Application of small sample BP neural network in quantitative analysis of EDXRF

Fei Li\textsuperscript{a, b}, Zhuoyao Tang\textsuperscript{b}, Liangquan Ge\textsuperscript{a, b,*}, Nanxing Wu\textsuperscript{b}, Feng Cheng\textsuperscript{b}, Kun Sun\textsuperscript{b}

\textsuperscript{a} Applied Nuclear Technology in Geosciences Key Laboratory of Sichuan Province, China
\textsuperscript{b} College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, China

* Corresponding author and e-mail: Liangquan Ge, glq@cdut.edu.cn

Abstract. Quantitative analysis of EDXRF is affected by matrix effect, randomness of nuclear radiation, interaction of elements, statistical fluctuation of radiation detection process, etc. Its algorithm needs to consider many factors, and the established functional relationship is often a complex non-linear function. Therefore, effective quantitative analysis method has always been the key research direction of spectral resolution technology. In this paper, we use MATLAB software to study the effect of BP neural network in quantitative analysis of EDXRF by establishing the non-linear relationship between the counting rate of each element and the content of a single element in a small sample. The experimental results show that the small sample neural network can establish a stable structure and be applied to quantitative analysis of EDXRF, but the number of samples restricts the accuracy of prediction results, most of which can only guarantee 20% to 30% relative error.

1. Introduction
In EDXRF quantitative analysis, the sample is stimulated by X-ray to produce characteristic X-ray and scattering ray. These characteristic X-rays and scattered rays enter the detector directly. Through the pre-circuit, linear pulse amplifier, single-channel or multi-channel pulse amplitude analyzer, the energy of various rays and their corresponding nuclides or elements are determined. The signals generated by certain energy rays are analyzed to determine the activity or content of nuclides. Therefore, the objective of quantitative analysis is achieved. In this process, spectral analysis techniques are needed, including spectral smoothing, peak searching, peak location determination, element identification, peak boundary path determination, background subtraction, peak area calculation, overlapping peak decomposition and content calculation. In 2019, Li Fei et al. [1] reviewed the related algorithms of spectral smoothness, background subtraction and overlapping peak decomposition in XRF spectral analysis technology. In EDXRF quantitative analysis, there are many elements in the sample to be measured, among which the elements except the elements to be measured are collectively called matrix. Matrix effect can lead to considerable errors in EDXRF analysis. The development of matrix correction can be divided into three stages: the research period of empirical correction equation, the establishment and maturity period of basic parameter method and theoretical $\alpha$ coefficient algorithm, and the development period characterized by automation and intellectualization.
At present, the existing matrix correction methods mainly include: theoretical influence coefficient method, basic parameter method, empirical coefficient method, experimental non-standard sample calculation, experimental correction method and so on. But these algorithms are complex and depend on a specific system. For different analysis objects, different algorithms have different analysis accuracy, thus introducing different errors.

Psychologist W. S. McCulloch and mathematical logician W. Pitts first proposed the mathematical model of simulating human neurons (MP model) in 1943, which opened the age of studying artificial neural networks. Artificial neural network does not need to consider the complex relationship between the concentration and strength of the elements to be measured. In recent 20 years, it has been widely used in EDXRF spectral data processing algorithms. Guo Panlin et al. [3] quantitatively analyzed the X-ray energy spectrum of Pt-Pd alloy by using artificial neural network and least square regression method respectively. The relationship between sample concentration and strength was obtained. The fitting degree of neural network was 0.1. Compared with the traditional least square method, the accuracy of neural network in non-linear fitting was relatively higher. In 2012, Xu Lipeng et al. [4] proposed an improved BP neural network model based on principal component analysis (PCA). The model was used to quantitatively analyze Fe and Ti elements in unknown geological samples and compared with chemical analysis. The experimental results show that the relative error between the predicted value and the chemical analysis value of the improved BP neural network is less than 3%, and the BP neural network has a good prediction effect. In 2015, Wang Jun et al. [5] proposed a hybrid algorithm of genetic algorithm (GA) to optimize BP neural network (GA-BP). The content of titanium and iron in five types of ore samples in Panzhihua mining area was quantitatively analyzed through overall prediction and classification prediction experiments. The experimental results show that the relative error of 76.7% samples is less than 2% by comparing the predicted value of GA-BP neural network with the chemical analysis value, which shows the effectiveness of this method in the correction of matrix effect. Principal Component Analysis (PCA), as a chemometric method, has a good effect in the application of EDXRF, but PCA's ability to deal with non-linear relations and generalization is inferior to that of neural networks. In 2017, J. J. Okonda et al. [6] compared PCA with BP network when using EDXRF to analyze trace bio-metals in soft tissues. In the experiment, the R2 value of the neural network ranges from 0.890 to 0.941, and the R2 value of PCA ranges from 0.680 to 0.871, which shows that the generalization ability of the neural network is better than that of PCA.

