Radial Power Distribution Reconstruction Method via Ex-core-detector Signals Based on Artificial Neural Networks

Xingfang Wang, Youqi Zheng*, Xiayu Wang, Bowen Xiao
Xi'an Jiaotong University, No. 28, Xianning West Road, Xi'an, Shaanxi 710049, China

Keywords: power distribution reconstruction, ex-core neutron detector, microreactor

1. Introduction

In the context of advancing toward a decarbonized society, nuclear energy assumes a pivotal role, offering an ecologically sustainable source of electricity. Small Modular Reactors (SMRs) are known for their exceptional flexibility, sustainability, safety, and cost-effectiveness. Consequently, SMRs have gained considerable acclaim as a viable solution to meeting various energy demands. Yet, there is a critical challenge in ensuring the safety and efficient operation of SMRs, which lies in monitoring core power distribution.

Currently, with developed methods and application cases, monitoring system based on in-core detectors such as BEACON [1] have demonstrated their capability to light water reactors. Nevertheless, the limited lifespan of these in-core detectors poses challenges for achieving online monitoring. Additionally, it is impractical to make the deployment of in-core detectors in SMRs due to the prevalent compact spatial arrangements and harsh environments.

In response to these exigencies and challenges, scholarly inquiries have been directed towards the prospect of reconstructing power distribution via ex-core detectors. However, as shown by the work of F Li [2], X Peng [3] and their contemporaries, predicting radial or three-dimensional power distribution of reactors like Pressurized Water Reactors (PWRs) and High-Temperature Gas-cooled Reactors (HGTRs) through ex-core neutron detectors did not get satisfactory results.

Nevertheless, SMRs exhibit unique characteristics compared to larger reactors, such as smaller radial sizes and shorter distances between ex-core detectors and fuel elements, as well as higher neutron energy resulting in longer neutron transport mean free paths and increased neutron leakage. These factors offer the potential for ex-core detectors to effectively reflect power distribution changes in fuel elements.

The process of predicting the fuel power distribution matrix from the signal matrix of ex-core detectors can be considered an inverse problem, thus, it is too elusive for deterministic methods to solve their mapping relationship. As an alternative, neural networks have gained prominence for their ability to capture complex non-linear relationships and adaptively learn from massive amounts of data that contain the corresponding relationships. Based on reactor power variation characteristics and ex-core detector spatial response function characteristics, this research aims to optimize the implementation of neural network methods during the reconstruction of radial power distribution from ex-core detector signals.

2. Methods and Results

2.1 Artificial Neural Network Model

In this study, the dimension of the input signal matrix (number of detectors) is significantly smaller than the dimension of the output matrix (number of fuel regions), hence a suitable model choice could indeed be a fully connected deep learning neural network. A fully connected neural network consists of three main parts: the input layer, hidden layers, and the output layer, as shown in Fig. 1. When the hidden layers have a large number of layers, the neural network is referred to as a deep neural network. Each hidden layer contains several neurons, and the neurons within the same layer are not connected. Each neuron in the subsequent layer is connected to all the neurons in the previous layer.

$$a_j^{(k)} = \sum_{i=1}^{m_{k-1}} \omega_{ji}^{(k)} z_i^{(k-1)} + b_i^{(k)} \tag{1}$$

Once the weighted signals from the previous layer are received by neuron i in the hidden layer k, this neuron processes the incoming signals through its activation function, transforming them into the output signal transmitted to the hidden layer k+1.

$$z_i^{(k)} = f\left(a_i^{(k)}\right) \tag{2}$$

ReLU is currently the most commonly used activation function in deep neural networks. It outputs the input value when it is positive and zero when it is negative. Its simplicity and efficiency have made it the preferred choice for deep neural networks.

$$f(x) = \max(0, x) \tag{3}$$

The learning process of a neural network (as shown in Fig. 2) involves gradually approximating the real input-output relationship. It entails adjusting the network's weight parameters step by step using backpropagation and gradient descent along the direction of the loss function to minimize it. The objective is to make the neural network model closely approximate the true

input-output relationship, allowing it to improve its performance through iterative adjustments.

$$\boldsymbol{\omega}^{(n+1)} = \boldsymbol{\omega}^{(n)} - \alpha \frac{1}{m} \frac{\partial E}{\partial \boldsymbol{\omega}^{(n)}} \Big|_{\boldsymbol{\omega}^{(n)}, \boldsymbol{b}^{(n)}} \tag{4}$$

$$\boldsymbol{b}^{(n+1)} = \boldsymbol{b}^{(n)} - \alpha \frac{1}{m} \frac{\partial E}{\partial \boldsymbol{b}} \Big|_{\boldsymbol{\omega}^{(n)}, \boldsymbol{b}^{(n)}} \tag{5}$$

in which, $\boldsymbol{\omega}, \boldsymbol{b}$ is weights and biases to be updated in each hidden layer of network, n *is* the number of iterations, α is learning rate, E *is* the loss function of network.

