Research Article

Chaos-generalized regression neural network prediction model of mine water inflow

Jianlin Li[1,2] · Luyang Wang[1] · Xinyi Wang[1,2] · Peiqiang Gao[1]

Received: 5 September 2021 / Accepted: 26 October 2021
Published online: 08 November 2021

Abstract
Artificial neural network (ANN) provides a new way for mine water inflow prediction. However, the effectiveness of prediction using ANN model would not be guaranteed if the influencing factors of water inflow are difficult to quantify or there are only a few observation data. Chaos theory can recover the rich dynamic information hidden in time series. By reconstructing water inflow time series in phase space, the multi-dimensional matrix could be obtained, with each column representing an influencing factor of water inflow and its value representing the change of the influencing factor with time. Therefore, a new prediction model of mine water inflow can be established by combining ANN with chaos theory when lacking data on the influencing factors of water inflow. In the present study, the No. 12 coal mine of Pingdingshan China was selected as the study site. The Chaos-GRNN model and Chaos- BPNN model of mine, water inflow were established by using the water inflow data from February 1976 to December 2013. The model was verified by using the water inflow values in the 24 months from 2014 to 2015. The number embedded dimension (*M*) of influencing factors of water inflow determined by phase space reconstruction was 7, meaning that there were 7 influencing factors of water inflow and 7 neurons in GRNN input layer, and the time delay was 13 months. The value of GRNN input layer neurons was determined accordingly. The maximum Lyapunov index was 0.0530, and the prediction time of GRNN was 19 months. The two models were evaluated by using four evaluation indices (*R*, RMSE, MAPE, NSE) and violin plot. It was found that both models can realize the long-term prediction of water inflow, and the prediction effectiveness of Chaos-GRNN model is better than that of Chaos-BPNN model.

Keywords Chaos theory · Phase space reconstruction · Generalized regression neural network · Input layer neuron · Long-term prediction of mine water

1 Introduction

In China, more than 200 water inrushes from Ordovician limestone have occurred in North China coalfields over the past 40 years, resulting in economic losses of more than 30 billion yuan and 1300 deaths [22, 25].Water inrush problems in deep mining become increasingly serious as the mining depth increases. Mine water inflow is one of basic technical conditions for coal mine development, which is the total amount of surface water, fissure water, karst water, old kiln water, and other water flowing into mine pit. Accurately forecast the water inflow of coal mines can help to prevent coal mine accidents and ensure the safety of coal mine production and efficient use of mine water [34, 40]. Therefore, scholars have gradually developed the prediction of mine water inflow from deterministic methods (such as hydrogeological analytical method, water balance method, numerical methods, etc.)

✉ Jianlin Li, ljl.wy@163.com | [1]Institute of Resources & Environment, Henan Polytechnic University, Jiaozuo 454000, China. [2]Collaborative Innovation Center of Coal Work Safety and Clean High Efficiency Utilization, Jiaozuo 454000, Henan Province, China.

to non-deterministic stochastic prediction (such as analog method, fuzzy mathematical model, gray system theory, etc.) [1, 24, 33, 41]. However, mine water inflow is related to many factors, such as strata pressure, mining disturbance, geologic structures, and water pressure, and these factors have obvious nonlinear characteristics [11, 16]. Moreover, with the increasing depth of coal mining, the hydrogeological conditions of mine water filling are becoming more and more complex, and the control factors of mine water inflow are increasing. Consequently, these prediction methods are usually time-consuming and less reliable.

Artificial neural network (ANN) model is able to learn the underlying relationship between input and output signals of a sequential process with no need to take explicit physical rules into onsideration [18]. So many scholars use ANN to predict mine water onflow [3], and these models have higher prediction accuracy. Generalized Regression Neural Network (GRNN) is a radial basis function network based on mathematical statistics [5]. It has the advantages of simple model structure, few parameters to adjust, strong approximation ability, fast learning speed, and high prediction accuracy, which has been widely used in many fields [5, 36]. Wang et al. [31] selected three main influencing factors of mine water inflow as the input neurons in GRNN to predict the mine water inflow. Using GRNN to predict the mine water inflow, the numbers of input layer neurons and their values are the key to determining the accuracy of the model, which are determined by the choice of the influencing factors of mine water inflow and their values. However, the data of many influencing factors of mine water inflow usually are difficult to collect and quantify, moreover, it is difficult to determine the key factors among all the influencing factors of mine water inflow.

Fortunately, chaos theory considers that there are a lot of information about related factors in the time series of mine water inflow, and the number and value of related factors can be obtained through the reconstructed phase space [19, 29]. Chaos theory is an effective tool for studying complex systems [13], which can restore the rich dynamic information hidden in a single variable of the system [27].

So, this study combined chaos theory with GRNN, using reconstructed phase space of mine water inflow time series to determine the number of GRNN input layer neurons and their values (reducing the subjectivity and limitations of the original GRNN), establishing Chaos-Generalized Regression Neural Network (Chaos-GRNN) model to achieve a more effective prediction model of mine water inflow. Additionally, this model is used to predict the mine water inflow for a long time (24 months) to verify practicality of the model.

The paper is organized as follows. Section 2 briefly introduces the basic theory behind reconstructed phase space, and GRNN, outlines the proposed Chaos-GRNN model approach. Section 3 describes the application of Chaos-GRNN model to predict mine water inflow. Section 4 discusses the rationality, the range of parameters of the Chaos-GRNN model, and the comparison with the prediction results of the Chaos-BPNN model. Section 5 lists the main conclusions.

2 Methodology

2.1 Reconstructed phase space

Takens' time-delay embedding theorem (1981) paved the way for the analysis of chaotic time series. In the reconstructed m-dimensional space, the regular trajectory of the motive system can be restored [8, 10]. The key to reconstruct phase space is to determine the time delay τ and the embedding dimension m. In the paper, the embedding dimension m was determined by the Cao method [2], the time delay τ was determined by the autocorrelation function method [7, 38].Further details about chaos theory can be found in Karunasingha et al. [13] and Huang et al. [12].

