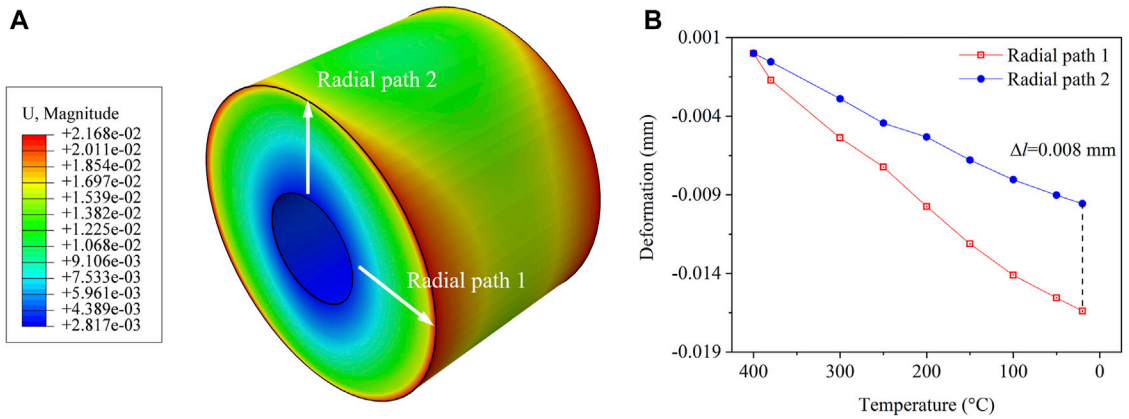


FIGURE 9 | (A) Contour of the asymmetric deformation of the designed model; **(B)** deformation along two radial paths in glass with the thickest and the thinnest metal shell.

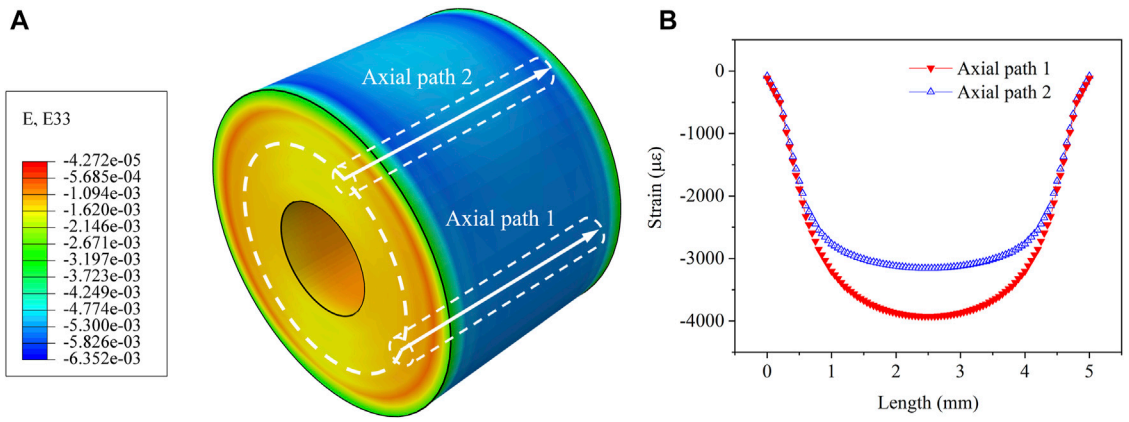


FIGURE 10 | (A) Contour of the compressive strain of the designed model; **(B)** large-gradient nonuniform strain distribution along two radial paths in glass with the thickest and the thinnest metal shell.

and path 4 (**Figure 5A**), respectively. The gradient of the strain distribution increased with the thickness of the metal shell. Based on the theoretical measuring resolution in **Section 4.1**, the asymmetric deformation was feasible to be monitored by the distributed FBG sensors embedded in the hermetic material.

A total of four FBGs were embedded simultaneously to monitor the asymmetric deformation during the experiment. The details of four FBGs are shown in **Table 3**. The AFBGs were embedded in paths 1~3, and the single UFBG was located in path 4. The two kinds of FBGs were arranged to provide comparisons of the chirped spectrum with nonuniform strain distribution. The distributed monitoring results are shown in **Figure 11**. The spectra signified with a purple frame were FBGs located in the hermetic material, which generated obvious distortions influenced by the strain variations. The spectra with a yellow frame were FBGs settled near the hermetic material, and the Bragg wavelength shifted as the temperature cooled down without the broadening of the FWHM.

The distortions of spectra in the purple frame tended to vary with the thickness of the metal shell, and the feature parameters were extracted as shown in **Figure 12**. The Bragg wavelength of four FBGs in the hermetic material generated an obvious shift due to the induced compressive strain. However, because the strain distribution was nonuniform with a large gradient, the relationship between the Bragg wavelength shift and the average strain was not linear. In **Figure 12**, the Bragg wavelength of paths 1 and 2 (the thickness of the metal shell >5.5 mm) was considerably larger than those of paths 3 and 4 (the thickness of the metal shell <1 mm), of which the shift was 0.94 nm, so the asymmetric deformation of the hermetic material was characterized effectively via the Bragg wavelength shift. The FWHM increased with increasing metal shell thickness, and the bandwidth variation induced by the asymmetric deformation was up to 1.16 nm. Therefore, the broadening of the FWHM was feasible to monitor the distributed nonuniform strain distribution in the hermetic material.

TABLE 3 | Parameters of customized FBG array sensors.

FBG position	Refractive index	Array quantity	Origin wavelength (nm)	Origin FWHM (nm)
1–3	f(z) = exp [−10 × (z−L/2)2/L2]	2	1545, 1555	0.82, 0.65
4	10^{-4}	1	1550	0.28

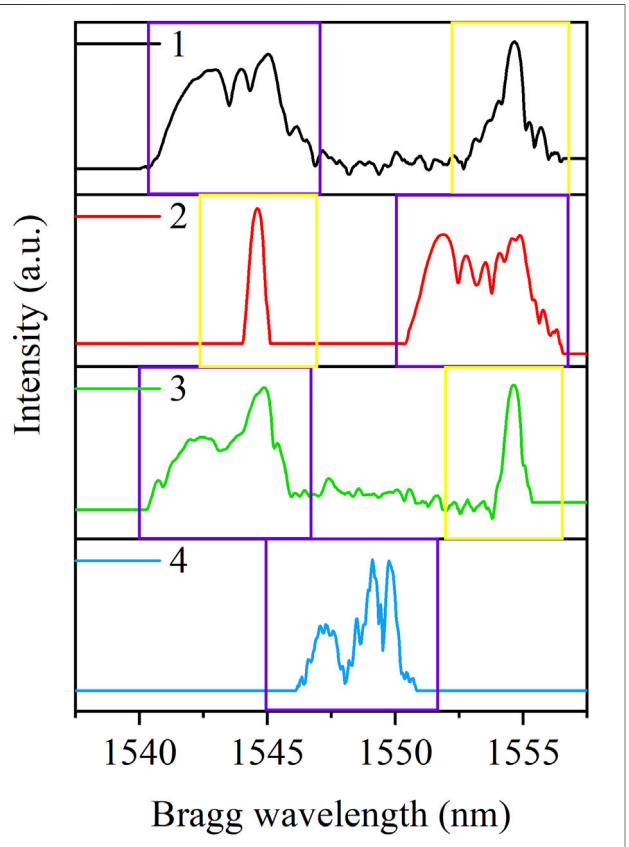


FIGURE 11 | Spectral characterization of the asymmetric deformation in glass-to-metal obtained by an interrogator.

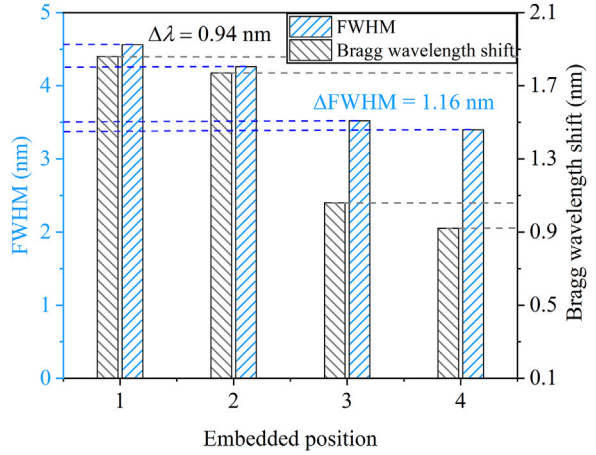


FIGURE 12 | Monitoring and discrimination of the asymmetric deformation reflected by the FWHM and Bragg wavelength shift of FBGs.

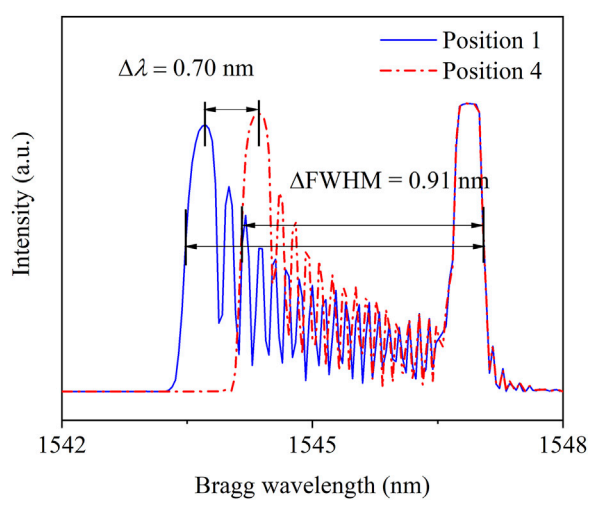


FIGURE 13 | Reconstruction spectrum by the TMM in position 1 and position 4 of the experimental model.

To verify the accuracy and reliability of the experimental results, the combined TMM and FEM method was carried out to reconstruct the spectrum variation of embedded FBGs in paths 1 and 4. The strain distribution along the special paths was obtained by numerical results in **Figure 13**, and then the spectra were reconstructed by the transfer matrix with the nonuniform strain distribution as shown in **Figure 10B**. The theoretical Bragg wavelength shift of FBGs in paths 1 and 4 was 0.70 nm, and the FWHM variation was about 0.91 nm. Compared with the experimental results, the deviation was around 20%, which was induced by the approximate strain distribution of the TMM and the boundary conditions of the FEM. The experimental results proved to remain consistent with the simulations in different paths of the asymmetric model. Thus, in summary, the spectrum would generate obvious distortions when the asymmetric deformation developed in glass-to-metal, and the Bragg wavelength and FWHM would increase as the deformation becomes more severe. Because the FBG spectrum remained stable under consistent operating conditions, the asymmetric deformation could be characterized with the variations in the spectrum distortion, Bragg wavelength, and FWHM. Therefore, the glass-to-metal sample produced with such defect would be replaced promptly to prevent hermetic failure under operating conditions.

5 CONCLUSION AND FUTURE WORK

This research investigated the spectrum characterization of FBG to non-uniform strain distribution and the derived monitoring and prediction method of the asymmetric deformation defect.

1) The TMM was demonstrated to be a feasible approach to reconstruct the FBG spectra with large-gradient nonuniform strain distributions in glass-to-metal equipment. The FWHM of the reconstructed spectrum tended to increase linearly with the gradient parameter a of strain distributions. The measuring resolution of strain variation in this research was 70 με.
2) The spectrum was feasible to monitor the asymmetric deformation defect of the special model by the distributed FBG array. The embedded FBGs generated distortions affected by the large-gradient nonuniform strain, and the outer FBGs remained as origin spectra with individual Bragg wavelength shift.
3) The FBG spectrum showed distinct variations in different paths of the experimental model. Both the Bragg wavelength shift and the variations in the FWHM proved that the distributed FBGs could measure the asymmetric deformation effectively during the manufacturing process.

Based on the proposed approach, the potential defect in the hermetic material was able to be monitored and prevented, and the quality of products could be inspected and improved after the manufacturing of glass-to-metal equipment.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

YH conceived and supported the study. FZ fabricated the experimental models and carried out the defect-monitoring experiments and the theoretical analysis. FZ, YH, and HZ prepared the manuscript. FZ, YH, HZ, and LJ contributed toward revising the manuscript.

FUNDING

This work was supported by the National S&T Major Project of China (ZX069), the project supported by the Science Foundation of Jimei University, China (ZQ2021053), and the Fujian Provincial Department of Education, China (JAT210222).

REFERENCES

Chen, S.-Z., Wu, G., and Xing, T. (2017). Deflection Monitoring for a Box Girder Based on a Modified Conjugate Beam Method. *Smart Mater. Struct.* 26, 085034. doi:10.1088/1361-665x/aa7973

Fan, Z., Diao, X., Liu, M., Zhang, Y., Huang, Z., and Yan, H. (2019). On-line Monitoring of Sealing Glass in Electrical Penetration Assembly Based on Femto-Laser Inscribed Fiber Bragg Grating Sensors. *Opt. Express* 27, 608–620. doi:10.1364/oe.27.000608

Fan, Z., Diao, X., Hu, K., Zhang, Y., Huang, Z., Kang, Y., et al. (2020). Structural Health Monitoring of Metal-To-Glass-Ceramics Penetration during Thermal Cycling Aging Using Femto-Laser Inscribed FBG Sensors. *Sci. Rep.* 10, 12330–12413. doi:10.1038/s41598-020-69282-7

Fazzi, L., Rajabzadeh, A., and Milazzo, A. (2019). "Analysis of FBG Reflection Spectra under Uniform and Non-uniform Transverse Loads," in Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2019 (SPIE), 109701, 589–597. doi:10.1117/12.2513795

Goossens, S., De Pauw, B., Geernaert, T., Salmanpour, M. S., Sharif Khodaei, Z., Karachalios, E., et al. (2019). Aerospace-grade Surface Mounted Optical Fibre Strain Sensor for Structural Health Monitoring on Composite Structures Evaluated against In-Flight Conditions. *Smart Mater. Struct.* 28, 065008. doi:10.1088/1361-665x/ab1458

Hill, K. O., and Meltz, G. (1997). Fiber Bragg Grating Technology Fundamentals and Overview. *J. Light. Technol.* 15, 1263–1276. doi:10.1109/50.618320

Jin, X., Yuan, S., and Chen, J. (2019). On Crack Propagation Monitoring by Using Reflection Spectra of AFBG and UFBG Sensors. *Sensors Actuators A Phys.* 285, 491–500. doi:10.1016/j.sna.2018.11.052

Kakei, A., Epaarachchi, J. A., Islam, M., and Leng, J. (2018). Evaluation of Delamination Crack Tip in Woven Fibre Glass Reinforced Polymer Composite Using FBG Sensor Spectra and Thermo-Elastic Response. *Measurement* 122, 178–185. doi:10.1016/j.measurement.2018.03.023

Kersey, A. D., Davis, M. A., Patrick, H. J., Leblanc, M., Koo, K. P., Askins, C. G., et al. (1997). Fiber Grating Sensors. *J. Light. Technol.* 15, 1442–1463. doi:10.1109/50.618377

Li S, S., Zhu, Q., Hu, K., Cai, Y., Liu, Z., Chen, F., et al. (2022). Determination of Compressive Stress in Glass-To-Metal Seals Using Photoluminescence Spectroscopy Technique. *Ceram. Int.* 48, 13379. doi:10.1016/j.ceramint.2022.01.219

Li J, J., Yan, J., Zhu, J., and Qing, X. (2022). K-BP Neural Network-Based Strain Field Inversion and Load Identification for CFRP. *Measurement* 187, 110227. doi:10.1016/j.measurement.2021.110227

Ling, H.-y., Lau, K.-t., Cheng, L., and Jin, W. (2006). Viability of Using an Embedded FBG Sensor in a Composite Structure for Dynamic Strain Measurement. *Measurement* 39, 328–334. doi:10.1016/j.measurement.2005.11.011

Majumder, M., Gangopadhyay, T. K., Chakraborty, A. K., Dasgupta, K., and Bhattacharya, D. K. (2008). Fibre Bragg Gratings in Structural Health Monitoring-Present Status and Applications. *Sensors Actuators A Phys.* 147, 150–164. doi:10.1016/j.sna.2008.04.008

Morana, A., Baghdasaryan, T., Girard, S., Marin, E., Geernaert, T., Thienpont, H., et al. (2019). Radiation-Induced Effects on Fiber Bragg Gratings Inscribed in Highly Birefringent Photonic Crystal Fiber. *IEEE Trans. Nucl. Sci.* 66, 120–124. doi:10.1109/tns.2018.2886167

Okabe, Y., Tsuji, R., and Takeda, N. (2004). Application of Chirped Fiber Bragg Grating Sensors for Identification of Crack Locations in Composites. *Compos. Part A Appl. Sci. Manuf.* 35, 59–65. doi:10.1016/j.compositesa.2003.09.004

Rajabzadeh, A., Heusdens, R., Hendriks, R. C., and Groves, R. M. (2019). Characterisation of Transverse Matrix Cracks in Composite Materials Using Fibre Bragg Grating Sensors. *J. Light. Technol.* 37, 4720. doi:10.1109/jlt.2019.2919339

Stathopoulos, N. A., Savaidis, S. P., Simos, H., Rigas, E., Correia, R., James, S. W., et al. (2019). Transmission Line Method for the Simulation of Fiber Bragg Gratings. *Appl. Opt.* 58, 353–360. doi:10.1364/ao.58.000353

Van Lancker, B., Hertelé, S., De Corte, W., Dispersyn, J., De Waele, W., and Belis, J. (2016). "Application of Digital Image Correlation in Linear Structural Adhesive Glass-Metal Connection Testing," in GlassCon Global (FCA Conferences, LLC), 305–313.

Wang, H., Xiang, P., and Jiang, L. (2019). Strain Transfer Theory of Industrialized Optical Fiber-Based Sensors in Civil Engineering: A Review on Measurement Accuracy, Design and Calibration. *Sensors Actuators A Phys.* 285, 414–426. doi:10.1016/j.sna.2018.11.019

Xiong, L., Guo, Y., Jiang, G., Jiang, L., and Zhou, X. (2019). Fiber Bragg Grating Displacement Sensor with High Measurement Accuracy for Crack Monitoring. *IEEE Sensors J.* 19, 1. doi:10.1109/jsen.2019.2930761

Zaghloul, M. A. S., Wang, M., Huang, S., Hnatovsky, C., Grobnic, D., Mihailov, S., et al. (2018). Radiation Resistant Fiber Bragg Grating in Random Air-Line Fibers for Sensing Applications in Nuclear Reactor Cores. *Opt. Express* 26, 11775–11786. doi:10.1364/oe.26.011775

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fenrg.2022.874194

Neural Network Acceleration of Genetic Algorithms for the Optimization of a Coupled Fast/Thermal Nuclear Experiment

John Pevey, Vlad Sobes and Wes. J. Hines*

University of Tennessee, Nuclear Engineering Department, Knoxville, TN, United States

Genetic algorithms (GA) are used to optimize the Fast Neutron Source (FNS) core fuel loading to maximize a multiobjective function. The FNS has 150 material locations that can be loaded with one of three different materials resulting in over 3E+71 combinations. The individual designs are evaluated with computationally intensive calls to MCNP. To speed up the optimization, convolutional neural networks (CNN) are trained as surrogate models and used to produce better performing candidates that will meet the design constraints before they are sent to the costly MCNP evaluations. A major hurdle in training neural networks of all kinds is the availability of robust training data. In this application, we use the data produced by the GA as training data for the surrogate models which combine geometric features of the system to predict the objectives and constraint objectives. Utilizing the surrogate models, the accelerated algorithm produced more viable designs that significantly improved the objective function utilizing the same computational resources.

Keywords: nuclear reactor, optimization, fast neutron source, convolutional neural networks, surrogate model, machine learning

OPEN ACCESS

Edited by:
Xingang Zhao,
Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Xinyan Wang,
Massachusetts Institute of Technology, United States
Yinan Cai,
Massachusetts Institute of Technology, United States

****Correspondence:***
John Pevey
jpevey@vols.utk.edu

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *11 February 2022*
Accepted: *28 April 2022*
Published: *17 June 2022*

Citation:
Pevey J, Sobes V and Hines WJ (2022) Neural Network Acceleration of Genetic Algorithms for the Optimization of a Coupled Fast/Thermal Nuclear Experiment. Front. Energy Res. 10:874194. doi: 10.3389/fenrg.2022.874194

1 INTRODUCTION

The optimization of nuclear problems can be a complex task often with multiple competing objectives and constraints. There is much research into the optimization of various aspects of nuclear reactors such as the initial design (Gougar, et al., 2010), fuel shuffling (Zhao, et al., 1998; Chapot, Da Silva and Schirru 1999), and shielding (Kim and Moon 2010; Tunes, De Oliveira and Schön 2017) optimizations. Due to the non-linear nature of these problems, optimization algorithms such as evolutionary and simulated annealing algorithms are often used. These methods do not guarantee that the optimal solution is found but can, with sufficient computational resources, clever heuristics, and the application of expert knowledge, often find solutions which are near-optimal.