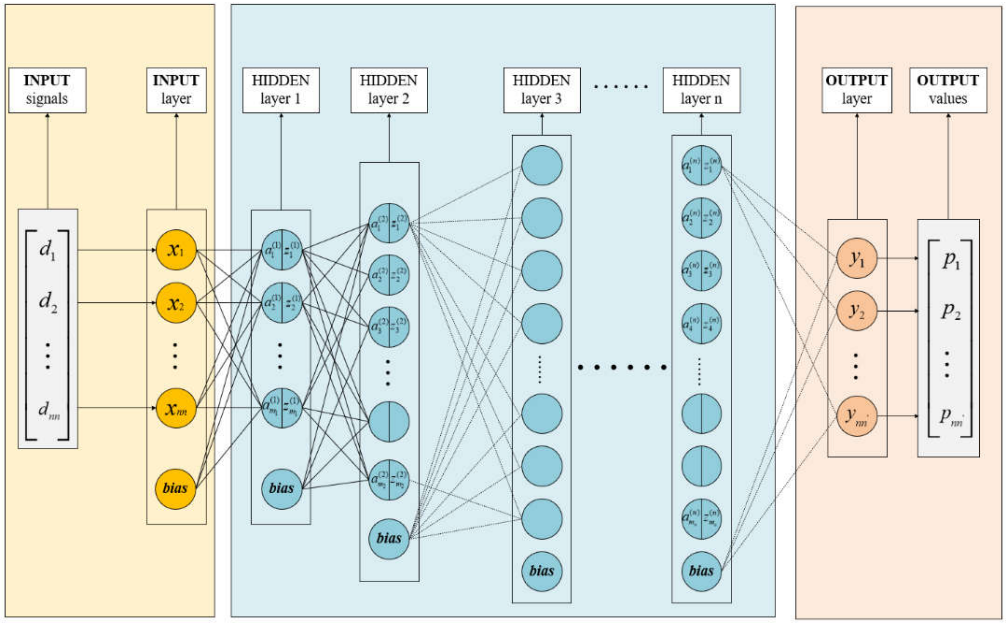

Fig. 1. Artificial neural network model diagram

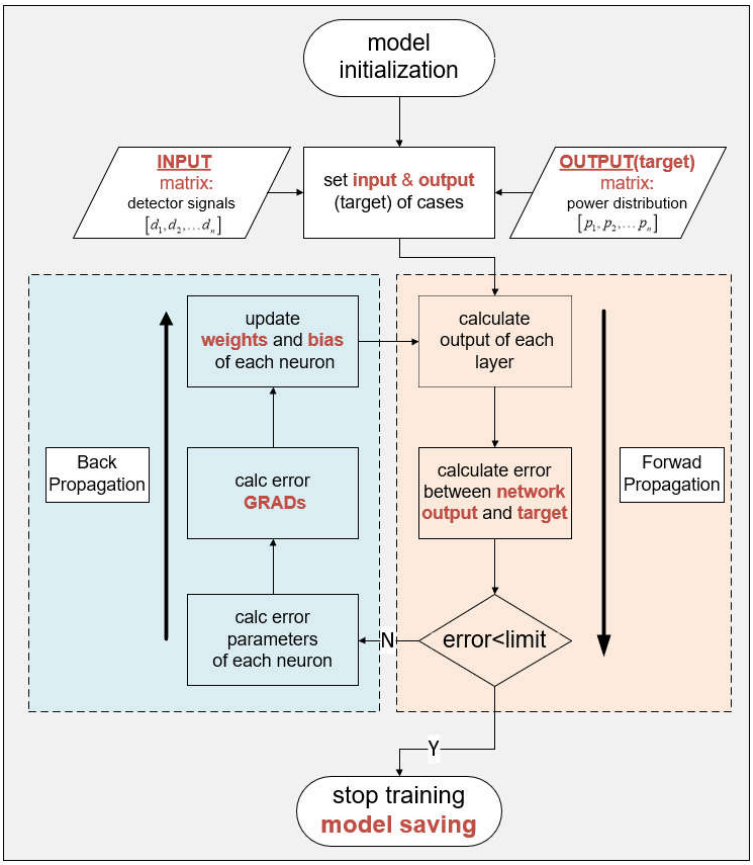

Fig. 2. Flow chart of machine learning iterative computation

2.3 Reactor Core Calculation

A reactor core model simplified from the TOPAZ-II reactor was investigated. The TOPAZ-II reactor is a thermal ion space reactor system, where the in-core thermal ion fuel elements can generate electric power ranging from 4.5 to 5.5 kW. The reactor comprises 37 fuel elements, with ZrH serving as the moderator material and Be as the reflector material. As shown in Fig.3, thermal power regulation, reactivity compensation, and emergency shutdown functions are accomplished by rotating 12 movable beryllium cylinders with boron carbide-shielded fan-shaped blades. These cylinders are installed within the radial reflector. They are divided into two groups: three safety drums and nine control drums. Due to the compact structure and limited number of fuel components, it is an ideal candidate for ex-core power monitoring research. Table I provides a detailed description of this reactor.

The SARAX code system [5-7] was used in the reactor core calculation, which was developed for the advanced reactor analysis at Xi'an Jiaotong University. As shown in Fig. 4, is capable of modeling external detectors [8]; accordingly, the SARAX program could simultaneously solve the power distribution of fuel elements and the neutron flux at various ex-detector locations during the progress of core eigenvalue transport calculation.

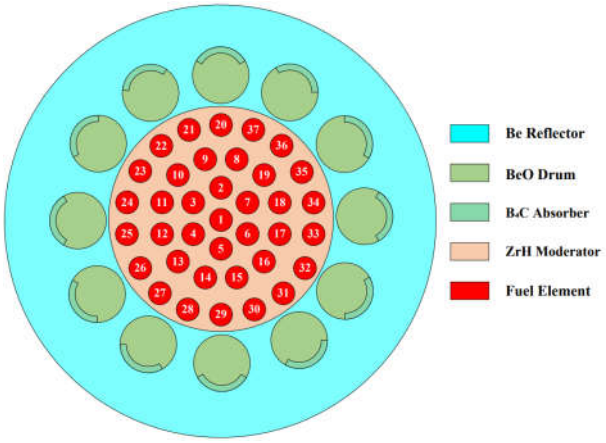

Fig. 3. TOPAZ-II radial layout

Table I: TOPAZ-II Parameters

Parameters	Value
Thermal power (kW_{th})	125
Electrical power (kW_e)	5
Reactor Neutron spectrum	Epithermal
Reactor Fuel	UO2
Fuel enrichment	96%
Number of fuel elements	37
Reactor core diameter (mm)	260
Reactor core height (mm)	375
Coolant type	NaK-78
Reactor control	9 control drums and 3 safety drums
Moderator	$ZrH_{1.85}$
Reflector	Beryllium
Movable absorber	Boron-Carbide
Ex-detector material	BF3
Number of ex-detectors	12

