Nuclear Science and Techniques (2025) 36:172
https://doi.org/10.1007/s41365-025-01749-6

β-ray induced X-ray spectroscopy for tritium analysis with back propagation neural network

Hong Huang[1] · Zhu An[1] · Jing-Jun Zhu[1]

Received: 22 January 2025 / Revised: 10 February 2025 / Accepted: 12 February 2025 / Published online: 1 July 2025

Abstract

β-ray-induced X-ray spectroscopy (BIXS) is a promising technique for tritium analysis that offers several unique advantages, including substantial detection depth, nondestructive testing capabilities, and ease of operation. For thin solid tritium-containing samples with substrates, the currently used BIXS analysis method can measure the tritium depth profile and content when the sample thickness is known. In this study, a backpropagation (BP) neural network algorithm was used to predict the tritium content and thickness of a thin solid tritium-containing sample with substrates and a uniformly distributed tritium profile. A semi-analytical method was used to generate datasets for training and testing the BP neural network. A dataset of β-decay X-ray spectra from 420 tritium-containing zirconium models with different known thicknesses and tritium-to-zirconium ratios was used as the input data. The corresponding zirconium thicknesses and tritium-to-zirconium ratios served as the output for training and testing the BP neural network. The mean relative errors (MREs) of the zirconium thickness in the training and test datasets were 0.56% and 0.42%, respectively, whereas the MREs of the tritium-to-zirconium ratio were 0.59% and 0.38%, respectively. Furthermore, the trained BP neural network demonstrates excellent predictive capability across various levels of statistical uncertainty. For the experimental β-decay X-ray spectra of two tritium-containing samples, the predicted zirconium thicknesses and tritium-to-zirconium ratios showed good agreement with the results obtained through the elastic backscattering spectrometry (EBS).

Keywords Tritium analysis · β-ray induced X-ray · Uniformly distributed tritium · Unknown thickness · Semi-analytical · Back propagation neural network

1 Introduction

Nondestructive detection techniques are widely used to measure the tritium content and its distribution. Tritium nondestructive detection techniques mainly include β particle counting [1], elastic backscattering spectrometry (EBS) [2], calorimetry [3], imaging plate analysis [4], and β decay-induced X-ray spectroscopy (BIXS) [5–11]. β particle counting and imaging plate analysis can only provide information on the surface distribution of tritium, while calorimetry can only determine the total tritium content; none of these methods can obtain tritium depth profiles. In contrast, EBS can obtain tritium depth profiles and tritium contents but requires large equipment, such as an accelerator. BIXS measures tritium depth profiles and tritium content by detecting X-rays produced by electrons resulting from tritium β decay in materials. This method has several notable advantages, including a large detection depth, nondestructive testing capabilities, and ease of operation [12]. BIXS analysis methods can be classified into analytical BIXS method [5] and Monte Carlo (MC)-based methods [9]. The analytical method proposed by Matsuyama [5] in 1998 was based on empirical formulas and did not consider the complicated

This work was supported by the National Natural Science Foundation of China (No. 12175158) and the Institute of Nuclear Physics and Chemistry, China Academy of Engineering Physics (No. HG2022022).

✉ Zhu An
anzhu@scu.edu.cn

Jing-Jun Zhu
zhujingjun@scu.edu.cn

[1] Key Laboratory of Radiation Physics and Technology of Ministry of Education, Institute of Nuclear Science and Technology, Sichuan University, Chengdu 610064, China

transport processes of electrons and photons in materials. However, MC BIXS, introduced by An et al. [9], uses Monte Carlo simulations (i.e., the PENELOPE code [13]) to model the tritium β-decay X-ray spectra and combines the simulated and experimental spectra to obtain tritium depth profiles and contents [14, 15]. Although MC BIXS is more accurate owing to its consideration of the complex geometry and electron and photon transport in materials, it is time-consuming and requires sufficient statistical accuracy in the simulated X-ray spectra. Therefore, we developed a semi-analytical BIXS method that combines MC simulations with analytical calculations [16]. This approach offers a 73 times improvement in computational efficiency compared to MC BIXS and simultaneously maintains high accuracy, for example, the difference in tritium content obtained by the semi-analytical BIXS and MC BIXS for the same tritium-containing sample was only 0.82% [16].

For thin solid tritium-containing samples with substrates, which were the type of samples often encountered in the application of BIXS [14, 15], the present BIXS methods required prior knowledge of the thickness of the sample, and the details of the BIXS analysis have been described in Refs [14, 15]. The sample thickness needed in BIXS analysis is often obtained by weighing during sample preparation or by EBS. Currently, in the practical application of BIXS for sample testing, a need has been proposed by the BIXS user; that is, without prior knowledge of the sample's thickness, the tritium content and sample thickness can be obtained simultaneously using BIXS for a thin solid tritium-containing sample with a substrate.

In some cases, a good linear relationship between tritium content and X-ray intensities can be obtained [17, 18]; for example, Matsuyama et al. discovered that the intensities of characteristic X-rays Ar(Kα) exhibited a strong linear correlation with the total tritium content in tritium-containing graphite plates [17]; therefore, the tritium content can be derived from the X-ray intensities by interpolation. However, for thin solid tritium-containing samples with substrates, different combinations of sample thickness and tritium content can yield identical X-ray intensities, making it impossible to simultaneously determine both the thickness of the sample and the tritium content through simple interpolation. In such cases, the shape of the X-ray spectrum must be considered. First, we studied the case in which tritium was uniformly distributed in the sample. To simultaneously obtain both the thickness and tritium content of a uniformly distributed tritium sample, in this study, we propose a reconstruction approach for BIXS based on an artificial neural network (ANN) algorithm, in which a large dataset must be constructed to train the ANN; the fast semi-analytical model developed by us [16] to calculate the X-ray spectrum of the tritium-containing sample allows it to build a training dataset. An ANN is highly fault-tolerant, fast, and scalable, with excellent parallel processing capabilities [19–22]. ANN have been widely applied in nuclear science and technology, including in neutron spectrum unfolding [23], nuclear power plant dynamic behavior prediction [24], nuclear spectral analysis [25], and Rutherford backscattering spectrum analysis [26–29].

The backpropagation (BP) neural network developed by Rumelhart et al. [30] is a specific implementation of an ANN, particularly for training multilayer feed-forward networks. It consists of an input layer, hidden layers, and an output layer, with the neurons in each layer fully connected only to adjacent neurons. The simple structure and stability of a BP neural network render it effective for high-precision nonlinear fitting [31–35]. BP networks have been successfully applied to tasks, such as simple classification [36], neutron spectrum resolution [37], nuclide identification [38], and pulse shape discrimination [39]. Most recently, Zhao et al. [40] used a BP neural network to reconstruct tritium depth profiles in materials in a simulation study of BIXS, with the analysis depth limited to 20 μm for a sample of 1 mm thick (i.e., equivalent to a semi-infinite sample).

The remainder of this paper is organized as follows: Section 2 introduces the methods used in this work, including the construction of the tritium β-decay X-ray spectra for BP network training and the construction process of the BP network. Section 3 presents and discusses the results, including the detailed optimization process, test, and generalizability of the BIXS BP network. The application of the BIXS BP network to experimental X-ray spectra is discussed. Finally, Sect. 4 concludes the paper.

2 Methods

2.1 Semi-analytical BIXS X-ray spectrum

A large dataset of BIXS spectra is required to train the ANN, including the X-ray spectra induced by electrons from β decays of tritium in the sample, corresponding tritium depth profiles, and sample thicknesses. In this study, the semi-analytical model developed in Ref. [16] for calculating the X-ray spectrum of tritium-containing sample was employed to generate the dataset. The experimental setup of the BIXS, based on a silicon drift detector (SDD), is shown in Figure 1 and identical to the setup described in detail in [16]. A 5.01 μm-thick aluminum film was used as the β-ray stopping layer, and tritium-containing samples (i.e., zirconium films in this study) were supported by 1 mm-thick molybdenum substrates.