When an optimization is of a function that is prohibitively expensive (such as solving the neutron transport equation), a surrogate model (Sobester, Forrester and Keane 2008) is produced and optimized instead. Surrogate models generally trade the computation expense of the original function for less accurate, but less expensive functions. The surrogate model is optimized instead of the original function. These can be as simple as linear functions or as complex as deep neural networks. In nuclear optimizations, surrogate models have been built to approximate expensive functions such as finite-element structural analysis (Prabhu, et al., 2020), computational

fluid dynamics (Hanna, et al., 2020), and neutron transport (Faria and Pereira, 2003; Hogle 2012; El-Sefy et al., 2021; Sobes, et al., 2021) related objectives.

In this paper, an optimization of the Fast Neutron Source at the University of Tennessee was performed using objectives calculated by solving the neutron transport equation. A surrogate model for these objectives is presented, and a genetic algorithm with and without acceleration using that surrogate model are compared. The following sections include overviews of both the Fast Neutron Source and of the neural network architecture used as the surrogate models. Two genetic algorithm optimizations are presented, the first uses the Non-dominated Sorting Algorithm-II and Monte Carlo N-Particle transport code (MCNP) (Goorley, et al., 2012) to solve the objective and constraint functions. The second optimization uses surrogate models with the NSGA-II algorithm, after which the individuals in the Pareto Front are evaluated with MCNP.

1.1 Fast Neutron Source

The Fast Neutron Source (FNS) (Pevey, et al., 2020) will be a platform for sub-critical integral cross section experiments at the University of Tennessee. It will be driven by 2.5 MeV neutrons produced by a deuterium-deuterium (DD) neutron generator and feature a flexible construction which will produce sub-critical benchmark experiments targeting specific nuclear data needs of next generation reactors.

The reduction of the uncertainty on next-generation reactor designs is a need for the expected rapid deployment of next generation reactors. Nuclear data uncertainty is propagated to all nuclear-related figures of merit of reactors such as k-eff, void, temperature and power reactivity coefficients and reactivity worth's. In a recent assessment of the nuclear data needed for advanced reactors, it was found that for several next-generation reactors need better resolved nuclear data (Bostelmann, et al., 2021). For example, in a sensitivity analysis of the Advanced Burner Reactor 1000 MWth Reference Concept, it was shown that the uncertainty in important nuclear characteristics is driven by uncertainties in uranium, plutonium, iron and sodium. In this concept, the uncertainty on k-eff, temperature coefficients of reactivity and Na void worth were 0.900%, 8.397% and 13.483%, respectively. These uncertainties require added margin in designs and can lead to less-than-optimal designs to account for these uncertainties.

Systematic integral data assimilation can be done to decrease these uncertainties. In this type of analysis, a suite of known benchmark models is collected, and a sensitivity analysis is completed for each. The known experimental values (and associated biases in the computational models) are then used to further decrease the uncertainty in the relevant quantities of interest. The FNS will be a source of these types of benchmarks which targets the reduction of uncertainties in neutronics calculation due to nuclear data uncertainty. The goal of a given configuration of the FNS then is to maximize the relevance of the experiment to some target advanced reactor concept and to maximize the total flux produced by the configuration to reduce the required FNS run-times.

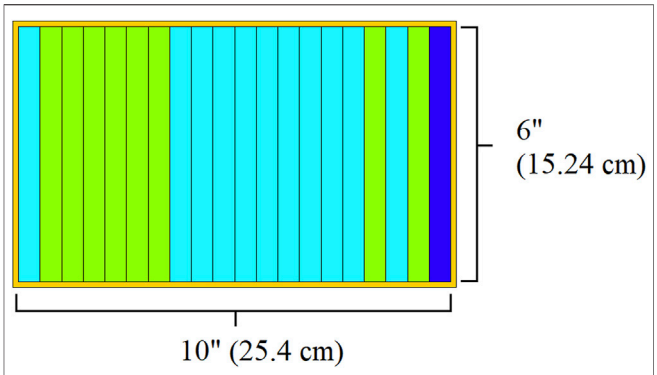


FIGURE 1 | Fast neutron source single cassette MCNP geometry.

The most basic geometry unit of the FNS are six″ x six″ x 0.5″ plates which can be one of three different materials in this study. Up to 20 of these plates are combined into aluminum cassettes (See **Figure 1**). Twenty-five of these cassettes are then combined into a 5 × 5 array called a zone (See **Figure 2**). There are three zones in the FNS, as seen in **Figure 2A**. In each zone, due to rotational symmetry, there are up to six unique cassette patterns (labeled I-N in **Figure 2B**). Note that in the work in this paper only the interior three cassette patterns (I, J, K) in each zone are optimized. The corner cassettes (N) are filled with stainless steel in this optimization and the cassettes labelled L and M are either the target coolants material (Zone A, B) or the thermal moderator (Zone C). The center cassette pattern, labeled I, is a variable sized that can be between 0-30 plates (up to an interior length 15″). The length of this cassette is a function of the number of plates within the cassette, with the experiment volume moving along with the cassette's changing length. The DD neutron source is in Zone C in a fixed location and is modelled as an isotropic 2.5 MeV neutron source. Other features of the FNS in the MCNP model are the stainless-steel reflector (F), the concrete pedestal (H), and the B_4C plates (G) which ensure subcriticality when inserted.

2 MATERIALS AND METHODS

2.1 Non-Dominated Sorting Algorithm-II

Genetic algorithms are a class of optimization algorithms which implement natural selection to optimize what may otherwise be intractable optimization problems. In simplest terms, a genetic algorithm takes an initial generation of individuals which are evaluated with respect to one or more objectives, and a subset of these individuals is selected and then combined to produce a unique individual and/or mutated randomly. How exactly each of the steps is accomplished is part of the art of a well-designed genetic algorithm. Genetic algorithms can also be augmented by using user-defined heuristics in each step to further increase the effectiveness of the algorithm to produce a suite of individuals required by the analyst.

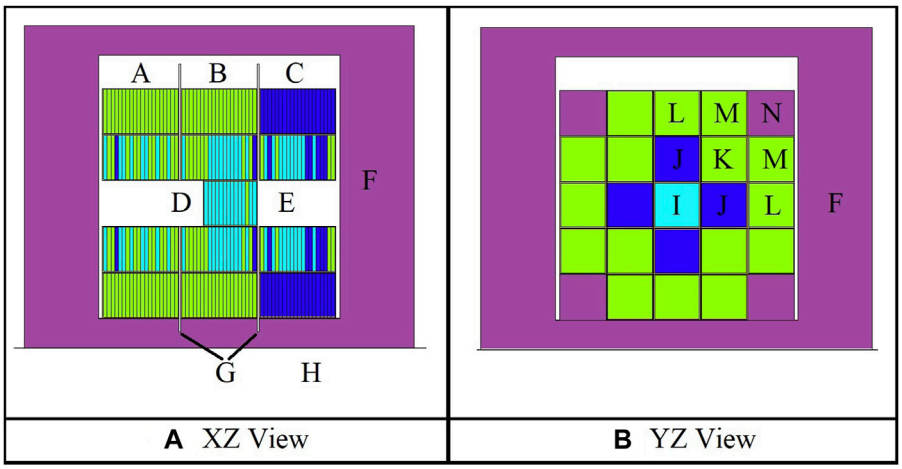


FIGURE 2 | Fast Neutron Source MCNP Geometry **(A)** XZ and **(B)** YZ View.

The Non-Dominated Sorting Algorithm-II (Deb et al., 2002) is a well-known variation on the standard genetic algorithm which incorporates the heuristics of elitism, non-dominated sorting and crowding distance to select the parents of the next generation. Elitism is simply the idea that parents are compared to the children of the current generation. This ensures that no progress made during the optimization is lost from one generation to the next. Non-dominated sorting is a heuristic for selecting which individuals are selected as parents of the next generation, which ranks individuals by the number of other individuals which dominate it (i.e.: have a better evaluation of an objective function). If no other individual has at least one objective which is better than a given individual, then that individual is non-dominated and is assigned rank 1. Subsequent individuals are assigned a non-dominated rank based on how many and which individuals dominate it. Rank 2 individuals are only dominated by rank 1 individuals, etc. Parents of the next generation are selected based on their non-dominated rank. If there are more individuals in each rank than there are available slots for parents, then the crowding distance metric is applied. Crowding distance is a heuristic which calculates the volume around each individual in the objective space. Individuals are selected as parents first based on the minimum and maximum for each objective function, and then by which occupy the largest volume in the objective space until the parent population is full. This heuristic is meant to both preserve the maximum and minimum individuals for each objective function and to preferentially select individuals which are in a less populated section of the objective space as parents with the goal of increasing the genetic diversity of the population.

The NSGA-II Algorithm is presented in **Figure 3**. It is adapted from the original paper describing the algorithm, A Fast and Elitist Multiobjective Genetic Algorithm NSGA-II (Deb et al., 2002):

2.2 Convolutional Neural Networks

A convolutional neural network (CNN) is a type of artificial neural network which can approximate a function in which not just the input values are important, but the relative positional information of the inputs is also important. CNNs are used primarily in machine vision tasks (Krizhevsky, Sutskever and Hinton 2012) and language processing tasks (Kalchbrenner, Grefenstette and Blunsom 2014). In machine vision tasks, the magnitude, relative position, and combinations of pixels are important for predicting what the pixels represent. In machine language tasks the relative positions of words are an important aspect to producing accurate translations.

CNN architectures include several layer types such as convolutional layers, pooling layers, non-linearity layers and fully connected layers (Albawi et al., 2017). In the convolutional layer, the namesake of the neural network architecture, is the convolution operation performed on the input to the layer and one or more learned kernels to produce a feature map. The convolution operation is a mathematical function that describes how one function modifies another as it is shifted over it. In practical terms, with a 2D input and kernel, the dot product between both functions is calculated and stored in the feature map. The kernel is then shifted by some number of columns and the dot product is calculated again. A non-linear function, such as a Rectified Linear Unit, is applied to the outputs of the convolutional layers.

The next layer type is the pooling layers. In these layers, a fixed filter is applied to the input to the layer. Commonly used filters are averaging and maximum pooling filters which return the average of a subset of the feature map or the maximum value within a subset. Unlike the convolutional layer, the stride of this operation is generally equal to the width of the filter. Commonly in machine-vision tasks a 2×2 filter is used. A 2×2 max pooling layer would reduce the size of the feature map by producing a new feature map which is composed of the maximum values in each 2×2 grid in the input feature map.

$R_t = P_t \cup Q_t$	Combine parent and offspring population
$F = fast - non - dominated - sort(R_t)$	$F = (F_1, F_2, I)$, all nondominated fronts of R_t
$P_{t+1} = \emptyset \; and \; i = 1$	
$until \; \lvert P_{t+1}\rvert + \lvert F_i\rvert \leq N$:	Until the parent population is filled
Crowding-distance-assignment(F_i)	Calculate crowding-distance in F_i
$P_{i+1} = P_{i+1} \cup F_i$	Include *i*th nondominated front in the parent population
$i = i + 1$	Check the next front for inclusion
Sort($F_i \prec_n$)	Sort in descending order using $\prec_n$
$P_{t+1} = P_{t+1} \cup F_i[1:(N - \lvert P_{t+1}\rvert)]$	Choose the first $(N - \lvert P_{t+1}\rvert)$ elements of F_i
$Q_{t+1} = make - new - pop(P_{t+1})$	Use selection, crossover, and mutation to create a new population Q_{t+1}
$t = t + 1$	Increment the generation counter

FIGURE 3 | The Non-Dominated Sorting Genetic Algorithm-II (Deb et al., 2002).

The last layer type commonly used in convolutional neural networks are fully connected layers. These final layers take as input the flattened final feature maps from the previous layers and non-linearly combine them into a prediction. The weights and biases in these layers, along with the kernels of the convolutional layers, are trained by backpropagation algorithm used to train other neural networks.

In nuclear applications, the relative position of materials to each other is important information when predicting nuclear quantities of interest such as k-eff. The k-eff of a given configuration of fissile, moderating, and absorbing materials is a function of where these materials are in 3-D space relative to each other, along with their respective nuclear data. CNNs produce predictions based on combinations of features which incorporate this 3-D data. There are other deep neural network architectures, such as recurrent (Liang and Hu, 2015) and transformer networks (Han et al., 2022), that have similarly been applied to machine vision tasks and therefore may also be able to predict nuclear related figures of merit such as k-eff, representativity, etc.

3 RESULTS AND DISCUSSION

This section discusses the initial optimization of the FNS using the NSGA-II algorithm and the subsequent optimization using the CNN-based surrogate models for the objective and constraint functions.

3.1 Optimization of the FNS by NSGA-II

The target of this FNS optimization is a generic sodium cooled fast reactor spectra. The objectives of this optimization are the maximization of the neutron flux per source particle in the experiment volume, the maximization of the representativity of the flux spectra in the experiment volume and the maximization of the change in k-eff when placing the target material in the experiment volume. These objectives are used as heuristics in place of the true objective of the FNS, which is to produce configurations which minimize the uncertainty on a target reactor concept propagated from nuclear data. Maximizing the total flux per source particle would mean reducing the total time required to complete a FNS experiment to sufficient statistical certainty. Representativity, or the E similarity coefficient in the SCALE manual (Rearden and Jessee 2018), is the angle between two n-length vectors in n-dimensional space. A value of 0 means that the two vectors are perpendicular to each other, a value of 1.0 would correspond to the two vectors pointing in the same direction and are therefore proportional to each other. The integral k-eff objective seeks to maximize the delta between the FNS experiment and an integral experiment where the entire experiment volume is filled with the material of interest. This last objective approximates a potential use-case of the FNS to perform integral experiment where the reactivity worth of the target moderator in the system is being measured before and after insertion. Maximizing the Δk-eff of that experiment would make the practical matter of measuring the reactivity difference easier.

These objectives were calculated by an MCNP source calculation with a total uncertainty on the experimental volume flux tally converged to <0.005% standard error. In addition, a constraint on k-eff was enforced which required all parents to have a k-eff below 0.95. This constraint was calculated using MCNP and to a standard uncertainty of at least 0.00150 dk-eff. An increasingly strict constraint on

representativity based on the idea of simulated annealing was also enforced. This constraint increased linearly over the first 50 generations of the optimization to E > 0.95. This constraint was enforced such that if there were not enough individuals in the parent population which met it, then it would be relaxed until at least 20 individuals met the constraint. Enforcing the constraint on representativity later in the optimization allows the algorithm to explore areas on the design space that would not be allowed by a strict constraint.

This optimization used three plate types: 9.75% enriched uranium metal, polyethylene, and sodium metal. The initial optimization of the FNS was run on the NECLUSTER at the University of Tennessee. This NSGA-II algorithm was implemented with the parameters described in **Table 1**. A total of 8,100 potential patterns of the FNS were evaluated in this optimization. Of these, 3,145 individual patterns met the k-eff criteria and were evaluated for the neutron flux-based objectives. The stopping criteria used for this optimization was number of generations, which was selected due to taking approximately 3 days (wall time) to complete. The MCNP calculations of the objective functions required the most computational time.

Figure 4 shows the average k-eff, representativity, total flux, and the integral k-eff value (times 10). The integral value is multiplied by 10 to show more detail in the data. Some features of this figure are that the linearly increasing constraint on representativity is obvious from both the linearly increasing average representativity and the decreasing average total flux over generations 34 to 53. Thereafter, the average total flux of the parents increases slightly but plateaus at generation 61.

This optimization produced a Pareto front of FNS designs which can be seen in **Table 2**. This set of individuals represent the trade-off between the objective functions of representativity, total flux and integral k-eff as found by the optimization algorithm. The representativity of these individuals ranges from 0.9510 to 0.9789. The total flux in the experiment volume ranges from 0.00146 to 0.0041, and the Δk of the integral experiment ranges from -0.00592 to +0.00195. Like previous analysis of a simplified approximation of the FNS, the representativity of the flux in the experiment volume and the total of that flux are negatively correlated (-0.858) while the flux and Δk-eff are positively correlated (0.810). An increase in Δk-eff will increase the total flux in the experimental volume but at the expense of the ability of the flux spectra to match the primarily fast target spectra as measured by representativity.

3.2 Surrogate Model Optimization of the FNS

The surrogate model optimization of the FNS used the same objectives and constraint as the optimization described in **Section 3.1**, but with the CNN-based surrogate models as solvers for the MCNP k-eff and source calculations required to evaluate the objectives. The architectures of these networks were found by the application of the Keras Tuner Python library (O'Malley, et al., 2019) using the data produced by the NSGA-II optimization. The hyper parameters found by this optimization can be found in **Table 3**. Further discussion of the method of optimizing the surrogate models will be presented in a forthcoming PhD dissertation at the University of Tennessee by the lead author.

The surrogate-based optimization algorithm is as follows:

1. Initialize population of 100 individuals and evaluate for objectives and constraint with MCNP.
2. Train surrogate models if more than 100 individuals have been evaluated.
3. Run NSGA-II algorithm (Deb et al., 2002) with the parameters described in **Table 1**, using the CNN-based surrogate models as the objective and constraint solvers.
4. Evaluate the final Pareto front (80 individuals) from the surrogate-based optimization with MCNP
5. If the total number of generations equals the stopping criteria, exit.
6. Return to Step #2

The data produced by the outer loop of the optimization, where MCNP is used to calculate the objectives and the constraint functions for every individual, was used for the training of the surrogate models used to evaluate the objectives of the inner loop of the optimization. A minimum of 100 valid individuals are required for training to occur, so after the first generation only the k-eff surrogate model was trained. The objective functions were trained after generation 2. The surrogate models were trained for 500 epochs using 90% of the available training data with 10% used for validation. The weights and biases of the epoch which

TABLE 1 | Description of NSGA-II and surrogate NSGA-II hyperparameters.

Description	NSGA-II hyperparameters	Surrogate NSGA-II hyperparameters
Stopping Criteria	100 Generations after initial generation	10 Generations after initial generation
Parent Population	20	80
Child Population	80	1,920
Crossover Rate	50%	50%
Crossover Type	Single point crossover	Single point crossover
Mutation Rate	10%, per plate	10%, per plate
Mutation Type	Single plate material change	Single Plate Material Change
Initial Population	100 randomly created	80 Pareto front individuals selected from all individuals evaluated with MCNP plus 1,920 randomly created
Objective/Constraint Solver	MCNP	CNN surrogate

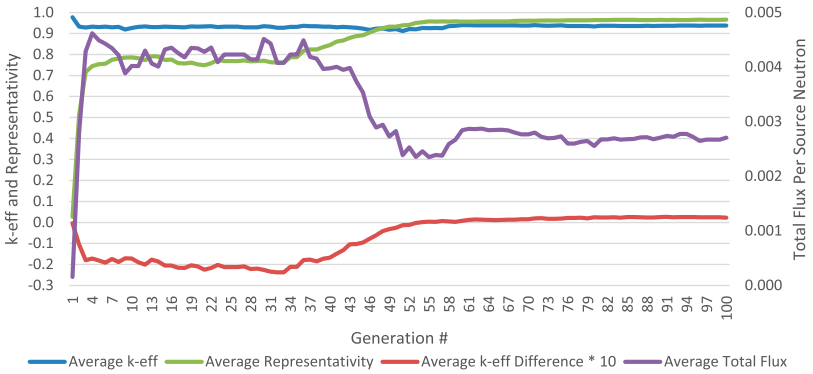


FIGURE 4 | Average parent objective function and constraint evaluation during NSGA-II optimization.

TABLE 2 | Final 20 individuals NSGA-II optimization.

Individual #	k-eff	Representativity	Total flux per source particle	Delta k-eff
1	0.94695	0.97893	0.002218	0.00002
2	0.93427	0.97825	0.002068	0.00144
3	0.91207	0.97750	0.001461	0.00191
4	0.91730	0.97548	0.001545	0.00405
5	0.94904	0.97460	0.002740	0.00178
6	0.91365	0.97293	0.001757	0.00468
7	0.93771	0.97225	0.002202	–0.00195
8	0.93805	0.96988	0.002228	0.00393
9	0.93879	0.96869	0.002518	0.00296
10	0.94841	0.96764	0.002966	0.00374
11	0.94323	0.96562	0.002991	0.00086
12	0.92581	0.96504	0.002105	0.00592
13	0.94596	0.96129	0.003270	–0.00027
14	0.92903	0.96045	0.002396	0.00471
15	0.94987	0.95876	0.003172	0.00266
16	0.94908	0.95574	0.003736	0.00122
17	0.94713	0.95333	0.003174	0.00268
18	0.94721	0.95227	0.003905	0.00150
19	0.94859	0.95187	0.004131	–0.00151
20	0.94371	0.95095	0.003571	0.00450

minimized the validation error were used as the surrogate model for that generation. The models were initialized with random weights at the beginning of the optimization.

The ability of the surrogate models to predict their respective values can be quantified by calculating the mean squared error (MSE) of the predictions to the true values as defined by,

$$Mean\,Squared\,Error = \frac{1}{n}\sum_{n=1}^{n}(x_i - \hat{x}_i)^2.$$

Where, x_i and $\hat{x}_i$ are the true and predicted values.