(1) Time delay τ

Let x_i denotes the observed mine water inflow, with $\overline{x}$ denoting their means, and n denoting the number of data points considered, τ denoting the number of time moving. Autocorrelation functiont $C(\tau)$ can be expressed as

$$C(\tau) = \frac{\frac{1}{n-\tau}\sum_{i=1}^{n-\tau}(x_i - \overline{x})(x_{i+\tau} - \overline{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{1}$$

The curve draws with τ as the independent variable and $C(\tau)$ as the dependent variable. When $C(\tau)$ drops to (1–1/e) of the initial value, the corresponding τ is the time delay of reconstructing space.

(2) Embedding dimension m

The distance change $\alpha(i, m)$ of the nearest point in phase space under different embedding dimensions is calculated by formula (2) firstly; then the mean $E(m)$ of $\alpha(i, m)$ is calculated by formula (3), $E_1(m)$ is calculated by formula (4). Finally, to draw the curve between $E_1(m)$ and m. When the change of $E_1(m)$ gradually becomes stable, the m at the stable position is what is required.

$$\alpha(i,m) = \frac{\left\|Y_{\eta(n)} - Y_n\right\|_{\infty}^{(m+1)}}{\left\|Y_{\eta(n)} - Y_n\right\|_{\infty}^{m}} \tag{2}$$

where $\|\|_{\infty}^{(m+1)}$ is the $m+1$ dimensional space-norm; Y_n is the closest vector.

$$E(m) = \frac{1}{n - m\tau} \sum_{i=1}^{n-m\tau} \alpha(m,i) \tag{3}$$

$$E_1(m) = \frac{E(m)}{E(m+1)} \tag{4}$$

(3) Reconstructed phase space

After τ and m are determined, phase space reconstruction is performed by the mine water inflow time series $X=\{x_1, x_2,\ldots,x_n\}$. Getting a m-dimensional vector sequence:

$$Y_i = \left\{x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}\right\} \tag{5}$$

where $i=1,2, \ldots, M$, $M=n-(m-1)\tau$.

2.2 *The maximum Lyapunov index λ_{max}*

The maximum Lyapunov exponent $\lambda_{\max}$ not only is the identification parameter of the chaotic feature (if $\lambda_{\max}>0$, then the system has chaotic characteristics), but also can determine the step size R ($R=1/\lambda_{\max}$) of prediction model [30, 32].

After reconstructing the phase space, to calculate the distance (L_i) between two adjacent points $Y(j)$ and $Y(i)$in the m-dimensional space by the formula (6):

$$L_i = \min\left[\|Y(j) - Y(i)\|\right], i,j = 1,2,\ldots,M \tag{6}$$

Then, $\lambda_{\max}$ can be obtained by the formula (7):

$$\lambda_{\max} = \frac{1}{M-1}\frac{1}{\tau}\sum_{i=1}^{M-1} \ln\frac{L_{i+1}}{L_i} \tag{7}$$

The step size R of prediction model can be obtained by the formula (8):

$$R = 1/\lambda_{\max} \tag{8}$$

2.3 Generalized regression neural networks (GRNN)

Generalized regression neural networks (GRNN) were first proposed by Specht [26], which was a radial basis function network [5].

The function $f(x, y)$ is defined as the joint probability density function (PDF) of two random variables x and y. When $f(x, y)$ is known, the condition mean of y on a given x_0 can be calculated by

$$\hat{y}(x_0) = E(y|x_0) = \frac{\int_{-\infty}^{+\infty} yf(x_0,y)\mathrm{d}y}{\int_{-\infty}^{+\infty} f(x_0,y)\mathrm{d}y} \tag{9}$$

In fact, $f(x, y)$ is usually unknown in many systems. GRNN uses a group of measured x and y values to estimate $f(x, y)$ based on the consistent estimators proposed by Parzen [26, 36]. For a given measured sample data point $\{x_i, y_i|I=1, 2,\ldots,N\}$, these timated PDF can be expressed as

$$f(x_0,y) = \frac{1}{n(2\pi)^{\frac{m+1}{2}}\sigma^{m+1}} \sum_{i=1}^{n} e^{-d(x_0,x_i)}e^{-d(y_0,y_i)} \tag{10}$$

$$d(x_0,x_i) = (x_0 - x_i)^T(x_0 - x_i)/2\sigma^2 \tag{11}$$

$$d(y_0,y_i) = (x_0 - x_i)^2/2\sigma^2 \tag{12}$$

where n is the number of measured samples, m is the dimension of vector random variable x, and σ is the smooth parameter of the Gauss function (also called the spread factor).

By substituting the estimated PDF in formula (10, 11 and 12) into the condition mean in formula (9) and interchanging the order of integration and summation, the desired condition mean of y for a given x_0 can be yielded:

$$\hat{y}(x_0) = \frac{\sum_{i=1}^{n} y_i e^{-d(x_0,x_i)}}{\sum_{i=1}^{n} e^{-d(x_0,x_i)}} \tag{13}$$

GRNN is structurally composed of an input layer, a pattern layer, a summation layer, and an output layer (Fig. 1).

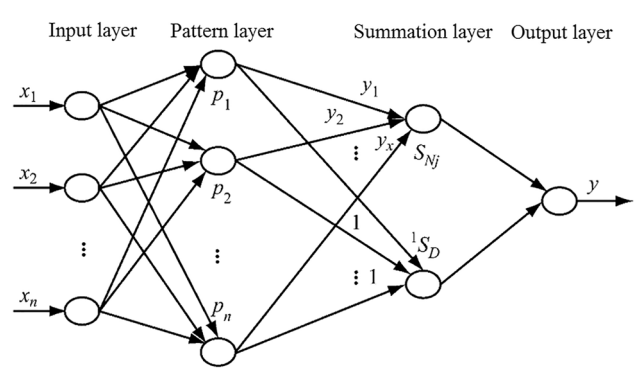

Fig. 1 Structural diagram of GRNN

The training set consists of values of inputs x, each with a corresponding value of an output y. This regression method produces the estimated value of y, which minimizes the squared error [14].The network has the advantages of simple structure, few parameters, and fast prediction speed, however, there is also a disadvantage that the selection of input layer neurons is subjective.

The MATLAB software (R2015b) is utilized to implement analysis of chaotic characteristics and GRNN for mine water inflow estimation.

2.4 Chaos-GRNN model

Giving full play to the respective advantages of chaos theory and GRNN, and combining them organically to establish a Chaos-GRNN prediction model for mine water inflow (Fig. 2).