The accuracy of the results of the neural network strongly depends on the training samples. The larger the amount of data, the more accurate the prediction results are. However, in the actual production, there are occasions when the number of samples is insufficient. At this time, the evaluation of the reliability of the neural network is a key problem. This paper mainly studies the prediction ability and accuracy evaluation of small sample BP neural network in EDXRF.

2. Material and Method

2.1. The Principle of Neural Network

BP neural network is a kind of feed-forward neural network with error back-transferring (see Figure 1). BP network establishes the relationship between input nodes and output nodes through initial weights and thresholds, and obtains the initial output of the network through the forward propagation of data along the network. The specific steps are as follows:

1. Input: \( x_1, x_2, x_3, \ldots, x_n \) are \( n \) inputs for neurons, while the total input was expressed by \( X = [x_1, x_2, x_3, \ldots, x_n]^T \).

2. Network weights and thresholds: The network weights and thresholds can be adjusted in the neural network, which is also an important part of the characteristic regulation of the neural network, and plays a key role in the training and learning process of the neural network. \( w_1, w_2, w_3, \ldots, w_n \) represents the weight of the network representing the proportion of the input; \( b \), a scalar of \( 1^*1 \), represents the threshold representing a network weight whose input is constant to 1.
\[ \text{net} = \sum_{k=1}^{n} x_k w_{i,k} + b \]

③ Summation unit: Weighted summation of input signals.

④ Transfer function: As the main way of neuron processing, transfer function calculates the function of sum and unit to get the output of neuron.

⑤ Output: The final output is processed by weighted summation and transfer function.

Then, on the premise of defining the expected output, the network corrects the weights and thresholds layer by layer according to the error function between the calculated value and the expected value, so that the error function decreases along the negative gradient. Learning stops until the network error is less than the expected error or the maximum number of training times is reached. The network will record the current weight matrix and threshold matrix for effective calculation and prediction of data to be processed. [7]

![BP neural network diagram](image)

**Figure 1.** BP neural network

### 2.2. Experimental Design

The purpose of this experiment is to study the effect of small sample BP neural network in dealing with non-linear problems in EDXRF quantitative analysis. Using BP neural network, the non-linear relationship between the counting rate of each element and the content of a single element is established. The function relationship that needs to be determined can be expressed as follows:

\[ C_i = f(w_1A_i + b_1) + f(w_2A_2 + b_2) + \ldots + f(w_nA_n + b_n) \]

In the above formula, \( C_i \) represents the content of an element, \( A_i \) represents the counting rate of an element, \( w \) and \( b \) represents the weight and threshold respectively. \( f(\cdot) \) represents the activation function. The tangent S-type function is used in this experiment; \( n \) represents the number of elements to be processed, and the number of experimental data is 6.

In this experiment, three neural network models of target elements (Fe, Cu, Pb) were established to determine the relationship between element content and the counting rate of each element. Input experimental samples for training, the network adjusts and corrects the weights and thresholds of the above function relations continuously until the error reaches below the preset error, then stop training. At this time, BP neural network can be identified as stable, and error analysis and prediction can be carried out. The specific steps of the experiment are as follows:

1. The EDXRF data are sorted and grouped for training set samples of neural network training, including the counting rate and standard content data of 15 groups of samples of 6 elements. The prediction set after network training is detected, including three sets of counting rate data of six elements and corresponding actual content data, for comparison with the predicted output value.
The BP neural network is constructed by MATLAB. The counting rate data of each reference element is taken as input vector, the content data is taken as target output, the initial parameters of the network are set, and the training of the neural network is carried out.