The MSE of each of the surrogate models initially are relatively large and decrease as the number of training examples increases. The surrogate model predicting k-eff is the first to stabilize at a value of approximately 1.0E-01 at generation 10. During the optimization the training data for this model increases by 80 in each step. If an individual does not have a k-eff below 0.95, then the three objective functions are not calculated for it. **Figure 5** shows the MSE of the prediction of the objective and constraints versus the true, MCNP-calculated, values for all individuals produced in each generation. The MSE of the surrogate models of the objective functions plateau at approximately

TABLE 3 | FNS CNN surrogate model hyperparameters.

Hyper parameter	k-eff	Total flux	Representativity	Integral K-Eff
1st Conv. Width	32	40	56	16
2nd Conv. Width	64	64	16	32
3rd Conv. Width	64	40	48	40
4th Conv. Width	40	24	16	24
5th Conv. Width	N/A	N/A	N/A	40
6th Conv. Width	N/A	N/A	N/A	16
Kernel Size	4	4	3	3
Dense Layer Width	32	256	32	96
Max Pool Size	9	4	10	7
# of Hidden Conv. Layers	3	3	3	5
Dropout Percentage	0.5	0.2	0.5	0.4
Learning Rate	0.0208839	0.0160466	0.0317394	0.0018139

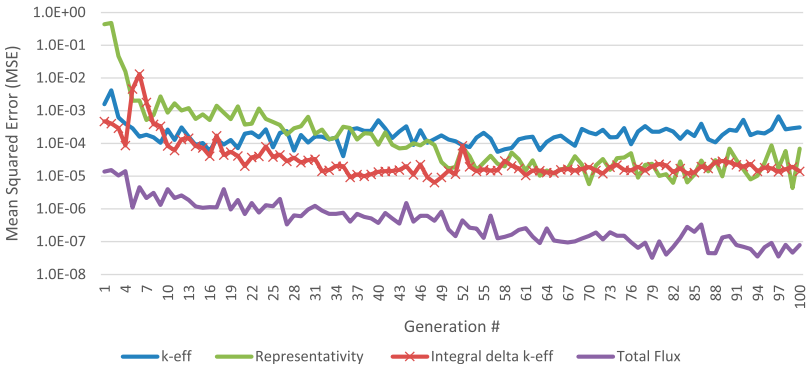


FIGURE 5 | MSE of surrogate models during optimization.

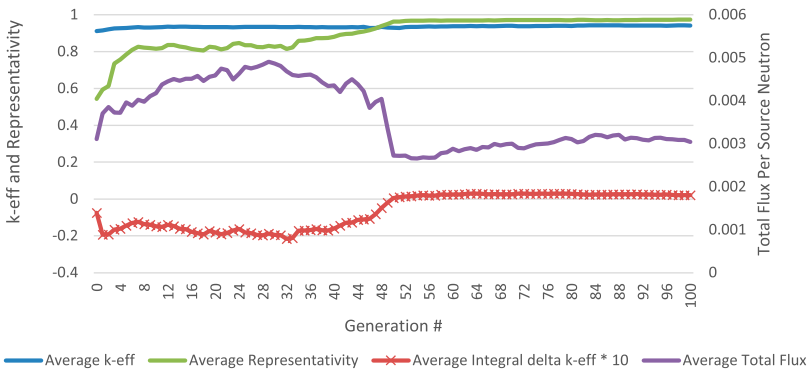


FIGURE 6 | Average objective and constraint evaluation during surrogate NSGA-II optimization for all parents in each generation.

generation 36 for the experimental k-eff value surrogate, generation 65 for the representativity surrogate and generation 80 for the total flux surrogate.

During this optimization, a total of 8,100 individuals were evaluated with MCNP and 2,120,000 individuals were evaluated with the surrogate models. Of the 8,100 patterns evaluated with MCNP a total of 5,868 individuals met the k-eff constraint and therefore evaluated with a MCNP source calculation. The average objective function and k-eff of all the individuals in the parent population is plotted in **Figure 6**. Like the NSGA-II algorithm before, the constraint on representativity that is maximally applied at generation 50 has a visible effect on the parent population.

The objective and constraint evaluations of the Pareto front individuals are provided in **Table 4**. The representativity of these individuals ranged between 0.95075 and 0.99502, the total flux per source neutron between 0.00077 and 0.00512 and the maximum and minimum integral Δk-eff values were +0.00717 and -0.00490, respectively. Like in the NSGA-II calculation, representativity and total flux per source neutron had a strong negative correlation (-0.95096).

3.3 Comparison of the Optimizations

The surrogate-based NSGA-II optimization of the FNS produced a Pareto front of potential FNS designs which outperformed those produced by the standard NSGA-II algorithm. **Table 5** presents the average, maximum and minimum of the objectives and constraint functions of the final Pareto front of each calculation. The surrogate-based optimization produced individuals with higher representativity, total flux per source particle and both a larger positive and negative integral Δk-eff. The final Pareto front of both calculations plotted by representativity vs. total flux per source particle and integral experimental k-eff vs. representativity is presented in **Figure 7** and **Figure 8**. These figures show that the surrogate-model based optimization produced a suite of individuals which better optimize the objectives.

A total of 8,100 potential FNS patterns were evaluated in during the both the NSGA-II and NSGA-II surrogate optimizations. In the standard NSGA-II calculation a total of 3,145 potential patterns met the k-eff constraint of 0.95. During the NSGA-II with surrogate model, a total of 5,845 potential patterns met the k-eff constraint. This is an increase of over 85% more viable FNS configurations evaluated in the surrogate accelerated calculation.

TABLE 4 | Final 100 individuals CNN-surrogate NSGA-II optimization.

Ind. #	k-eff	Rep	Total flux	Int Δk-eff	Ind. #	k-eff	Rep	Total flux	Int Δk-eff
1	0.86656	0.99502	0.00085	−0.00058	41	0.94480	0.97322	0.00303	0.00381
2	0.84565	0.99483	0.00077	−0.00128	42	0.94430	0.97188	0.00280	0.00645
3	0.90486	0.99476	0.00110	0.00037	43	0.94883	0.97118	0.00352	0.00157
4	0.90453	0.99462	0.00106	0.00148	44	0.94627	0.97090	0.00307	0.00400
5	0.88307	0.99409	0.00092	0.00303	45	0.94724	0.97068	0.00339	0.00256
6	0.94573	0.99409	0.00176	0.00003	46	0.94332	0.97064	0.00317	0.00430
7	0.94834	0.99308	0.00192	0.00116	47	0.94785	0.96974	0.00344	0.00199
8	0.94824	0.99297	0.00185	−0.00179	48	0.94906	0.96917	0.00357	0.00157
9	0.94705	0.99280	0.00187	−0.00385	49	0.94334	0.96841	0.00322	0.00327
10	0.94807	0.99187	0.00201	0.00203	50	0.94213	0.96836	0.00324	0.00393
11	0.94601	0.99114	0.00197	0.00304	51	0.94366	0.96778	0.00309	0.00459
12	0.94664	0.99079	0.00208	0.00229	52	0.94655	0.96706	0.00338	0.00465
13	0.94800	0.98978	0.00200	−0.00348	53	0.93790	0.96665	0.00303	0.00529
14	0.94374	0.98886	0.00172	−0.00448	54	0.94861	0.96638	0.00346	0.00184
15	0.94591	0.98876	0.00230	0.00208	55	0.94789	0.96625	0.00344	0.00261
16	0.88478	0.98852	0.00110	0.00502	56	0.94633	0.96557	0.00350	0.00167
17	0.94551	0.98746	0.00239	0.00278	57	0.94892	0.96512	0.00356	0.00307
18	0.93909	0.98695	0.00181	0.00415	58	0.94965	0.96450	0.00413	0.00157
19	0.94767	0.98596	0.00225	−0.00348	59	0.94463	0.96399	0.00357	0.00343
20	0.94870	0.98494	0.00251	0.00296	60	0.94700	0.96272	0.00419	0.00030
21	0.94806	0.98472	0.00252	−0.00003	61	0.94698	0.96081	0.00430	−0.00065
22	0.93854	0.98468	0.00188	−0.00374	62	0.94770	0.96068	0.00435	0.00118
23	0.94867	0.98407	0.00253	0.00192	63	0.94942	0.96004	0.00430	0.00121
24	0.94735	0.98385	0.00240	0.00301	64	0.94336	0.95989	0.00432	0.00135
25	0.92275	0.98374	0.00172	0.00576	65	0.94886	0.95958	0.00456	0.00107
26	0.94248	0.98346	0.00207	−0.00490	66	0.94951	0.95917	0.00458	0.00181
27	0.94835	0.98342	0.00237	0.00348	67	0.94756	0.95866	0.00452	0.00330
28	0.94765	0.98300	0.00266	0.00090	68	0.94824	0.95809	0.00479	0.00130
29	0.94594	0.98173	0.00233	0.00524	69	0.91854	0.95791	0.00279	0.00717
30	0.94690	0.98116	0.00266	0.00408	70	0.94885	0.95771	0.00459	0.00290
31	0.94897	0.97944	0.00282	0.00201	71	0.94605	0.95724	0.00434	0.00338
32	0.94878	0.97897	0.00267	0.00249	72	0.94886	0.95590	0.00433	0.00344
33	0.94429	0.97866	0.00280	0.00213	73	0.94709	0.95555	0.00473	0.00371
34	0.94766	0.97851	0.00283	0.00327	74	0.94565	0.95527	0.00476	0.00243
35	0.94557	0.97816	0.00282	0.00388	75	0.94688	0.95517	0.00448	0.00508
36	0.94740	0.97655	0.00288	−0.00150	76	0.94799	0.95497	0.00486	0.00351
37	0.94984	0.97601	0.00315	−0.00178	77	0.94526	0.95436	0.00458	0.00443
38	0.94662	0.97463	0.00286	0.00258	78	0.94195	0.95367	0.00453	0.00575
39	0.94811	0.97421	0.00304	0.00197	79	0.94922	0.95197	0.00512	0.00101
40	0.94567	0.97381	0.00266	0.00411	80	0.94733	0.95075	0.00472	0.00513

TABLE 5 | Average, Minimum and Maximum Constraint and Objective Functions of Final Pareto Front of NSGA-II and NSGA-II Surrogate Calculations. Mean values are italicized with the minimum and maximum values given in brackets below.

	NSGA-II	Surrogate NSGA-II	% Difference
k-eff	*0.93829, 0.91207, 0.94987*	*0.94102, 0.84565, 0.94984*	*0.003% [−7.3%, 0.3%]*
Representativity	*0.96557, 0.95095, 0.97893*	*0.99502, 0.95075, 0.97402*	*1.6% [−0.02%, 0.9%]*
Total Flux Per Source Particle	*0.00271, 0.00146, 0.00413*	*0.00304, 0.00077, 0.00512*	*24% [−47%, 12%]*
Integral k-eff	*0.00592, -0.00195, 0.00224*	*0.00209, -0.00490, 0.00717*	*−6.7% [151%, 21%]*

The trade-off of this increase in both calculational efficiency and in the more optimized Pareto Front is computational time for both the increased number of MCNP calculations and for the training and utilizing of the surrogate models. The increased number of MCNP calculations is a by-product of more effectively producing potential FNS designs and could be resolved by decreasing the number of the parent population. In the first time step the training of the surrogate models required an average of 2.3 min. By the final time step this increased to a maximum of 33 min. This increase in time requirement could be offset by capping the total number of examples used for surrogate model training or by reducing the size and complexity of the surrogate models themselves. Once trained, the surrogate models required on average a total of approximately 7.5 s to evaluate the 1,920 unique children in each interior GA step.

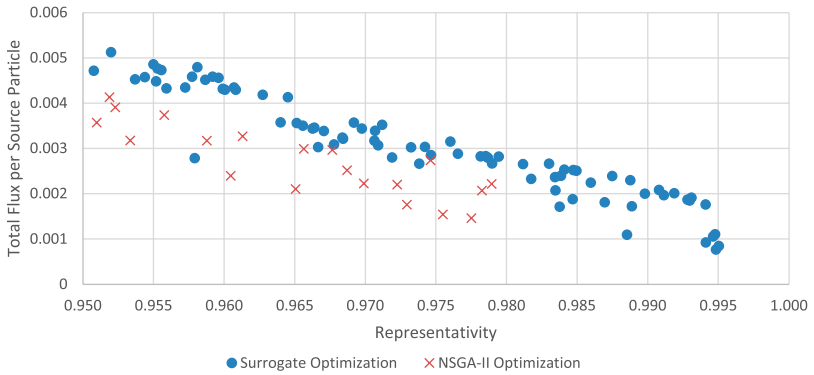


FIGURE 7 | Comparison of Pareto Front Individuals from Both Optimizations, Total Flux vs. Representativity.

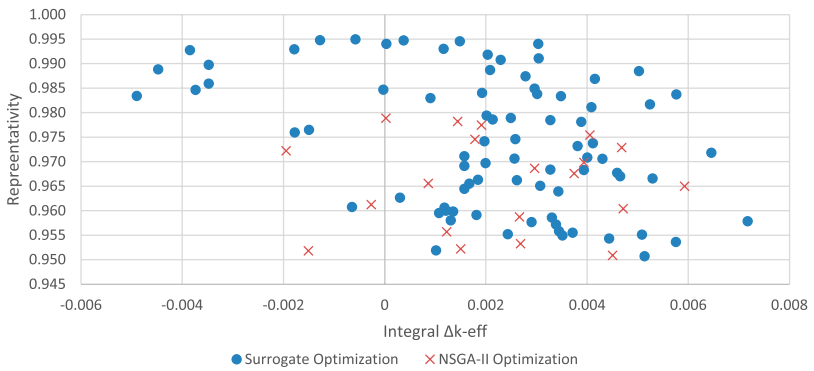


FIGURE 8 | Comparison of Pareto Front Individuals from Both Optimizations, Representativity *vs*. Integral Experiment Δk-eff.

4 CONCLUSION

The work in this paper presents the benefit of accelerating a genetic algorithm used for a multi-objective optimization of a nuclear experiment using convolutional neural network surrogate models comparing to a standard benchmark genetic algorithm. These surrogate models are trained in-line during the genetic algorithm and allow the evaluation of an increased number of potential designs, which leads to an increase in all objective functions. The architectures for the k-eff, representativity, neutron flux and integral k-eff experiment surrogate models are presented along with the methodology for producing them. Future work includes further expanding the use of the surrogate models for other useful objectives relevant for selecting FNS designs and producing more surrogates for other FNS designs targeting uncertainties in next-generation reactor designs.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

JP: Student research lead and lead author. VS: Committee member and primary research advisor. WH: Graduate committee chair and research lead.

FUNDING

This work is funded by the University of Tennessee, Knoxville Department of Nuclear Engineering.

ACKNOWLEDGMENTS

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenrg.2022.874194/full#supplementary-material

REFERENCES

Albawi, S., Tareq, A. M., and Al-Zawi, S. (2017). "Understanding of a Convolutional Neural Network," in 2017 International Conference on Engineering and Technology (ICET), 1–6. doi:10.1109/icengtechnol.2017.8308186

Bostelmann, R., Ilas, G., Celik, C., Holcomb, A. M., and Wieselquist, W. (2021). *Nuclear Data Assessment for Advanced Reactors*. Oak Ridge, TN (United States): Tech. Rep., Oak Ridge National Lab. doi:10.2172/1840202

Chapot, J. L. C., Carvalho Da Silva, F., and Schirru, R. (1999). A New Approach to the Use of Genetic Algorithms to Solve the Pressurized Water Reactor's Fuel Management Optimization Problem. *Ann. Nucl. Energy* 26, 641–655. doi:10.1016/s0306-4549(98)00078-4

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6, 182–197. doi:10.1109/4235.996017

El-Sefy, M., Yosri, A., El-Dakhakhni, W., Nagasaki, S., and Wiebe, L. (2021). Artificial Neural Network for Predicting Nuclear Power Plant Dynamic Behaviors. *Nucl. Eng. Technol.* 53, 3275–3285. doi:10.1016/j.net.2021.05.003

Faria, E. F., and Pereira, C. (2003). Nuclear Fuel Loading Pattern Optimisation Using a Neural Network. *Ann. Nucl. Energy* 30, 603–613. doi:10.1016/s0306-4549(02)00092-0

Goorley, T., James, M., Booth, T., Brown, F., Cox, L. J., Durkee, J., et al. (2012). Initial MCNP6 Release Overview. *Nucl. Technol.* 180, 298–315. doi:10.13182/nt11-135

Gougar, H. D., Ougouag, A. M., Terry, W. K., and Ivanov, K. N. (2010). Automated Design and Optimization of Pebble-Bed Reactor Cores. *Nucl. Sci. Eng.* 165, 245–269. doi:10.13182/nse08-89

Han, K., Yunhe, W., Hanting, C., Xinghao, C., Jianyuan, G., Zhenhua, L., et al. (2022). "A Survey on Vision Transformer," in IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE).

Hanna, B. N., Dinh, N. T., Youngblood, R. W., and Bolotnov, I. A. (2020). Machine-Learning Based Error Prediction Approach for Coarse-Grid Computational Fluid Dynamics (CG-CFD). *Prog. Nucl. Energy* 118, 103140. doi:10.1016/j.pnucene.2019.103140

Hogle, S. (2012). *Optimization of Transcurium Isotope Production in the High Flux Isotope Reactor*. PhD Thesis. University of Tennessee.

Kalchbrenner, Nal., Grefenstette, Edward., and Blunsom, Phil. (2014). "A Convolutional Neural Network for Modelling Sentences," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (Baltimore, Maryland: Association for Computational Linguistics), 655–665. doi:10.3115/v1/p14-1062

Kim, B. S., and Moon, J. H. (2010). Use of a Genetic Algorithm in the Search for a Near-Optimal Shielding Design. *Ann. Nucl. Energy* 37, 120–129. doi:10.1016/j.anucene.2009.11.014

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* 25. doi:10.1145/3065386

Liang, M., and Xiaolin, H. (2015). "Recurrent Convolutional Neural Network for Object Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3367–3375.

O'Malley, T., Elie, B., Long, J., Franccios, C., Haifeng, J., and Luca, I. (2019). *Keras Tuner*. https://github.com/keras-team/keras-tuner.

Pevey, J., Chvála, O., Davis, S., Sobes, V., and Hines, J. W. (2020). Genetic Algorithm Design of a Coupled Fast and Thermal Subcritical Assembly. *Nucl. Technol.* 206, 609–619. doi:10.1080/00295450.2019.1664198

Prabhu, S. R., Pandey, M. D., Christodoulou, N., and Leitch, B. W. (2020). A Surrogate Model for the 3D Prediction of In-Service Deformation in CANDU Fuel Channels. *Nucl. Eng. Des.* 369, 110871. doi:10.1016/j.nucengdes.2020.110871

Rearden, B. T., and Jessee., M. A. (2018). *SCALE Code System*. Oak Ridge, TN (United States): Tech. Rep., Oak Ridge National Lab. doi:10.2172/1426571

Sobes, V., Hiscox, B., Popov, E., Archibald, R., Hauck, C., Betzler, B., et al. (2021). AI-Based Design of a Nuclear Reactor Core. *Sci. Rep.* 11, 19646–19649. doi:10.1038/s41598-021-98037-1

Sobester, A., Forrester, A., and Keane, A. (2008). *Engineering Design via Surrogate Modelling: A Practical Guide*. John Wiley & Sons.

Tunes, M. A., De Oliveira, C. R. E., and Schön, C. G. (2017). Multi-Objective Optimization of a Compact Pressurized Water Nuclear Reactor Computational Model for Biological Shielding Design Using Innovative Materials. *Nucl. Eng. Des.* 313, 20–28. doi:10.1016/j.nucengdes.2016.11.009

Zhao, J., Knight, B., Nissan, E., and Soper, A. (1998). FUELGEN: A Genetic Algorithm-Based System for Fuel Loading Pattern Design in Nuclear Power Reactors. *Expert Syst. Appl.* 14, 461–470. doi:10.1016/s0957-4174(98)00005-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 19 July 2022
doi: 10.3389/fenrg.2022.851231

Surrogate Model of Predicting Eigenvalue and Power Distribution by Convolutional Neural Network

Jinchao Zhang[1]*, *Yufeng Zhou*[1], *Qian Zhang*[2]*, *Xiang Wang*[1] and *Qiang Zhao*[1]

[1]Fundamental Science on Nuclear Safety and Simulation Technology Laboratory, College of Nuclear Science and Technology, Harbin Engineering University, Harbin, China, [2]Department of Physics, Laboratory for Advanced Nuclear Energy Theory and Applications, Zhejiang Institute of Modern Physics, Zhejiang University, Hangzhou, China

During loading pattern (LP) optimization and reactor design, a lot of time consumption spent on evaluation is one of the key issues. In order to solve this issue, the surrogate models are investigated in this paper. The convolutional neural network (CNN) and fully convolutional network (FCN) are adopted to predict the eigenvalue and the assembly-wise power distribution (PD) for a simplified pressurized water reactor (PWR) during depletion, respectively. For the eigenvalue prediction during depletion, the error in the begin of cycle (BOC) and middle of cycle (MOC) is higher than that in the end of cycle (EOC). For the BOC and MOC, the samples with discrepancy over 500 pcm are less than 1%, except four burnup points. For the EOC, the fraction of samples with error over 500 pcm is less than 1%. As for the error of assembly power, the average absolute error is on the same level for all test cases. The average absolute relative error in the center region and the peripheral region is higher than that in the inter-ring region. The prediction results indicate the capability of neural network to predict core parameters.