The steps to establish the Chaos-GRNN model are as follows:

First, by chaos theory to construct the sample set of GRNN. To calculate the time delay τ and the embedding dimension m and the maximum Lyapunov exponent $\lambda_{\max}$ of the time series$\{x_1, x_2,\ldots,x_n\}$. By formula (5), the time series$\{x_1, x_2,\ldots,x_n\}$will transform into the sample set: $Y=(Y_1,Y_2,\ldots,Y_M)^T$, $M=n-(m-1)\tau$. And the sample set will be divided into the training samples and test samples.

Second, to determine input layer neuron sand pattern layer neurons of GRNN. m is the number of factors affecting the mine water inflow, which is equal the number of both input layer neurons and pattern layer neurons of the Chaos-GRNN model; the value of the input layer neuron is determined by formula (5): $Y_i=(x_i, x_{i+\tau},\ldots, x_{i+(m-1)\tau})$.

Third, to optimize the smooth parameter of GRNN. Based on the neural network toolbox in MATLAB, the cross-validation method is used to optimize the optimal Spread parameter value of newgrnn() function to obtain the smoothing factor σ.

Fourth, to set up model and forecast water inflow. Combine the trained input layer, mode layer data and optimization parameters to construct an optimized Chaos-GRNN model and import the data to predict. The step size of Chaos-GRNN model is determined by formula (8).

Last, to evaluate the model. The model evaluation is based on four performance metrics described below. Let $x_o(i)$ and $x_e(i)$ denote, respectively, the estimated and observed mine water inflow, with $\overline{x_o}$, $\overline{x_e}$ denoting their means, and n denoting the number of data points considered.

(1) The coefficient of correlation (R) can be expressed as

$$R=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_o(i)-\overline{x_o})(x_e(i)-\overline{x_e})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_o(i)-\overline{x_o})^2}\cdot\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_e(i)-\overline{x_e})^2}} \tag{14}$$

where the closer the coefficient R is to1,the more perfect is the positive linear relationship.

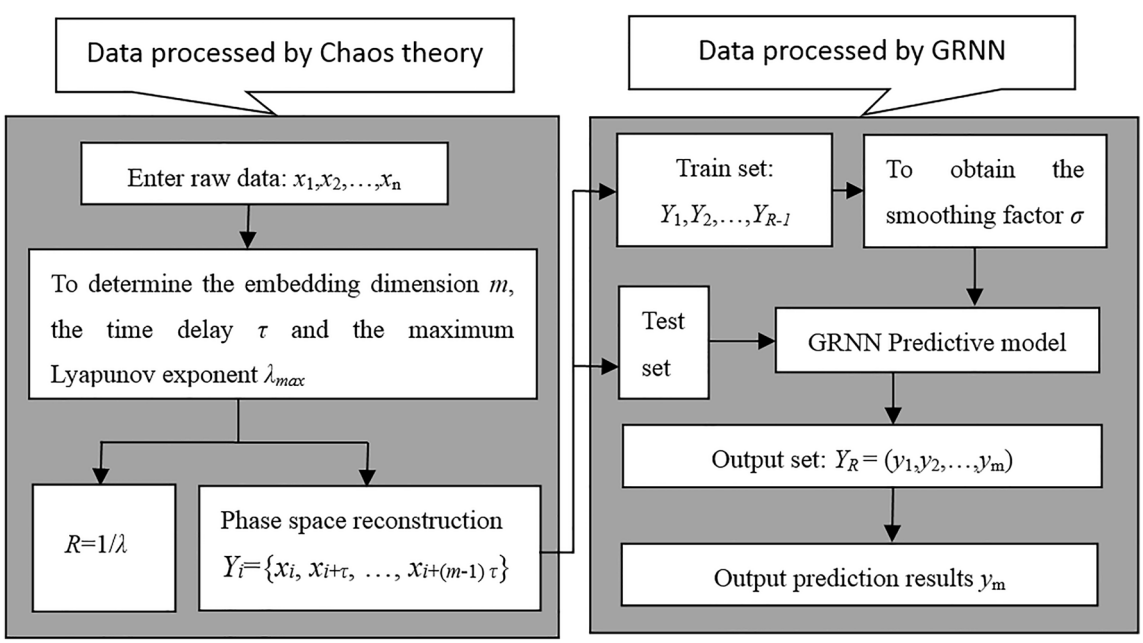

Fig. 2 Structural diagram of Chaos-GRNN model

(2) The root mean squared error (RMSE) can be expressed as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_e(i) - x_o(i))^2} \tag{15}$$

(3) The mean absolute percentage error (MAPE) can be expressed as:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{x_o(i) - x_e(i)}{x_o(i)}\right| \times 100 \tag{16}$$

The closer the values of RMSE or MAPE are to 0, the more reliable is the model performance.

(4) Nash–Sutcliffe efficiency (NSE) can be expressed as:

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(x_o(i) - x_e(i))^2}{\sum_{i=1}^{n}(x_o(i) - \overline{x_o})^2} \tag{17}$$

The value of NSE is $-\infty$ to 1. NSE is close to 1, indicating high reliability of the model; NSE is close to 0, indicating that the model is generally reliable, but the process simulation error is large; NSE is far less than 0, indicating that the model is not credible.

There are many performance metrics to evaluate a model. In this study, the above four indices were chosen, because these four indices are most basic and used most widely [6].

3 Application and results

3.1 Sources of mine water inflow of the study sites

Pingdingshan No.12 Coal Mine is located in Pingdiangshan city of Henan province, China (Fig. 3).The study area, approximately 12.87 km^2, is part of Pingdingshan Mining Area. So far, there have been 9 water inrush accidents in Pingdingshan No.12 Coal Mine.

At present, the normal mine water inflow is about 129 m^3/h. The direct sources of mine water inflow are the water in the goaf of the early mining, the sandstone water in the roof of the coal seam, and the limestone water in the coal floor. The indirect sources of mine water inflow are the atmospheric precipitation and surface water.