To calculate the X-ray spectrum of the tritium-containing sample using the semi-analytical model, internal bremsstrahlung (IB), external bremsstrahlung (EB), and characteristic X-rays were considered. The detailed

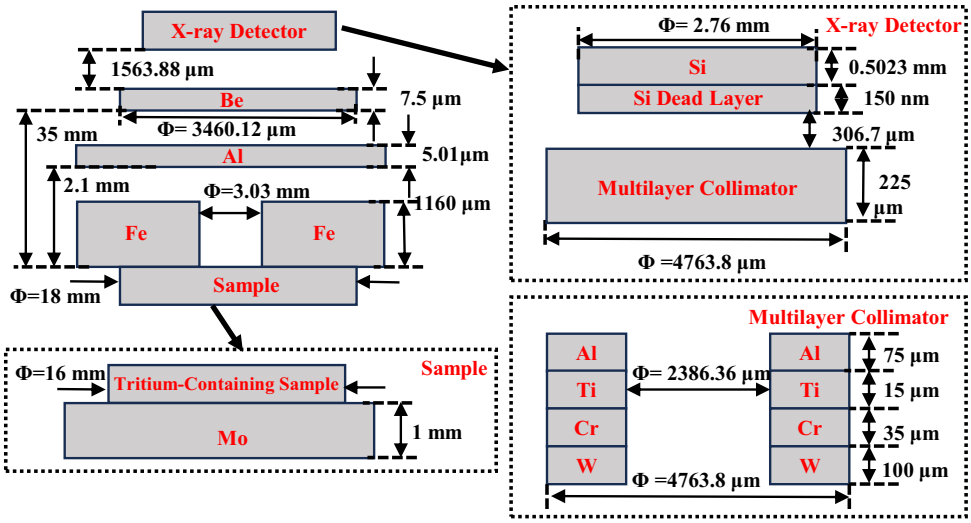

Fig. 1 The geometrical model for the BIXS setup

calculation process is described in Refs. [16], and a brief description is provided here. The total X-ray fluence, which is the differential in the photon energy k per β electron from tritium decay per solid angle, can be expressed by the following formula [16]:

$$N^{\mathrm{TOT}}(k,\hat{\Omega}_\gamma;Z) = H(k,\hat{\Omega}_\gamma)\times \left(N^{\mathrm{CH}}(k,\hat{\Omega}_\gamma;Z)+ N^{\mathrm{IB}}(k,\hat{\Omega}_\gamma;Z)+N^{\mathrm{EB}}(k,\hat{\Omega}_\gamma;Z)\right), \quad (1)$$

where $N^{\mathrm{CH}}(k,\hat{\Omega}_\gamma;Z)$, $N^{\mathrm{EB}}(k,\hat{\Omega}_\gamma;Z)$, and $N^{\mathrm{IB}}(k,\hat{\Omega}_\gamma;Z)$ represent the corresponding fluences of the characteristic X-rays, EB photons, and IB photons, respectively, and $H(k,\hat{\Omega}_\gamma)$ is the attenuation of X-rays in the filters (i.e., the Be window and Si dead layer). Z is the atomic number, $\hat{\Omega}_\gamma$ is the emission direction of the photons. The depth distributions of characteristic X-rays in materials were simulated using the modified MC PENEPMA code [41] and used to calculate the $N^{\mathrm{CH}}(k,\hat{\Omega}_\gamma;Z)$. The electron distributions in materials, including both energy and angular distributions, were simulated with the modified MC PENELOPE code [13] and used to calculate the $N^{\mathrm{EB}}(k,\hat{\Omega}_\gamma;Z)$, which can be determined as follows [16]:

$$N^{\mathrm{EB}}(k,\hat{\Omega}_\gamma;Z) = n\int_0^{Dz} \mathrm{d}s \int_k^{E_{\max}} \mathrm{d}E \int \mathrm{d}\hat{\Omega}_{\mathrm{e}} N^{\mathrm{e}}(Dz,E,\hat{\Omega}_{\mathrm{e}};Z) \times \frac{\mathrm{d}^2\sigma_{\mathrm{br}}(k,\hat{\Omega}_\gamma\cdot\hat{\Omega}_{\mathrm{e}};Z,E)}{\mathrm{d}k\mathrm{d}\hat{\Omega}_\gamma} f(Dz,k,\hat{\Omega}_\gamma;Z), \quad (2)$$

where $E_{\max}$ is the maximum kinetic energy of electrons from tritium β-decay, $N^{\mathrm{e}}(Dz,E,\hat{\Omega}_{\mathrm{e}};Z)$ is the electron distributions in materials with energy E moving in the direction $\hat{\Omega}_{\mathrm{e}}$ at depth Dz generated by an electron from tritium β-decay, n is the atomic or molecular density with a unit of number of atoms or molecules per cubic centimeter, $f(Dz,k,\hat{\Omega}_\gamma;Z)$ is the self-absorption of the target, $\frac{\mathrm{d}^2\sigma_{\mathrm{br}}(k,\hat{\Omega}_\gamma\cdot\hat{\Omega}_e;Z,E)}{\mathrm{d}k\mathrm{d}\hat{\Omega}_\gamma}$ is the double differential bremsstrahlung cross sections [42, 43]:

$$\frac{\mathrm{d}^2\sigma_{\mathrm{br}}(k,\hat{\Omega}_\gamma\cdot\hat{\Omega}_{\mathrm{e}};Z,E)}{\mathrm{d}k\mathrm{d}\hat{\Omega}_\gamma} = \chi\frac{Z^2}{\beta_1^2}\frac{1}{k}S = \chi\frac{Z^2}{\left(1-\frac{1}{(1+\frac{E}{m_{\mathrm{e}}c^2})^2}\right)}\frac{1}{k}S, \quad (3)$$

where $m_{\mathrm{e}}c^2$ is the electron rest energy, S is the shape function of the bremsstrahlung angular distribution, the Kissel-Quarles-Pratt (KQP) [44] theory was used in this work, and χ is the scaled cross-section differential in k. The other symbols have the same definitions as those in Eqs. 1 and 2. The Knipp-Uhlenbeck-Bloch (KUB) model proposed by Knipp and Uhlenbeck [45] and Bloch [46] was used to calculate $N^{\mathrm{IB}}(k,\hat{\Omega}_\gamma;Z)$. Each X-ray spectrum consists of 200 energy bins at interval of 0.093 keV.

2.2 Construction of BP network dataset

A total of 420 tritium-containing zirconium samples of different thicknesses and tritium-to-zirconium ratios were used to generate a dataset of tritium β-decay X-ray spectra. The zirconium thicknesses before absorbing tritium ranged between 3 μm and 5 μm, with the assumption that tritium was uniformly distributed throughout the sample.

The zirconium thickness was divided into 21 groups with an interval of 0.1 µm, whereas the tritium-to-zirconium ratio ranged from 0.1 to 2, divided into 20 groups with an interval of 0.1. The time required by the semi-analytical method to obtain a tritium β-decay X-ray spectrum for each combination of sample thickness and tritium-to-zirconium ratio was approximately 1 h [16]. Thus, the total time required to obtain 420 X-ray spectra is approximately 19 d [16]. From Ref. [16], it is noted that the semi-analytical X-ray spectra were expressed in unit of "counts per keV per β decay". To ensure consistency between the semi-analytical and experimental spectra, the unit of the semi-analytical spectra was converted to the unit of the experimental spectra, that is, counts per keV per second, based on the number of tritium atom decays, N_T, within a unit time t:

$$N_T = N_{Zr} \times R \times \left(1 - \exp\left(-\frac{\ln 2 \times t}{T}\right)\right), \tag{4}$$

where N_{Zr} is the number of zirconium atoms, R is the tritium-to-zirconium ratio, and T is the half-life of tritium (approximately 12.25 years.) The semi-analytical X-ray spectra were convoluted using the Gaussian response function of the SDD [47]. The full width at half maximum (FWHM) of the SDD in the BIXS experimental setup was 185 eV at 5.89 keV, as measured using a standard Fe-55 radioactive source:

$$FWHM = \sqrt{8WFE\ln 2 + \triangle E_{elec}^2} = 2\sqrt{2\ln 2}\sigma, \tag{5}$$

where W is the average energy for electron–hole creation (3.62 eV), F is Fano factor (0.12), E is the X-ray energy, $\triangle E_{elec}$ is the electronic noise (141.55 eV), and σ is the standard deviation of the Gaussian distribution. Figure 2 shows the semi-analytical tritium β decay X-ray spectra before and after convolution for a sample with a zirconium thickness of 3 µm and a tritium-to-zirconium ratio of 0.1.