After the network training is completed, the element counting rate imported into the prediction set is compared with the actual content data, and the error is analyzed and evaluated.

Fifteen groups of standard counting rates and standard content data of six elements Zn, As, Fe, Ti, Cu and Pb were selected as training set samples of the neural network before using MATLAB to construct the neural network (see Table 1 and Table 2).

**Table 1.** Counting rate of each element in training sample

| Number | Zn    | As    | Fe    | Ti    | Cu    | Pb    |
|--------|-------|-------|-------|-------|-------|-------|
| 1      | 3.87634 | 0.376302 | 199.647 | 0.333595 | 3.31702 | 9.57401 |
| 2      | 5.36762 | 0.077832 | 126.214 | 0.292757 | 1.74736 | 10.1929 |
| 3      | 3.3076  | 0.159167 | 76.1615  | 0.443485 | 1.44727 | 22.2033 |
| 4      | 1.36034 | 0.17394 | 404.633 | 0.22188 | 2.65149 | 4.89463 |
| 5      | 0.7455  | 1.64159 | 483.343 | 0.00222 | 6.61986 | 2.35262 |
| 6      | 1.03209 | 4.69056 | 320.782 | 0.318703 | 5.19784 | 2.64562 |
| 7      | 16.5059 | 0.313031 | 144.679 | 0.132521 | 1.83673 | 11.1681 |
| 8      | 10.3162 | 0.292433 | 167.151 | 0.165875 | 1.86819 | 10.0682 |
| 9      | 6.22    | 0.483652 | 149.821 | 0.324392 | 1.59367 | 11.9402 |
| 10     | 3.17609 | 0.399014 | 153.92  | 0.39988 | 1.62931 | 6.7775 |
| 11     | 13.2366 | 0.414737 | 168.244 | 0.050899 | 2.70346 | 12.0419 |
| 12     | 11.3321 | 0.591143 | 142.52  | 0.259835 | 0.896827 | 10.3352 |
| 13     | 5.46654 | 0.231044 | 203.736 | 0.457959 | 1.88377 | 8.35871 |
| 14     | 3.64284 | 0.41217 | 246.55  | 0.053573 | 3.30441 | 6.18569 |
| 15     | 1.7316  | 0.518939 | 213.743 | 0.43542 | 3.17259 | 5.31308 |
Table 2. Element Content of Training Sample

| Number | Zn   | As   | Fe   | Ti   | Cu   | Pb   |
|--------|------|------|------|------|------|------|
| 1      | 1.72 | 21   | 3.63 | 20.4 | 98   | 155  |
| 2      | 2.36 | 16.3 | 2.46 | 19.4 | 20   | 187  |
| 3      | 1.27 | 11.4 | 1.40 | 12   | 26   | 380  |
| 4      | 0.26 | 40   | 7.21 | 64   | 58   | 77   |
| 5      | 0.1  | 144  | 8.83 | 40   | 552  | 42   |
| 6      | 0.16 | 390  | 5.66 | 53   | 314  | 39   |
| 7      | 8.27 | 24.3 | 3.14 | 31.5 | 21   | 236  |
| 8      | 5    | 25   | 3.36 | 33   | 25   | 172  |
| 9      | 2.62 | 19   | 2.92 | 26   | 22   | 226  |
| 10     | 1.33 | 21.4 | 2.95 | 25.4 | 24.7 | 182  |
| 11     | 5.83 | 29   | 3.30 | 32   | 19   | 240  |
| 12     | 5    | 21.6 | 2.88 | 28.5 | 21.6 | 195  |
| 13     | 2.45 | 27.4 | 3.72 | 33   | 31   | 152  |
| 14     | 1.53 | 37   | 4.51 | 41   | 38   | 115  |
| 15     | 0.4  | 32   | 3.81 | 27.4 | 61   | 68   |

The counting rate of another 10 groups of samples is selected as the prediction set of the neural network (see Table 3). The output of the network operation is compared with the real value, and the error magnitude and reason are analyzed.

Table 3. Element Counting Rate of Prediction Set

| Number | Zn         | As         | Fe         | Ti         | Cu         | Pb         |
|--------|------------|------------|------------|------------|------------|------------|
| 1      | 3.56897    | 0.054897*  | 236.4      | 0.23597    | 2.46358    | 4.89463    |
| 2      | 2.65847    | 0.077832   | 135.68     | 0.5948*    | 1.5798     | 3.6897     |
| 3      | 1.5987     | 0.19894    | 89.564     | 0.0857     | 1.44727    | 21.568     |
| 4      | 1.2867     | 0.15896    | 386.59     | 0.18697    | 1.4589     | 3.5897     |
| 5      | 0.7455     | 1.64159    | **498.6**  | 0.00222    | 6.61986    | 2.35262    |
| 6      | **0.6359** | **5.1648** | 331.548    | 0.218703   | **6.92649**| **2.03256**|
| 7      | 14.5889    | 0.28978    | 120.265    | 0.132521   | 1.76458    | 10.5898    |
| 8      | 8.6976     | 0.189747   | 145.654    | 0.12354    | 1.86819    | 10.0682    |
| 9      | 5.49       | 0.35978    | 126.56     | 0.35947    | 1.2564     | 10.2659    |
| 10     | 2.8975     | 0.2945     | 112.56     | 0.37584    | 1.2359     | 5.4777     |

The data with a * sign in the table indicates that the value is outside the training sample range.
3. Results and Discussions

The experiment uses MATLAB software to realize the neural network operation, imports the data of training set samples, takes the standard counting rate of each element as Input Data, and takes the standard content as Target Data. Then the neural network model is constructed for each element quantitative analysis in turn. In the network window, the type of neural network is selected as Feed-forward BackProp (BP Neural Network). Only one hidden layer, one input layer and one output layer are set. The neurons in the hidden layer are set to eight. An input and target are selected. The activation function is selected as TANSIG (tangent S function), and the neural network model is generated (see figure 2).

![Neural Network Model](image)

**Figure 2. Neural Network Model**

Some network parameters can be initialized by train window. Through repeated training experiments, the training step length is 1000, the learning time is 120 seconds, the maximum error number is 6, and the weight change rate is 0.001. At this time, the training effect is the best.

When the network is trained, the input data (i.e. counting rate data) are classified. Most of the data are used as training group for network training and learning, and the functional relationship between counting rate and content is grasped. A small number of data are extracted as test and validation groups, which are used to test and verify the functional relationship after the network completes training. Finally, the total (all) training data are fitted, and the operating chart and fitting chart of Fe, Cu and Pb are obtained respectively.
Figure 3. (a) Operating Chart of Fe; (b) Fitting Chart of Fe; (c) Operating Chart of Cu; (d) Fitting Chart of Cu; (e) Operating Chart of Pb; (f) Fitting Chart of Pb;

The prediction set data of Fe, Cu and Pb are simulated in turn and compared with the actual data to make a comparison error table. The results are as follows:
Table 4. Data Error Table of Fe, Cu, Pb Prediction Set

| Number | Fe Prediction Error Analysis Table | Cu Prediction Error Analysis Table | Pb Prediction Error Analysis Table |
|--------|-----------------------------------|-----------------------------------|-----------------------------------|
|        | absolute error | absolute error | absolute error | absolute error | absolute error | relative error |
| 1      | -0.90435        | 15.3019           | 37.9205          | 37.9205          | -0.90435        | -0.1957        |
| 2      | -0.75999        | 15.6572           | 3.3605           | 3.3605           | -0.75999        | -0.2969        |
| 3      | -0.45971        | -9.5444           | 6.6806           | 6.6806           | -0.45971        | -0.2873        |
| 4      | 0.045206        | 10.1086           | 12.5479          | 12.5479          | 0.045206        | 0.0067         |
| 5      | 0.79023         | -1.9446           | 0.0080129        | 0.0080129        | 0.79023         | 0.0827         |
| 6      | -4.3299         | -11.0009          | -265.4928        | -265.4928        | -4.3299         | -1.0260        |
| 7      | -0.28057        | 1.2599            | -3.0464          | -3.0464          | -0.28057        | -0.1087        |
| 8      | -0.32935        | 31.4553           | -2.1854          | -2.1854          | -0.32935        | -0.1272        |
| 9      | -0.91627        | -108.1182         | -2.0254          | -2.0254          | -0.91627        | -0.5763        |
| 10     | 0.53373         | 105.1649          | -1.1202          | -1.1202          | 0.53373         | 0.2765         |

From the error analysis table, it can be seen that the prediction errors fluctuate between 0.6% and 60%. There are five groups of data errors below 20%, three groups of data errors above 30%, and two groups of data errors between 20% and 30%. The overall accuracy of Pb element content prediction is general. From the error analysis table, it can be seen that the prediction error fluctuates between 0% and 90%. Six sets of data errors are less than 20%. Three sets of data errors are more than 30%. Only one set of data errors is between 20% and 30%. The overall accuracy of Fe element content prediction is relatively poor. From the error analysis table, it can be seen that the prediction error fluctuates between 0% and 90%. Six sets of data errors are less than 20%. Three sets of data errors are more than 30%. Only one set of data errors is between 20% and 30%.

In the operation of neural network, the overgeneralization of input data may lead to the error of final prediction results. This overgeneralization refers to the upper and lower bounds of the training set data used in the training process of the neural network. When the network training is finished and the new data is predicted and simulated, if the input data is not within the upper and lower bounds of the training set, the generalization problem will arise. If this generalization problem occurs on multiple input data, the accuracy of the final prediction results of the network will be greatly reduced. The effect of neural network prediction and simulation depends on the range of original data. The neural network used in this experiment has some limitations in the application of actual prediction and measurement data.

4. Conclusion
In this paper, the quantitative analysis of EDXRF based on small sample BP neural network algorithm is studied. Through experiments and data analysis, it is considered that the application of neural network in quantitative analysis of EDXRF can be realized and has certain advantages. However, when the number of samples is relatively small, there are some drawbacks. Because the amount of data involved in fitting is insufficient, it is easy to lose some of the original features of the data after training. In addition, small samples can easily lead to overgeneralization of the neural network, resulting in great error in experimental prediction. In practical applications, neural networks often rely too much on the data to be processed. Although small samples can reduce the time of data pre-processing, they can not guarantee the accuracy of analysis in EDXRF.
Acknowledgement
This work was funded by the National Natural Science Foundation of China (No.41774147), the National Key R&D Project (No.2017YFC0602100) and Science and technology activity project for overseas personnel in Sichuan province.

References
[1] Li Fei, Ge Liangquan, etc. 2019. Recent developments on XRF spectra evaluation. Applied Spectroscopy Reviews.
[2] Gi ang, Zhuo Shangjun, Le Guohui. 2011. Energy dispersive X-ray fluorescence spectroscopy. Science Press.
[3] Guo Panlin, Wang Jiqing, etc. 1999. Application comparison of artificial neural network and least squares regression in quantitative analysis of XRF. Nuclear technology. P725-730.
[4] Xu Lipeng, Ge Liangquan, etc. 2013. Application of EDXRF analysis based on PCA-BP neural network in determination of iron and titanium in geological samples. Spectroscopy and spectral analysis. P1392-1396.
[5] Wang Jun, Liu Mingzhe, etc. 2015. BP neural network optimized by genetic algorithm is used to predict the content of ferrotitanium in EDXRF. Atomic energy science and technology. P1143-1148.
[6] Okonda, J.J., Angeyo, K. H.. A nested multivariate chemometrics based calibration strategy for direct trace biometal analysis in soft tissue utilizing Energy Dispersive X-Ray Fluorescence (EDXRF) and scattering spectrometry. Applied Radiation and Isotopes. P49–56.
[7] Yang Huijun, Feng Keming, etc. 2019. Improved GPT2w model based on BP neural network and its global precision analysis. Systems Engineering and Electronic Technology. P500-508.