3.2 Characteristics of water inflow

Monthly mine water inflow data are obtained from Pingdingshan No.12 Coal Mine between February 1976 and

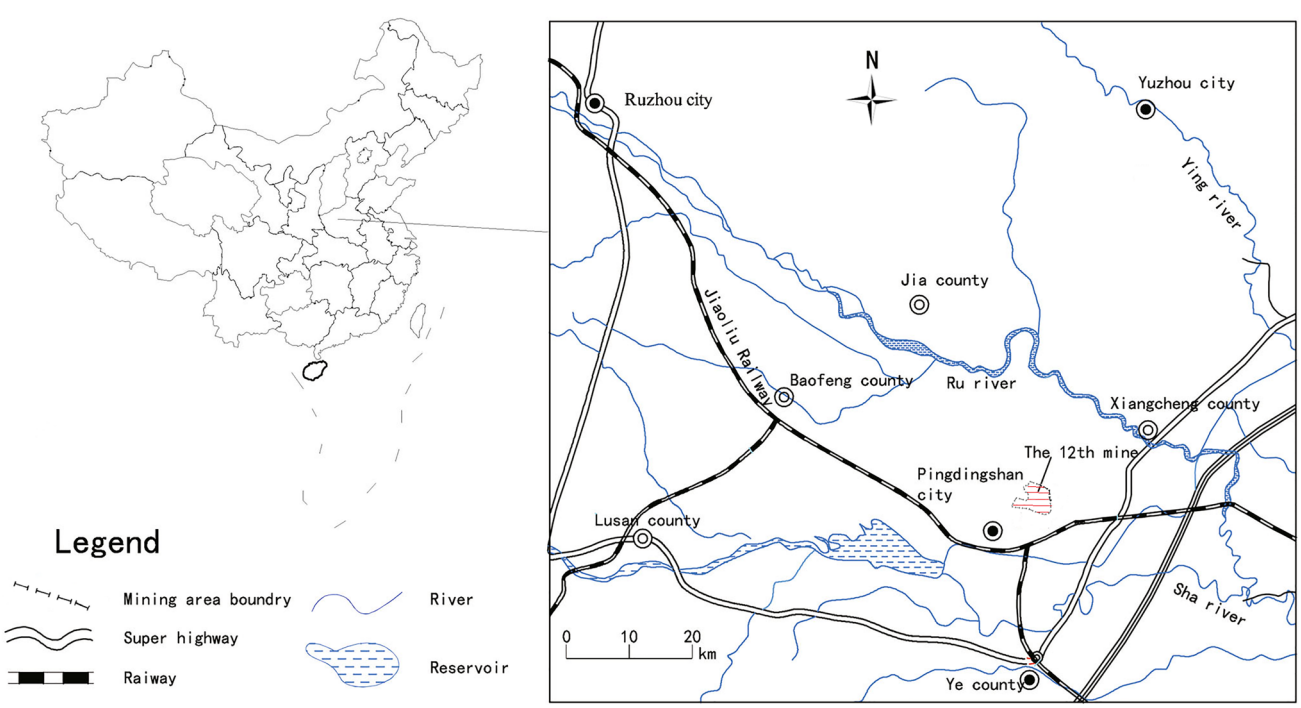

Fig. 3 Location of Pingdingshan No. 12 Coal Mine, China

December 2015. Thus, the time series data set contains 479 points (Fig. 4).

The data of mine water inflow from February 1976 to December 2013 (No.1–No.455) are used to reconstruct phase space of the mine water inflow time series. The data from July 2012 to December 2013 (No.438–No.455) are used to set up Chaos-GRNN model; the data from January 2014 to December 2015(No.456–No.479) are used to verify model.

3.3 Phase space reconstruction

The time delay of the water inflow time series, $\tau = 13$ months, is determined by formula (1) (Fig. 5a); the embedding dimension $m = 7$, is obtained by formulas (2, 3 and 4) (Fig. 5b).Using formulas (6) and (7), the maximum Lyapunov exponent λ_{max} is equal to 0.053.

The embedding dimension $m = 7$ indicates that there are 7 factors affecting the mine water inflow, so as to determine the number of input layer neurons in the Chaos-GRNN model is 7. $\lambda_{max} = 0.053 > 0$, indicates that the mine water inflow time series has chaotic characteristics. $R = 1/\lambda_{max} = 18.86 \approx 19$, indicates that prediction step size of the Chaos-GRNN model is 19 steps.

The phase space of mine water inflow time series in Pingdingshan No. 12 Coal Mine is reconstructed to obtain a 7-dimensional vector space Y with a time delay of 13 months:

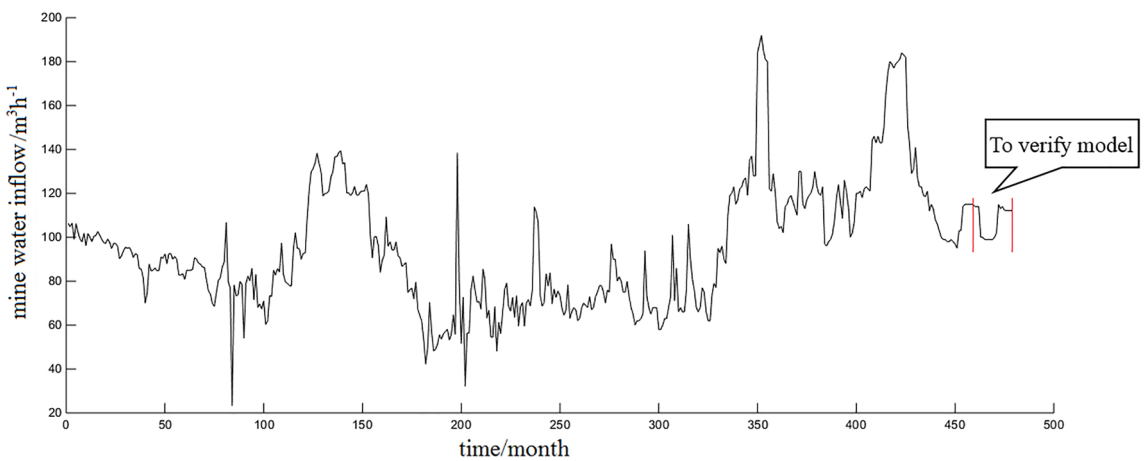

Fig. 4 Monthly mine water inflow curve of Pingdingshan No.12 Coal Mine

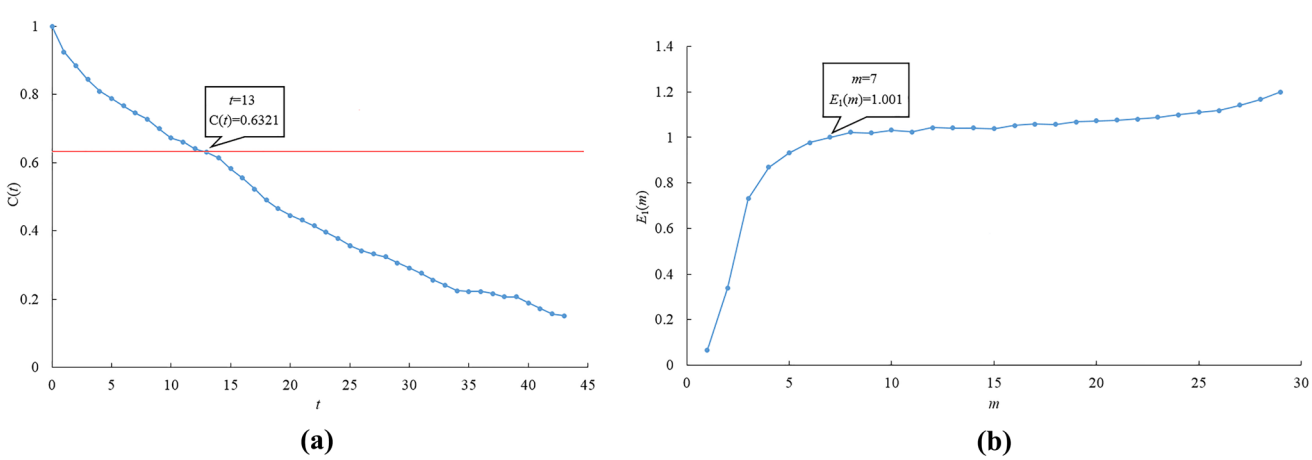

Fig. 5 **a** Relationship between time delay τ and autocorrelation coefficient C(τ); **b** The relationship between m and $E_1(m)$ of mine discharge series

$$Y=\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_{377} \\ Y_{378} \\ \vdots \\ Y_{401} \end{pmatrix}=\begin{pmatrix} x_1 & x_{14} & x_{27} & x_{40} & x_{53} & x_{66} & x_{79} \\ x_2 & x_{15} & x_{28} & x_{41} & x_{54} & x_{67} & x_{80} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{377} & x_{390} & x_{403} & x_{416} & x_{429} & x_{442} & x_{455} \\ x_{378} & x_{391} & x_{404} & x_{417} & x_{430} & x_{443} & \boxed{x_{456}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{401} & x_{414} & x_{427} & x_{440} & x_{453} & x_{466} & \boxed{x_{479}} \end{pmatrix} \tag{18}$$

3.4 Water inflow forecast

Due to the test samples using the water inflow data from January 2014 to December 2015, so need to assume that the values of x_{456}, …, x_{479} are unknown. That is, in the formula (18), the last column number of $Y_{378} \sim Y_{401}$ is unknown, so the training samples must be selected before Y_{378}. Additionally, the step size of the Chaos-GRNN model is 19steps. So the detailed procedure to predict mine water inflow is shown as follows:

Step one: To forecast x_{456}.

Selecting $(Y_{360}, Y_{361},\ldots,Y_{377})^T$ as the training sample in GRNN. Output layer output is $\hat{Y}_{378}$, that is, the predicted value of Y_{378} is:

$$\hat{Y}_{378} = \left(\hat{x}_{378}, \hat{x}_{391}, \hat{x}_{404}, \hat{x}_{417}, \hat{x}_{430}, \hat{x}_{443}, \hat{x}_{456}\right) \tag{19}$$

In $\hat{Y}_{378}$, $\hat{x}_{456}$ is predictive value of x_{456}.

Step two: To forecast x_{457}.

Let $Y'_{378} = \left(x_{378}, x_{391}, x_{404}, x_{417}, x_{430}, x_{443}, \hat{x}_{456}\right)$ Selecting $(Y_{361}, Y_{362},\ldots,Y_{377}, Y'_{378})^T$ as the training sample in GRNN. Output layer output is $\hat{Y}_{379}$, that is, the predicted value of Y_{379} is:

$$\hat{Y}_{379} = \left(\hat{x}_{379}, \hat{x}_{392}, \hat{x}_{405}, \hat{x}_{418}, \hat{x}_{431}, \hat{x}_{444}, \hat{x}_{457}\right) \tag{20}$$

In $\hat{Y}_{379}$, $\hat{x}_{457}$ is predictive value of x_{457}.

Step three: To forecast x_{458}.

Let $Y'_{379} = \left(x_{379}, x_{392}, x_{405}, x_{418}, x_{431}, x_{444}, \hat{x}_{457}\right)$ Selecting $(Y_{362}, Y_{363},\ldots,Y_{377}, Y'_{378}, Y'_{379})^T$ as the training sample in GRNN. Output layer output is $\hat{Y}_{380}$, that is, the predicted value of Y_{380} is:

$$\hat{Y}_{380} = \left(\hat{x}_{380}, \hat{x}_{393}, \hat{x}_{406}, \hat{x}_{419}, \hat{x}_{432}, \hat{x}_{445}, \hat{x}_{458}\right) \tag{21}$$

In $\hat{Y}_{380}$, $\hat{x}_{458}$ is predictive value of x_{458}.

Step four: Repeat the above steps to obtain the predicted values of x_{459},…,x_{479}, namely, the verified values of mine water inflow from January 2014 to December 2015 in Pingdingshan No. 12 Mine.

In the prediction process, the initial value of the smoothing factor σ is set to 0.1, the step size is 0.1, and the final value is 2. Under MATLAB, the optimal smoothing factors for the 24 groups of models selected by the cross-validation method are from 0.1to 0.9 (Table 1 and Fig. 6).