2.3 Construction of BP network

Figure 3 illustrates the structure of a backpropagation (BP) network designed to predict the zirconium thickness and tritium-to-zirconium ratio. The structure consists of three layers: input, hidden, and output layers. The number of hidden layers shown in this figure differs from that of the optimized configuration adopted for our application. Only the energy region of the X-ray spectrum between 1 keV and 15 keV was used to train the BP neural network to avoid noise in the low-energy region and poor statistics in the high-energy region of the experimental X-ray spectrum. The input layer consisted of 150 neurons, corresponding to 150 energy bins of the tritium β decay X-ray spectrum in the 1 keV–15 keV range. The output layer contains two neurons that represent the zirconium thickness and tritium-to-zirconium ratio. To enhance the accuracy and computational efficiency of the

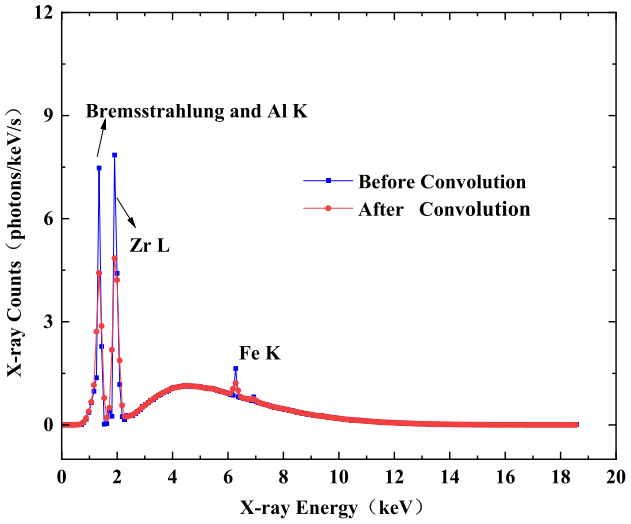

Fig. 2 Tritium semi-analytical β-decay X-ray spectrum before and after convolution for a sample with a zirconium thickness of 3 µm and a tritium-to-zirconium ratio of 0.1. Bremsstrahlung and Al K, Zr L, and Fe K represent the bremsstrahlung plus Al K-shell characteristic X-ray, Zr L-shell characteristic X-ray, and Fe K-shell characteristic X-ray, respectively

neural network model, a scaling transformation of the input and output data were performed [26], and the β-decay X-ray spectra were treated as follows:

$$C1(k) = \frac{\log_{10}(C(k))}{10}, \tag{6}$$

where $C(k)$ and $C1(k)$ are the photon counts with energy k before and after processing, respectively. The thickness of zirconium was divided by 10 and the tritium-to-zirconium ratio was normalized as follows:

$$N_{in} = \frac{N_i - N_{min}}{N_{max} - N_{min}}, \tag{7}$$

where N_i and N_{in} are the actual and scaled tritium-to-zirconium ratios, respectively. N_{min} is the minimum tritium-to-zirconium ratio (e.g., 0.1) and N_{max} is the maximum tritium-to-zirconium ratio (e.g., 2.0).

The network was implemented and trained in Python 3.7 using the PyTorch library [48], with the error backpropagation algorithm employed for training. The mean squared error (MSE) [49] was used as the loss function, which was commonly used in regression tasks. It is calculated by summing the squared differences between the predicted and true values.

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2, \tag{8}$$

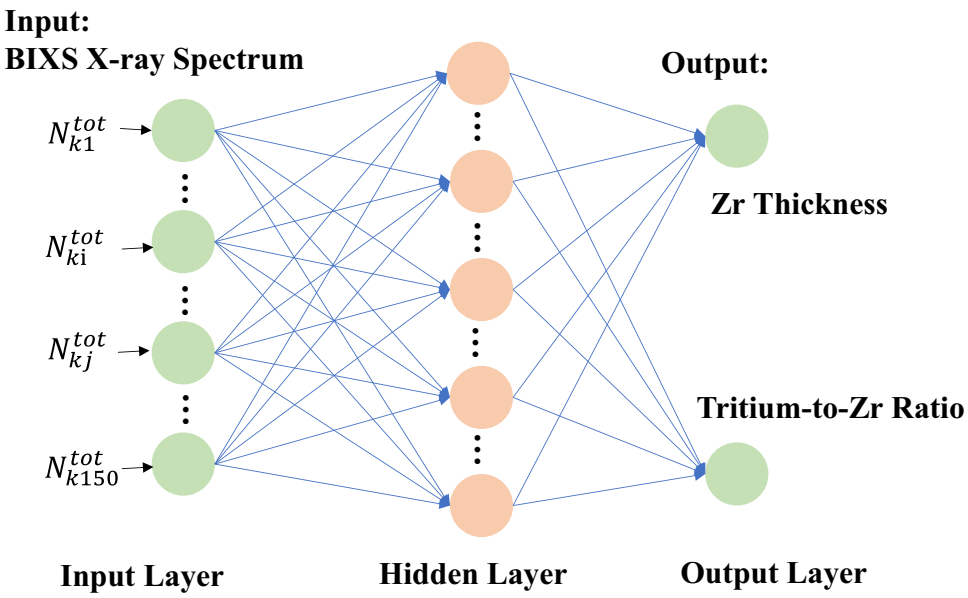

Fig. 3 The structure of BP network used to predict the zirconium thickness and tritium-to-zirconium ratio

where y_i and $\hat{y}_i$ represent the true and predicted values, respectively, and m is the number of y_i. With the same learning rate, the AdamW optimizer [50] consistently demonstrated faster convergence and smaller test errors. Currently, AdamW has been widely adopted and selected as the optimization algorithm for the BIXS neural network structure. The training process was repeated five times, resulting in five individual ANNs to evaluate the precision in terms of reproducibility. This yielded the MSE values for the two ANN outputs (tritium-to-zirconium ratio and zirconium thickness). The training and testing datasets were randomly split in a ratio of 0.8:0.2. The training process was performed iteratively to optimize network parameters, such as weights and biases, allowing for accurate pattern recognition and feature extraction. An iteration is referred to as an "epoch", indicating that the neural network processes the entire training dataset once. The number of epochs was set to 10,000.

3 Results and discussion

3.1 Optimization of BP network structure

To select the optimal network structure, the numbers of hidden layers and neurons were optimized. To achieve nonlinear transformations, the ReLU activation function [51] was used for all the hidden layers. Table 1 shows the MSEs as functions of the numbers of hidden layers and neurons. From the results in Table 1, it can be observed that when the number of neurons exceeds 10, the MSEs tend to decrease as the