Keywords: surrogate model, convolutional neural network, reactor design, eigenvalue prediction, power distribution prediction

OPEN ACCESS

Edited by:
Jun Wang,
University of Wisconsin-Madison, United States

Reviewed by:
Chunpeng Wu,
Duke University, United States
Jiankai Yu,
Massachusetts Institute of Technology, United States
Han Bao,
Idaho National Laboratory (DOE), United States

***Correspondence:**
Jinchao Zhang
13935397912@hrbeu.edu.cn
Qian Zhang
zhangqian0515@zju.edu.cn

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *09 January 2022*
Accepted: *31 May 2022*
Published: *19 July 2022*

Citation:
Zhang J, Zhou Y, Zhang Q, Wang X and Zhao Q (2022) Surrogate Model of Predicting Eigenvalue and Power Distribution by Convolutional Neural Network. Front. Energy Res. 10:851231. doi: 10.3389/fenrg.2022.851231

1 INTRODUCTION

One of the key issues during loading pattern (LP) optimization and reactor design is time consumption for evaluating millions of LPs. The purpose of evaluation is to give out the fitness of each LP, which is commonly represented by the core key parameters. The conventional evaluation method gives the fitness by executing core calculation repeatedly, and it is the main source of time consumption. Therefore, a surrogate model, which rapidly produces the core key parameters, is desired.

In the past research, the artificial neural network (ANN) has been used in predicting core key parameters. Due to the constraints of computing resources, earlier studies apply the multi-layer perceptron (MLP) as the prediction model. The linearized parameter or macro data in the core are used as the input. Early examples of research into the model include the prediction of power peak factor (Mazrou and Hamadouche, 2004; Souza and Moreira, 2006; Niknafs et al., 2010; Saber et al., 2015), eigenvalue (Mazrou and Hamadouche, 2004; Saber et al., 2015), departure from nucleate boiling ratio (Lee and Chang, 2003), and core reload program (Kim et al., 1993a; Kim et al., 1993b; Hedayat et al., 2009). However, previous studies with the MLP model have failed to find any link between the input data and the environment. Loss of spatial information is an inherent problem of MLP, which is caused by the linearization of input parameters.

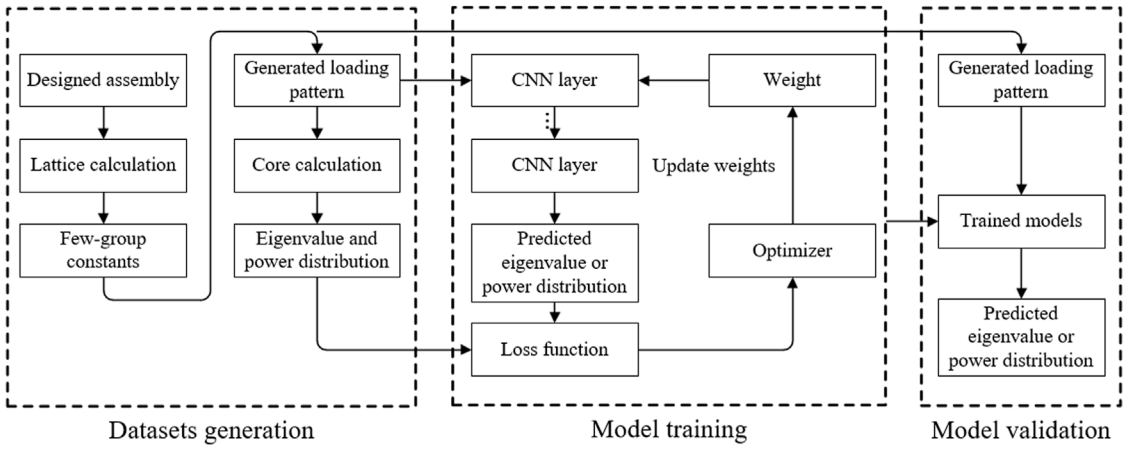


FIGURE 1 | Computing framework of core parameter prediction.

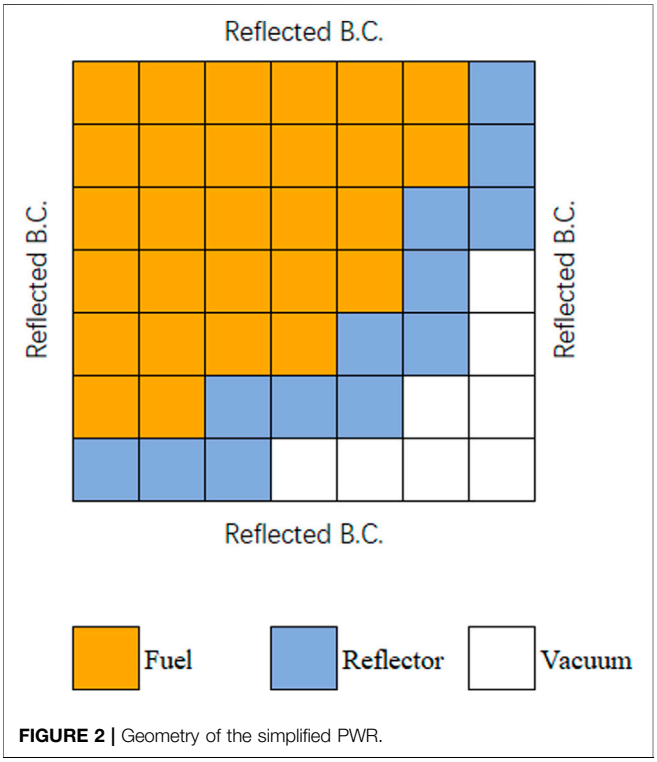


FIGURE 2 | Geometry of the simplified PWR.

Recently, researchers have shown an increased interest in predicting core key parameters with the convolutional neural network (CNN) (Krizhevsky et al., 2012). Unlike the MLP neural network, the CNN directly uses the information related to the problem as its input. In this way, the CNN avoids the inherent problem caused by the linearization of input parameters. Besides, the CNN uses the convolutional kernel as its base unit, which is beneficial to the learning of local features. Thus, in core parameter prediction, the CNN has a higher potential than the MLP, which is composed of dense layers. Surveys such as that conducted by Jang and Lee (2019) have shown that the CNN has higher accuracy than the conventional neural network, when they use the LP information as input to predict the peak factoring and cycle lengths. Further research (Jang and Lee, 2019) reveals that there is still some potential of the CNN model. When regularization and normalization are used, the prediction accuracy could be improved. Unlike Jang and Lee (2019) and Jang (2020), Zhang (2019) predicted the eigenvalue by the CNN with the assembly cross sections (XSs) as its input. The results indicate that the single freedom of XS as the input of the CNN has better performance than the multiple of that. In addition to lumped parameter prediction, Lee et al. (2019) used the macroscopic XSs as input and predicted the assembly-wise power distribution (PD). The results show that the CNN model has better performance for the problems similar to training data than the dissimilar problems. This phenomenon could be mitigated by the involvement of adversarial training data. Besides, the same padding setting is used to ensure the unchanged data size before and after through the convolutional layer. In the field of PD prediction, Whyte and Parks (2020) took the LP information as input to predict the pin-wise PD. Different from the former, they achieved PD prediction by reshaping normal CNN output to the LP size. However, to predict the PD by the CNN, the original network needs some special settings or changes. For example, Lee et al. (2019) involved the

TABLE 1 | Burnup points.

BU (GWd/t)	27.5, 30, 32.5, 35, 37.5, 40, 42.5, 45, 47.5, 50, 52.5, 55, 57.5, 60, 62.5, 65, 67.5, 70, 72.5, 75, 77.5, 80, 82.5, 85, 87.5, 90, 92.5, 95, 97.5, 100

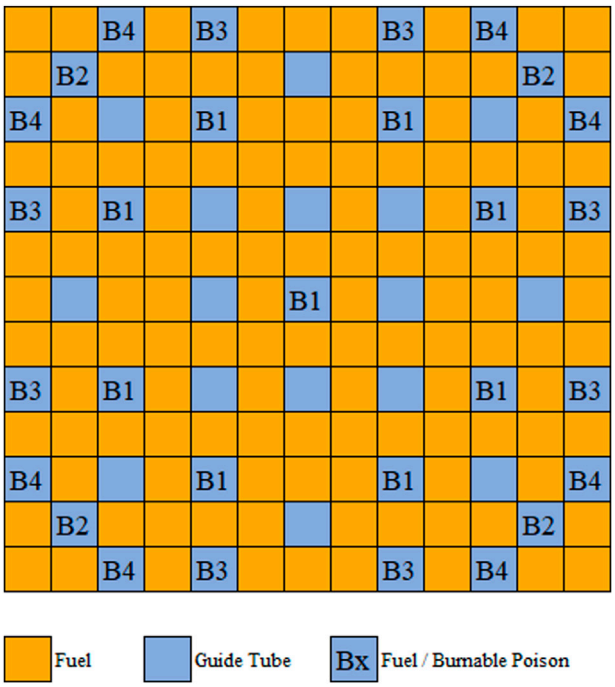
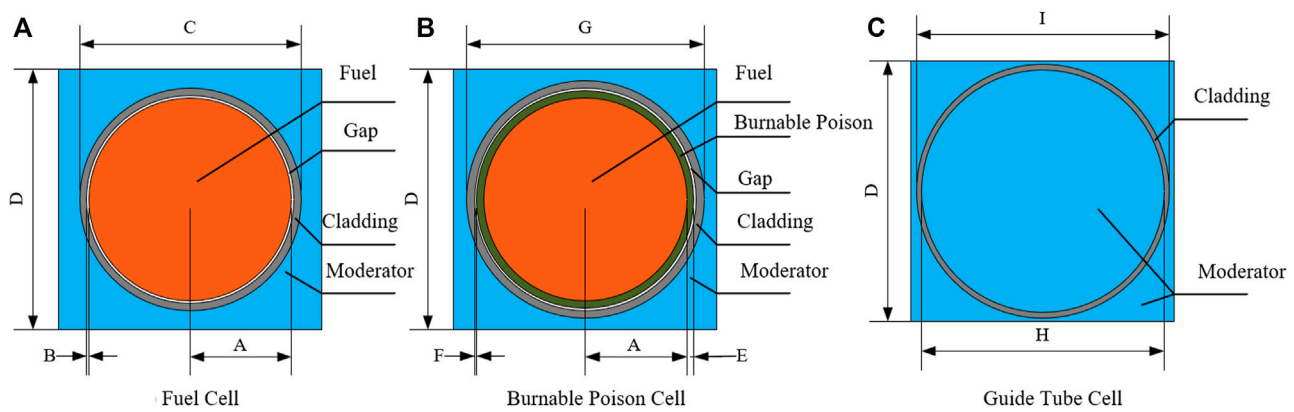


FIGURE 3 | Assembly geometry.

same padding setting. Whyte and Parks (2020) reshaped the normal output. To avoid further machinery, Long et al. (2015) designed a fully convolutional network (FCN) to achieve pixel-to-pixel prediction. For PD prediction, the pixel-to-pixel predicting process is similar to the conventional core calculation. They both use the pixel-like information as input and output. The main differences include two aspects. The first one is that the CNN uses the LP information as its input, while the core calculation uses the assembly XSs as its input. Another one is that the CNN predicts the PD by a neural network, while the core calculation gets it by solving partial differential equations. Therefore, the FCN is a natural choice for assembly-wise PD prediction with the LP information as its input. Zhang et al. (2020) modified the FCN to predict the PD and flux distribution with the assembly XSs as its input. Compared to the MLP, the FCN shows better performance. This research reveals that the FCN has potential in distributed parameter prediction. But there has been minimal investigation of predicting the PD with the FCN during depletion.

Therefore, in this study, the CNN and FCN are implemented to predict eigenvalue and assembly-wise PD during depletion, respectively. And a simplified PWR problem with 13 burnup points is used to assess the performance of the trained models. To simplify the input of the neural networks, we choose the single freedom as the model input. The LP is encoded as the index matrix to serve as input of the models.

The remainder of this paper is organized as follows. The methodology is introduced in **Section 2**. The numerical results are presented in **Section 3**. Finally, **Section 4** gives conclusions.



Abbreviations	Description	cm
A	Fuel radius	0.3378
B	Gap thickness in fuel cell	0.005
C	Cladding outer diameter in fuel cell	0.8076
D	Fuel rod pitch	1.1628
E	Poison thickness	0.015
F	Gap thickness in poison cell	0.005
G	Cladding outer diameter in poison cell	0.8376
H	Inner diameter in Guide tube cell	0.9676
I	Outer diameter in Guide tube cell	1.0896

FIGURE 4 | Geometry configuration of different cells.

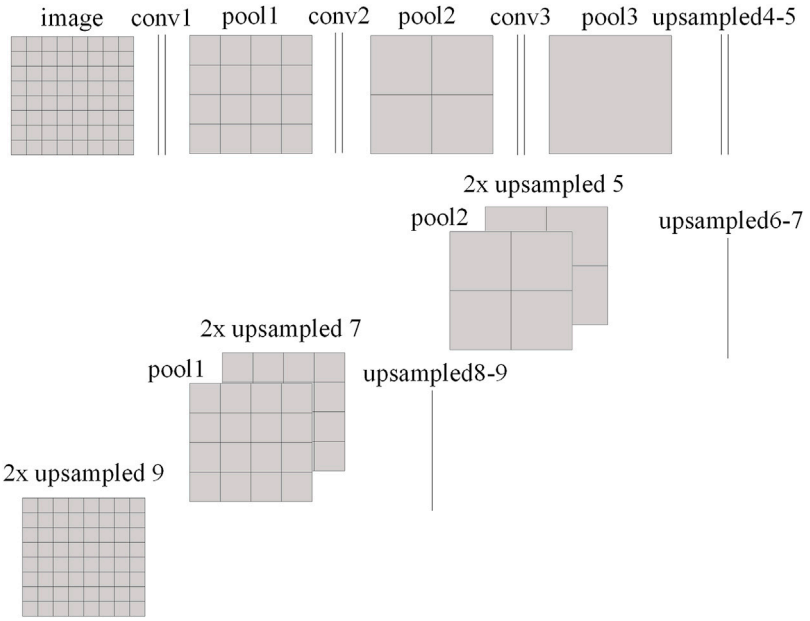


FIGURE 5 | Schematic diagram of the FCN structure.

TABLE 2 | Model parameters.

	CNN	FCN
Activation functions	ReLU (rectified linear unit)	ReLU
Loss function	MSE (mean squared error)	MSE
Optimizer	Adam (adaptive moment estimation)	Adam
Learning rate	1e-3	1e-4
Batch size	512	512
Epoch	700	700

2 METHODOLOGY

2.1 Computing Framework Based on Neural Network

In the evaluation process, the conventional method gets the eigenvalue and PD by performing core calculation. But the method used in this article predicts them with the CNN and FCN models. It is the main difference between this research and before. Neural network models are generated by training with the datasets. The general computing framework is shown in **Figure 1**, which comprises three parts:

(1) Dataset generation. First, multiple fuel assemblies with different enrichment fuel and burnable poison rod quality are designed and labeled with a unique ID. Second, the few-group constants are generated by lattice calculation with the Monte Carlo code Serpent (Leppänen et al., 2015). Third, the core LP is generated by the random method. Finally, the core calculation is performed with the in-house diffusion code to generate the training and validation datasets. The details are described in **Section 2.2**.

(2) Model training. For a neural network model, it achieves learning knowledge by adjusting the parameters in its network. The learning process is named training. The architecture of the network decides the degree of learning. In this study, the CNN and the FCN are adopted as prediction models. They are introduced in **Section 2.3**.

(3) Model verification. The verification is performed to verify the efficiency of the trained models. And **Section 3** gives the results.

2.2 Dataset Generation

The eigenvalue and the PD during burnup are determined by the initial LP. In this study, the reflector is fixed. The LP is randomly generated in a simplified PWR geometry in **Figure 2**, and the repetitive one will be abandoned. This generation process stops until the dataset size is reached. Then, the Serpent code is used to generate assembly few-group constants with the reflective boundary condition. Finally, the core calculation is performed with the in-house diffusion code for these LPs to obtain the eigenvalue and PD in each burnup point.

In order to preserve spatial information, the LP is encoded as a two-dimensional matrix, which is comprised of assembly IDs. Different assemblies have different fuel enrichments and burnable poison rod quantities. They include 25 enrichments varying from 12% to 18% divided into a constant interval of 0.25% and six different poison rod numbers including 0, 9, 13, 17, 21, and 25. The poison rods have the same ^{10}B enrichment, which is 95%. Through the arrangement and combination of these settings, 150 different assemblies are formed. The temperature of the different problems is fixed as 900K. The fuel assembly is depleted to 100GWd/t, and the specific burnup steps are listed in **Table 1**. The basic power density of the fuel assembly is 0.5 MW/kg. In the designed assemblies, the

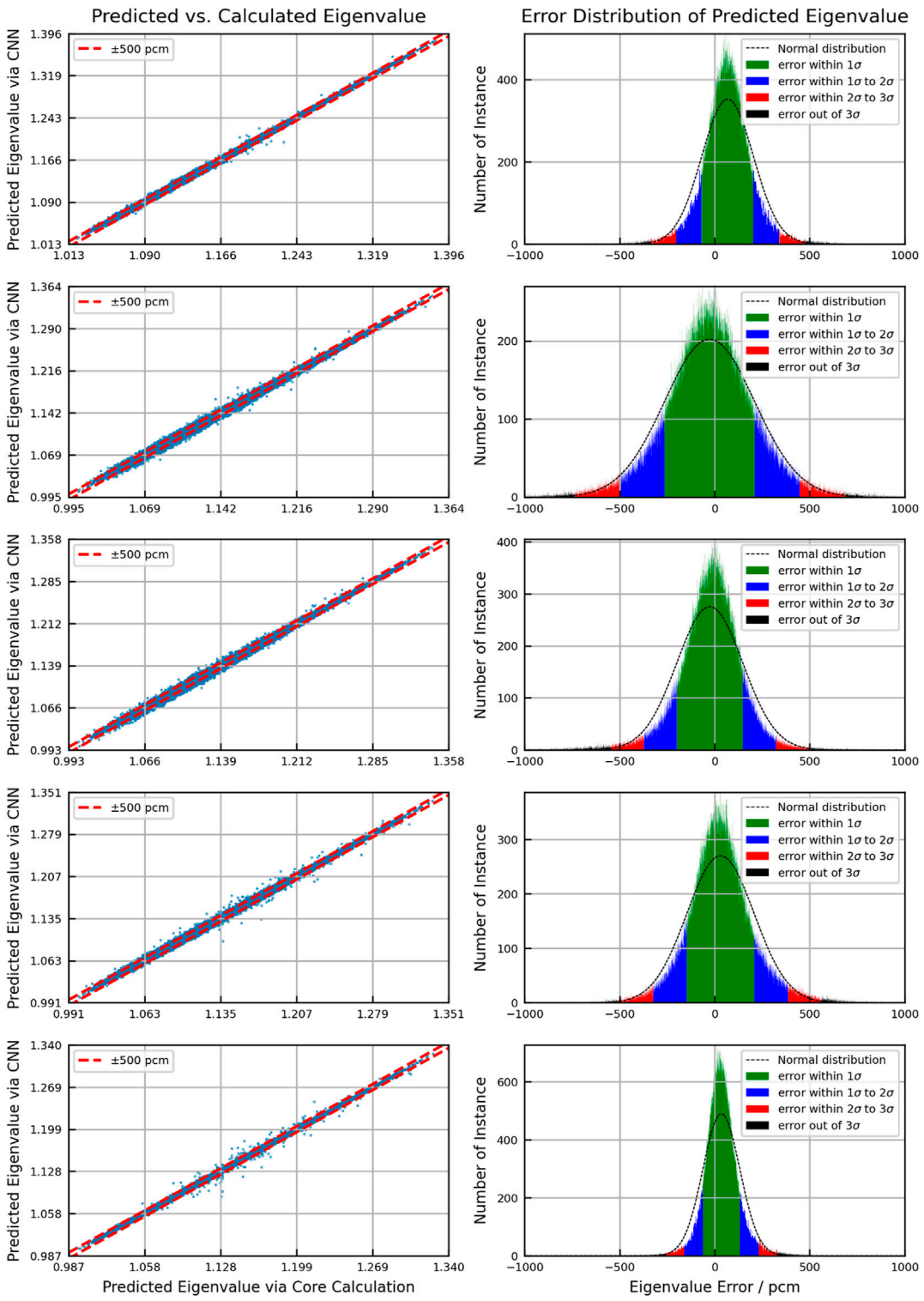


FIGURE 6 | Eigenvalue predicted accuracy of the BOC.