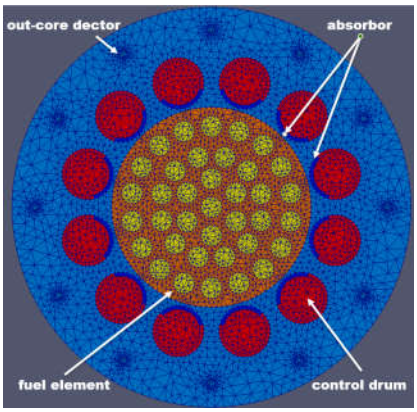

Fig. 4. SARAX simulation computation model for simplified TOPAZ-II

As shown in Fig. 5, while control drums rotate to different angles, the radial relative power distribution of fuel elements changes dramatically due to the neutron absorber.

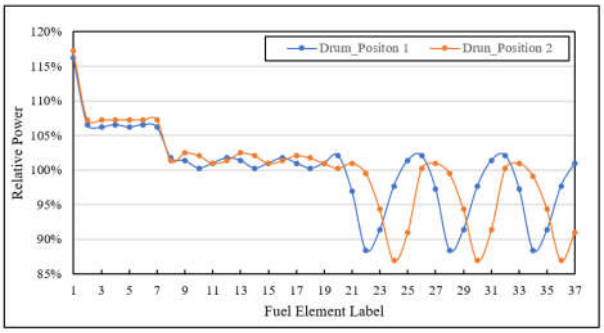

Fig. 5. Relative power distribution under various drum positions

Meanwhile, the neutron flux at the external neutron detector also changes. The relative neutron flux level is lower in the ex-core, it can be assumed that the nuclides density and reaction cross-section of the detector remain constant during the monitoring system operation period. The output electric current value I of the detector can be considered proportional to the nuclear reaction rate R at the ex-core detector location. During practical engineering conditions, the reactor control system uses the calibrated current signals from various detectors to monitor core power. However, in numerical simulation, it is hard to directly use the measured values of the external current signals. Instead, the nuclear reaction rates at each external detector obtained from the core physics simulation program can be used as an approximation to the required current signals from each detector.

In an attempt to use the neural network method simulating the mapping correlation between the detector signal matrix and the radial power distribution of fuel elements, it is essential to acquire an extensive dataset encompassing diverse drum positions.

To simulate the practical operational processes of the reactor and generate datasets, the safety drums are fully extended, and the control drums are grouped into symmetrical positions, rotating together in sets of three. This entails filling a substantial amount of randomly generated drum position data into the input cards of the SARAX. Upon completing the transport calculation of feature values, this integration enables the simultaneous acquisition of accurate power distribution solutions for all fuel elements and response signals from ex-core detectors, which are represented by the absorption reaction rates.

The neural network, in its learning process, establishes a mapping relationship between the detector response signal matrix and the core power distribution obtained from program calculations. This equips the network with the capability to predict reactor power with precision.

When the mapping relationship features in the training set samples are insufficient compared to the intricate feature conditions inherent in the problem, the neural network encounters difficulty learning features that don't exist within the training set. As a result, maintaining robust performance across diverse test sets becomes a challenge. Consequently, creating a dataset tailored for this specific issue necessitates generating a substantial array of randomly varied drum positions. This essential requirement further highlights the necessity for significant parallel computing resources to concurrently manage the resulting computations.

Using the BSCC high-performance supercomputing platform, the SARAX program undertook parallel computations encompassing 14,500 distinct operational conditions. Subsequent calculations yielded B-10 absorption reaction rate matrices at the detectors and radial power distributions for each of these scenarios.

To facilitate real-time model monitoring during training and to conduct comprehensive closed testing post-training, it becomes crucial to segment these data samples into three distinct subsets: the training set, validation set, and testing set. The training set is employed to update model parameters through the gradient descent algorithm. The validation set monitors potential overfitting and enables early stopping strategies. Meanwhile, the testing set remains independent from the training process, and is dedicated solely to the final evaluation of network performance.

Considering the available information, which includes responses solely from 12 detectors, resulting in each data sample featuring only 12 input attributes, a deep neural network model is employed.

Nonetheless, the neural network model still exhibits considerable deviations and inaccuracies in predicting the power distribution for the entire array of 37 fuel elements, as opposed to the SARAX calculating result.

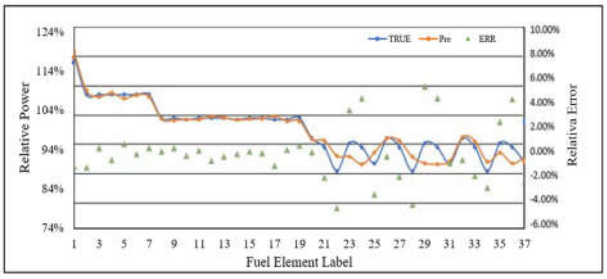

Fig. 6. Optimal prediction result of relative power distribution

The optimal result of predicting the power distribution for the entire array of 37 fuel elements is shown in Fig.6. The maximum relative error could reach up to approximately 6%. As depicted by the blue and orange lines, there is an evident distortion in the power distribution outcome. The results lead to a discernible realization that the performance of the artificial neural network model, in terms of its direct prediction of the power distribution across the entire set of 37 fuel elements, falls short of satisfactory.

This situation can be attributed to the direct influence of neutron absorbers positioned along the outermost periphery of the fuel elements, consequently causing notable power distribution fluctuations of outermost fuel elements. It is, therefore, imperative to optimize the structure of neural networks.

Under various drum positions within the simplified TOPAZ-II reactor, the most significant relative power fluctuations are observed exclusively within the outermost ring of fuel elements. Hence, this issue can be divided into two parts:

Problem 1, prediction of the relative power distribution for outer ring fuel elements (i.e., elements labeled from 20 to 37).

Problem 2, prediction of the relative power distribution for inner ring fuel elements (i.e., elements labeled from 1 to 19).

As random drum positions are set, symmetric control is followed, with the drums rotating synchronously at the same angle. Consequently, the relative power of the three fuel elements in symmetric positions is closely comparable, allowing for an approximation of equality. Thus, the prediction of the relative power distribution for outer ring fuel elements can be simplified to the forecast of relative power for six non-adjacent fuel elements.

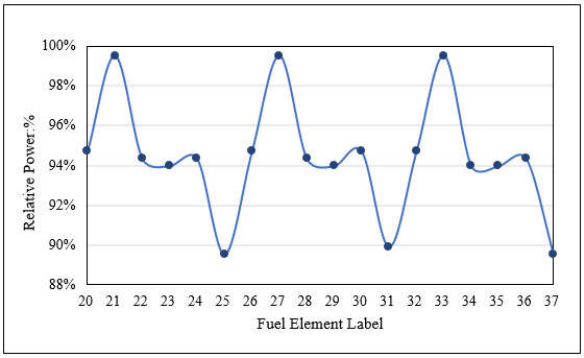

Fig. 7. Relative power distribution of outermost ring of elements

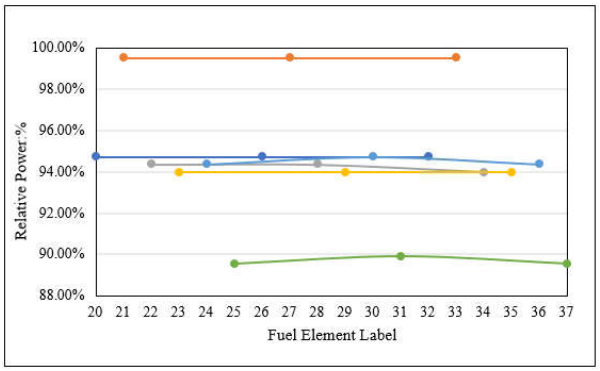

Fig. 8. Relative power distribution of outermost ring of elements

As depicted in Fig. 7 and Fig. 8, in contrast to the significant fluctuations in relative power observed among the non-symmetrically positioned fuel elements in the outermost ring (with relative power variations approaching nearly 10%), the relative power values for the three symmetrically positioned elements can be reasonably approximated as consistent (with variations primarily within 1%).

2.4 Model Construction

Regarding problem 1, the network input layer comprises 12 nodes, while the output layer comprises 6 nodes. To mitigate the risk of overfitting, an architecture of 8 hidden layers has been established, each containing a varying number of neurons: 512, 1024, 2048, 1024, 512, 128, 8, and 6. Appendix A and B respectively present the detailed network parameters provided by Keras. Following min-max normalization and data type conversion, the samples are fed into the deep neural network. The chosen activation function for the network is the Rectified Linear Unit (ReLU) function [9], and the regularization method employed is BatchNormalization [10]. The network optimization is carried out using the Adam parameter update algorithm. The objective function is Root Mean Squared Error (RMSE), while the performance metric is evaluated using Mean Absolute Error (MAE). The expressions of RMSE and MAE are as follows:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \tag{6}$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|\, y_i - \hat{y}_i \,| \tag{7}$$

in which, m is the number of fuel elements, y_i - relative power values calculated by SARAX, $\hat{y}_i$ -- relative power values predicted by the neural network. During the training process of problem 1, the training dataset was subjected to 769 epochs of training cause that the loss function of the validation set did not show a continuous decrease for 20 consecutive rounds. As a preventive measure against overfitting, the training iteration was automatically terminated by implementing EarlyStopping [11]. In the final epoch, the parameters are presented in the Table II:

Table II: Training and validation parameters of the last epoch

	RMSE	MAE
Training Set	0.0278	0.0234
Valadition Set	0.0350	0.0297

The training process for Problem 2 follows a similar pattern and does not require further elaboration. The progression of training and validation set loss functions, as well as the monitoring metrics, during training, is illustrated in Fig.9 and Fig.10.

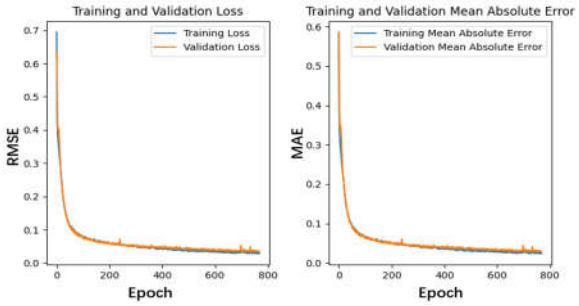

Fig. 9. Parameters of the training and validation set in Problem 1

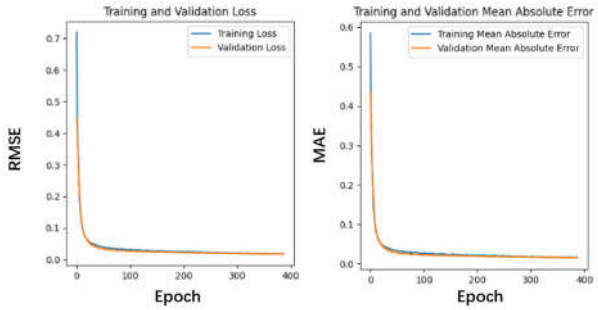

Fig. 10. Parameters of the training and validation set in Problem 2

As the number of training steps increases, both the training set and validation set loss function values initially exhibit a rapid decrease, followed by a gradual decline toward stability. Throughout the training process, the metrics of the training set and validation set change in synchronization, without any notable occurrence where the validation set metrics significantly surpass the training set metrics. This serves as evidence of the network model's strong generalization capability. After the completion of training, all weight parameters and biases of the model are saved by the network for preparation for performance testing on a closed test set in the future.

2.5 Result

The expression of mean relative error (MRE) is as follows:

$$MRE = \frac{1}{37}\sum_{i}^{37}\left(\frac{y_i - \hat{y}_i}{y_i}\right) \quad (8)$$

Based on 500 testing samples, the distribution of MRE is shown in the Fig. 11.

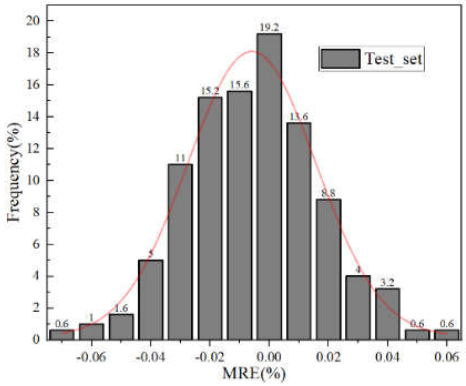

Fig. 11. Frequency distribution map of test set MRE

The frequency of data samples of different MRE in the prediction results is provided. It can be observed that MRE of the network's predictions for the test dataset input scenarios is predominantly concentrated within the range of [-0.04-0.04%]. The results of relative errors approximately conform to a normal distribution, with an expected value of μ = -5.92E-5 and a standard deviation of σ = 4.34E-4. The maximum MRE is below 0.0692%.

Subsequently, in the test set, the mean relative errors for each fuel element are taken as the reference sample error column, denoted as RefSeq. Following this, the reference error column is compared with the errors of the result having the maximum MRE (MAX), the result with an approximate average MRE (AVE), and the result with a minimum MRE (MIN). Fig.12-Fig.17 illustrates these three groups of data.

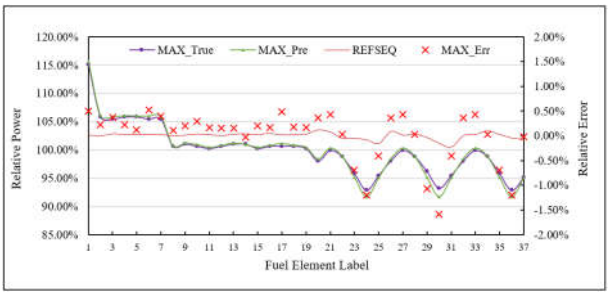

Fig. 12. Predicted and calculated relative power distribution of the maximum MRE group

As shown in Fig.12, among the 37 fuel elements of the maximum MRE group, the maximum positive error of relative power is 0.51%, and the minimum negative error is -1.60%. There is no evident distortion in the relative power prediction results for all fuel elements. As shown in Fig.13, the power distribution shape of the prediction outcomes closely aligns with the calculated result, effectively fulfilling the requirements for core power monitoring.

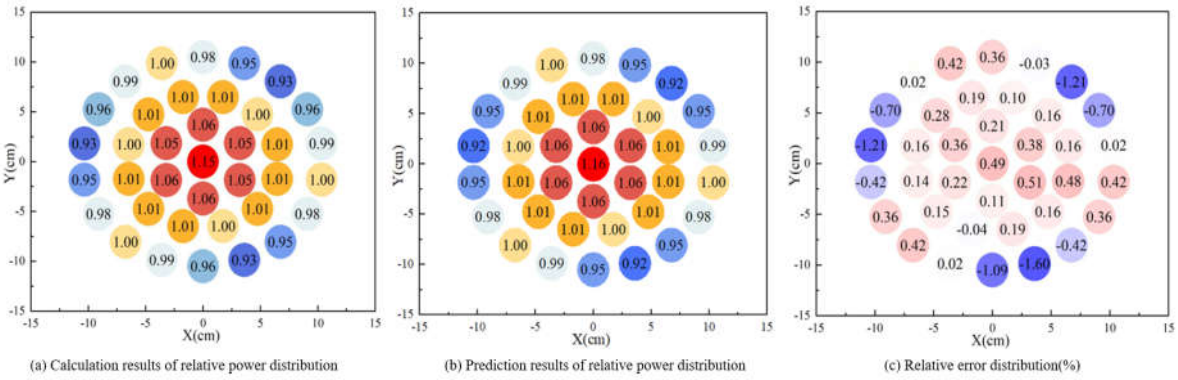

Fig. 13. Relative power distribution and relative error of the maximum MRE group

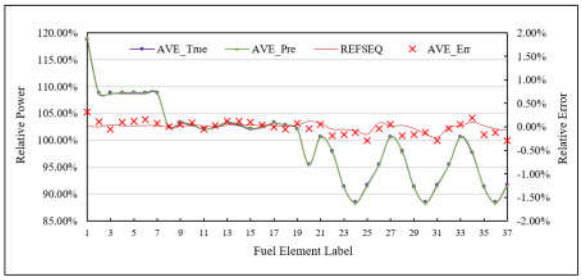

Fig. 14. Predicted and calculated relative power distribution of the average MRE group

As shown in Fig.14, among the 37 fuel elements of the average MRE group, the maximum positive error of relative power is 0.32%, and the minimum negative error is -0.28%. There is no evident distortion in the relative power prediction results for all fuel elements. As shown in Fig.15, the power distribution shape of the prediction outcomes closely aligns with the calculated result, reaffirming the model's robust predictive capacity.

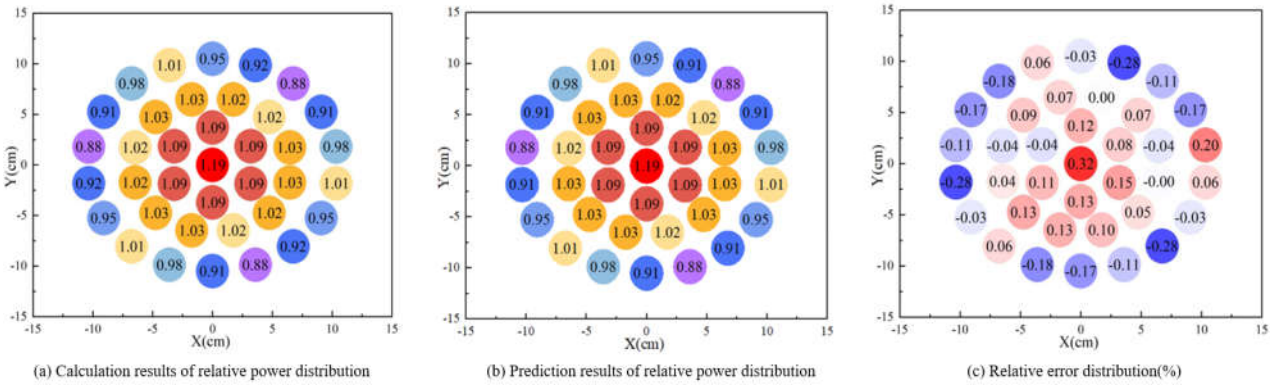

Fig. 15. Relative power distribution and relative error of the average MRE group

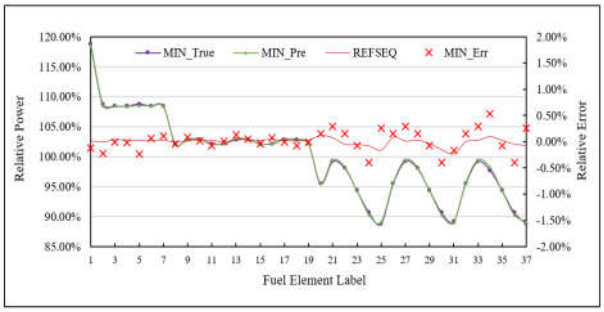

Fig. 16. Predicted and calculated relative power distribution of the minimum MRE group

As shown in Fig.16, among the 37 fuel elements of the minimum MRE group, the maximum positive error of relative power is 0.51%, and the minimum negative error is -0.42%. There is no evident distortion in the relative power prediction results for all fuel elements. As shown in Fig.17, the power distribution shape of the prediction outcomes closely aligns with the calculated result, effectively fulfilling the requirements for core power monitoring.

In conclusion, these 3 group of results highlight the model's noteworthy accuracy and capability in characterizing power distribution dynamics among diverse fuel elements.

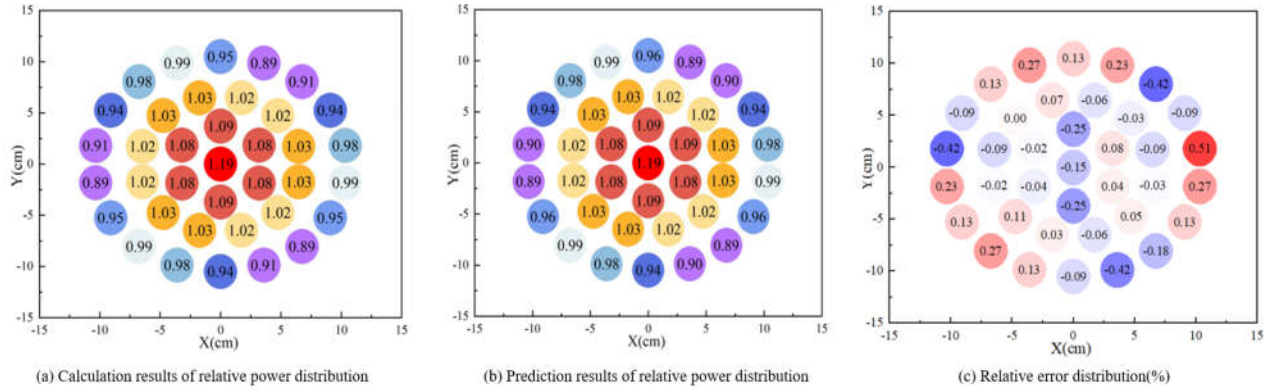

Fig. 17. Relative power distribution and relative error of the minimum MRE group

To simulate realistic current signal fluctuations from external detectors, a 1% amplitude noise signal conforming to a standard normal distribution is added to the response signals of the detectors in the test dataset. The distribution of MRE for the network's predicted results is shown in Fig 18.

The frequency of data samples of different MRE in the prediction results is provided. It can be observed that MRE of the network's predictions for the test dataset input scenarios is predominantly concentrated within the range of [-0.4~0.4%]. The results of relative errors approximately conform to a normal distribution, with an expected value of μ = -4.36 E-04 and a standard deviation of σ = 1.98 E-03. The maximum MRE is above 0.6%. Furthermore, as shown in Fig.19 and Fig.20, the group with the maximum MRE has exhibited distortion.

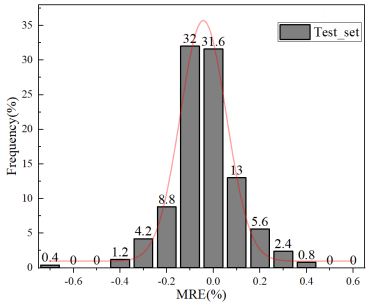

Fig. 18. Frequency distribution map of MRE for the test set with 1% noise influence added to the input signals

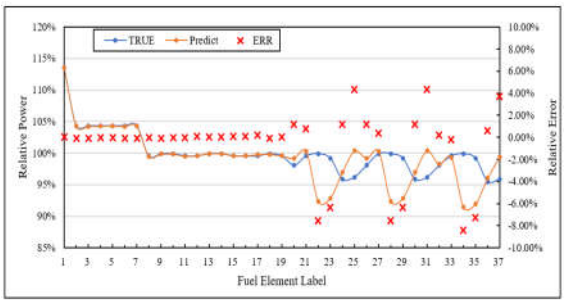

Fig. 19. Predicted and calculated relative power distribution of the maximum MRE group

As depicted in Figure 19, within the group of 37 fuel elements with the maximum MRE, the maximum positive error in relative power amounts to 4.31%, while the least negative error reaches -7.26%. As illustrated in Figure 20, a pronounced distortion is observable in predicting the power distribution of the outermost fuel elements. The predicted power distribution significantly diverges from the computed outcome. This discrepancy implies that the neural network model struggles to provide accurate and comprehensible predictions when input signals incorporate noise, ultimately falling short of meeting the requirements for effective core power monitoring.

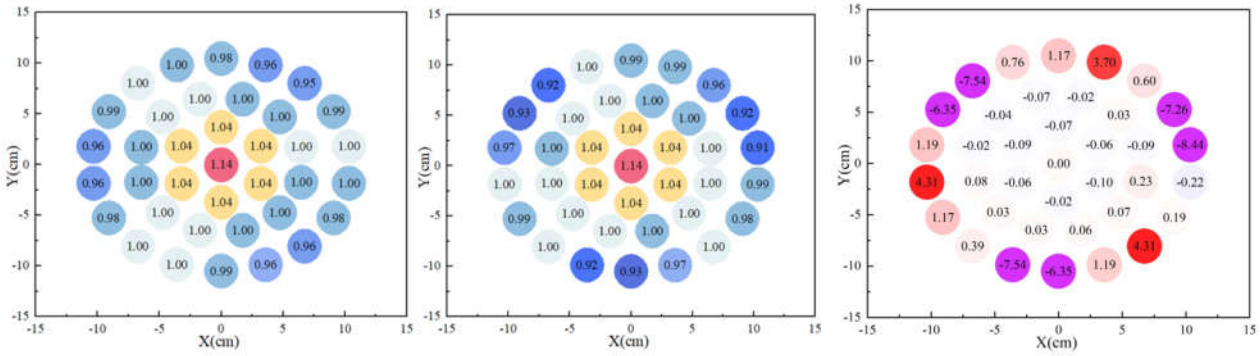

Fig. 20. Relative power distribution and relative error of the maximum MRE group

To address the issues of elevated neural network prediction errors and distorted fitting of relative power distribution for certain operational conditions when noise is introduced, a solution involves training the network for robustness by using noisy signals as input samples. This can be achieved by adding noise to 10,000 sets of samples within the training dataset, resulting in a combined total of 20,000 samples for training. Subsequently, the network's resistance to interference can be verified using the aforementioned noisy test dataset.

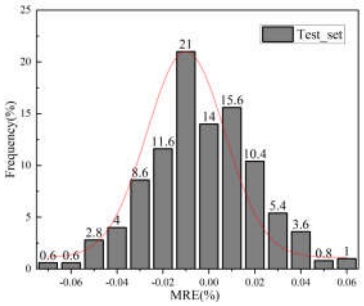

Fig. 21. Frequency distribution map of MRE of prediction results after training the neural network model with noisy signals.

In Fig.21, the frequency of data samples of different MRE in the prediction results is provided. It can be observed that MRE of the network's predictions for the test dataset is predominantly concentrated within the range of [-0.06~0.06%]. The results of relative errors approximately conform to a normal distribution, with an expected value of μ = -3.36 E-03 and a standard deviation of σ = 4.31E-2. The maximum MRE is below 0.07%.

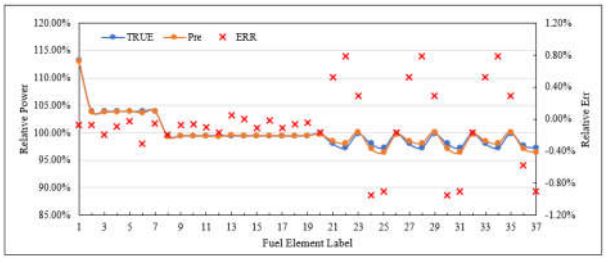

Fig. 22. Predicted and calculated relative power distribution of the maximum MRE group

As depicted in Fig.22, within the group of 37 fuel elements with the maximum MRE, the maximum positive error in relative power amounts to 0.786%, while the least negative error reaches -0.952%. As depicted in Fig. 23, it is evident that the prediction results of power distribution exhibit minimal distortion. The projected power distribution closely aligns with the computed outcome, indicating a high level of consistency. This consistency further suggests that the neural network model excels in providing accurate and comprehensive predictions, even when the input signals contain noise. Ultimately, this capability meets the requirements for efficient core power monitoring.

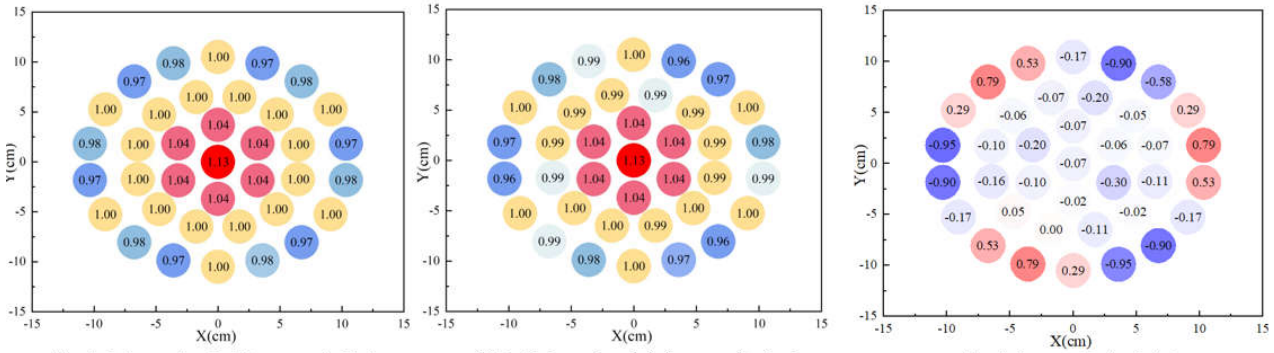

Fig. 23. Relative power distribution and relative error of the maximum MRE group

3. Conclusions

This study introduces a methodology for the reconstruction of microreactor radial power distribution using artificial neural networks, employing ex-core-detector signals. The focus of this investigation centers on the examination of the simplified TOPAZ-II microreactor. Herein, a deep fully connected neural network is deployed to reconstruct the radial power distribution of fuel elements across distinct drum positions. The efficacy of the neural network model is examined in the context of input signals affected by noise and its resilience to noise interference throughout the training process.

The outcomes of the prediction and the associated errors are thoroughly deliberated upon, subsequently contributing to the evaluation of the precision of this approach for core power monitoring, with the utilization of signals emanating from the ex-core nuclear measurement system. The salient findings of this study are succinctly outlined as follows:

(1) In the case of microreactors characterized by a reduced number of fuel assemblies or elements as well as smaller core dimensions, the neural network approach utilizing signals from external neutron detectors demonstrates a high degree of accuracy in reconstructing the radial power distribution within the core.

(2) When the neural network model undertakes the direct prediction of the relative power of all fuel elements, the ensuing prediction outcomes are marred by distortions and noteworthy errors. Conversely, when predictions are made independently for the inner and outer fuel elements, the neural network model achieves a commendable level of accuracy in its prediction results.

(3) The neural network model consistently yields precise prediction outcomes in circumstances where the input signals remain devoid of noise interference.

(4) In scenarios involving input signal noise, the direct prediction of accurate relative power distribution by the neural network model becomes impeded. Consequently, the prediction outcomes are susceptible to distortion.

(5) Through the adoption of training procedures involving noisy input and output data, the neural network can be notably enhanced in terms of its resilience against interference. Under scenarios where input signals are subjected to noise, the neural network can still proficiently achieve accurate reconstructions of core power distribution.

REFERENCES

[1] W. A. Boyd and R. W. Miller, The BEACON On-Line Core Monitoring System: Functional Upgrades and Applications, Organisation for Economic Co-Operation and Development, pp.115-124, 1997.

[2] F. Li, X. H. Zhou, D. Y. Wang, et al, Monitoring of In-core Power Distribution by Ex-core Detectors, Nuclear Power Engineering, Vol. 31, pp.92-96, 2010.

[3] X Peng J, K Wang, Q Li. A new power mapping method based on ordinary kriging and determination of optimal detector location strategy. Annals of Nuclear Energy, 2014, 68:118-123.

[4] M Gori: Machine Learning. Elsevier Ltd:2018-01-01.

[5] S Zhou, C Wu, L Cao, et al. LAVENDER: A steady-state core analysis code for design studies of accelerator driven subcritical reactors. Nuclear Engineering and Design, 2014, 278: 434-444.

[6] Y. Zheng, X. Du, Z. Xu, et al, SARAX: A New Code for Fast Reactor Analysis Part I: Methods, Nuclear Engineering and Design, Vol. 340, pp.421-430, 2018.

[7] Y. Zheng, L. Qiao, Z. Zhai, et al, SARAX: A New Code for Fast Reactor Analysis Part II: Verification, Validation and Uncertainty Quantification, Nuclear Engineering and Design, Vol. 331, pp.41-53, 2018.

[8] B Xiao, Y. Zheng, Y. Wang, et al. Application of SARAX Code System on the Calculation of Complex Unstructured Geometry Core. Nuclear Techniques. 46.06(2023):115-122.

[9] Strong, A Christopher, Wu, et al. Global optimization of objective functions represented by ReLU networks. Machine Learning, 2021. doi: 10.1007/S10994-021-06050-2

[10] S Ioffe, C Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015.

[11] L Prechelt. Automatic early stopping using cross-validation: quantifying the criteria. Neural Networks, 1998,11(4): 761-767.

APPENDIX A

The Structure of deep neural network model for Problem 1 printed by Keras.

Layer	(type)	Dense Num	Param#
flatten	(Flatten)	/	0
batch_normalization	(Batch)	/	48
dense	(Dense)	512	6656
dense_1	(Dense)	1024	525312
batch_normalization_1	(Batch)	/	4096
dense_2	(Dense)	2048	2099200
batch_normalization_2	(Batch)	/	8192
dense_3	(Dense)	4096	8392704
batch_normalization_3	(Batch)	/	16384
dense_4	(Dense)	2048	8390656
batch_normalization_4	(Batch)	/	8192
dense_5	(Dense)	1024	2098176
batch_normalization_5	(Batch)	/	4096
dense_6	(Dense)	512	524800
batch_normalization_6	(Batch)	/	2048
dense_7	(Dense)	128	65664
batch_normalization_7	(Batch)	/	512
dense_8	(Dense)	8	1032
dense_9	(Dense)	6	54

Total params: 22,147,822

APPENDIX B

The Structure of deep neural network model for Problem 2 printed by Keras.

Layer	(type)	Dense Num	Param
flatten	(Flatten)	/	0
batch_normalization	(Batch)	/	48
dense	(Dense)	64	832
batch_normalization_1	(Batch)	/	256
dense_1	(Dense)	128	8320
batch_normalization_2	(Batch)	/	512
dense_2	(Dense)	256	33024
batch_normalization_3	(Batch)	/	1024
dense_3	(Dense)	512	131584
batch_normalization_4	(Batch)	/	2048
dense_4	(Dense)	1024	525312
batch_normalization_5	(Batch)	/	4096
dense_5	(Dense)	512	524800
batch_normalization_6	(Batch)	/	2048
dense_6	(Dense)	256	131328
batch_normalization_7	(Batch)	/	1024
dense_7	(Dense)	32	8224
dense_8	(Dense)	19	627

Total params: 1,375,107