Table 1 MSEs as a function of the number of hidden layers and neurons

Hidden layers		Number of neurons					
		10	30	50	70	100	150
Training	1	7.43×10^{-3}	4.75×10^{-4}	3.82×10^{-4}	2.59×10^{-4}	1.68×10^{-4}	1.15×10^{-4}
	2	4.43×10^{-4}	7.75×10^{-5}	4.53×10^{-5}	3.48×10^{-5}	2.48×10^{-5}	2.18×10^{-5}
	3	5.57×10^{-3}	3.64×10^{-5}	1.79×10^{-5}	1.74×10^{-5}	1.36×10^{-5}	1.02×10^{-5}
	5	6.56×10^{-5}	1.80×10^{-5}	1.26×10^{-5}	6.83×10^{-6}	6.54×10^{-6}	4.06×10^{-6}
	7	9.48×10^{-5}	1.44×10^{-5}	9.07×10^{-6}	8.35×10^{-6}	8.54×10^{-6}	6.40×10^{-6}
	10	1.88×10^{-3}	1.31×10^{-5}	1.91×10^{-5}	3.74×10^{-5}	3.74×10^{-5}	1.28×10^{-5}
Test	1	7.07×10^{-3}	4.33×10^{-4}	3.28×10^{-4}	2.01×10^{-4}	1.23×10^{-4}	8.39×10^{-5}
	2	3.93×10^{-4}	6.04×10^{-5}	3.80×10^{-5}	2.74×10^{-5}	1.91×10^{-5}	1.73×10^{-5}
	3	4.66×10^{-3}	3.20×10^{-5}	1.46×10^{-5}	1.41×10^{-5}	1.16×10^{-5}	1.00×10^{-5}
	5	6.61×10^{-5}	1.66×10^{-5}	1.08×10^{-5}	5.99×10^{-6}	7.70×10^{-6}	6.45×10^{-6}
	7	8.71×10^{-5}	1.29×10^{-5}	9.72×10^{-6}	8.77×10^{-6}	7.70×10^{-6}	6.45×10^{-6}
	10	2.09×10^{-3}	1.21×10^{-5}	2.08×10^{-5}	3.23×10^{-5}	3.23×10^{-5}	1.16×10^{-5}

number of hidden layers and neurons increases. When the number of hidden layers was between 2 and 10, and the number of neurons was between 30 and 150, the MSEs of the neural network for predicting the zirconium target thickness and tritium-to-zirconium ratio became small, ranging from approximately 5.99×10^{-6} to 7.75×10^{-5}.

To simplify the neural-network structure while maintaining a low MSE, we selected a relatively simple five-layer neural network with three hidden layers, each containing 150 neurons. The total MSEs for the training and testing datasets were 1.02×10^{-5} and 1.00×10^{-5}, respectively. The mean relative errors (MREs) [52] for the zirconium thickness and tritium-to-zirconium ratio in the training dataset of the neural network were 0.56 and 0.42%, respectively, whereas those in the test dataset, the MREs were 0.59 and 0.38%, respectively. The MRE [52] was calculated as follows:

$$MRE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{y_i - \hat{y}_i}{y_i}\right|, \tag{9}$$

where the definitions of y_i, $\hat{y}_i$ and m are identical to those in 8.

3.2 Optimization of activation functions

In this study, the activation functions used in the hidden layers were optimized. To simplify the optimization process, the same activation function (i.e., ReLU [51], sigmoid [53], or tanh [54]) was applied to all the hidden layers. Table 2 shows the MSEs of the BP network for different activation functions in the hidden layers. It can be observed that the MSEs vary with the choice of activation function. We selected the activation function that resulted in the best MSE, i.e., the ReLU function. The MSEs for the training and test datasets are 1.02×10^{-5} and 1.00×10^{-5}, respectively.

Figure 4 presents the MSEs of the BP network for both the training and test datasets as a function of the number of epochs. It can be observed that the BIXS neural network converged rapidly, with sufficient convergence achieved within 2000 iterations. Moreover, the MSEs for both datasets exhibited minimal divergence as the number of epochs increased, indicating that overfitting did not occur during training. Figure 5 presents the MREs of the training dataset and test dataset for each training period. For the tritium-to-zirconium ratio and zirconium thickness, the MRE differences obtained from each training were considerably small, indicating the good stability of the system. Figure 6 shows a comparison between the true and predicted values for the zirconium thickness and tritium-to-zirconium ratio for all 420 sets of data. Excellent predictions were obtained with the true and predicted values in close agreement. The MREs for the 420 datasets used to train and test the BP network are listed in Table 3.

Table 2 MSEs of the BP network for different activation functions in the hidden layers

Parameters	ReLU	Tanh	Sigmoid
MSE for training	1.02×10^{-5}	3.71×10^{-5}	1.62×10^{-5}
MSE for test	1.00×10^{-5}	3.01×10^{-5}	1.31×10^{-5}

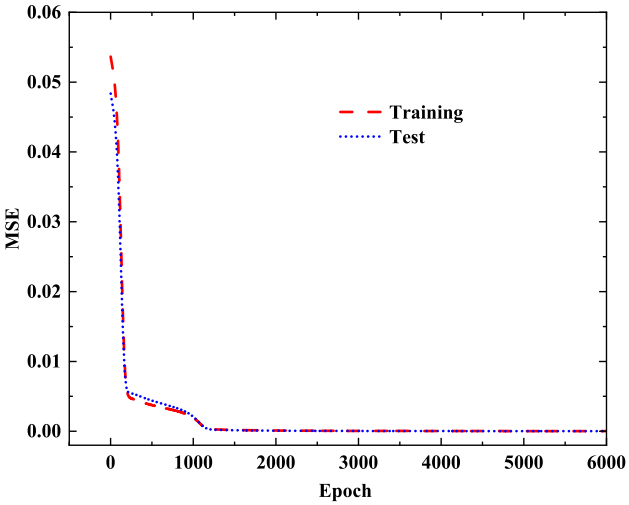

Fig. 4 The MSEs of the BP network for both the training and test datasets as a function of the number of epochs

3.3 Effect of statistical uncertainty

The experimental BIXS X-ray spectrum may exhibit varying degrees of statistical uncertainty, which can affect the prediction accuracy. To assess the effect of statistical uncertainty, the uncertainties ranging from 0.5 to 3% were randomly with Gaussian distribution added to the 84 sets of test data of X-ray spectra, which were randomly selected from the 420 data, as described in the section "C. Construction of BP network". Figure 7 compares the true and predicted values for the 84 sets of test data without statistical uncertainty. It can be observed that the true and predicted results are in close agreement. Table 4 lists the MREs for 84 sets of test data under different levels of uncertainty. The results demonstrate that uncertainty can affect the accuracy of neural network predictions to some extent. Generally, the greater the uncertainty, the larger the relative error in the prediction result. When the statistical uncertainty of the test set is 3%, the neural network maintains higher prediction accuracy for both zirconium thickness and the tritium-to-zirconium ratio, with average relative errors of 5.48% and 5.08%, respectively.

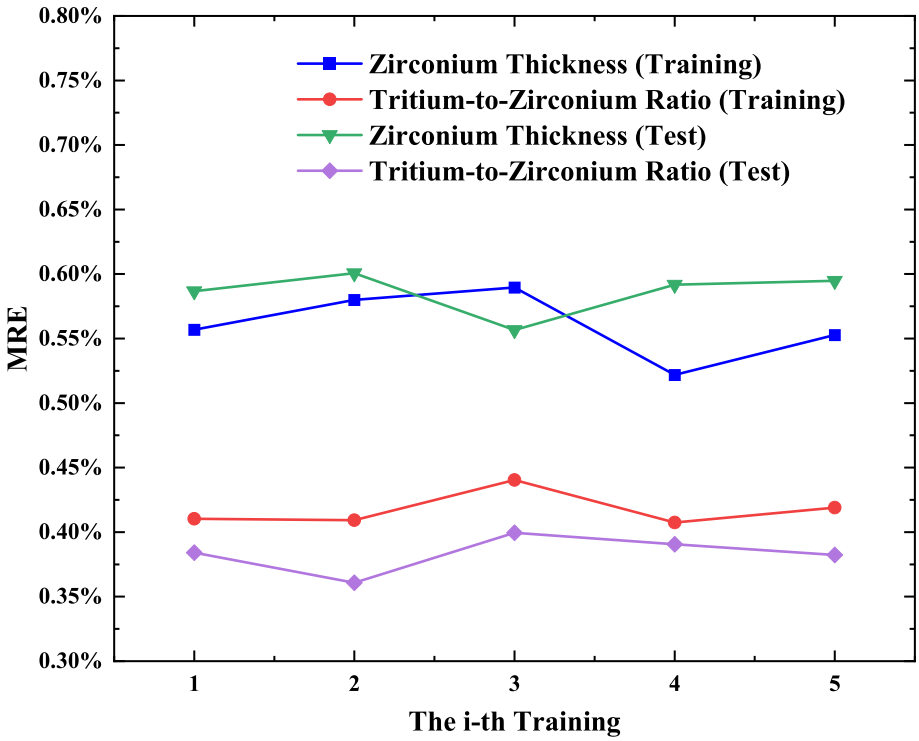

Fig. 5 The MREs of the training dataset and the test dataset in each repeated training

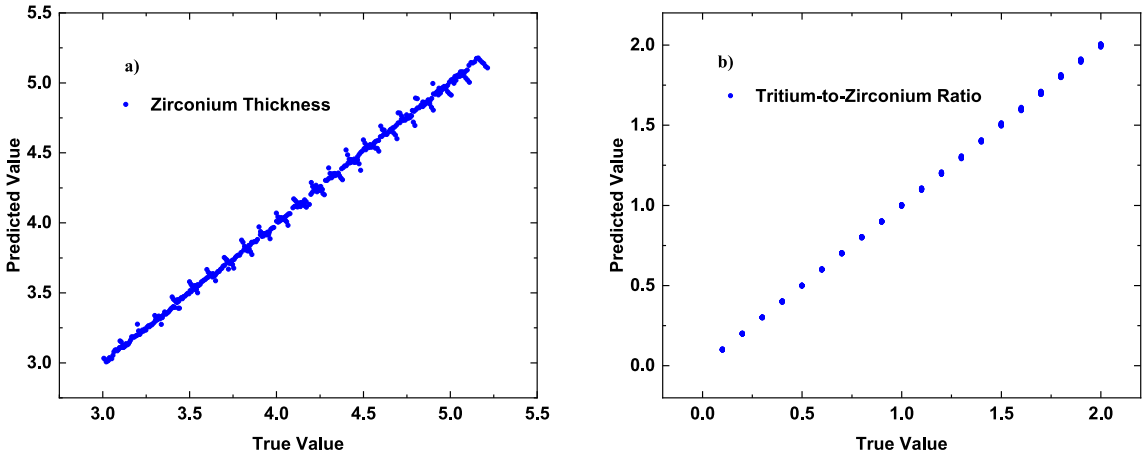

Fig. 6 Comparison between the true and predicted values of **a** zirconium thickness and **b** tritium-to-zirconium ratio for all 420 sets of data

Table 3 MREs for the 420 sets of data used to train and test the BP-network

Parameters	Training (%)	Test (%)
Zirconium thickness	0.56	0.59
Tritium-to-Zirconium ratio	0.42	0.38

3.4 Generalizability of BP network

Additional 22 sets of data, with the zirconium thicknesses outside the BP network training range from 3 µm to 5 µm, were used to preliminarily test the generalization ability of the BP neural network (i.e., the capability of extrapolation). The zirconium thicknesses and tritium-to-zirconium ratios for the 22 datasets are listed in Table 5, and the MREs of the predicted results are listed in Table 6. From Table 6, it can be observed that the BP neural network demonstrates good prediction capability, even for the dataset outside the training range. For the 22 sets of data, the MREs of the predicted zirconium thicknesses and tritium-to-zirconium ratios are 7.61 and 9.33%, respectively. For a dataset outside the training parameter range, the prediction

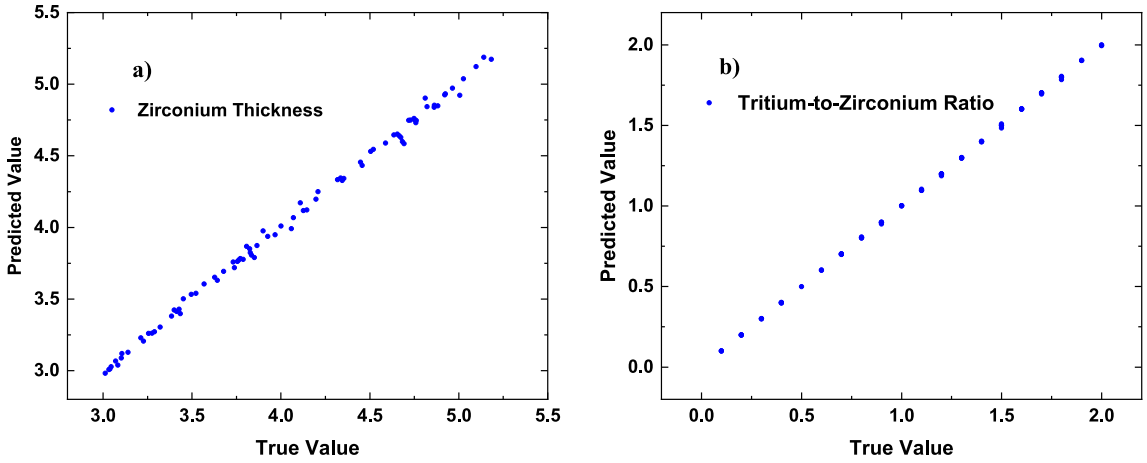

Fig. 7 Comparison between the true and predicted values of **a** zirconium thickness and **b** tritium-to-zirconium ratio for the 84 sets of test data without statistical uncertainty

Table 4 MREs for the 84 sets of test data under the different levels of uncertainty

Relative uncertainties (%)	Zr thickness (%)	Tritium-to-Zr ratio (%)
0	0.59	0.38
0.5	1.20	1.09
1	1.85	1.64
1.5	3.32	2.48
2	3.44	3.15
3	5.48	5.08

Table 5 Zirconium thicknesses and tritium-to-zirconium ratios of the 22 sets of data used to test the generalization of the BP neural network

Tritium-to-Zr ratio (Zirconium thickness (μm))			
0.2 (1.5)	0.5 (5.5)	1.5 (1.5)	1.63 (5.5)
0.2 (1.7)	0.5 (6.0)	1.5 (1.8)	1.63 (6.0)
0.2 (2.2)	0.5 (6.5)	1.5 (2.0)	1.63 (6.5)
0.2 (2.4)	0.5 (6.8)	1.5 (2.3)	1.63 (6.8)
0.2 (2.8)	0.5 (7.0)	1.5 (2.5)	1.63 (7.0)
1.0 (2.5)	1.5 (5.5)		

Table 6 MREs for the 22 sets of data used to test the generalization of the BP neural network

Prediction parameters	MRE (%)
Zirconium thickness	7.61
Tritium-to-Zirconium ratio	9.33

accuracy of the neural network decreases, which is consistent with the conclusion in Ref. [27].

3.5 Application of BIXS BP neural network

3.5.1 Sample

Two tritium-containing samples were prepared, labeled No. 26 and No. 42, in which zirconium films were deposited onto smooth molybdenum substrates (approximately 1 mm thick) using the electron beam evaporation technique. The processes for preparing the tritium-containing samples are the same as those described in Refs.[15, 55].

The thicknesses of the zirconium films and the tritium depth profiles of samples 26 and 42 were measured using the elastic backscattering spectrometry (EBS) method. The EBS experiment was conducted using a 3 MV tandetron accelerator at Sichuan University [56]. The incident proton energy was 3 MeV with a current intensity of approximately 2 nA. The proton beams impacted the sample surface vertically, and a Si detector (ORTEC U-012-050-100) was placed at 165° with respect to the proton beam direction. The experimental details are presented in Refs.[15].

The experimental spectra of the EBS were analyzed using the SIMNRA program [57] with the same processes as in Refs.[14, 15] The parameters used in SIMNRA were consistent with those in Ref. [15]. The uncertainties for the tritium contents and zirconium thicknesses obtained using the EBS were approximately 11.7% and 5%, respectively, mainly arising from uncertainties in the proton elastic scattering cross-section data, stopping power, and statistical uncertainties [14, 15]. For example, Figure 8 shows the EBS experimental spectrum and SIMNRA simulation for sample 26, where the tritium distribution in the SIMNRA simulation was assumed to be uniform. The background spectrum was obtained by

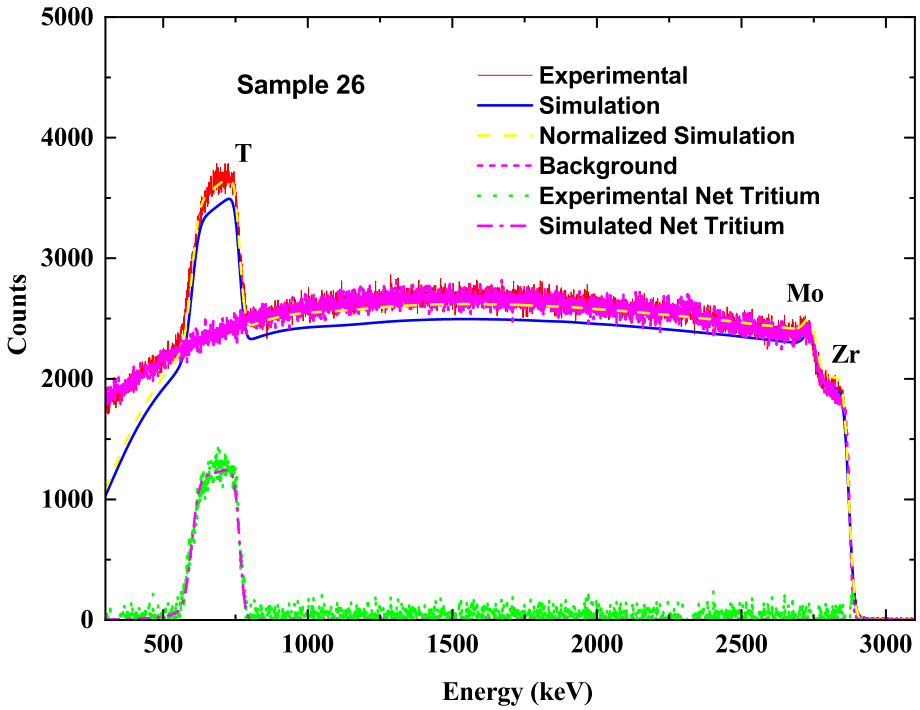

Fig. 8 The experimental spectrum of EBS and the SIMNRA simulation for sample 26. The normalized simulation spectrum represents the simulation spectrum that was normalized to the experimental tritium counts (see the text for other details)

measuring a hydrogen-containing Zr film sample with a Mo substrate, whose geometric dimensions and hydrogen isotope-zirconium ratio were approximately the same as those of samples 26 and 42; the EBS experimental conditions were the same. The experimental net tritium spectrum was obtained by deducting the background spectrum from the experimental spectrum of the tritium-containing samples. The measured results are summarized in Table 7. The EBS results indicate that the tritium distributions for samples 26 and 42 were uniform.

3.5.2 Predictions of BP network

The experimental BIXS spectra of samples 26 and 42 were obtained using the same experimental setup as in Refs. [14, 15] and corrected for the signal pile-up effect by using Monte Carlo method [15], as shown in Figure 9. Corrections for the pile-up effect were less than 1%. The trained BP neural network was used to analyze the experimental BIXS spectra after correcting for the intrinsic detection efficiency of the SDD. The predicted values of the zirconium thicknesses and tritium-to-zirconium ratios are listed in Table 7. The predicted tritium-to-zirconium ratios have been converted to the tritium-to-zirconium ratios on the day of the EBS experiment.

For sample 26, the predicted zirconium thickness is 1.60 µm, which deviates by 2.43% from the EBS result, whereas for sample 42, the predicted zirconium thickness is 1.63 µm, showing a 5.23% difference compared to the EBS result. The predicted tritium-to-zirconium ratios for samples 26 and 42 are 1.85 and 1.91, with the relative deviations of 1.07% and 2.14% relative to the EBS results, respectively.

Considering that the predicted values from the BP neural network are consistent with the measured results from the EBS within experimental uncertainties, we can conclude that the BP neural network can be used to predict the thicknesses and tritium contents simultaneously with higher accuracy for thin solid tritium-containing samples with substrates and uniform tritium distribution. Previously, we employed the MC BIXS method for tritium analysis, which required prior information about the sample thickness, and its accuracy was affected by the reconstruction algorithm [9, 14, 15, 58, 59]. However, the present BIXS BP neural network approach overcomes the difficulties inherent in traditional regularization BIXS analysis methods.

4 Summary

In this study, an artificial neural network (ANN) algorithm was employed to predict the tritium content and thickness of thin solid tritium-containing samples with substrates and uniform tritium distributions. The semi-analytical method developed earlier for calculating the X-ray spectrum for tritium-containing samples was used to generate the dataset for training and testing the BIXS BP neural network. The neural network was optimized in several aspects, including the number of hidden layers, neurons, and activation function.

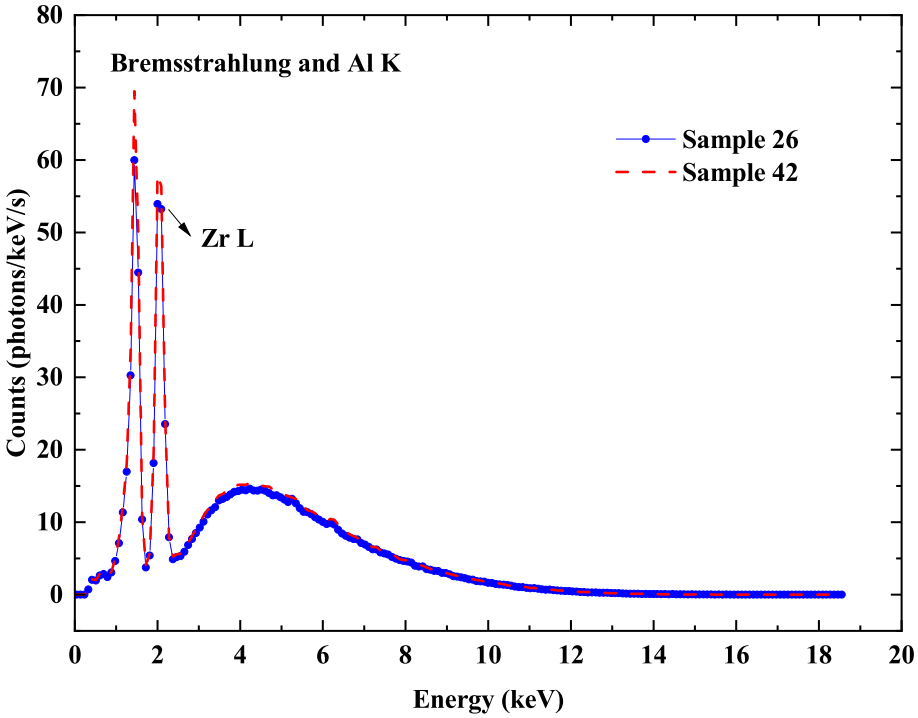

Fig. 9 The experimental BIXS spectra for samples 26 and 42, which were corrected for the signal pile-up effect by employing Monte Carlo method

The well-trained BIXS BP neural network delivers accurate predictions for the parameters (i.e., the zirconium thickness and tritium-to-zirconium ratio) within the training range as well as demonstrates strong prediction performance outside the training range. For the thickness of zirconium, the MREs for the training dataset and test dataset are 0.56 and 0.59%, respectively. For the tritium-to-zirconium ratio, the MREs for the training dataset and test dataset are 0.42% and 0.38%, respectively. For parameters outside the training range, the MREs for the zirconium thickness and tritium-to-zirconium ratio are 7.61% and 9.33%, respectively. The trained BP neural network shows excellent predictive capability across various levels of statistical uncertainty.

The BIXS BP neural network successfully predicted the zirconium thicknesses and tritium-to-zirconium ratios from the experimental X-ray spectra obtained in BIXS experiments using two tritium-containing samples with substrates and uniform tritium distributions, which were in good agreement with the EBS results. This work demonstrates the applicability of the BP neural network in the BIXS method for analyzing thin solid samples with substrates and uniform tritium distributions without the need for prior knowledge of sample thicknesses.

Table 7 Results of zirconium thickness and tritium-to-zirconium ratio obtained by EBS and BP-Network

Samples	Parameters	EBS	BP-network
26	Zirconium thickness (μm)	1.64	1.60
	Tritium-to-Zirconium ratio	1.87	1.85
42	Zirconium thickness (μm)	1.72	1.63
	Tritium-to-Zirconium ratio	1.87	1.91

Acknowledgements The authors are grateful to the staff at the Institute of Nuclear Science and Technology, Sichuan University, for the operation of the 3 MV tandetron accelerator and Wei Ding at the Institute of Nuclear Physics and Chemistry, China Academy of Engineering Physics, for providing help during this study.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Hong Huang, Zhu An and Jing-Jun Zhu. The first draft of the manuscript was written by Hong Huang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data Availability The data that support the findings of this study are openly available in Science Data Bank at https://cstr.cn/31253.11.sciencedb.20902 and https://doi.org/10.57760/sciencedb.20902.

Declarations

Conflict of interest Zhu An is an editorial board member for Nuclear Science and Techniques and was not involved in the editorial review, or the decision to publish this article. All authors declare that there are no Conflict of interest.

References

1. M. Matsuyama, S. Ueda, K. Watanabe, In situ observation of tritium interactions with Pd and Zr by beta-ray induced X-ray

spectrometry. Fusion Eng. Des. **49**, 885–891 (2000). https://doi.org/10.1016/S0920-3796(00)00326-4
2. T. Fu, Z. An, J.J. Zhu et al., Tritium and helium analyses in thin films by enhanced proton backscattering. Chin. Phys. C **38**, 088203 (2014). https://doi.org/10.1088/1674-1137/38/8/088203
3. J.A. Mason, G. Vassallo, Tritium measurement by isothermal calorimetry. Fusion Technol. **21**, 425–429 (1992). https://doi.org/10.13182/FST92-A29782
4. Y. Hatano, M. Hara et al., Measurement of tritium concentration in water by imaging plate. Fusion Eng. Des. **87**, 965–968 (2012). https://doi.org/10.1016/j.fusengdes.2012.02.057
5. M. Matsuyama, K. Watanabe, K. Hasegawa, Tritium assay in materials by the bremsstrahlung counting method. Fusion Eng. Des. **39**, 929–936 (1998). https://doi.org/10.1016/S0920-3796(98)00232-4
6. M. Röllig, T. Aso, M. Hara et al., Galet–Benchmark of a Geant4 based application for the simulation and design of beta induced X-ray spectrometry systems. Fusion Eng. Des. **143**, 91–98 (2019). https://doi.org/10.1016/j.fusengdes.2019.03.086
7. Y. Yang, Z. Chen, R. Wang et al., Effects of tritium 2-D distribution on tritium depth profile reconstruction in BIXS measurements. Fusion Eng. Des. **130**, 142–147 (2018). https://doi.org/10.1016/j.fusengdes.2018.03.034
8. Z.L. Chen, P. Huang, Y. Yang et al., Theoretical investigation of tritium concentration quantification method for DT fuel system using β-ray induced X-rays. Fusion Eng. Des. **184**, 113303 (2022). https://doi.org/10.1016/j.fusengdes.2022.113303
9. Z. An, Q. Hou, J.J. Long, Reconstruction of depth distribution of tritium in materials by β-ray induced X-ray spectrometry. Nucl. Instrum. Methods Phys. Res. B **266**, 3643–3646 (2008). https://doi.org/10.1016/j.nimb.2008.06.020
10. L. Mao, Z. An, Q.Q. Wu et al., Effect of geometrical parameter's uncertainty of BIXS experimental setup for tritium analysis. Nucl. Instrum. Methods Phys. Res. B **289**, 52–55 (2012). https://doi.org/10.1016/j.nimb.2012.08.005
11. M. Matsuyama, T. Tanabe, N. Noda et al., Nondestructive measurement of surface tritium by β-ray induced X-ray spectrometry (BIXS). J. Nucl. Mater. **290**, 437–442 (2010). https://doi.org/10.1016/S0022-3115(00)00581-X
12. L. Mao, Z. An, J.H. Liang et al., Effects of internal bremsstrahlung of tritium β-decay and surface roughness in the BIXS method. Nucl. Instrum. Methods Phys. Res. B **269**, 105–110 (2011). https://doi.org/10.1016/j.nimb.2010.10.024
13. F. Salvat, J.M. Fernandez-Varea, J. Sempau, *Penelope-2008: A code system for Monte Carlo simulation of electron and photon transport* (OECD/NEA Data Bank, Paris, France, 2008)
14. H. Chen, W. Ding, Z. An et al., BIXS for tritium analysis with Ar gas and Al thin film as β-ray stopping layers and comparison with EBS. Radiat. Phys. Chem. **174**, 108931 (2020). https://doi.org/10.1016/j.radphyschem.2020.108931
15. B. Liu, W. Ding, Z. An et al., Tritium analysis in zirconium film with BIXS and EBS: generality test of Al thin film as the β-ray stopping layer in BIXS. Fusion Eng. Des. **172**, 112751 (2021). https://doi.org/10.1016/j.fusengdes.2021.112751
16. H. Huang, Z. An, J.J. Zhu, A semi-analytical model for calculating the X-ray energy spectrum of thin tritium-containing sample in BIXS analysis. IEEE Trans. Nucl. Sci. **72**, 82-92 (2025). https://doi.org/10.1109/TNS.2024.3523106
17. M. Matsuyama, T. Murai, K. Watanabe, Quantitative measurement of surface tritium by β-Ray-Induced X-Ray Spectrometry (BIXS). Fusion Sci. Tech **41**, 505–509 (2002). https://doi.org/10.13182/FST02-A22640
18. M. Matsuyama, Y. Torikai, M. Hara et al., New technique for non-destructive measurements of tritium in future fusion reactors. Nucl. Fusion **47**, S464 (2007). https://doi.org/10.1088/0029-5515/47/7/S09
19. M. Buscema, An brief overview and introduction to artificial neural networks. Subst. Use Misuse **37**, 1093-1148 (2002). https://doi.org/10.1081/JA-120004171
20. Y.Y. Cao, J.Y. Guo, B. Zhou, Predictions of nuclear charge radii based on the convolutional neural network. Nucl. Sci. Tech. **34**, 152 (2023). https://doi.org/10.1007/s41365-023-01308-x
21. L.Y. Zhou, H. Zha, J.R. Shi, A non-invasive diagnostic method of cavity detuning based on a convolutional neural network. Nucl. Sci. Tech. **33**, 94 (2022). https://doi.org/10.1007/s41365-022-01069-z
22. Y.S. Hao, Z. Wu, Y.H. Pu et al., Research on inversion method for complex source-term distributions based on deep neural networks. Nucl. Sci. Tech. **34**, 195 (2023). https://doi.org/10.1007/s41365-023-01327-8
23. S.A. Hosseini, Neutron spectrum unfolding using artificial neural network and modified least square method. Radiat. Phys. Chem. **126**, 75–84 (2016). https://doi.org/10.1016/j.radphyschem.2016.05.010
24. M. El-Sefy, A. Yosri, W. El-Dakhakhni et al., Artificial neural network for predicting nuclear power plant dynamic behaviors. Nucl. Eng. Tech. **53**, 3275–3285 (2010). https://doi.org/10.1016/j.net.2021.05.003
25. P.E. Keller, L.J. Kangas, G.L. Troyer et al., Nuclear spectral analysis via artificial neural networks for waste handling. IEEE Trans. Nucl. Sci. **42**, 709–715 (1995). https://doi.org/10.1109/23.467888
26. R. da, S. Guimarães, T.F. Silva, C.L. Rodrigues et al., Processing of massive Rutherford back-scattering spectrometry data by artificial neural networks. Nucl. Instrum. Methods Phys. Res. B **493**, 28–34 (2021). https://doi.org/10.1016/j.nimb.2021.02.010
27. G. Magchiels, C.B. Mtshali, L. Kotsed et al., A machine learning approach to self-consistent RBS data analysis and combined uncertainty evaluation. Nucl. Instrum. Methods Phys. Res. B **551**, 165354 (2024). https://doi.org/10.1016/j.nimb.2024.165354
28. D.D. Cohen, J. Crawford, Machine learning techniques to determine elemental concentrations from raw IBA spectra. Nuclear Instr. Methods Phys. Res. Sect. B: Beam Interact. Mater. Atoms. **546**, 165169 (2024). https://doi.org/10.1016/j.nimb.2023.165169
29. G. Magchiels, N. Claessens, J. Meersschaut et al., Enhanced accuracy through machine learning-based simultaneous evaluation: a case study of RBS analysis of multinary materials. Sci. Rep. **14**, 8186 (2024). https://doi.org/10.1038/s41598-024-58265-7
30. D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors. Nature **323**, 533–536 (1986). https://doi.org/10.1038/323533a0
31. L. Tang, S. Zhou, K.B. Shi et al., Application of a neural network model with multi-model fusion for fluorescence spectroscopy. Nucl. Sci. Tech. **35**, 178 (2024). https://doi.org/10.1007/s41365-024-01528-9
32. B.C. Wang, C.X. Tang, M.T. Qiu et al., A machine learning approach to TCAD model calibration for MOSFET. Nucl. Sci. Tech. **34**, 192 (2024). https://doi.org/10.1007/s41365-023-01340-x
33. G. Cybenko, Approximation by superpositions of a sigmoidal function. Math. Control Signals Syst. **2**, 303–314 (1989). https://doi.org/10.1007/BF02551274
34. H.L. Zheng, X.G. Tuo, S.M. Peng et al., Determination of gamma point source efficiency based on a back-propagation neural network. Nucl. Sci. Tech. **29**, 61 (2018). https://doi.org/10.1007/s41365-018-0410-4
35. K. Hornik, M. Stinchcombe, H. White, Multilayer feed forward networks are universal approximators. Neural Netw. **2**, 359–366 (1989). https://doi.org/10.1016/0893-6080(89)90020-8
36. J.D. Paola, R.A. Schowengerdt, A review and analysis of back propagation neural networks for classification of remotely-sensed multi-spectral imagery. Int. J. Remote Sens. **16**, 3033–3058 (1995). https://doi.org/10.1080/01431169508954607

37. H.R. Vega-Carrillo, V.M. Hernández-Dávila, E. Manzanares-Acuña et al., Neutron spectrometry using artificial neural networks. Radiat. Meas. **41**, 425–431 (2006). https://doi.org/10.1016/j.radmeas.2005.10.003
38. Y.Q. Li, B.J. Zhu, Y. Lv et al., Artificial neural network algorithm for pulse shape discrimination in 2α and 2β particle surface emission rate measurements. Nucl. Sci. Tech. **34**, 153 (2013). https://doi.org/10.1007/s41365-023-01305-0
39. D. Lian, P. Gong, X. Tang et al., Rapid nuclide identification algorithm based on convolutional neural network. Ann. Nucl. Energy **133**, 483–490 (2019). https://doi.org/10.1016/j.anucene.2019.05.051
40. C. Zhao, W. Jin, Y. Shi et al., Utilizing BP neural networks to accurately reconstruct the tritium depth profile in materials for BIXS. Nucl. Sci. Tech. **36**, 15 (2025). https://doi.org/10.1007/s41365-024-01569-0
41. X. Llovet, F. Salvat, *PENEPMA-2018, a Monte Carlo code for the simulation of x-ray emission spectra using PENELOPE* (Universitat de Barcelona, Spain, 2020)
42. S.M. Seltzer, M.J. Berger, Bremsstrahlung energy spectrum from electrons with kinetic energy 1 keV-10 GeV incident on screened nuclei and orbital electrons of neutral atoms with Z=1-100. Atomic Data Nucl. Data Tables **35**, 345–418 (1986). https://doi.org/10.1016/0092-640X(86)90014-8
43. A. Omar, P. Andreo, G. Poludniowski, A model for the energy and angular distribution of x rays emitted from an x-ray tube. Part I. Bremsstrahlung production. Med. Phys. **47**, 4763–4774 (2020). https://doi.org/10.1002/mp.14359
44. L. Kissel, C.A. Quarles, R.H. Pratt, Shape functions for atomic-field bremsstrahlung from electrons of kinetic energy 1–500 *keV* on selected neutral atoms $1 < Z < 92$. Atomic Data Nucl. Data Tables **28**, 381–460 (1983). https://doi.org/10.1016/0092-640X(83)90001-3
45. J.K. Knipp, G.E. Uhlenbeck, Emission of gamma radiation during the beta decay of nuclei. Physica **3**, 425–439 (1936). https://doi.org/10.1016/S0031-8914(36)80008-1
46. F. Bloch, On the continuous γ-radiation accompanying the β-decay. Phys. Rev. **50**, 272–278 (1936). https://doi.org/10.1103/PhysRev.50.272
47. H. Zhang, Z. An, J.J. Zhu et al., Bremsstrahlung of 5–25 *keV* electrons incident on $MoSi_2$, TiB_2 and ZrB_2 thick solid conductive compounds. Nucl. Instrum. Methods Phys. Res. B **553**, 165407 (2024). https://doi.org/10.1016/j.nimb.2024.165407
48. S. Imambi, K.B. Prakash, G.R. Kanagachidambaresan, *Programming with TensorFlow: Solution for edge computing applications* (Springer, India, 2021), pp.87–104
49. H. Park, S.H. Park, H.B. Kong et al., Weighted sum MSE minimization under per-BS power constraint for network MIMO systems. IEEE Commun. Lett. **16**, 360–363 (2012). https://doi.org/10.1109/LCOMM.2012.010512.112300
50. Z. Zhuang, M. Liu, A. Cutkosky et al., Understanding AdamW through proximal methods and scale-freeness. arXiv:2202.00089 (2022). https://doi.org/10.48550/arXiv.2202.00089
51. D. Yarotsky, Error bounds for approximations with deep ReLU networks. Neural Netw. **94**, 103–114 (2017). https://doi.org/10.1016/j.neunet.2017.07.002
52. T.O. Hodson, Root means square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci. Model Dev. Discuss. **15**, 5481-5487 (2022). https://doi.org/10.5194/gmd-15-5481-2022
53. J. Han, C. Moraga, The influence of the sigmoid function parameters on the speed of backpropagation learning. Int. Workshop Artif. Neural Networks, Berlin, Heidelberg: Springer Berlin Heidelberg, **930**, 195–201 (1995)
54. E. Fan, Extended tanh-function method and its applications to nonlinear equations. Phys. Lett. A **277**, 212–218 (2000). https://doi.org/10.1016/S0375-9601(00)00725-8
55. H.L. Zhang, W. Ding, R.R. Su et al., Depth profiles of D and T in Metal-hydride films up to large depth. Nucl. Instrum. Methods Phys. Res. B **371**, 174–177 (2016). https://doi.org/10.1016/j.nimb.2015.11.030
56. J.F. Han, Z. An, G.Q. Zheng et al., An ion beam facility based on a 3 MV tandetron accelerator in Sichuan University. Nucl. Instrum. Methods Phys. Res. B **418**, 68–73 (2018). https://doi.org/10.1016/j.nimb.2018.01.002
57. M. Mayer, *SIMNRA User's Guide, Report IPP 9/133* (Max-Planck-Institut für Plasmaphysik, Garching, Germany, 1997)
58. J.J. Long, Z. An, Comparison of reconstruction methods of depth distribution of tritium in materials based on BIXS. Nucl. Instrum. Methods Phys. Res. B **267**, 1852–1855 (2009). https://doi.org/10.1016/j.nimb.2009.02.066
59. W.G. Zhang, H.W. Sun, F.Y. Zeng et al., Tritium analysis in titanium films by the BIXS method. Nucl. Instrum. Methods Phys. Res. B **275**, 20–23 (2012). https://doi.org/10.1016/j.nimb.2012.01.003