moderator is fixed as the 561K water without the void and boron. The cladding material is fixed as the 600K stainless steel. The gap is fixed as 600K oxygen. The compositions of the above materials are given in **Supplementary Appendix Table SA1**. **Figure 3** and **Figure 4** describe the geometry of fuel assembly and the configuration of pin cell, respectively. As shown in

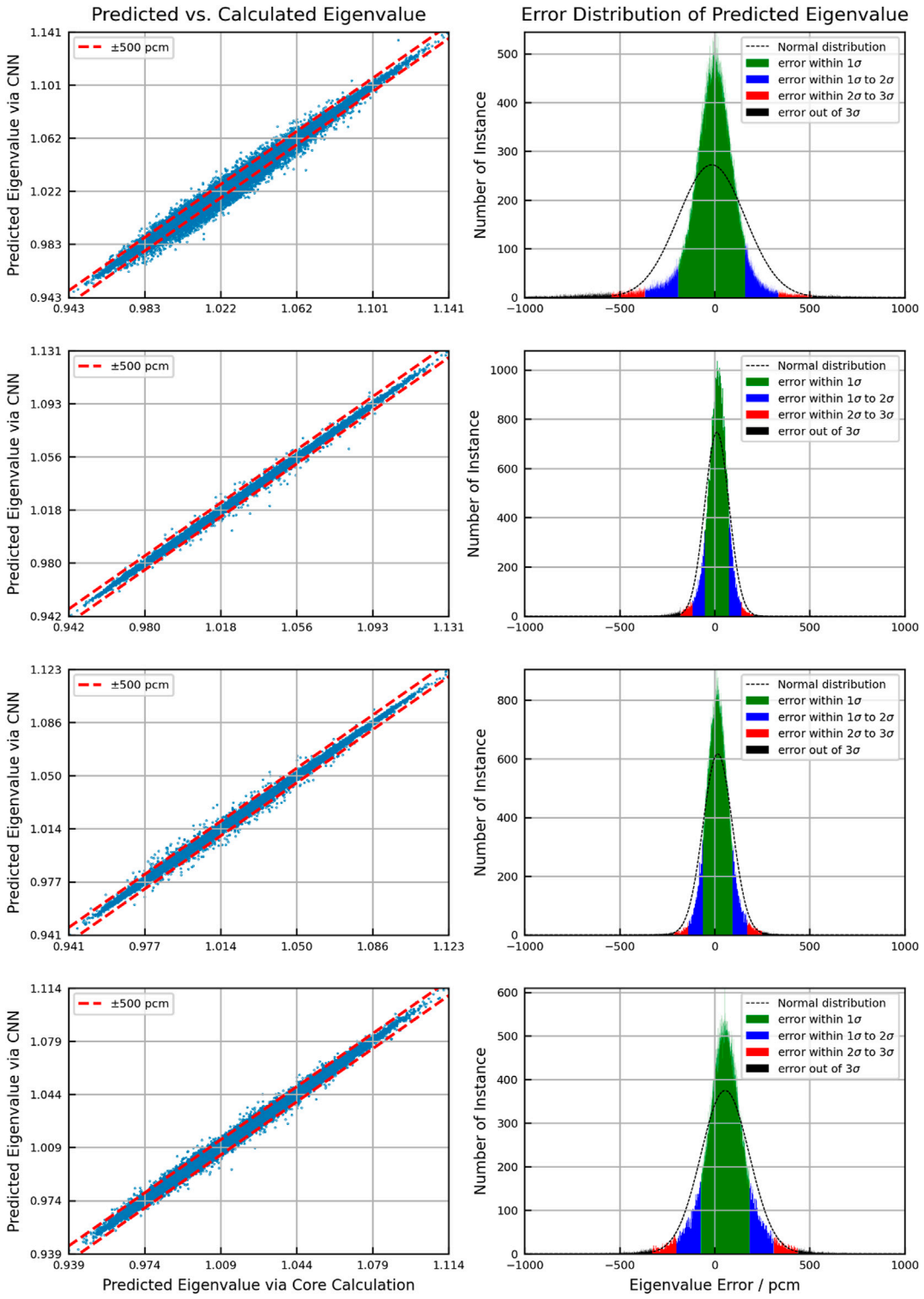


FIGURE 7 | Eigenvalue predicted accuracy of the MOC.

Figure 3, there are four different groups of poison rod locations. For the assembly with nine poison rods, the location of the poison rod is marked with B1. For the assembly with thirteen poison rods, the poison rods are placed not only in B1 but also in B2. For the assembly with seventeen poison rods, the poison rods are placed in B1 and B3. For the assembly with twenty-one

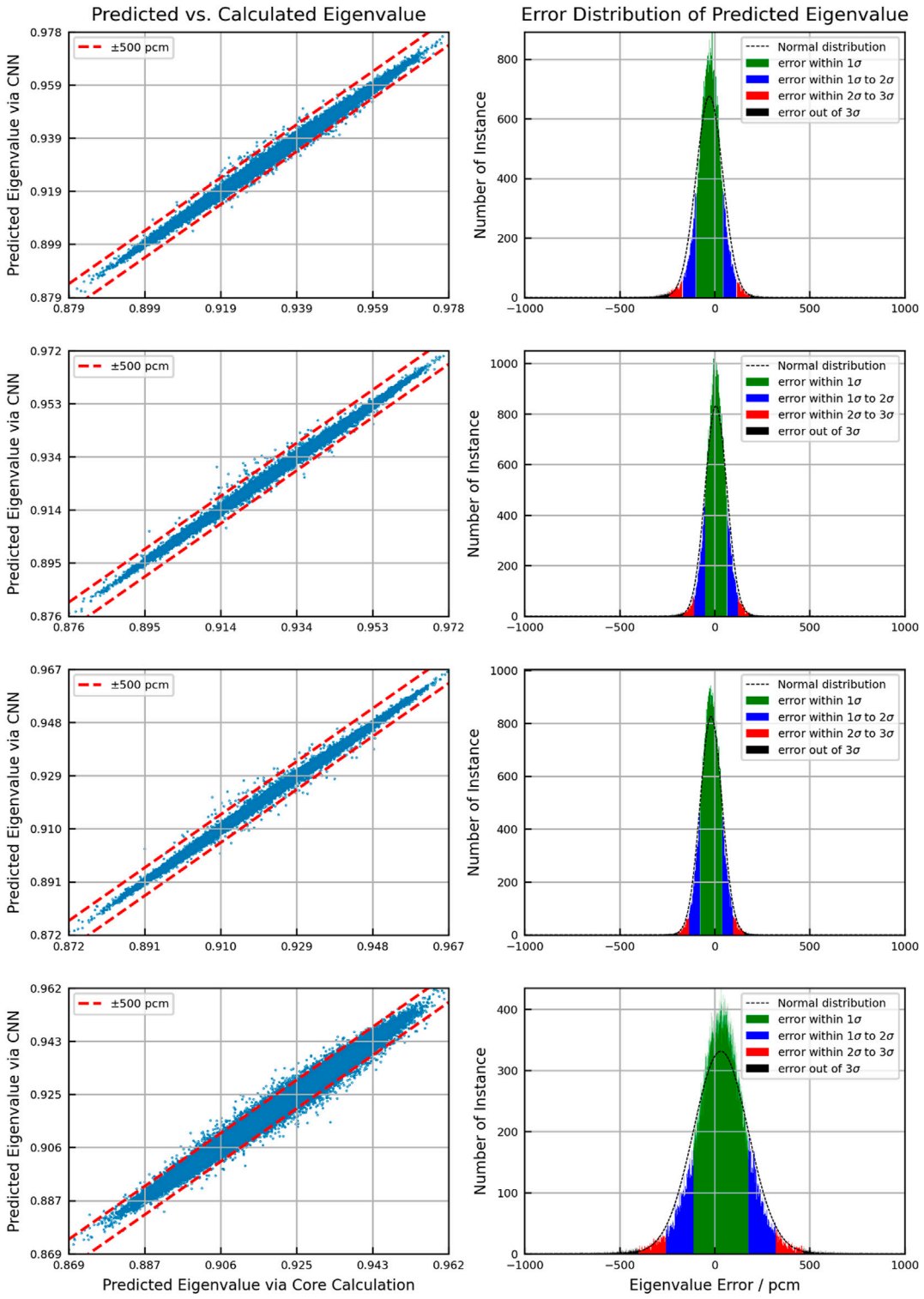


FIGURE 8 | Eigenvalue predicted accuracy of the EOC.

poison rods, the poison rods are placed in B1, B2, and B3. For the assembly with twenty-five poison rods, the poison rods are placed in B1, B3, and B4.

In addition, considering that the eigenvalue in different burnup points is not prior data, the eigenvalue is not normalized.

TABLE 3 | Eigenvalue prediction error of trained models.

Depletion (GWd/t)		e_{avg}[a]	e_{std}[b]	Frac. with e^c<200pcm	Frac. with e^c>500pcm
BOC	0.0	68	136	84.52	0.68
	0.2	−27	237	64.84	4.40
	0.5	−26	174	81.62	1.80
	1.0	31	177	78.86	1.56
	2.0	35	98	94.99	0.15
MOC	35.0	−16	176	88.42	2.61
	37.5	12	64	99.00	0.07
	40.0	16	78	98.08	0.16
	42.5	55	128	87.68	0.52
EOC	92.5	−27	71	98.20	0.07
	95.0	7	58	99.51	0.04
	97.5	−19	58	99.58	0.04
	100.0	33	144	84.90	0.55

a = average error (pcm).
b = standard deviation (pcm).
c = fraction of the eigenvalue with error (%).

2.3 Model Description

Different problems need different models with different architectures. In this study, two types of neural network models are used. We performed some primary sensitivity analysis on the network architectures used in this research. The results indicate that the current networks are the best. Any adjustment to them will worsen the prediction accuracy. Besides, it is difficult to find the law between the adjustment and the model performance. Thus, we do not report it in this manuscript. A complete sensitivity analysis of the network architecture requires a lot of iterations, which are difficult to complete in this research and are considered for the future. Then, the models used in this research are introduced. The first type is the CNN, which is adopted to predict the eigenvalue. Its structure comprises convolutional, pooling, and fully connected layers, which are shown in **Supplementary Appendix Table SA2**. The convolutional layer aims to take the spatial data into account. The pooling layer summarizes the feedback of the whole neighborhood and improves the efficiency of the network. The network finishes up with the fully connected layer, which connects the network and the object.

The second type is the FCN, which is a variant of CNN and is presented diagrammatically in **Figure 5**. It is noticed that, at the end of conventional CNN, several upsampling layers and concatenate layers are added to achieve backward stride convolution. They combine different feature layers and generate the output of corresponding size to the input. Thus, the FCN can predict the assembly-wise PD with the LP as its input. The FCN structure, which is adopted in this study, is shown in **Supplementary Appendix Table SA3**. **Table 2** gives the parameters used in the above models.

Furthermore, there are two points that need attention. First, the models with different burnup points are trained separately, which means that different burnup points have different neural networks. Second, due to the computational load for training 50 neural networks, this study chooses several representative burnup points. Thirteen burnup points are selected as representatives, including 0, 0.2, 0.5, 1.0, 2.0, 35.0, 37.5, 40.0, 42.5, 92.5, 95, 97.5, 100.0 GWd/t, which represent the begin of cycle (BOC), middle of cycle (MOC), and end of cycle (EOC).

In this study, the Keras framework (Chollet, 2015) is used to establish neural network structures upon TensorFlow. The CPU and GPU used in this work are 3.3 GHz Intel Core i9-7900X and Nvidia GeForce GTX 2080Ti, respectively.

3 NUMERICAL RESULTS

3.1 Eigenvalue Prediction

In this section, the CNNs are trained to predict the eigenvalue in different burnup points. The architecture shown in **Supplementary Appendix Table SA2** was used. 1 million samples (0.8 million for training and 0.2 million for validation, with no overlap between the two datasets) were generated to train and validate the CNNs.

Figures 6–8 show the scattering plot of eigenvalue and the error distribution of predicted eigenvalue. In the scattering plot, the red line is the mean absolute error (MAE) ±500pcm. In the error distribution histogram, the black line is the normal distribution curve based on the prediction results. The σ symbol represents the standard deviation of error distribution. The green region, the blue region, the red region, and the black region are the normal distribution range of 1σ, 2σ, and 3σ and the region out of 3σ, respectively. **Table 3** summarizes the detailed results.

As a result of validation, in different burnup stages, the error distribution of the discrepancy between the predicted eigenvalue and the reference is close to normal distribution. However, compared to the normal distribution, the eigenvalue prediction error of CNN models is higher within the 1σ range. This means that the distribution of prediction error is more concentrated around the average error. But in the region of error exceeding 1σ, the distribution is wider than the normal distribution. The average error in all cases is within 100pcm. There is no obvious peak shift. Besides, the samples with the absolute

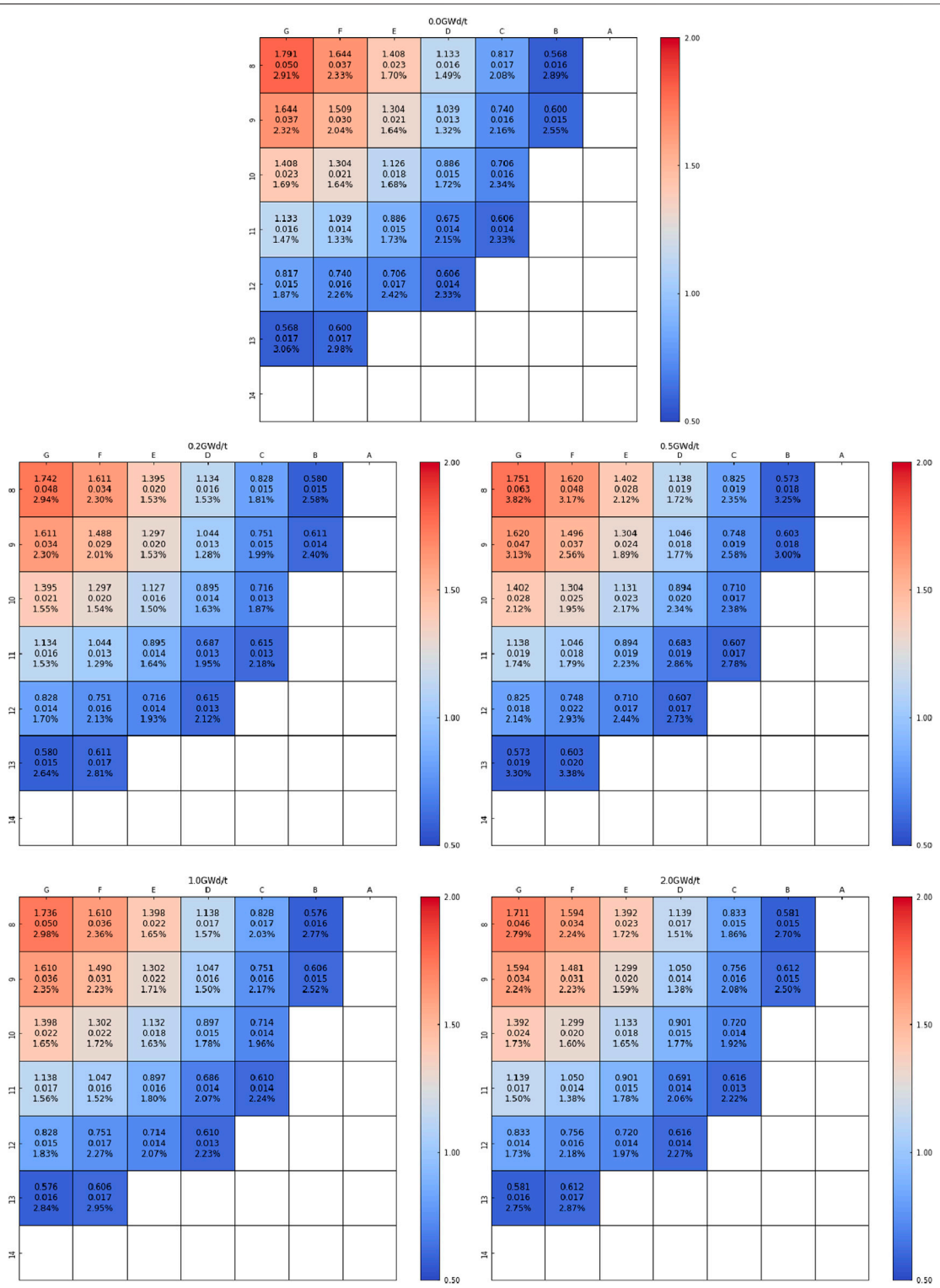


FIGURE 9 | Power distribution prediction result of the BOC. First line: assembly average power. Second line: mean absolute error. Third line: average value of absolute relative error.

error over 500 pcm are less than 3%, except for the second burnup point in the BOC. In general, the error in the BOC and MOC is higher than that in the EOC. This is because the reactivity of different assemblies varies greatly with poison depletion, which is greatly affected by the location. It increases the difficulty of the prediction process.

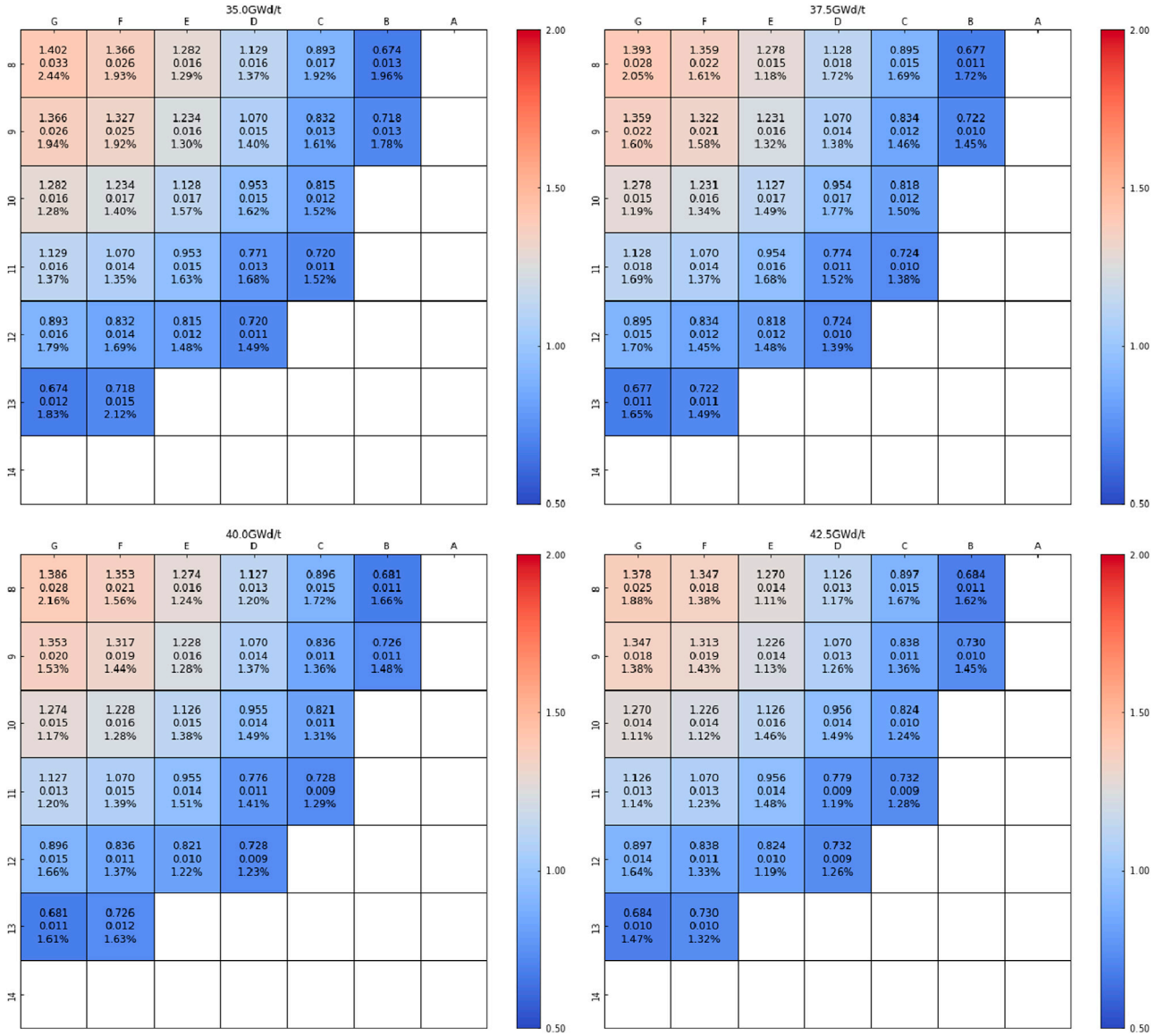


FIGURE 10 | Power distribution prediction result of the MOC. First line: assembly average power. Second line: mean absolute error. Third line: average value of absolute relative error.

3.2 Power Distribution Prediction

As a further test of neural network, the FCNs are trained to predict PD. The architecture, shown in **Supplementary Appendix Table SA3**, was used. 1 million PD samples, generated with the eigenvalue, were used to train and validate the FCNs.