Table 1 Chaos-GRNN model prediction results

Time	Measured values (m^3/h)	σ	Predictive values (m^3/h)	Time	Measured values (m^3/h)	σ	Predictive values (m^3/h)
2014.01	115	0.7	114.997	2015.01	99	0.4	111.417
2014.02	115	0.2	114.993	2015.02	99	0.3	112.850
2014.03	115	0.3	114.996	2015.03	100	0.5	112.795
2014.04	115	0.6	114.999	2015.04	102	0.1	111.994
2014.05	114	0.1	114.997	2015.05	115	0.6	112.573
2014.06	114	0.4	114.996	2015.06	113	0.4	112.722
2014.07	114	0.4	114.996	2015.07	114	0.8	112.972
2014.08	100	0.8	111.883	2015.08	112.3	0.9	112.972
2014.09	100	0.8	111.847	2015.09	112.2	0.7	112.943
2014.10	99	0.6	111.849	2015.10	112.2	0.1	112.945
2014.11	99	0.3	111.792	2015.11	112.1	0.6	113.075
2014.12	99	0.5	111.792	2015.12	112.3	0.9	112.543

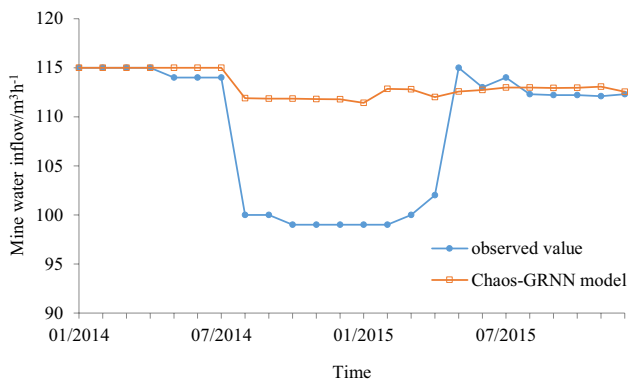

Fig. 6 Comparison between measured value and predicted value of mine water inflow

4 Discuss

4.1 Chaos Parameters of mine water inflow time series

Parameters like time delay τ, embedding dimension m and Maximum Lyapunov exponent λ_{max} are very important for describing the chaotic characteristics and predicting of mine water inflow. In this paper, we have calculated the chaotic parameters of mine water inflow in some coal mines China (Table2).

In Table 2, we can see that:

(1) All the maximum Lyapunov exponent λ_{max} in different coal mines are greater 0. By chaos theory, $\lambda_{max} > 0$ means that the mine water inflow time series has chaotic characteristics. Therefore, this result lays the foundation to use the Chaos-GRNN model for mine water inflow prediction.

(2) τ, m, λ_{max} of the mine water inflow time series are different, which reflects the diversity and complexity of factors related to mine water inflow. According to chaos theory, the embedding dimension m of mine water inflow time series is equal to the number of factors affecting the mine water inflow [15, 19, 29]. In Table 2, the value of m is between 4 and 8, which indicating that the influencing factors of mine water inflow in different coal mines are roughly 4–8. This result reflects the difficulty of accurately predicting the amount of mine water inflow. On the other hand, it also shows the advantage of chaos theory in mine water inflow prediction (influencing factors of mine water inflow and their values are quantified by Y_i).

(3) However, the value of Y_i is still the amount of mine water inflow, not the value of each influencing factor of mine water inflow, so this quantification is not a complete quantification. This also led to the slow application of chaos theory.

ANN simulates the way the human brain nerves process information. It has powerful self-adaptation and self-learning functions for samples, and just has the ability to process this information. Therefore, the chaos theory and ANN are coupled to establish the Chaos-ANN model (Fig. 7), which provides a powerful tool for the effective prediction of mine water inflow.

4.2 Comparison of Chaos-BPNN and Chaos-GRNN prediction models

Artificial neural network (ANN), which owns powerful adaptive and self-learning functions for samples, can effectively obtain output from input neurons. Therefore, the chaos theory and ANN can be organically combined (the number and value of input neurons and predicted

Table 2 The chaotic parameters of mine water inflow in some coal mines China

Research area	Coalfield	Calculation scale	τ	m	λ_{max}	References
Doulishan Mine	Hunan China	Month	6	8	> 0	Xiao [35]
Enkou Mine	Hunan China	Month	6	7	> 0	Xiao [35]
Qiaotouhe Mine	Hunan China	Month	5	6	> 0	Xiao [35]
Fangshan Mine	North China	Month	–	4	–	Huang [8, 10]
Hebi No.4 Mine	North China	Month	–	4	–	Huang [9]
Liuqiao No.2 Mine	Huaibei China	Month	3	8	0.1427	Yang [39]
Wutongzhuang Mine	North China	Month	2	5	0.1396	Tang [28]
Yangcun Mine	North China	Month	7	4	0.1679	Qiao [21]
Gujiao Mine	North China	Month	4	7	0.2053	Qi [20]
A certain Mine	–	Day	27	5	0.1094	Wang [30]
Acertain Mine	–	Month	4	4	0.053	Chen [4]
Pingdingshan No.12 Mine	North China	Month	13	7	0.053	This paper

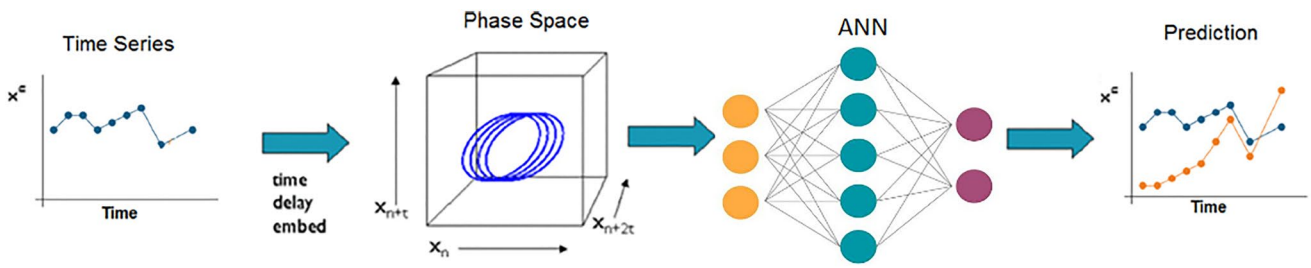

Fig. 7 Schematic representation of the proposed Chaos-ANN model (modified by Fig. 1, Mohammad [17])

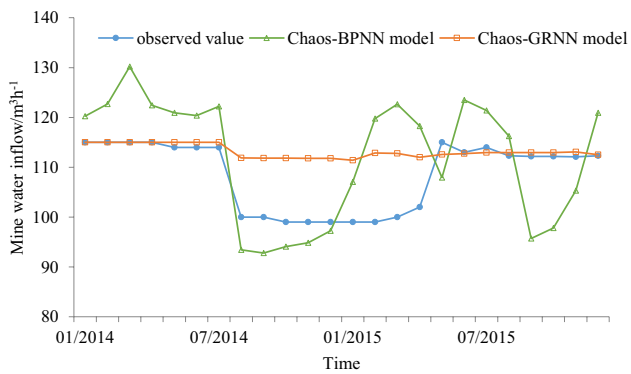

Fig. 8 Comparison of prediction results of Chaos-GRNN model and Chaos-BPNN model

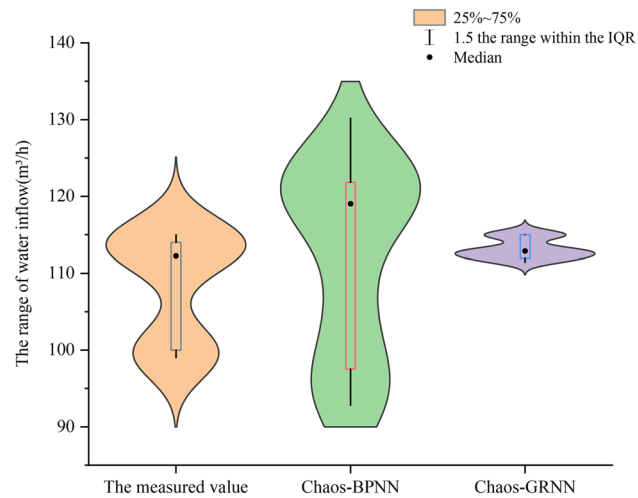

Fig. 9 Violin plots of two predictive models

duration of ANN model can be determined by phase space reconstruction) to predict mine inflow.

Among the artificial neural networks, BP neural network (BPNN) is the most widely used now. So in order to compare with the prediction effect of the Chaos-GRNN model, we also establish a Chaos-BPNN model for predicting the water inflow in Pingdingshan No. 12 Coal Mine.

The Chaos-BPNN model consists of three parts: the input layer, the hidden layer, and the output layer, and the layers are connected by weights. Since the embedding dimension $m = 7$, the Chaos-BPNN model sets 7 input elements, and 1 output element. The number of hidden layer neurons is determined by empirical Equation [42] and trial and error method [23, 37]. The Chaos-BPNN model sets 7 hidden layer elements.

After normalizing the data, we conduct network training by MATLAB. The prediction results are shown in Fig. 8.

The models evaluation are based on *R*, RMSE, MAPE, NSE, and violin plot (Fig. 9). Using formulas (14, 15, 16 and 17), the values of these four indices are calculated (Table 3).

It can be seen from Table3: (1) The value of *R* of Chaos-GRNN model and Chaos-BPNN model is 0.755 and 0.546, respectively. The value of *R* of Chaos-GRNN model is more closer to 1. (2) The value of RMSE, MAPE, and NSE of Chaos-GRNN model is 7.63, 5.02%, and − 0.233, respectively. The values of these three indices are closer to 0 than those of Chaos-BPNN model.

Table 3 Prediction results of Chaos-BPNN model

	R	RMSE	MAPE	NSE
Chaos-GRNN model	0.755	7.63	5.02%	− 0.233
Chaos-BPNN model	0.546	10.75	8.71%	− 1.450

It can be seen from Fig. 9: (1) The median of Chaos-GRNN model is close to the median of measured value, while the median of Chaos-BPNN model is obviously larger. (2) The discreteness of Chaos-BPNN model is much greater than that of Chaos-GRNN model. (3) The density function of Chaos-GRNN model is more concentrated than that of Chaos-BPNN model. Moreover, the maximum value of the density function of Chaos-GRNN model is more closer to the maximum value of the density function of the measured value.

By comparing the four indices (*R*, RMSE, MAPE, NSE) and violin plot of the two models, the forecast result of

Chaos-GRNN model is better than that of Chaos-BPNN model.

5 Conclusions

Chaos theory can recover the rich dynamic information hidden in time series. By reconstructing water inflow time series in phase space, the multi-dimensional matrix could be obtained, with each column representing an influencing factor of water inflow and its value representing the change of the influencing factor with time. Therefore, a new prediction model of mine water inflow is established by combining ANN with chaos theory when lacking data on the influencing factors of water inflow.

The two models were evaluated by using four evaluation indices (*R*, RMSE, MAPE, NSE) and violin plot. It was found that both models can realize the long-term prediction of water inflow, and the prediction effectiveness of Chaos-GRNN model is better than that of Chaos-BPNN model.

However, whether all the time series of mine water inflow has chaotic characteristics, there is no clear answer yet. Therefore, before using the Chaos-GRNN model, it is necessary to verify the chaotic characteristics of the water inflow time series firstly. Once the water inflow time series does not have chaotic characteristics, other predict methods need to be used.

Acknowledgements This work was financially supported by Henan Provincial China Natural Science Foundation Project (182300410155): Study of chaos characteristics of mine hydrological systems and mine inflow prediction model, the National Natural Science Foundation of China (Grant 41672240 and 41573095), and Henan Province's Technological Innovation Team of Colleges and Universities (Grant 15IRTSTHN027), and Key scientific research projects of colleges and universities in Henan Province in 2022(22A170009).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Bahrami S, Ardejani FD, Aslani S et al (2014) Numerical modelling of the groundwater inflow to an advancing open pit mine: Kolahdarvazeh pit, Central Iran. Environ Monit Assess 186:8573–8585
2. Cao LY (1997) Practical method for determining the minimum embedding dimension of a scalar time series. Physica D 110:43–50
3. Cao QK, Zhao F (2011) Forecast of water inrush quantity from coal floor based on genetic algorithm-support vector regression. J China Coal Soc 36(12):2097–2111
4. Chen YH, Yang YG, Peng GH (2008) Chaotic time series analysis and prediction for mine water inflow. Coal Geol Explor 36(4):34–36 (**in Chinese, with English abstract**)
5. Feng Y, Cui NB, Gong DZ et al (2017) Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration. Agric Water Manag 193:163–173
6. Hong M, Wang D, Wang YK et al (2016) Mid-and long-term runoff predictions by an improved phase-space reconstruction model. Environ Res 148:560–573
7. Hu ZY, Zhang C, Luo GP et al (2013) Characterizing cross-scale chaotic behaviors of the runoff time series in an inland river of Central Asia. Quatern Int 311:132–139
8. Huang CH, Chen KK, Li ZH et al (2016) Mine water inrush prediction based on phase space reconstruction and support vector machines. J Henan Polytech Univ (Nat Sci) 35(2):202–205 (**in Chinese, with English abstract**)
9. Huang CH, Feng T, Wang WJ et al (2010) Mine water inrush prediction based on fractal and support vector machines. J China Coal Soc 35(5):806–810 (**in Chinese, with English abstract**)
10. Huang FM, Yin KL, Zhang GR et al (2016) Landslide displacement prediction using discrete wavelet transform and extreme learning machine based on chaos theory. Environ Earth Sci 75:1376–1394
11. Huang YH, Yang SQ (2019) Mechanical and cracking behavior of granite containing two coplanar flaws under conventional triaxial compression. Int J Damage Mech 28:590–610
12. Huang Z, Law KT, Liu H, Jiang T (2009) The chaotic characteristics of landslide evolution: a case study of Xintan landslide. Environ Geol 56:1585–1591
13. Karunasingha DSK, Liong SY (2018) Enhancement of chaotic hydrological time series prediction with real-time noise reduction using Extended Kalman Filter. J Hydrol 565:737–746
14. Kisi O (2008) The potential of different ANN techniques in evapotranspiration modelling. Hydrol Process 22:2449–2460
15. Lv JH, Lu JA, Chen SH (2002) Chaotic time series analysis and its applications. Wuhan University Press, Wuhan
16. Ma D, Duan HY, Li WX et al (2020) Prediction of water inflow from fault by particle swarm optimization-based modified grey models. Environ Sci Pollut Res 27:42051–42063
17. Mohammad AG, Rahman K, Ali DM et al (2018) Chaos-based multigene genetic programming: a new hybrid strategy for-river flow forecasting. J Hydrol 562:455–467
18. Meng XM, Yin MS, Ning LB et al (2015) A threshold artificial neural network model for improving runoff prediction in a karst watershed. Environ Earth Sci 74:5039–5048
19. Peitgen HO, Jürgens H, Saupe D (2004) Chaos and fractals: new frontiers of science, 2nd edn. Springer, New York (**ISBN 0 387 20229 3**)
20. Qi YM, Wu JN, Zhao PW (2010) Dynamic chaos effect and prediction of complicated mine water inrush. Coal Eng 10:83–86 (**in Chinese, with English abstract**)

21. Qiao W, Li WP, Hu G et al (2008) Local method prediction of chaotic time series for groundwater level in mining area. Mining P&D 28(6):66–69 (**in Chinese, with English abstract**)
22. Qiu M, Han J, Zhou Y, Shi LQ (2017) Prediction reliability of water inrush through the coal mine floor. Mine Water Environ 36(2):217–225
23. Qiu RJ, Wang YK, Wang D et al (2020) Water temperature forecasting based on modified artificial neural network methods: two cases of the Yangtze River. Sci Total Environ 737:139729
24. Qu XY, Qiu M, Liu JH et al (2019) Prediction of maximal water bursting discharge from coal seam floor based on multiple nonlinear regression analysis. Arab J Geosci 12:567
25. Shi LQ, Qiu M, Wang Y et al (2019) Evaluation of water inrush from underlying aquifers by using a modified water-inrush coefficient model and water-inrush index model: a case study in Feicheng coalfield. China Hydrogeol J 27:2105–2119
26. Specht DF (1991) A general regression neural network. IEEE Trans Neural Netw 2(6):568–576
27. Takens F (1981) Detecting strange attractors in turbulence. In: Rand DA, Young LS (eds) Dynamical systems and turbulence, lecture notes in mathematics. Springer-Verlag, Berlin, pp 366–381
28. Tang L, Yang YG (2010) Chaotic time series analysis and its application research. J Wuhan Univ Technol 32(19):189–192 (**in Chinese, with English abstract**)
29. Thiel M, Kurths J, Romano MC (2010) Nonlinear dynamics and chaos: advances and perspectives understanding complex systems. Springer, Heidelberg
30. Wang W, Luo ZQ, Wang YW et al (2013) Research into mine water inflow forecast based on chaotic theory. China Saf Sci J 23(4):52–57 (**in Chinese, with English abstract**)
31. Wang XD, Dong H (2014) Prediction of mine water inflow based on generalized regression neural network. China Saf Prod Sci Technol 10(11):90–93 (**in Chinese, with English abstract**)
32. Wolf A, Swift JB, Winney HL (1985) Determining Lyapunov exponents from a time series. Physica D 16:285–317
33. Wu C, Wu X, Zhu G et al (2019) Predicting mine water inflow and groundwater levels for coal mining operations in the Pangpangta coalfield. China Environ Earth Sci 78:130
34. Wu Q, Liu YZ, Luo LH et al (2015) Quantitative evaluation and prediction of water inrush vulnerability from aquifers overlying coal seams in Donghuantuo Coal Mine. China Environ Earth Sci 74(3):1429–1437
35. Xiao J, Shao JL, Cui YL et al (2005) Computation and identification of Lyapunov exponent for the chaotic character of time series of mine hydraulic discharge. Site Investig Sci Technol 23(5):12–15 (**in Chinese, with English abstract**)
36. Xie YY, Li CJ, Lv YJ et al (2019) Predicting lightning outages of transmission lines using generalized regression neural network. Appl Soft Comput J 78:438–446
37. Yan H, Zhang JX, Zhou N et al (2020) Application of hybrid artificial intelligence model to predict coal strength alteration during CO_2 geological sequestration in coal seams. Sci Total Environ 711:135029
38. Yang XH, Mei Y, She DX et al (2011) Chaotic Bayesian optimal prediction method and its application inhydrological time series. Comput Math Appl 61:1975–1978
39. Yang YG, Chen YH (2009) Chaotic characteristics and prediction for water inrush in mine. Earth Sci J China Univ Geosci 34(2):258–262 (**in Chinese, with English abstract**)
40. Yohannes Y, Andrew P (2017) Predicting open pit mine inflow and recovery depth in the Durvuljinsoum, Zavkhan Province, Mongolia. Mine Water Environ 36:114–123
41. Zhang K, Cao B, Lin G et al (2017) Using multiple methods to predict mine water inflow in the Pingdingshan no. 10 coal mine. China Mine Water Environ 36:154–160
42. Zhang YW, Guo HS, Tu H et al (2019) Gas concentration prediction based on neural network with random hidden weight. Comput Eng Sci 41(04):699–707