For the relative power of each assembly in different burnup stages, **Figures 9–11** give the assembly average power, the assembly average absolute error, and the average value of the absolute relative error. The color of figures is given according to the assembly average power.

The average power of assembly decreases with the increase of the distance from the core center. The average absolute error shows the same trend. The error in the core center is higher than that in the core periphery. But the error is in the same level. For the average value of the absolute error, similar to eigenvalue prediction, the error in the BOC and MOC is higher than that in the EOC. In a specific burnup point, the error in the center region and peripheral region is higher than that in the inter-ring region. This phenomenon is caused by the spatial self-shielding effect. It means that there is different performance in different positions, even if the configuration is the same. Besides, the boundary condition, the reflector, and the void region exist in the core center and peripheral regions, which increases the difficulty of learning in these regions.

3.3 Discussion on Efficiency

In the process of evaluating the accuracy of models, the efficiency was also tested. During the diffusion calculation,

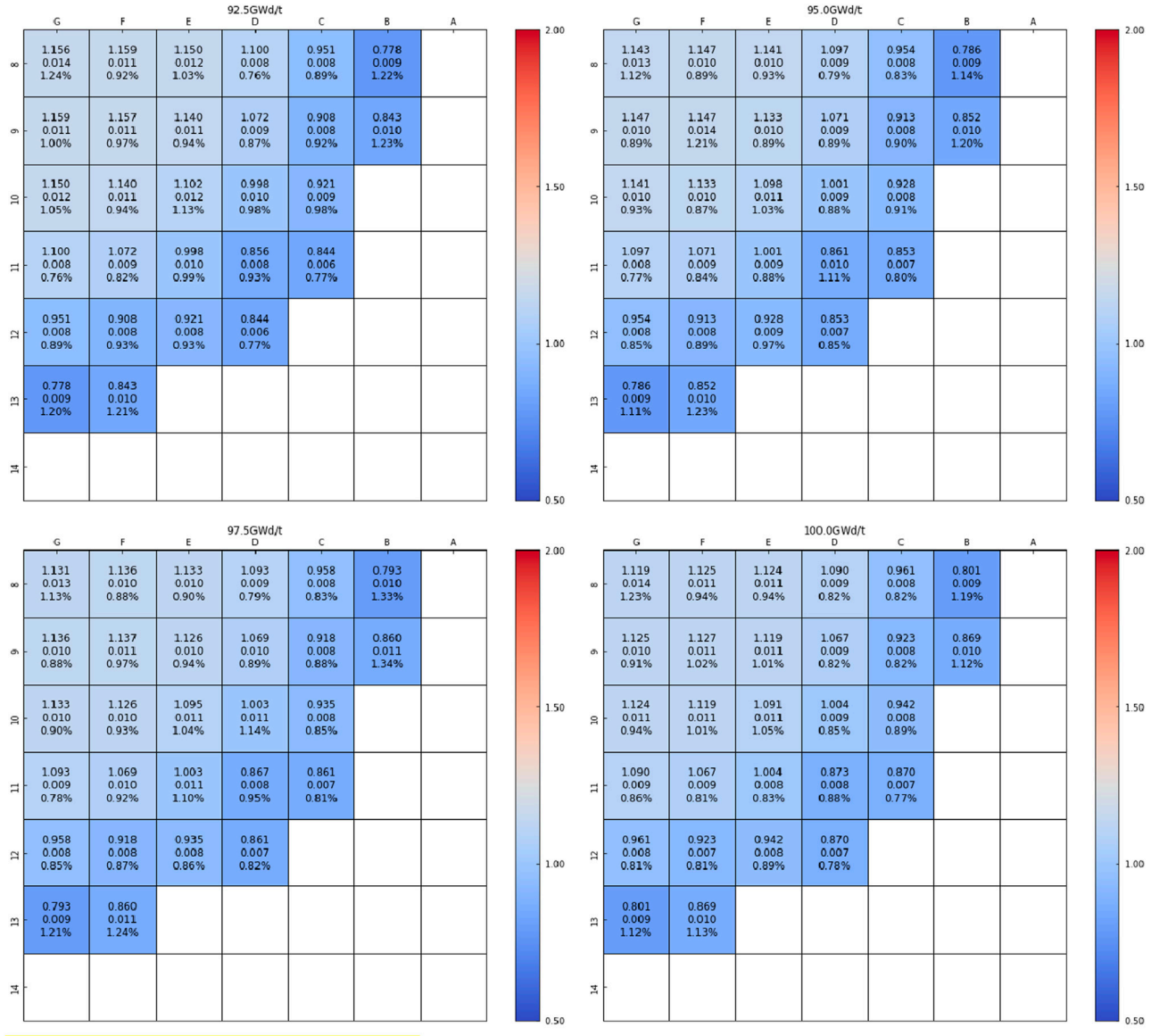


FIGURE 11 | Power distribution prediction result of the EOC. First line: assembly average power. Second line: mean absolute error. Third line: average value of absolute relative error.

TABLE 4 | Time comparison.

	Numbers (million)	Time
Dataset generation	1	2 days
CNN/training/one epoch	0.8	23 s
CNN/validation	0.2	18 s
FCN/training/one epoch	0.8	40 s
FCN/validation	0.2	14 s

each node was considered an assembly, and two-group constants were used. A steady-state diffusion calculation takes a few seconds, and the burnup calculation time can be neglected. It takes nearly 2 days to generate 1 million samples. However, for the neural network models obtained with training, it takes 18 s and 14 s to generate 0.2 million eigenvalues and PDs. The calculation efficiency is remarkably improved. For the CNN and FCN models, learning once (epoch) costs 23 s and 40 s, respectively. The detailed time comparison is listed in **Table 4**.

Similar to the application of neural network in other fields, the generalization of trained models is an issue. In the process of training, each assembly was labeled with a unique ID. The existing IDs cannot identify any new assembly. It directly leads to the lack of generalization ability. This is the limitation of neural networks in this study.

4 CONCLUSION

In this study, the CNN and the FCN were adopted to predict the eigenvalue and the power distribution for a simplified core during burnup, respectively. The loading pattern is encoded as a two-dimensional matrix as the models' input. As a result of validating for 0.2 million samples, the performance of the trained models for the EOC is better than that for the BOC and MOC. For eigenvalue prediction, the fraction of the eigenvalue with error more than 500 pcm is less than 1% for all burnup points in the EOC. But there are third burnup points in the BOC and one burnup point in the MOC, where the fraction is over 1%. This is caused by the changes of reactivity balance between poison and fuel in the BOC and MOC. However, in the EOC, the differences between different assemblies become small as the poison nuclide depletes to a negligible level. With regard to the power distribution, the performance of the trained models for the EOC is better than that for the BOC and MOC, too. The mean absolute error of assembly power follows that the error decreases with the increasing distance from the core center. But the error is in the same level. The average value of absolute relative error in the center and peripheral regions is larger than that in the inter-ring region. It is caused by the self-shielding effect, which leads to the different performance in different positions, even if the configuration is the same. Besides, the presence of boundary condition and the fact that the same information is shared among different fuel positions also increase the difficulty of learning.

This investigation indicates that the neural network has the capability to predict core key parameters, such as the eigenvalue and power distribution. Compared with the conventional diffusion calculation, the introduction of neural network as the surrogate model significantly reduces the computation time. However, the error is influenced by the depletion and the assembly location. This indicates that it is not appropriate to directly use the ID of assembly as input in this research. The selection of the model input needs to be further analyzed.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, and further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

JZ: Writing-original draft, Visualization, Simulation. YZ: Simulation. QZ (3rd author): Writing-review, Idea, Supervision. XW: Resources. QZ (5th author): Resources.

FUNDING

This work was supported by the National Natural Science Foundation of China (12105063), the Science and Technology on Reactor System Design Technology Laboratory (HT-KFKT-02-2019004), the Stability Support Fund for Key Laboratory of Nuclear Data (JCKY2021201C154), and the Natural Science Foundation of Heilongjiang Province of China (LH2020A001).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenrg.2022.851231/full#supplementary-material

REFERENCES

Chollet, F. (2015). Keras. Available at: https://keras.io.

Hedayat, A., Davilu, H., and Sepanloo, K. (2009). Estimation of Research Reactor Core Parameters Using Cascade Feed Forward Artificial Neural Networks. *Prog. Nucl. Energy* 51 (6-7), 709–718. doi:10.1016/j.pnucene.2009.03.004

Jang, H. (2020). *Application of Convolutional Neural Network to Fuel Loading Pattern Optimization by Simulated Annealing*. Korea: Transactions of the Korean Nuclear Society Autumn Meeting.

Jang, H., and Lee, H. (2019). *Prediction of Pressurized Water Reactor Core Design Parameters Using Artificial Neural Network for Loading Pattern Optimization*. Korea: Transactions of the Korean Nuclear Society Spring Meeting.

Kim, H. G., Chang, S. H., and Lee, B. H. (1993a). Optimal Fuel Loading Pattern Design Using an Artificial Neural Network and a Fuzzy Rule-Based System. *Nucl. Sci. Eng.* 115 (2), 152. doi:10.13182/NSE93-A28525

Kim, H. G., Chang, S. H., and Lee, B. H. (1993b). Pressurized Water Reactor Core Parameter Prediction Using an Artificial Neural Network. *Nucl. Sci. Eng.* 113 (1), 70.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet Classification with Deep Convolutional Neural Networks. *Adv. neural Inf. Process. Syst.* 25, 1097.

Lee, G.-C., and Heung Chang, S. (2003). Radial Basis Function Networks Applied to DNBR Calculation in Digital Core Protection Systems. *Ann. Nucl. Energy* 30 (15), 1561–1572. doi:10.1016/s0306-4549(03)00099-9

Lee, J., Nam, Y., and Joo, H. (2019). *Convolutional Neural Network for Power Distribution Prediction in PWRs*. Jeju, Korea: Transactions of the Korean Nuclear Society Spring Meeting.

Leppänen, J., Pusa, M., and Valtavirta, V. (2015). The Serpent Monte Carlo Code: Status, Development and Applications in 2013. *Ann. Nucl. Energy* 82, 142. doi:10.1016/j.anucene.2014.08.024

Long, J., Shelhamer, E., and Darrell, T. (2015). "Fully Convolutional Networks for Semantic Segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 3431–3440.

Mazrou, H., and Hamadouche, M. (2004). Application of Artificial Neural Network for Safety Core Parameters Prediction in LWRRS. *Prog. Nucl. Energy* 44 (3), 263–275. doi:10.1016/s0149-1970(04)90014-5

Niknafs, S., Ebrahimpour, R., and Amiri, S. (2010). Combined Neural Network for Power Peak Factor Estimation. *Aust. J. Basic Appl. Sci.* 4 (8), 3404.

Saber, A. S. (2015). "Nuclear Reactors Safety Core Parameters Prediction Using Artificial Neural Networks," in 2015 11th International Computer Engineering Conference (ICENCO) (IEEE).

Souza, R. M. G., and Moreira, J. M. (2006). Neural Network Correlation for Power Peak Factor Estimation. *Ann. Nucl. Energy* 33 (7), 594. doi:10.1016/j.anucene.2006.02.007

Whyte, A., and Parks, G. (2020). "Surrogate Model Optimization of a 'micro Core' PWR Fuel Assembly Arrangement Using Deep Learning Models," in

Proc. Int. Conf. Physics of Reactors 2020 (PHYSOR 2020), Cambridge, United Kingdom.

Zhang, Q. (2019). "A Deep Learning Model for Solving the Eigenvalue of the Diffusion Problem of 2-D Reactor Core," in Proceedings of the Reactor Physics Asia 2019 (RPHA19) Conference, Osaka, Japan, December 2–3, 2019.

Zhang, Q., Zhang, J. C., Liang, L., Li, Z., and Zhang, T. F. (2020). "A Deep Learning Based Surrogate Model for Estimating the Flux and Power Distribution Solved by Diffusion Equation," in Proc. Int. Conf. Physics of Reactors 2020 (PHYSOR 2020), Cambridge, United Kingdom.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Energy Research

ORIGINAL RESEARCH
published: 15 August 2022
doi: 10.3389/fenrg.2022.903464

An Evaluation of the Data-Driven Model for Bubble Maximum Diameter in Subcooled Boiling Flow Using Artificial Neural Networks

Xiaomeng Dong[1], Haoxian Chen[1], Changwei Li[1], Ming Yang[1], Yang Yu[2]* and Xi Huang[1]*

[1]Institute for Advanced Study in Nuclear Energy & Safety, College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen, China, [2]Nuclear Power Institute of China, Chengdu, China

In the subcooled boiling flow under low-pressure conditions, bubble characteristic diameter is of great influence on the surface heat transfer coefficient. However, large errors are still found in calculations using traditional mechanistic models or empirical correlations, especially for wide experimental condition. In this paper, we propose a widely applicable data-driven model using artificial neural networks (ANN) to predict the bubble maximum diameter and investigate the effect of experimental conditions. After a series of analyses on structural parameters and input parameters, the ANN model is established and validated based on six available experimental databases. The result shows that the relative error is around 14%. Uncertainty analysis is carried out for the four experimental conditions and two structural conditions. The results show the measuring accuracy of pressure is one of the most sensitive parameters on the prediction of bubble maximum diameter in the subcooled boiling flow under 1.0 MPa, especially for the bubble sizes larger than 0.5 mm. According to the results of uncertainty analysis, a new correlation is proposed for coefficients C and φ, which are used to express the effect of pressure and fluid dynamic. The new correlation works well for all the experimental databases, and the error for bubble datasets of large size is also modified. Furthermore, another independent validation with a low relative error to 14% is provided to prove the accuracy of the new correlation.

Keywords: bubble maximum diameter, data-driven model, subcooled boiling flow, artificial neural networks, sensitivity analysis

OPEN ACCESS

Edited by:
Jun Wang,
University of Wisconsin-Madison, United States

Reviewed by:
Luteng Zhang,
Chongqing University, China
Zhongchun Li,
Nuclear Power Institute of China (NPIC), China
Dongxu Ji,
The Chinese University of Hong Kong, China

****Correspondence:***
Xiaomeng Dong
dong_xiaomeng@yeah.net
Yang Yu
yuzhaoyang1987@sina.com

Specialty section:
This article was submitted to Nuclear Energy, a section of the journal Frontiers in Energy Research

Received: *24 March 2022*
Accepted: *06 June 2022*
Published: *15 August 2022*

Citation:
Dong X, Chen H, Li C, Yang M, Yu Y and Huang X (2022) An Evaluation of the Data-Driven Model for Bubble Maximum Diameter in Subcooled Boiling Flow Using Artificial Neural Networks. Front. Energy Res. 10:903464. doi: 10.3389/fenrg.2022.903464

1 INTRODUCTION

In a typical pressurized water reactor, saturation of boiling phenomenon is not allowed for its impairment to the fuel element. However, subcooled boiling flow may potentially occur in some special positions, such as the outlet of the fuel assembly or corners beside the spacer grid. The phase change could bring about a large increase in heat transfer efficiency of the heated surface. In complicated experiments involving a reactor core, the heat flux of the cladding surface is difficult to control precisely. Boiling crises determined by critical heat flux (CHF) may occur under some accident conditions. The CHF is a complicated two-phase flow phenomenon, characterized by a heat transfer mechanism change that rapidly decreases the efficiency of the heat transfer performance and increases the temperature of the heater

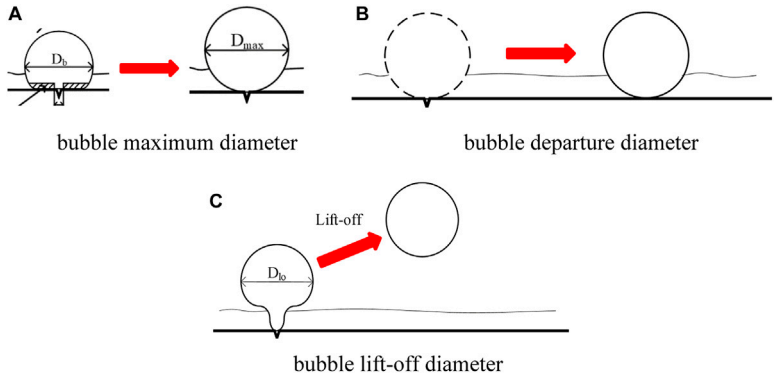


FIGURE 1 | The sketch map of bubble characteristic size. **(A)** Bubble maximum diameter, **(B)** bubble departure diameter, **(C)** bubble lift-off diameter.

surface. The high temperature could melt the fuel cladding and significantly damage the reactor core.

In general, the gas phase is always present in the form of bubbles during the subcooled boiling process. It features a series of action according to the effects of heat and fluid dynamics, such as generation, growth, sliding, and lift-off. In the subcooled boiling flow, the coefficient of the convective heat transfer is influenced by the characteristic parameters and dynamic behaviors of bubbles, their size and shape in particular. To describe the heat exchange and the heat flux partition on the heated surface in a theoretical view, the Rensselaer Polytechnic Institute (RPI) model (Kurul and Podowski, 1991) is widely used in the simulation of subcooled boiling flow. In this mechanistic model, the characteristic bubble diameter is considered one of the most important parameters, and it should be calculated accurately.

In the subcooled boiling flow, the characteristic bubble size has three main features, the maximum, departure, and lift-off diameter, which is shown in **Figure 1**. The maximum bubble diameter describes the limited size of growth for bubbles adjacent to the heated surface, while the other two parameters refer to the patterns of movement of the bubbles during the growth process.

There are few differences among these three parameters. However, the maximum bubble diameter and lift-off diameter are usually considered to be the same for the low and medium pressure conditions. Furthermore, the three bubble sizes show less of a difference for the non-slip bubble in low pressure and velocity (Hoang et al., 2016; Yoo et al., 2018). In calculations using computational fluid dynamics (CFD) method, all of these three characteristic sizes can be used in the calculation of the void fraction, heat and mass transfer, wall temperature, and quenching process of the different conditions (Tu and Yeoh, 2002; Krepper et al., 2013; Cheung et al., 2014; Gu et al., 2017).

Many approaches have been developed to calculate the bubble diameter, and these can be divided into three classifications, namely, empirical correlation, heat balance model, and force analysis model. Empirical correlations are proposed using the fitting of experimental databases. These usually contain several non-dimensional numbers, gradients, and ratios. Continuous multiplication and polynomial structure are a common functional form in the correlations, such as the model proposed by Prodanovic et al. (2002) and Brooks and Hibiki (2015). Certainly, these models show good performance for their own database. As the data expand and reach a larger scope, not all of the models could give the accurate results. Murallidharan et al. (2018) developed a bubble growth model that includes infinite bubble growth, wall effect multiplier, and bulk effect multiplier. Most of the coefficients in this model are fitted into polynomial structures. Although the model shows a good applicability over a wide range of conditions, it lacks any convenience of application and suitability for new experimental databases.

Apart from the empirical correlation, force balance and heat balance theories are used to calculate the characteristics bubble sizes from a theoretical view. According to the basic principle, the force balance model (Klausner et al., 1993; Situ et al., 2005; Yeoh et al., 2008) is suitable for calculating the departure and lift-off diameters of the bubbles, and the heat balance model (Ünal, 1976) is used to calculate the maximum diameter of the bubble. As can be seen from the application (Dong and Zhang, 2021) of these two models, the relative errors of the wide experimental databases reach around 40%. In reality, not all of the experiments provide local wall superheating as a result of immature measurement. The Chen correlation is always used for the calculation of wall superheat, especially in the heat balance model (Hoang et al., 2016; Dong and Zhang, 2021). This is the most accurate model for wall superheating, and its relative error is around 20%, which is a large proportion of the total error in the maximum bubble diameter.

Three submodules beyond wall superheating reflect the influence of pressure, velocity, and local subcooling. The former two are empirical correlations, while a partially mechanistic model is provided for the effects of local subcooling (Ünal, 1976). Because they were proposed in Ünal's original heat balance model, almost no modification has been made for over 40 years. Dong and Zhang (2021) used Reynolds numbers instead of velocity values, which decreases the relative error by about 5–10%. After the investigation of several experimental databases, a more accurate model should be proposed to describe the effects of pressure on the characteristic bubble sizes in the future.

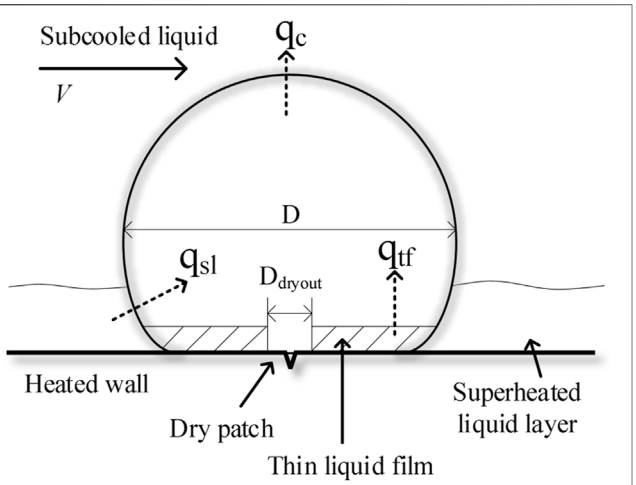


FIGURE 2 | Heat flux partition for a growing bubble in the subcooled boiling flow.

Thanks to the successful application of data mining techniques, data-driven theory has been used in many areas of industrial knowledge, such as fluid dynamics and intelligent manufacturing. In the fluid dynamics and thermodynamics, many data-driven models have been used to identify of flow regimes (Salgado et al., 2010; Affonso et al., 2020; Aarabi Jeshvaghani et al., 2021) and predict boiling crises (Greenwood et al., 2017; Yan et al., 2021) using artificial neural networks (ANN). All the results show good performance in a wide experimental condition. The relative error of CHF is around 20% which is better than that of empirical correlation or mechanistic models. Jung et al. (2020) investigated the bubble size distribution in turbulent air-water bubbly flows by using multi-layer ANNs. Compared to the 20% error of traditional theoretical models, the results of the use of ANNs show average relative error of 4.98% for the given experimental datasets.

This study discusses the application of ANNs for the calculation of maximum bubble diameter. We also try to investigate the influence of several experimental conditions based on the trained ANN model. The conclusions of uncertainty analysis are helpful to correct the correlation of coefficients to reflect the effects of pressure and mass flux in a more accurate way.

2 MECHANISTIC AND ARTIFICIAL NEURAL NETWORKS MODELS OF BUBBLE MAXIMUM DIAMETER

2.1 Mechanistic Models

Although there are differences among the proposed heat balance models, the key theoretical points are shared; these describe an equivalence between absorbed and released heat bubbles. As **Figure 2** shows, the bubble is regarded as an approximate sphere, while the growth process is determined by the heat from thin liquid film q_{tf}, from superheated liquid layer q_{sl}, and heat q_c, dissipated through the subcooled liquid at the upper-half surface. Although a dry patch at which the vapor is contacted with heated wall directly exists, its heat flux is neglected when the bubble reaches its maximum value.

According to heat balance theory, the mechanistic model can be developed as **Equation 1**.

$$\rho_g h_{lg} \frac{d}{dt}\left(\frac{\pi D^3}{6}\right) = q_{sl}A_{sl} + q_{tf}A_{tf} - q_c A_c \tag{1}$$

The heat flux components used in **Equation 1** are described in detail (Levenspiel, 1959; Zuber, 1961; Sernas and Hooper, 1969) and are given as follows.

$$\begin{cases} q_{sl} = \dfrac{k_l \Delta T_c}{\sqrt{\pi \alpha_l t}} \\ q_{tf} = \dfrac{k_l \Delta T_w \gamma}{\sqrt{\pi \alpha_l t}} \\ q_c = h_c \Delta T_{sub} \end{cases} \tag{2}$$

The model of q_{tf} was developed by Sernas and Hoopers (1969) and was used in Ünal's heat balance theory. Parameter γ is used to reflect the influence of the thermal properties of the heated wall. This is shown in **Equation 3**.

$$\gamma = \sqrt{\frac{k_s \rho_s C_s}{k_l \rho_l C_{pl}}} \tag{3}$$

h_c is the heat transfer coefficient for condensation. Based on the model developed by Levenspiel (1959), Ünal (1976) derived a model for condensation heat transfer coefficient, as follows:

$$h_c = \frac{C \varphi h_{lg} D}{2\left(1/\rho_g - 1/\rho_l\right)} \tag{4}$$

where C and φ are determined by the pressure and velocity of liquid phase, respectively.

$$C = \begin{cases} 65 - 5.69 \times 10^{-5}\left(P - 10^5\right) & 0.1MPa \le P \le 1MPa \\ 0.25 \times 10^{10} P^{-1.418} & 1MPa < P \le 17.7MPa \end{cases} \tag{5}$$

$$\varphi = \max\left[1, \left(v/0.61\right)^{0.47}\right] \tag{6}$$

Dong and Zhang (2021) provided a new model (**Equation 7**) for parameter φ using the Reynolds number instead of velocity. Re_l is widely used to determine the flow status, which has a decisive role on the characteristic temperature distribution of superheated liquid layer. In addition, it can reflect the shape effect of flow channel.

$$\varphi = \max\left(0.141, \left(Re_l/39300\right)^{1.43}\right) \tag{7}$$

The surface area ratio of each heat flux is also of great importance for the heat balance model. The surface area A_{total} of the generated bubble can be divided into four parts, including the area of thin liquid film A_{tf}, area of superheated liquid layer A_{sl}, area of heat dissipation A_c, and area of dry patch A_{dryout}. These five parameters can be described by the following equations.

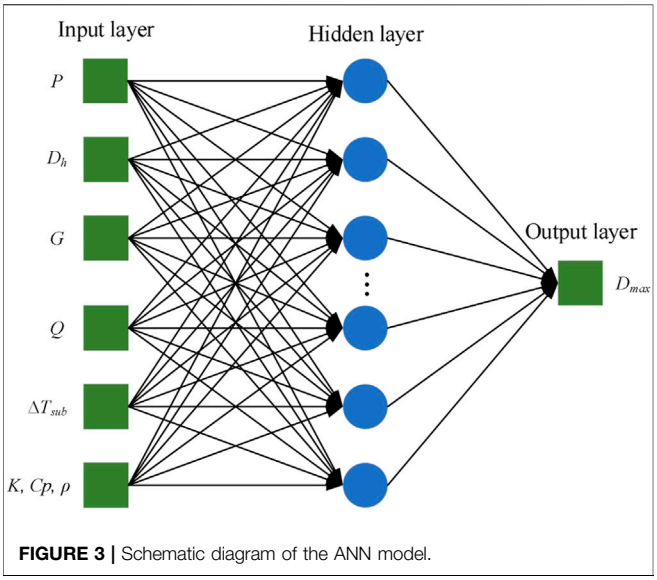


FIGURE 3 | Schematic diagram of the ANN model.

$$\begin{cases} A_{total} = \pi D^2 \\ A_{dryout} = \dfrac{\pi D_{dryout}^2}{4} \\ A_{tf} = L\pi D^2 - \dfrac{\pi D_{dryout}^2}{4} \\ A_{sl} = m\pi D^2 \\ A_c = n\pi D^2 \end{cases} \tag{8}$$

In **Equation 8**, L, m, and n refer to the fractions of surface area for thin liquid film, superheated liquid layer, and condensation. Under the assumption that the dry patch shrinks and disappears when the diameter of the bubble reaches the maximum value, the sum of L, m and n should be equal to 1.

By substituting **Eqs. 2–5**, **7**, and **8** into **Equation 1**, we obtain the basic equation for bubble diameter, as follows:

$$\frac{dD(t)}{dt} = \frac{2fk_l\Delta T_w}{\rho_g h_{lg}\sqrt{\pi\alpha_l t}}m + \frac{2k_l\Delta T_w\gamma}{\rho_g h_{lg}\sqrt{\pi\alpha_l t}}L - \frac{C\varphi D}{1-\rho_g/\rho_l}\Delta T_{sub}n \tag{9}$$

Using Ünal's method and the Taylor series (Ünal, 1976), the approximate solution of bubble maximum diameter can be written as follows:

$$\begin{cases} D_{max} = \dfrac{1.20724A}{\sqrt{B}} \\ t_{max} = \dfrac{1}{1.46B} \\ A = [fm + \gamma L]\dfrac{2k_l\Delta T_w}{\rho_g h_{lg}\sqrt{\pi\alpha_l}} \\ B = \dfrac{C\varphi n\Delta T_{sub}}{1-\rho_g/\rho_l} \end{cases} \tag{10}$$

where f is a coefficient of the characteristic temperature used to reflect an inhomogeneous of superheated liquid layer.

2.2 Artificial Neural Networks

In addition to the mechanistic model, an ANN model is proposed to predict the maximum diameter of a bubble using the BP (Back-Propagation) algorithm. The ANN model is a two-layer feed-forward network with sigmoid hidden neurons and linear output neurons. All of the calculations are carried out on the MATLAB platform.

As we can see from **Figure 3**, the ANN model consists of one input layer, one or more hidden layers and one output layer. Several parameters are chosen as input elements after analysis of the experimental conditions. These elements should reflect the effects of experimental conditions, such as pressure, mass flow, hydraulic diameter, heat flux, local subcooling, and the physical thermal properties of a heated surface.

For each neuron, an activation function is necessary to transforms the input value to the next hidden layer or the output layer. Commonly used activation functions include sigmoid/logistic, tansig, ReLU, and ELU. The main purpose of the activation function is to increase the non-linear ability of the ANN model.

The training algorithm is another important element of an ANN model. It is used to train the network and form a fixed model. However, it is very difficult to know which training algorithm will be most suitable for a given problem. This depends on many factors, including the complexity of the problem, the number of data points in the training set, and the number of weights and biases in the network. Different training algorithms should be compared to gain a good accuracy in the calculation.

The last key parameter is the coefficient of determination R, which represents the fitting degree between the experimental value and the calculated value. It varies from 0 to 1, with larger values being better. The function is expressed as follows:

$$R = \frac{\sum_{i=1}^{n}(y_i - \bar{y})\left(\left(\hat{y}_i - \bar{\hat{y}}\right)\right)}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2\sum_{i=1}^{n}\left(\hat{y}_i - \bar{\hat{y}}\right)^2}} \tag{11}$$

where n is the quantity of dataset, $\hat{y}_i$ represents the calculated value by ANN model, and y_i is the experimental value.

3 EXPERIMENTAL DATABASES

As **Table 1** shows, six experimental databases are used for the validation of ANN model. All of the experimental conditions are kept at low pressure and low velocity. The number of data points sum to 366, which is sufficient to execute an ANN.

The experimental data published by Situ et al. (2005), Brooks et al. (2015) Brooks and Hibiki (2015), Ahmadi et al. (2012), Okawa et al. (2007) and Xu et al. (2014) include bubble lift-off diameters. These datasets of the experiment are measured using a high-speed video camera. In the original heat balance theory, Ünal assumed that the bubbles would not leave the heated surface when they reached their maximum diameter. In this paper, bubble lift-off diameter is regarded as same as maximum bubble diameter.

TABLE 1 | Experimental databases for the validation.

Database	Prodanovic et al. (2002)	Situ et al. (2005)	Brooks et al. (2015), Brooks and Hibiki (2015)	Ahmadi et al. (2012)	Okawa et al. (2007)	Xu et al. (2014)
N	54	90	92	54	28	48
Fluid	Water	Water	Water	Water	Water	Water
Heated material	Stainless steel	Stainless steel	Stainless steel	Stainless steel	ITO film	Copper block
Geometry	Annulus	Annulus	Annulus	Rectangle	Rectangle	Rectangle
P (KPa)	105, 200, 300	101	150, 300, 450	98–860	121–125	101
D_h (m)	0.0093	0.019	0.019	0.01333	0.003	0.0032
q_w (KW/m^2)	100–1,200	60.7–206	100–492	81–611	67–549	26.3–215.4
V (m/s)	0.08–0.83	0.487–0.939	0.246–1.03	0.175–1.25	0.09–1.49	0.15–0.75
ΔT_{sub} (K)	10–60	1.38–19.88	5.4–39.8	4.0–29.7	9.2–20.8	6.6–27.4
D_b (mm)	0.37–3.24	0.145–0.605	0.046–0.338	0.02–3.90	0.50–3.02	0.091–0.245

Analyzing the original heat balance theory and using **Equation 3**, we can identify a power relation between the maximum bubble maximum diameter and the other four important parameters of the mechanistic model. The thermal properties of heated surface are not used as an impact parameter for the limited data of surface material. As can be seen in **Equation 5**, the direct influence of pressure P is worked through the parameter C. After integration, the order x of P should be in the range (-0.5, 0]. Considering that the vapor density is sensitive to the variation of pressure in the subcooled boiling flow, we also check the power relationship of ρ_g and the maximum bubble diameter. The pressure has a greater effect through changing the vapor density.

$$D_{max} \sim f\left(P^x, q^{1.0}, \Delta T_{sub}^{-0.5}, G^{-0.235}, \ \rho_g^{-1.0}, D_h^{-0.715}\right) \qquad (12)$$

In the original research, φ is only considered to be a function of velocity itself. However, we find that an equation incorporating the Reynolds number is more accurate for calculating the maximum diameter of the bubble in the former investigation (Dong and Zhang, 2021). The hydraulic diameter in the Reynolds number can reflect the effect of the channel dimension on bubble growth, which is more suitable for use in the calculation than velocity is.

4 VALIDATION AND DISCUSSION

In general, the key parameters of the ANN model include the number of neurons, number of hidden layers, activation function, and training algorithm. After a sensitivity analysis, we choose to use two hidden layers with eight neurons as the basic structure. The related R value can be increased to 0.9916, which is sufficiently accurate for the necessary calculations. In addition, the activation function and the training algorithm are also determined after comparing the accuracy of multiple combinations. The results show that the combination of logsig and tansig activation functions with the trainbr training algorithm gave the best performance, with a relative error of 13.50%. The trainbr algorithm, using Bayesian regularization back propagation, is a network training function that updates the weight and bias values according to Levenberg-Marquardt optimization. Trainbr can train any network, so long as its weight, net input, and transfer functions have derivative functions. It can minimize a linear combination of squared errors and weights and modify the linear combination so that at the end of training. the resulting network has good generalization qualities, even for different, small, or noisy datasets. Drawing on the above structure of the ANN model, more than 150 iterations of training are carried out for this prediction of maximum bubble diameter. The coefficient of determination R for nearly all the training sets is higher than 0.9, and 82% are higher than 0.95. This shows that the ANN model has a high confidence level, which increases the credibility of the calculated results.

4.1 Input Parameters in the Artificial Neural Networks Model

Appropriate input parameters are necessary structural elements in the ANN model. The parameters cannot be chosen arbitrarily and should reflect the effects of experimental conditions. For the prediction of maximum bubble maximum diameter in the subcooled boiling flow, experimental conditions, which include pressure, mass flow, hydraulic diameter, heat flux, local subcooling, and the physical and thermal properties of a heated surface should be paid more attention. Several dimensional parameters are chosen to test the relevance with bubble maximum diameter. The results are shown in **Table 2**. Multiple R and F refer to the significance level of relevance. It is concluded that dimensional dynamic viscosity μ_l/μ_v has little relevance to bubble maximum diameter, while the other parameters have a stronger or weaker relationship. However, some of the parameters have a direct connection with the others. ρ_l/ρ_v, Pe, Eo and S are chosen as the input parameters of ANN model. Considering the mechanism model and experimental conditions, subcooling ΔT_{sub} and heat flux q_w are added as a supplement.

4.2 Validation of Artificial Neural Networks Model

After determining the structural elements of ANN model, experimental databases are used to train the ANN model. For

TABLE 2 | Relevance analysis of potential input parameters.

Input Parameter	Description	Multiple R	Significance F
ρ_l/ρ_v	The ratio of liquid density and vapor density	0.299	5.71E-09
$Re = \frac{\rho_l u D_h}{\mu}$	Reynolds number used to describe flow status	0.404	7.95E-16
Pr	Prandtl number	0.316	5.97E-10
Pe = RePr	Peclet number which is used to describe the velocity ratio of convection and diffusion phenomenon	0.404	7.75E-16
$Eo = \frac{g(\rho_l-\rho_v)D_h^2}{\sigma}$	Eotvos number is used to describe the bubble shape in continuous liquid flow. It is the ratio of buoyancy force and surface tension	0.404	7.75E-16
$Mo = \frac{g(\rho_l-\rho_v)\mu_l^4}{\rho_l^2\sigma^3}$	Moton number works together with Eo to describe the bubble shape in continuous liquid flow	0.303	3.46E-09
μ_l/μ_v	Ratio of liquid dynamic viscosity and vapor dynamic viscosity	0.033	5.26E-01
$S = k_s\rho_s C_s$	Product of thermal conductivity, density, and heat capacity of solid heated surface	0.215	3.50E-05

TABLE 3 | Arithmetic-mean errors of different models for the databases.

Model	Prodanovic et al. (%)	Situ et al. (%)	Brooks et al. (%)	Ahmadi et al. (%)	Okawa et al. (%)	Xu Jianjun et al.
Ünal	30.75	164.71	248.54	240.67	34.16	—
Hoang	63.16	32.68	40.38	52.13	72.00	—
Levin	97.23	81.53	83.51	80.29	96.81	—
Previous mechanistic model (Dong and Zhang,2021)	42.40	31.32	35.66	36.90	32.52	—
ANN model	4.02	20.90	10.95	20.38	4.86	12.80%

Note: Because wall superheat data of Xu's experiment cannot be predicted by Chen's correlation, so this database is unable to calculate through mechanistic method. The superheat data is also a shortcoming of prediction and the main contribution of relative error in numerical calculation.

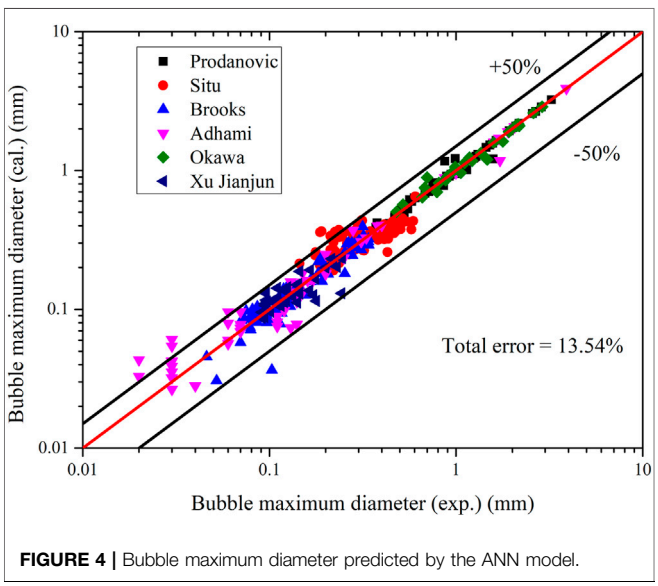


FIGURE 4 | Bubble maximum diameter predicted by the ANN model.

all 366 sets of bubble maximum diameter, 70% of databases are randomly selected as training data, and 15% are used as the test and validation respectively. The results are shown in **Table 3** and **Figure 4**; the average error of all the experimental databases is 13.54% which is much lower than for other mechanistic models. Furthermore, the ANN model increases the accuracy of calculations of large bubbles from Prodanovic's database.

However, due to the inner black box features of ANNs, the trained ANN model can only be used for prediction, with the experimental conditions covered by the trained database. Once one of the experimental parameters exceeds the range used by training, the calculated maximum bubble diameter could show a partial large relative error.

4.3 Uncertainty Analysis of Experimental Conditions

Considering installation error and measurement error, the uncertainty analysis of the experimental conditions is carried out in this section. The errors include the measurement accuracy of measuring instruments and the uncertainty of measuring position, while the other environmental influences are not considered. The errors of experimental conditions and their influence to input parameters are shown in **Table 4**. Because the maximum coarseness of the common pressure gauge reaches level 4, the relative error is 4%, so the measurement error for pressure is set to ±5% in this paper. Pressure fluctuation would change the density, Prandtl number, and surface tension. Dimensional density, *Pe*, and *Eo* are affected. In general, the average heat flux of the experiment is calculated based on the heating power and the scale of heated surface. The local heat flux of subcooled flow boiling can be measured with a non-contacting heat flow meter. Taking the measuring accuracy of the equipment into account, the average relative error of heat flux is set to ±5%. The flowmeter has a better accuracy the other instruments of measurement, that is, no more than 2% for various types. In addition, there is no position error in the measurement of mass flux. On the contrary, the main contribution of measurement error for local subcooling is the position error. In the heat balance theory, the local subcooling is used for the calculation of bubble condensation. It is therefore important to measure the subcooled

TABLE 4 | The potential measuring error for different experimental conditions.

Experimental Condition	Measurement Error	Affected Input Parameter	Relative Error
Pressure	±5%	ρ_l/ρ_v	±5%
		Pe	−0.7–0.25%
		Eo	−0.7–0.6%
Heat flux	±5%	q_w	±5.0%
Mass flux	±2%	*Pe*	±2.0%
Local subcooling	±3%	ΔT_{sub}	±3.0%
Hydraulic diameter	±3%	*Pe*	±3.0%
		Eo	±6.1%
Thermal properties of heated surface	±3%	*S*	±3.0%

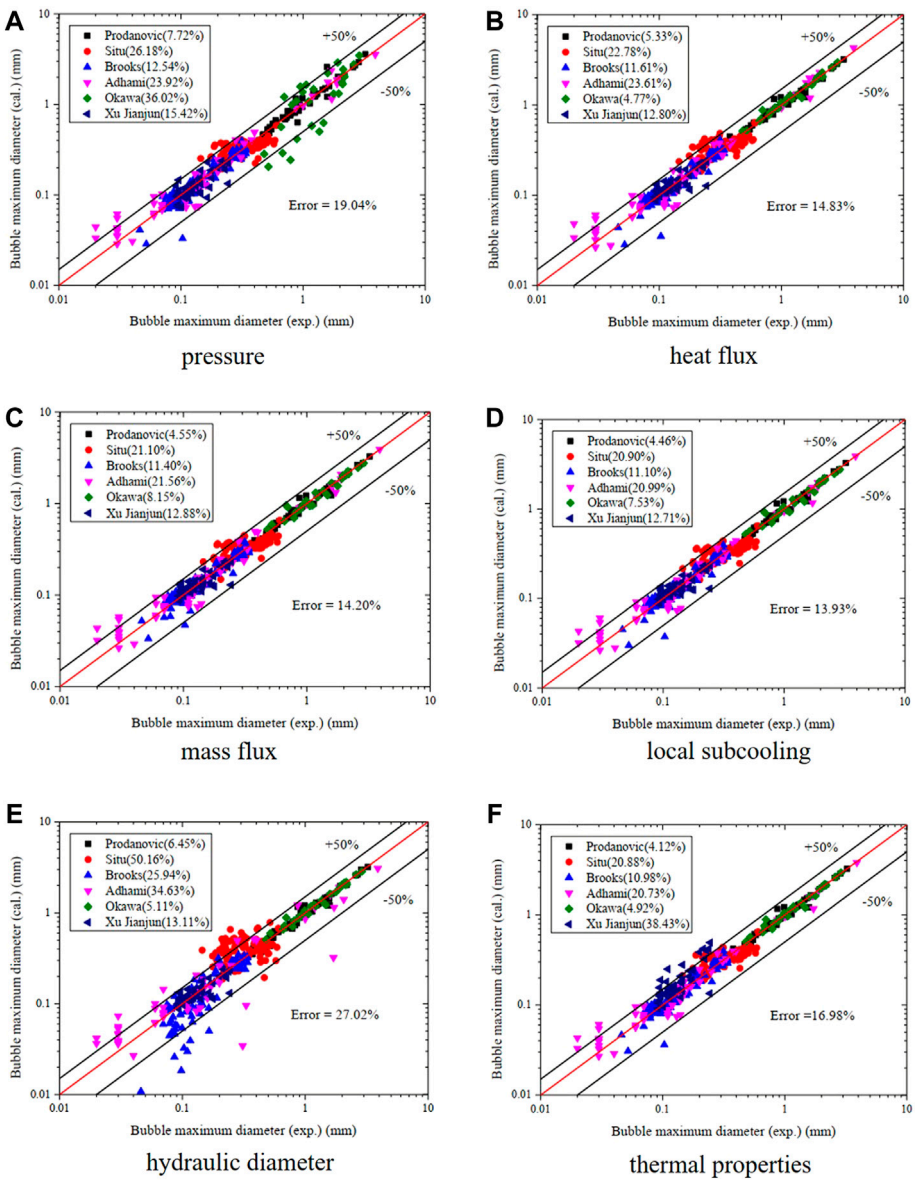


FIGURE 5 | Results of the ANN model under the fluctuations of experimental conditions. **(A)** Pressure, **(B)** heat flux, **(C)** mass flux, **(D)** local subcooling, **(E)** hydraulic diameter, **(F)** thermal properties.

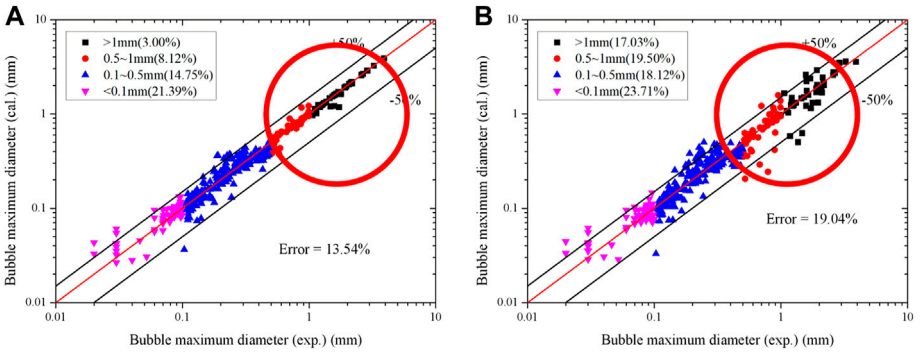


FIGURE 6 | The distributions of bubble maximum diameter for different ranges of size. **(A)** Excluding the measurement error of *P*, **(B)** including the measurement error of *P*.

temperature of the position that the bubble generates through non-contacting measurement instruments. Considering the potential position error of this parameter, the measuring error of local subcooling is set to ±3%. Apart from these four parameters, the hydraulic diameter and thermal properties of heated surface are also given ±3% as a measurement error in this sensitivity analysis. However, the trend of data deviation induced by these two reasons should be the same for everything in the same experimental facility.

From the given measurement error and position error, new values of experimental conditions are chosen according to a uniform distribution in the value range. Then the input parameters are calculated and are used to form a series of new matrix. The maximum bubble diameters shown in **Figure 5** are calculated by the trained ANN model based on the new databases of input parameters. As can be seen from the first four figures in **Figure 6**, only the fluctuations of pressure in four experimental conditions show appreciable effect on the bubble maximum diameter. The relative error of the total data set is enlarged from 13.54% to 19.04% when the pressure varies in a ±5% range. Therefore, pressure is considered to be the most sensitive parameter in the experimental conditions. This conclusion is consistent with the power relationship shown in **Equation 10**. It is also concluded that the ANN model has a good accuracy and robustness when facing the measurement error of experimental conditions. From the last two figures in **Figure 6**, the two experimental conditions related to experimental facility produce a large effect on the results, especially the hydraulic diameter. Furthermore, these two parameters remain the same through each experiment. Thus, it is necessary to determine the parameters related to the experimental section in a more accurate way before the experiment starts.

In the further analysis of pressure, the results are divided into four zones, according to the size of bubble maximum diameter. As can be seen in **Figure 6**, pressure has a larger influence on the zone where the bubble maximum diameter larger than 0.5 mm. The relative error increases about 12% while that of the other two zones varies only little. From a physical view, the larger bubble a larger amount of vapor inside. The fluctuation of pressure would influence larger bubble more through changing the vapor density. To increase the accuracy of the prediction, the correlation with pressure should be observed in both the experiments and in the mechanistic theoretical analysis.

4.4 New Correlation of Coefficients C and φ

The experimental condition pressure shows great sensitivity to the maximum bubble diameter, and its measurement should be prioritized. In the mechanistic model, we should also check the direct description of pressure, which is parameter C, used on the condensation part. In Ünal's original paper, the correlation (5) of C proposed by Ünal is fitted from no more than 30 experimental datasets, as well as parameter φ. This may have the largest contribution to the error of heat balance models. This error is especially remarkable for the large maximum bubble value under low-pressure conditions.

Parameter C is only related to pressure, and parameter φ is only related to liquid velocity or *Re* in the original heat balance model (Ünal, 1976) and its modification (Dong and Zhang, 2021). In this paper, the product of C and φ is regarded as one element $C\varphi$. Drawing on the conclusions of the sensitivity analysis of ANN model, further modifications are proposed for a mechanistic model that uses the input parameters ρ_l/ρ_v, *Pe*, and *Eo* instead of pressure and *Re* to form the new correlation of C and φ. After a statistical regression analysis, the new correlation of $C\varphi$ is shown below. It can be used directly in the coefficient B for **Equation 3**.

$$C\varphi = 2.454 \times 10^{-3} \left(\frac{\rho_l}{\rho_g}\right)^{-0.185} Pe^{0.890} Eo^{0.323} \tag{13}$$

The results shown in **Figure 7** are compared between the new correlation and original model (Dong and Zhang, 2021). The new correlation shows better performance for the bubble larger than 1 mm, in which the relative error is decreased from 51.39% to 33.19%. Other ranges of bubble size can also reach the same level or have greater accuracy.

Beyond the experimental data above, another calculation based on Kaiho et al.'s experiment (Kaiho et al., 2017) is carried out for an independent validation of the improved mechanistic model. The pressure range is 107–143 kPa, nearly

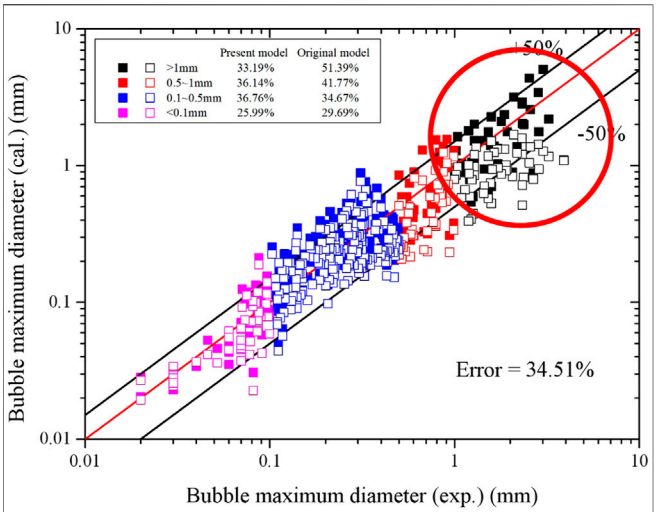


FIGURE 7 | Comparison of new correlation and original model for different bubble sizes.

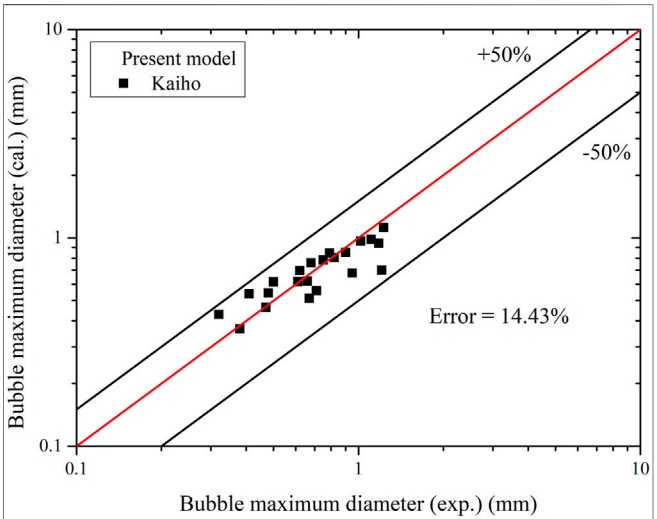


FIGURE 8 | Independent validation of present model for Kaiho's experiment.

atmospheric condition. The heat flux of the heated surface is around 175–617 kW/m^2 while the mass flux is 159–700 kg/m^{2}s. The subcooling of water is 10–30 K. The experimental results from Kaiho's research include the arithmetic mean value and the volume average value of maximum bubble diameter. In this validation, average volume is chosen as the experimental data, which is closer to the assumptions of heat balance theory. The calculated results are shown in **Figure 8**. The total relative error is 14.43%, which shows good performance for the present model.

5 CONCLUSION

Using ANNs, a data-driven model is proposed for the evaluation of bubble maximum diameter in subcooled boiling flow. After a basic sensitivity analysis is done on neuron number, two hidden layers with eight neurons each are used to develop the data-driven model. In addition, the activation function and training algorithm are screened out for the ANN model. Through the training using several experimental databases, the data-driven model shows good performance, with a relative error of around 14%. Sensitivity analysis is also proposed for the four experimental conditions and two structural conditions. The results identify the accuracy and robustness of the ANN model. It is also concluded that the measuring accuracy of pressure is of the most sensitivity on the bubble maximum diameter in the subcooled boiling flow under low-pressure conditions, especially for bubble sizes larger than 0.5 mm. A regression analysis of parameters C and φ, a new correlation is developed for the mechanistic model. This new model functions well for all the experimental databases and the large bubble datasets. Another independent validation also proves the accuracy of the improved mechanistic model. To sum up, the modified mechanistic model covers a wide range of subcooled boiling flow under low-pressure conditions. In the next step, additional attention will be paid to increasing the generalization performance of ANN model for larger experimental conditions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, and further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

XD proposed the research algorithm for the data-driven model and wrote the original draft. HC used MATLAB to implement the research method. CL discussed the accuracy and its modification of ANN model. MY helped develop the structural analysis of ANN model. YY collected and analyzed the experimental databases. XH helped develop the modification of mechanistic model and the additional validation.

FUNDING

This research was funded by the Guang Dong Basic and Applied Basic Research Foundation (No. 2021A1515110922). Thanks to Shenzhen University and Nuclear Power Institute of China for the support.

REFERENCES

Aarabi Jeshvaghani, P., Khorsandi, M., and Feghhi, S. A. H. (2021). Flow-Rate Prediction Independent of the Regime in a Dynamic Two-Phase Flow System Using a Simple Pulse Height Spectrum of a Detector and Artificial Neural Networks. *Nucl. Instrum. Meth. A* 1017, 165794. doi:10.1016/j.nima.2021.165794

Affonso, R. R. W., Dam, R. S. F., Salgado, W. L., Silva, A. X. d., and Salgado, C. M. (2020). Flow Regime and Volume Fraction Identification Using Nuclear Techniques, Artificial Neural Networks and Computational Fluid Dynamics. *Appl. Radiat. Isot.* 159, 109103. doi:10.1016/j.apradiso.2020.109103

Ahmadi, R., Ueno, T., and Okawa, T. (2012). Bubble Dynamics at Boiling Incipience in Subcooled Upward Flow Boiling. *Int. J. Heat. Mass Transf.* 55, 488–497. doi:10.1016/j.ijheatmasstransfer.2011.09.050

Brooks, C. S., and Hibiki, T. (2015). Wall Nucleation Modeling in Subcooled Boiling Flow. *Int. J. Heat. Mass Transf.* 86, 183–196. doi:10.1016/j.ijheatmasstransfer.2015.03.005

Brooks, C. S., Silin, N., Hibiki, T., and Ishii, M. (2015). Experimental Investigation Ofwall Nucleation Characteristics in Flow Boiling. *J. Heat. Transf.* 137, 1–9. doi:10.1115/1.4029593

Cheung, S. C. P., Vahaji, S., Yeoh, G. H., and Tu, J. Y. (2014). Modeling Subcooled Flow Boiling in Vertical Channels at Low Pressures - Part 1: Assessment of Empirical Correlations. *Int. J. Heat. Mass Transf.* 75, 736–753. doi:10.1016/j.ijheatmasstransfer.2014.03.016

Dong, X., and Zhang, Z. (2021). Mechanism Study of Bubble Maximum Diameter in the Subcooled Boiling Flow for Low-Pressure Condition. *Int. J. Heat. Mass Transf.* 164, 120585. doi:10.1016/j.ijheatmasstransfer.2020.120585

Greenwood, M. S., Duarte, J. P., and Corradini, M. (2017). Presentation and Comparison of Experimental Critical Heat Flux Data at Conditions Prototypical of Light Water Small Modular Reactors. *Nucl. Eng. Des.* 317, 220–231. doi:10.1016/j.nucengdes.2016.12.030

Gu, J., Wang, Q., Wu, Y., Lyu, J., Li, S., and Yao, W. (2017). Modeling of Subcooled Boiling by Extending the RPI Wall Boiling Model to Ultra-High Pressure Conditions. *Appl. Therm. Eng.* 124, 571–584. doi:10.1016/j.applthermaleng.2017.06.017

Hoang, N. H., Chu, I.-C., Euh, D.-J., and Song, C.-H. (2016). A Mechanistic Model for Predicting the Maximum Diameter of Vapor Bubbles in a Subcooled Boiling Flow. *Int. J. Heat. Mass Transf.* 94, 174–179. doi:10.1016/j.ijheatmasstransfer.2015.11.051

Jung, H., Yoon, S., Kim, Y., Lee, J. H., Park, H., Kim, D., et al. (2020). Development and Evaluation of Data-Driven Modeling for Bubble Size in Turbulent Air-Water Bubbly Flows Using Artificial Multi-Layer Neural Networks. *Chem. Eng. Sci.* 213, 115357. doi:10.1016/j.ces.2019.115357

Kaiho, K., Okawa, T., and Enoki, K. (2017). Measurement of the Maximum Bubble Size Distribution in Water Subcooled Flow Boiling at Low Pressure. *Int. J. Heat. Mass Transf.* 108, 2365–2380. doi:10.1016/j.ijheatmasstransfer.2017.01.027

Klausner, J. F., Mei, R., Bernhard, D. M., and Zeng, L. Z. (1993). Vapor Bubble Departure in Forced Convection Boiling. *Int. J. Heat. Mass Transf.* 36 (3), 651–662. doi:10.1016/0017-9310(93)80041-r

Krepper, E., Rzehak, R., Lifante, C., and Frank, T. (2013). CFD for Subcooled Flow Boiling: Coupling Wall Boiling and Population Balance Models. *Nucl. Eng. Des.* 255, 330–346. doi:10.1016/j.nucengdes.2012.11.010

Kurul, N., and Podowski, M. Z. (1991). "On the Modeling of Multidimensional Effects in Boiling Channels," in Proceedings of the 27th National Heat Transfer Conference, Minneapolis, USA 1991, 28–31.

Levenspiel, O. (1959). Collapse of Steam Bubbles in Water. *Ind. Eng. Chem.* 51, 787–790. doi:10.1021/ie50594a045

Murallidharan, J. S., Prasad, B. V. S. S. S., and Patnaik, B. S. V. (2018). A Universal Wall-Bubble Growth Model for Water in Component-Scale High-Pressure Boiling Systems. *Int. J. Heat. Mass Transf.* 122, 161–181. doi:10.1016/j.ijheatmasstransfer.2018.01.070

Okawa, T., Kubota, H., and Ishida, T. (2007). Simultaneous Measurement of Void Fraction and Fundamental Bubble Parameters in Subcooled Flow Boiling. *Nucl. Eng. Des.* 237, 1016–1024. doi:10.1016/j.nucengdes.2006.12.010

Prodanovic, V., Fraser, D., and Salcudean, M. (2002). Bubble Behavior in Subcooled Flow Boiling of Water at Low Pressures and Low Flow Rates. *Int. J. Multiph. Flow.* 28, 1–19. doi:10.1016/s0301-9322(01)00058-1

Salgado, C. M., Pereira, C. M. N. A., Schirru, R., and Brandão, L. E. B. (2010). Flow Regime Identification and Volume Fraction Prediction in Multiphase Flows by Means of Gamma-Ray Attenuation and Artificial Neural Networks. *Prog. Nucl. Energ* 52 (6), 555–562. doi:10.1016/j.pnucene.2010.02.001

Sernas, V., and Hooper, F. C. (1969). The Initial Vapor Bubble Growth on a Heated Wall during Nucleate Boiling. *Int. J. Heat. Mass Transf.* 12 (12), 1627–1639. doi:10.1016/0017-9310(69)90097-0

Situ, R., Hibiki, T., and Ishii, M. M. (2005). Bubble Lift-Off Size in Forced Convective Subcooled Boiling Flow. *Int. J. Heat. Mass Transf.* 48 (25–26), 5536–5548. doi:10.1016/j.ijheatmasstransfer.2005.06.031

Tu, J. Y., and Yeoh, G. H. (2002). On Numerical Modelling of Low-Pressure Subcooled Boiling Flows. *Int. J. Heat. Mass Transf.* 45 (6), 1197–1209. doi:10.1016/s0017-9310(01)00230-7

Ünal, H. C. (1976). Maximum Bubble Diameter, Maximum Bubble-Growth Time and Bubble-Growth Rate during the Subcooled Nucleate Flow Boiling of Water up to 17.7 MN/m2. *Int. J. Heat. Mass Transf.* 19, 643–649. doi:10.1016/0017-9310(76)90047-8

Xu, J., Chen, B., and Xie, T. (2014). Experimental and Theoretical Analysis of Bubble Departure Behavior in Narrow Rectangular Channel. *Prog. Nucl. Energ* 77, 1–10. doi:10.1016/j.pnucene.2014.06.002

Yan, M., Ma, Z., Pan, L., Liu, W., He, Q., Zhang, R., et al. (2021). An Evaluation of Critical Heat Flux Prediction Methods for the Upward Flow in a Vertical Narrow Rectangular Channel. *Prog. Nucl. Energ* 140, 103901. doi:10.1016/j.pnucene.2021.103901

Yeoh, G. H., Cheung, S. C. P., Tu, J. Y., and Ho, M. K. M. (2008). Fundamental Consideration of Wall Heat Partition of Vertical Subcooled Boiling Flows. *Int. J. Heat. Mass Transf.* 51, 3840–3853. doi:10.1016/j.ijheatmasstransfer.2007.11.047

Yoo, J., Estrada-Perez, C. E., and Hassan, Y. A. (2018). Development of a Mechanistic Model for Sliding Bubbles Growth Prediction in Subcooled Boiling Flow. *Appl. Therm. Eng.* 138, 657–667. doi:10.1016/j.applthermaleng.2018.04.096

Zuber, N. (1961). The Dynamics of Vapor Bubbles in Nonuniform Temperature Fields. *Int. J. Heat. Mass Transf.* 2, 83–98. doi:10.1016/0017-9310(61)90016-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer ZL declared a shared affiliation with the author(s) YY to the handling editor at the time of review.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership