

Research Paper

High-precision chaotic radial basis function neural network model: Data forecasting for the Earth electromagnetic signal before a strong earthquake

Guocheng Hao^{a,c,*}, Juan Guo^a, Wei Zhang^a, Yunliang Chen^b, David A. Yuen^d

^a School of Mechanical Engineering and Electronic Information, China University of Geosciences, Wuhan 430074, China

^b School of Computer Science, China University of Geosciences, Wuhan 430074, China

^c Department of Mathematics, Duke University, Durham, NC 27708, USA

^d Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, USA

ARTICLE INFO

Article history:

Received 1 March 2021

Revised 8 August 2021

Accepted 29 August 2021

Available online 12 October 2021

Handling Editor: M. Yoshida

Keywords:

Earth's natural pulse electromagnetic field

Chaos theory

Radial Basis Function neural network

Forecasting model

ABSTRACT

The Earth's natural pulse electromagnetic field data consists typically of an underlying variation tendency of intensity and irregularities. The change tendency may be related to the occurrence of earthquake disasters. Forecasting of the underlying intensity trend plays an important role in the analysis of data and disaster monitoring. Combining chaos theory and the radial basis function neural network, this paper proposes a forecasting model of the chaotic radial basis function neural network to conduct underlying intensity trend forecasting by the Earth's natural pulse electromagnetic field signal. The main strategy of this forecasting model is to obtain parameters as the basis for optimizing the radial basis function neural network and to forecast the reconstructed Earth's natural pulse electromagnetic field data. In verification experiments, we employ the 3 and 6 days' data of two channels as training samples to forecast the 14 and 21-day Earth's natural pulse electromagnetic field data respectively. According to the forecasting results and absolute error results, the chaotic radial basis function forecasting model can fit the fluctuation trend of the actual signal strength, effectively reduce the forecasting error compared with the traditional radial basis function model. Hence, this network may be useful for studying the characteristics of the Earth's natural pulse electromagnetic field signal before a strong earthquake and we hope it can contribute to the electromagnetic anomaly monitoring before the earthquake.

© 2021 China University of Geosciences (Beijing) and Peking University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The Earth's natural pulse electromagnetic field (ENPEMF) refers to the primary and secondary electromagnetic fields that are generated by natural field sources that can be received at the surface (Hao et al., 2018). Malyshkov and Malyshkov (2011) proposed that the ENPEMF was not only generated by the atmosphere but also closely related to earthquakes. In a time period that ranges from a few hours to a few days before an earthquake, the number of pulses will suddenly increase and subsequently decrease (Malyshkov and Malyshkov, 2009). The number of pulses is correlated with the season, the time, the earthquake intensity, and the occurrence distance. The ENPEMF signal can be regarded as the transient disturbance of the Earth's natural changing magnetic

field. The seismogenic information generation mechanism of the ENPEMF signal is shown in Fig. 1. Based on the theory of "tectonic magnetism", the ENPEMF signal carries a substantial amount of useful information about the geological structure and its dynamics (Vorobyov, 1979). During a geological disaster, the ENPEMF will exhibit severe anomalies, which may be helpful for electromagnetic anomaly monitoring before earthquakes (Malyshkov and Dzhumabaev, 1987).

Earthquake forecasting research has great significance on scientific and social sectors, especially in areas with frequent earthquake disasters, such as Japan and the Philippines. In recent years, some Japanese scientists have carried out effective monitoring of earthquake electromagnetics in Low-Frequency (LF) electric fields and directional observations of High-Frequency (HF) multiple intersections. At the beginning of 2000, two dense geomagnetic arrays established on the Izu Peninsula and Boso Peninsula in Japan provided important source data for studying the evolution of electromagnetic signals during the Izu earthquake and volcanic activ-

* Corresponding author at: School of Mechanical Engineering and Electronic Information, China University of Geosciences, Wuhan 430074, China.

E-mail address: haogch@cug.edu.cn (G. Hao).

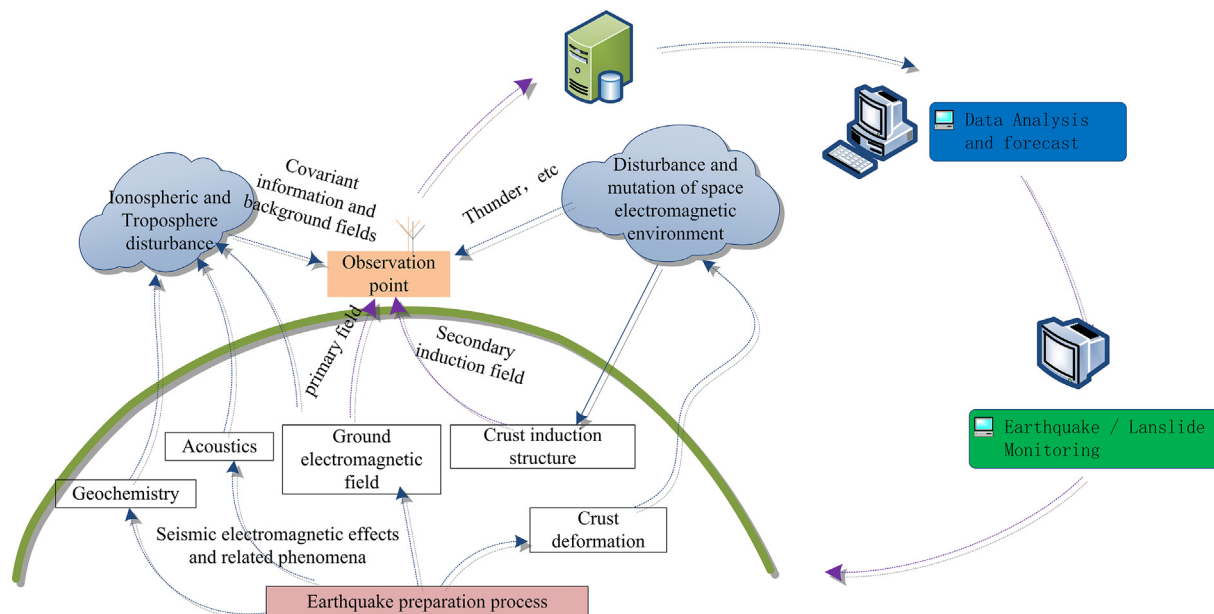


Fig. 1. The seismogenic information source mechanism.

ity. Preliminary results indicated that synchronous anomalous changes in some electromagnetic signals, including geoelectric and geomagnetic fields, were detected before the Lzu earthquake and volcanic activity in 2000. The internationally renowned geophysicist Ueda et al. (2014) discussed the development history and current status of earthquake forecasting insights from the electromagnetic phenomenon in Japan. Yuan et al. (2014) studied the progress of seismic electromagnetic observations and the electromagnetic precursor information focus on the United States region. The conclusion shows that observing in-depth the phenomenon of electromagnetic anomalies before the earthquake may be a possible means for exploring earthquake forecasting. Hattori et al. (2013) and Han et al. (2014, 2017) found that the pulse wave generated by the earthquake was extremely abnormal before the earthquake, which could be used to forecast the earthquake.

Forecasting the intensity trend of the ENPEMF signal is of substantial importance, and the related research is still in the early stages of scientific investigation. The most primitive ENPEMF signal intensity forecasting method uses mainly statistical methods for calculation and evaluation. The signal intensity before a strong earthquake is influenced by various complex factors, and the calculation processes of statistical methods are complex (Leung et al., 2001). With the development of information technology and big data, the forecasting model of the ENPEMF signal intensity before an earthquake is evolving into a highly interesting question, which may touch upon automation technology, artificial intelligence, and machine learning, among other areas.

In recent years, a variety of methods have been used for earthquake forecasting, such as the adaptive nonlinear filtering forecasting method (Zhang and Xiao, 2000), the support vector machine (SVM) method (Cui et al., 2004), and the neural network method (Chen et al., 2012). The radial basis function (RBF) neural network can simulate an artificial neural network and realize strong nonlinear approximation performance (Li and Lv, 2014). The RBF has been widely used in applications for radar, sea clutter, geology, dams, image processing, and other fields. Liu et al. (2020) utilized the backpropagation (BP) neural network algorithm and RBF neural network algorithm to analyze and control the collision response of the urban bridge adjacent beam and pier beam under a strong earthquake. Bagheri et al. (2019) developed three models including

logistic regression (LR), a multilayer perceptron artificial neural network (MLP), and an RBF artificial neural network to predict the possibility of seismic rockfalls on a regional scale. A comparative evaluation of Multilayer Feedforward Perceptron (MFP) and RBF neural networks' ability for instant estimation of r/c buildings' seismic damage level was done in a previous study. Wu et al. (2018) proposed the investigation of the Tikhonov regularization method in regional gravity field modeling by Poisson wavelets radial basis functions. Hong et al. (2018) presented a comprehensive area expansion prediction index method to apply the Global Navigation Satellite System (GNSS) for short-impending prediction of earthquakes. Xu et al. (2017) employed the chaotic characteristics of IPIX radar sea clutter data and studied the method of using an RBF-based self-adaptive fuzzy neural network for weak target detection against a chaotic background. Hou (2016) presented a method that is based on the space-time chaos of image sequences for the detection of small targets in sea clutter. The RBF neural network is used to reconstruct a dynamic model of sea clutter, and it is applied to forecast and cancel sea clutter. Dai and Chen (2016) developed a forecasting model for dam monitoring that utilizes a sequence wavelet RBF neural network that is based on chaos.

Forecasting of the ENPEMF signal variation tendency is challenging due to the chaotic and non-stationary characteristics. We propose a forecasting model that is based on chaos theory and the RBF neural network algorithm, which is named the chaotic RBF neural network. We obtain the effective embedded dimension and delay time for the reconstruction of the phase space of the ENPEMF data collected from the Lushan Ms7.0 earthquake. The obtained parameters are used as a basis for optimizing the RBF neural network. We employ the chaotic RBF neural network to forecast the ENPEMF intensity trend and compare the result with that of the traditional RBF neural network.

The Lushan Ms7.0 earthquake occurred on April 20, 2013. Its rupture zone is located in 29°28'N–30°56'N, 102°16'E–103°11'E, and the total area is 42,786.05 km². The location of the earthquake and the seismic station is indicated by the red point and the green point in Fig. 2. During this period, the GR-01 type equipment at the Wuhan Jiufeng Mountain Seismic Station had received the ENPEMF signal. We have placed three GR-01 receiving devices in the East-West direction (W–E) and North-South direction (N–S) of the

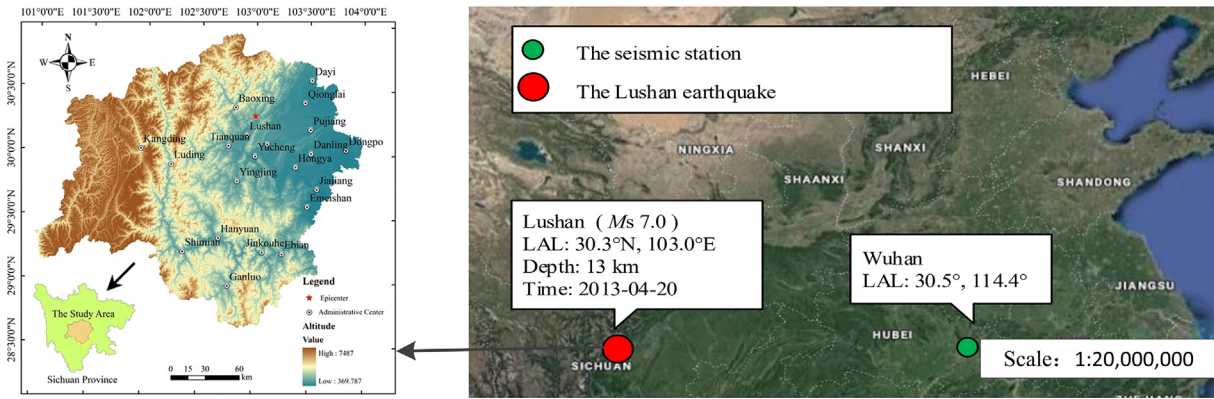


Fig. 2. Location of the Lushan earthquake and seismic station on April 20th, 2013.

Wuhan Jiufeng seismic station, which could collect the ENPEMF signals on the Earth surface. We set the appropriate amplification threshold according to the envelope size and curve shape from the data obtained, which is designed to avoid the saturation distortion of the waveform and reduce the difficulty of identification. In this section, the number of pulses (NH) data from the CN2 and CN3 from April 1 to April 28 is chosen as the analysis target. The sample rate of the signal is 1 s, there are 86,400 data in the whole day we adopted the squared average method to compress the data properly, which could help to keep the envelope and reduce the loss.

2. Chaos theory

2.1. Chaos parameter selection and calculation

With the development of chaos theory (Lorenz, 1969), in the early 1980s, Packard et al. (1980) proposed phase space reconstruction theory (PSRT). Since then, researchers have analyzed and studied chaotic temporal sequences. They proved that the phase space of the system can be reconstructed from a single time series. If the dimension of the delay coordinate is appropriate, we can find a suitable embedded dimension m for the phase space reconstruction of the signal. In the embedded dimension space, we can recover the regular trajectory attractor and preserve the differential homeomorphism with the motive power system (Kugiumtzis, 1996). PSRT is a prerequisite for the study of nonlinear dynamics of time series. The main processes of the chaotic system properties must be conducted in the phase space, such as the calculation of the parameters, the discrimination, and the establishment of the forecasting model. The performance of PSRT directly affects the subsequent analysis.

Recent studies have shown that the main factors affecting the quality of phase space reconstruction are the selection of delay time and embedding dimension. Kim et al. (1999) proposed the correlation integral calculation (C-C) method, which could simultaneously estimate the delay time τ and the width of the embedded window τ_w by using the correlation integral and then obtain the embedded dimension m according to $\tau_w = (m - 1)\tau$.

We assume the time series of ENPEMF is $\{q(t_i)\}$:

$$q(t_i) = q(t_0 + i\Delta t), i = 1, 2, \dots, N \quad (1)$$

where t_0 is the initial time and Δt is the sampling interval, N is the number of time-series data points. With suitable delay time τ and embedded dimension m , the intersequence $\{q(t_i)\}$ is extended to a phase distribution of m -dimensional phase space $\{Q(t_i)\}$:

$$Q(t_i) = q(t_i + (m-1)\tau), i = 1, 2, \dots, M \quad (2)$$

$$\{Q(t_i)\} = \{q(t_i), q(t_i + \tau), \dots, q(t_i + (m-1)\tau)\}, i = 1, 2, \dots, M \quad (3)$$

where τ is the delay time, m is the embedded dimension. Each column constitutes a phase point in the m -dimensional phase space. Each phase point has m components. The number of phase points is $M = N - (m - 1)\tau$, where N is the number of time-series data points. The line connecting the M phase points can describe the evolutionary trajectory of m -dimensional phase space, and the reconstructed phase space is equivalent to the original system in terms of topology.

The correlation integral of the embedded time series is defined as the following function:

$$C(m, N, r, t) = \frac{2}{M(M-1)} \sum_{1 \leq i < j \leq M} \theta(r - d_{ij}), r > 0 \quad (4)$$

where, $d_{ij} = \|Q(t_i) - Q(t_j)\|$. If $r - d_{ij} < 0$, then $\theta(r - d_{ij}) = 0$. If $r - d_{ij} > 0$, then $\theta(r - d_{ij}) = 1$.

In this method, we use the correlation integral of the series to construct the statistics, which represent the correlation of the non-linear time series. The relationship between the statistics and the delay time is used to determine the optimal delay time τ and the width of the embedded window τ_w , to determine the embedded dimension m . The statistic is defined as:

$$S(m, N, r, t) = C(m, N, r, t) - C^m(m, N, r, t) \quad (5)$$

The time series $\{q(t_i)\}, i = 1, 2, \dots, N$ is divided into t disjoint time series. When $t = 1$, it is a single time series itself; when $t = 2$, it is $\{q(t_1), q(t_3), \dots, q(t_{N-1})\}$ and $\{q(t_2), q(t_4), \dots, q(t_N)\}$, and the length is $N/2$. For the general natural number t , we can get:

$$\begin{aligned} &\{q(t_1), q(t_{t+1}), q(t_{2t+1}), \dots\} \\ &\{q(t_2), q(t_{t+2}), q(t_{2t+2}), \dots\} \\ &\dots \\ &\{q(t_t), q(t_{2t}), q(t_{3t}), \dots\} \end{aligned} \quad (6)$$

where the length is $l = \frac{N}{t}$. After dividing it into t disjoint subsequences, the statistic of each subsequence is defined as:

$$S(m, N, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, N/t, r, t) - C_s^m(1, N/t, r, t)] \quad (7)$$

When $N \rightarrow \infty$, the statistic of each subsequence is:

$$S(m, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, r, t) - C_s^m(1, r, t)], m = 2, 3, \dots \quad (8)$$

And then we define the maximum deviation with respect to r :

$$\Delta S(m, t) = \max\{S(m, r_j, t)\} - \min\{S(m, r_j, t)\} \quad (9)$$

In general, there is a certain range of choices for N , m , and r . When $2 \leq m \leq 7$, $\sigma/2 \leq r \leq 2\sigma$, $N \geq 500$ (σ refers to the mean square deviation or standard deviation of the time series), the asymptotic distribution can be well approximated by the finite series and $S(m, N, r, 1)$ can represent the correlation of the series. In specific calculation, we often take $m = 2, 3, 4, 5, 6, 7$, $r_i = i\sigma/2$, $i = 1, 2, 3, 4$, $N = 3000$. We calculate the following three statistics:

$$\bar{S}(t) = \frac{1}{16} \sum_{m=2}^7 \sum_{j=1}^4 S(m, r_j, t) \quad (10)$$

$$\Delta \bar{S}(t) = \frac{1}{4} \sum_{m=2}^7 \Delta S(m, t) \quad (11)$$

$$S_{cor}(t) = \Delta \bar{S}(t) + |\bar{S}(t)| \quad (12)$$

We utilize the C-C method to automatically search for the first zero of $\bar{S}(t)$ in Eq. (10) or the first minimum value of $\Delta \bar{S}(t)$ in Eq. (11), to find the first local maximum of the ENPEMF time series, which corresponds to the delay time τ . At the same time, we calculate the minimum value of $S_{cor}(t)$ in Eq. (12) to obtain the first overall maximum time window t of the time series, which is the optimal embedded window width τ_ω . Then, the value of the optimal embedded dimension m is determined by $\tau_\omega = (m - 1)\tau$.

A dynamical system with the same topological properties as the original system can be reconstructed in the phase space by selecting suitable delay time τ and embedded dimension m . Therefore, the internal nonlinear chaotic system of the ENPEMF can be restored via PSRT. Then, we use the chaotic RBF neural network to forecast the ENPEMF time series.

2.2. Chaotic recognition of ENPEMF time series

A chaotic time series is a univariate or multivariate time series that is generated by a chaotic dynamical system, which typically exhibits chaotic disorder features that are similar to random noise due to its inherent chaotic characteristics (Kim et al., 1999). However, a chaotic time series and the corresponding chaotic system follow an inherent correlation law, which is not as random as it seems to behave. The ENPEMF signal is a typical non-stationary signal due to its complex field source and internal information. To apply a chaotic time series analysis method to ENPEMF forecasting, it is extremely important to determine whether the ENPEMF sequence has chaotic characteristics. Rosenstein et al. (1993) improved the Wolf method based on the strategy of orbital tracking and proposed a small-data-volume method for calculating the maximum Lyapunov exponent, which is an important basis for judging whether the ENPEMF signal has chaotic characteristics.

The maximum Lyapunov exponent calculation steps by the small-data-volume method are as follows:

(1) For the ENPEMF time series, we employ the correlation integral calculation (C-C) method to calculate the delay time τ and the embedded dimension m , and we utilize the fast Fourier transform (FFT) method to analyze the signal spectrum and calculate the average period. The C-C method is to simultaneously determine the characteristic quantities τ and m , which can meet the requirements of phase space reconstruction of the system.

(2) We reconstruct the chaotic dynamical system $\{q_j, j = 1, 2, \dots, N\}$ based on the delay time and the embedded dimension.

(3) We find the nearest adjacent point q_j that is common to each point q_i in the phase space, and we restrict the separation distance as follows:

$$D_j(0) = \min \|q_i - q_j\|, |j - i| > \tau_p \quad (13)$$

where τ_p is the average period of the time series.

(4) For each point q_i in the reconstructed phase space, we calculate the distance $D_j(i)$ of the neighboring point q_j after i discrete time steps, as expressed in Eq. (14):

$$D_j(i) = \|q_{j+i} - q_{i+i}\|, i = 1, 2, \dots, \min(N - j, N - \hat{j}) \quad (14)$$

(5) We assume that the nearest neighbor of the i -th point in the phase space is approximately divergent at the maximum Lyapunov exponential rate:

$$D_j(i) = C_i e^{\lambda_1(j\Delta t)} \quad (15)$$

(6) We take the logarithm of both sides of Eq. (15) to obtain $\ln D_j(i) = \ln C_i + \lambda_1(j\Delta t)$, $i = 1, 2, \dots, \min(M - j, M - \hat{j})$. The equation represents a cluster of approximately parallel lines with a slope of λ_1 .

(7) Finally, we use the least-squares method to obtain a regression line, and the slope of the line is the maximum Lyapunov exponent λ_1 .

3. Forecasting model of chaotic RBF neural network

3.1. Chaotic RBF neural network

The radial basis function (RBF) neural network is a type of static forward network (Billings and Zheng, 1995), which has a structure that is similar to that of a multi-layer forward network. It includes an input layer, a hidden layer, and an output layer. There is a non-linear relationship between the input layer and the hidden layer, and a linear relationship between the hidden layer and the output layer, as illustrated in Fig. 3.

For the RBF neural network, the expression is as follows:

$$x(p+1) = f(x(p)) = \sum_{i=1}^k w_i \phi_i(\|x(p) - c_i\|) = W^T \phi \quad (16)$$

$$\phi_i(\|x(p) - c_i\|) = \exp\left(-\frac{\|x(p) - c_i\|^2}{2\tau_i^2}\right), i = 1, 2, \dots, k \quad (17)$$

where $x(p) \in R^k$ is the network input vector; $f \in R^1$ is the network output vector; $\phi(\cdot)$ is the Gauss function; $\phi = [\phi_1, \phi_2, \dots, \phi_k]^T$ is the hidden-layer output vector; k is the number of hidden-layer

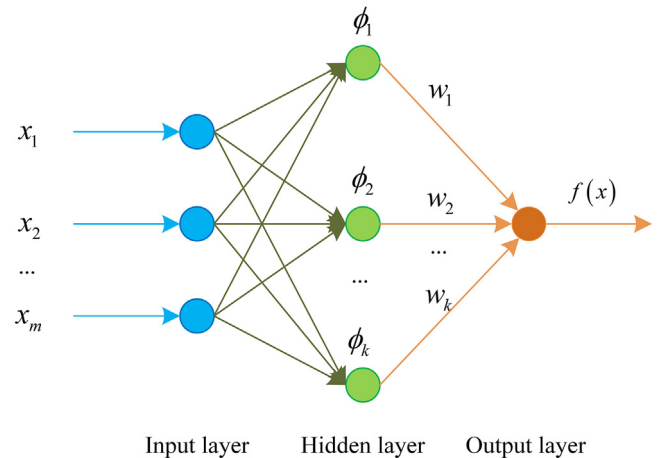


Fig. 3. RBF neural network structure of traditional study.

units; c_i and r_i are the center and width, respectively, of the Gauss function; and $W = [w_1, w_2, \dots, w_k]^T$ is the weight vector of the output layer of the network. Within the framework of the RBF neural network, the input-layer node transmits the signal to the hidden layer. Hidden-layer nodes are composed of radial action functions such as the Gaussian kernel functions, which apply a nonlinear transformation through the basis function to map the input space to a new space (Gan et al., 2009). The output-layer node realizes the linear weighting combination in the new space.

In this paper, we propose a new chaotic RBF neural network model by combining the chaotic optimization method with the RBF neural network, and we apply it to ENPEMF forecasting research. We use the PSRT principle to reconstruct the normalized ENPEMF data after normalization. Then, we utilize a subset of the reconstructed data to train the chaotic RBF neural network model and conduct forecasting simulation on the remaining reconstructed data.

Using the chaotic RBF neural network to forecast the ENPEMF time series, the number of neurons in each layer of the neural network depends on the scenario of the ENPEMF time series. The method for determining the structure of the chaotic RBF neural network is as follows:

(1) Determination of the input layer

Based on the chaotic characteristics of ENPEMF, the regularity of internal changes in the system can be fully demonstrated in high-dimensional space. Therefore, the embedded dimension m of the system is selected as the number of input layers of the model, and the point vector data in the phase space of the reconstructed system are used as the input data of the system. The input data contain the complete change law information of the system, which can satisfy the input requirements of the system. If the number of neurons in the input layer is equal to m , the forecasting effect is better (Lv et al., 2005; Zhang et al., 2007).

(2) Determination of the hidden layer

We select the Gaussian function $KMatrix = getKRBf(X, Y, \gamma)$ as the activation function of the hidden layer, where X and Y are horizontal and vertical coordinates of the data, $\gamma = 1/(2 \times \sigma^2)$ is a parameter. Then we invoke the neural network forecasting function $net = Newrb(P, T, GOAL, SPREAD, MN, DF)$, where P selects a phase-space reconstruction point vector matrix as an input sample, and the number of corresponding input layers is m ; T is the forecasting test sample; $GOAL$ is the mean-square error requirement of the forecasting model for ensuring convergence; $SPREAD$ is the extension speed; MN is the maximum number of hidden layers; DF is the number of hidden layers that are added during the establishment of two adjacent networks.

Based on the training samples, forecasting samples, and error target values, the number of hidden layers is continuously increased via a search method until that the error between the forecasted output value and the actual value falls below the error value set. Finally, we obtain the expected optimal number of hidden layers. Based on the Gaussian function, the output of the i -th hidden-layer node is expressed in Eq. (11):

$$R_i(p) = \exp\left[-\|p - c_i\|^2 / 2b_i^2\right], i = 1, 2, \dots, k \quad (18)$$

In real programming, the selection of the performance mean-variance (the minimum expected error) impacts the accuracy of the output. To realize suitable precision and avoid the “overfitting” phenomenon, we set the minimum expected error to 0.01 (Casdagli, 1989).

(3) Determination of the output layer

The forecasting model that is established in this paper is a single-step forecasting model, the received value is the next sample of the ENPEMF signal, and the number of corresponding output layers is 1.

3.2. Forecasting procedure of chaotic RBF neural network

In conclusion, the forecasting procedure of chaotic RBF neural network is as follows:

- (1) First, we conduct the previous smoothing and normalization processes on the ENPEMF time-series data to reduce the interference of the data factors with the forecasting results.
- (2) We judge the chaotic characteristics of the collected ENPEMF signal by calculating the Lyapunov exponent. The calculation steps by the small-data-volume method are in Section 2.2.
- (3) The C-C method is used to calculate the optimal embedded dimension m and the delay time τ of the ENPEMF time series.
- (4) In addition to using the optimal embedded dimension and delay time, we reconstruct the ENPEMF time series in the m -dimensional phase space to obtain the system-space point vector matrix.
- (5) According to the selection of the parameters of each layer, we determine the overall framework structure and construct the chaotic RBF neural network forecasting model.
- (6) We employ the training samples to train the constructed model, finally, we use the trained model to forecast the test samples to obtain the system's forecasting values. The workflow of the chaotic RBF neural network forecasting model of this study is shown in Fig. 4.

4. ENPEMF forecasting based on chaotic RBF

4.1. Experimental ENPEMF data

The GR-01 device produced by the Tomsk Branch of the Russian Academy of Sciences was adopted as the ground instrument to receive ENPEMF signals during the Lushan earthquake. We have installed three sets of equipment (Fig. 5) on the Jiufeng seismic station in Hubei province. The directions are W-E and N-S, and there are three channels: CN1, CN2, and CN3. The instrument records the number (NH) and amplitude (AH) of the ENPEMF signal pulses that exceed the set threshold. The data can be uploaded every 4 h to the Internet via GPRS with 6 documents per day. If the instrument receives abnormally large-scale pulses that differ from the envelope trajectory of the normal background pulses, they may contain information about abnormal changes in the earthquake. Finally, we convert the data into a .txt file. The operating frequency of the equipment is in the very-low-frequency (VLF) band, namely, 5–25 kHz, and the receiving frequency set in Wuhan is 14.5 kHz.

The ENPEMF is a non-periodic and non-stationary signal, and the output is the signal after digital quantization. The data storage format is the time-amplitude-pulse number (t -AH-NH) (Hao et al., 2018), and the time unit is seconds (S). AH is the quantized pulse amplitude that exceeds the set threshold. The amplitude unit corresponds to that of the data after amplification by the original millivolt (mv) signal, and the amplification factor is the same among the frequency bands. This serves as a reference for the size change of the envelope (Hao et al., 2016). NH is the number of quantized pulses that exceed a set threshold. Both AH and NH characterize the strength of the surface magnetic field. Table 1 shows part of the original NH data of the CN2 and CN3.

4.2. Forecasting results

We use the above C-C method to calculate the Lyapunov exponent, which is used to determine whether the ENPEMF time series has chaotic characteristics. Then we employ the above C-C

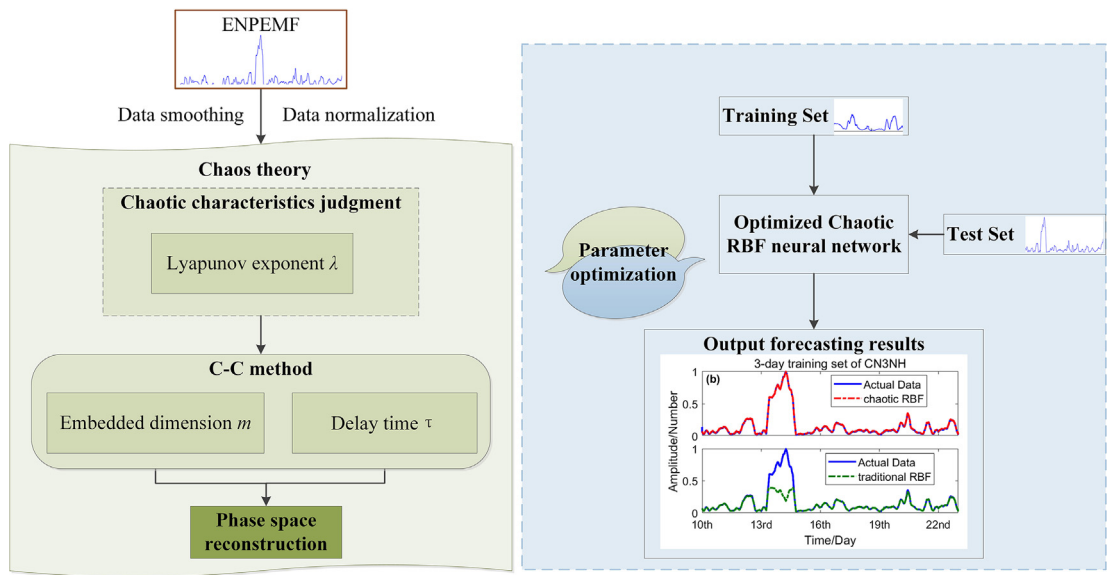


Fig. 4. The workflow of the chaotic RBF neural network forecasting model of this study.



Fig. 5. The ground instrument placed in seismic station.

Table 1
The original NH data of the CN2 and CN3.

18th April			19th April			20th April			21st April		
Time	CN2	CN3	Time	CN2	CN3	Time	CN2	CN3	Time	CN2	CN3
00:02:43	0	0	00:02:44	0	16	03:57:46	0	0	00:02:41	5	0
00:02:44	9	1	00:02:45	10	53	03:57:47	16	0	00:02:42	0	1
00:02:45	1	5	00:02:46	15	38	03:57:48	0	9	00:02:43	1	0
00:02:46	0	0	00:02:47	53	65	03:57:49	0	0	00:02:44	9	2
00:02:47	1	0	00:02:48	222	107	03:57:50	0	0	00:02:45	0	0
00:02:48	8	5	00:02:49	12	18	03:57:51	3	0	00:02:46	3	0
.....	...	...		...	...		...	...		...	...
23:59:58	0	50	23:59:58	4	5	23:59:58	4	0	23:59:58	0	19
23:59:59	40	62	23:59:59	1	0	23:59:59	2	8	23:59:59	0	0

algorithm to automatically search for delay time τ and the optimal embedded dimension m .

The first local minimum point of $\bar{S}(t)$ is 20, which corresponds to the optimal delay time of $\tau = 20$. The global minimum of $S_{cor}(t)$ is the optimal embedded window, namely, $\tau_{\omega} = 117$. Since $\tau_{\omega} = (m - 1)\tau$, we can obtain the embedded dimension $m = 7$

according to $\tau_{\omega} = (m - 1)\tau$. Therefore, from $\tau = 20$ and $m = 7$, the maximum Lyapunov exponent is calculated as 0.0213, which slightly exceeds 0; thus, the ENPEMF time series has chaotic characteristics and can be used for short-term forecasting.

To evaluate the forecasting performance of the established model, we program the chaotic RBF model forecasting process in

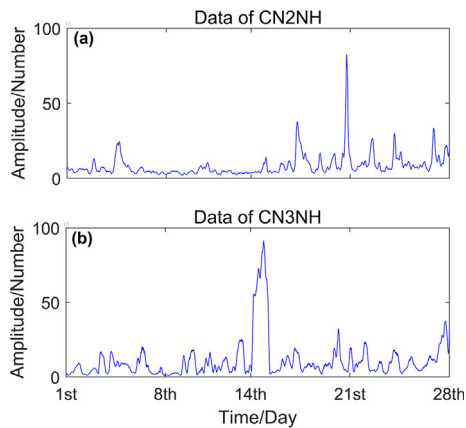


Fig. 6. Observation NH data from April 1st to 28th (a) CN2NH; (b) CN3NH.

the MATLAB 2018a environment. The data that are used were collected during the Ms7.0 Lushan earthquake (April 20th, 2013, Sichuan province, China). The NH data of the CN2 and CN3 from April 1st to 28th are selected, as shown in Fig. 6.

Fig. 6 depicts the variation tendency of the ENPEMF signal of CN2 and CN3 from April 1st to 28th. Before the earthquake, ENPEMF data fluctuated violently. We select the smoothed and normalized 28-day data as the experimental data. The average daily data volume of the ENPEMF consists of 84,467 time-domain data points. To assess the applicability of the chaotic RBF method, we establish two sets of experiments, using the 14-day ENPEMF data (simulation verification 1) and 21-day data (simulation verification 2) as test samples for forecasting, respectively. Meanwhile, in the experiments, the data of 3 days and 6 days were respectively utilized as training samples to train the RBF neural network and then forecast the test set, to verify the reliability and robustness of the method.

In simulation verification 1, we select the 14-day data (April 10th to 23rd) as the test sample. The training samples are 3-day

data (April 7th to 9th) and 6-day data (April 4th to 9th), respectively. We construct the chaotic RBF neural network structure based on PSRT. The input layer has 7 nodes, the output layer has 1 node, and the hidden layer has 10 nodes. The hidden-layer neuron transfer function is selected as the radial basis Gauss function, and the output is a linear function. Then, we utilize the established chaotic RBF neural network forecasting model and the traditional RBF neural network forecasting model to realize the single-step forecasting of the ENPEMF data. Fig. 7 summarizes the results of the proposed chaotic RBF neural network and the traditional RBF forecasting model for CN2NH and CN3NH data.

Fig. 7a and c is the forecasting results of the proposed chaotic RBF model and the traditional RBF through different training samples (3-day and 6-day) and 14-day test samples of CN2NH data. Meanwhile, Fig. 7b and d shows the result of applying the same settings to the CN3NH data. Both the chaotic RBF and traditional RBF models can track and fit the fluctuation trend of actual ENPEMF signal strength. However, the traditional RBF model fails to fit the actual value changes well at the moment of drastic data fluctuation, resulting in a large forecasting error. The proposed chaotic RBF algorithm has good tracking performance and a better fitting effect for the overall ENPEMF signal fluctuation trend in comparison with the traditional RBF model.

To evaluate the forecasting effect of the proposed chaotic RBF model, we select the absolute error as the evaluation index of forecast accuracy for ENPEMF data, as shown in Fig. 8. The overall forecasting error of the proposed chaotic RBF algorithm is smaller than that of the traditional RBF model.

In simulation verification 2, we select the 21-day data (April 8th to 28th) as the test sample with the 3-day data (April 5th to 7th) and 6-day data (April 2nd to 7th) as the training sample, respectively. Then, we compare the established chaotic RBF model and the traditional RBF forecasting model, which are presented in Fig. 9.

Fig. 9a and c shows the forecasting results by the proposed chaotic RBF model and the traditional RBF through different training samples (3-day and 6-day) and 21-day test samples of CN2NH

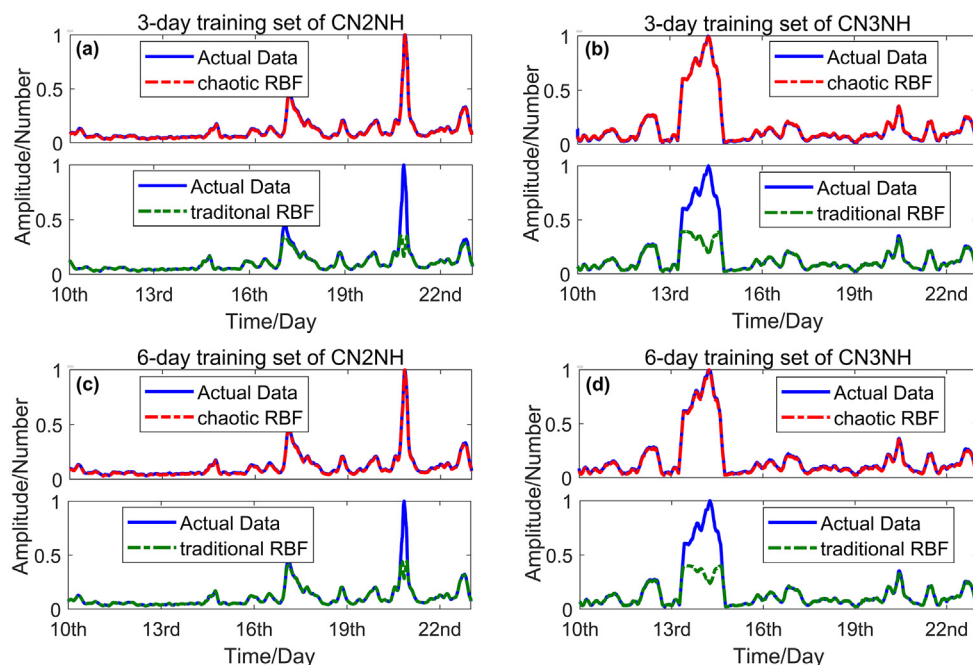


Fig. 7. Results of the proposed chaotic RBF model and the traditional RBF through different training samples and 14-day test samples (a) 3-day training samples of CN2NH; (b) 3-day training samples of CN3NH; (c) 6-day training samples of CN2NH; (d) 6-day training samples of CN3NH.

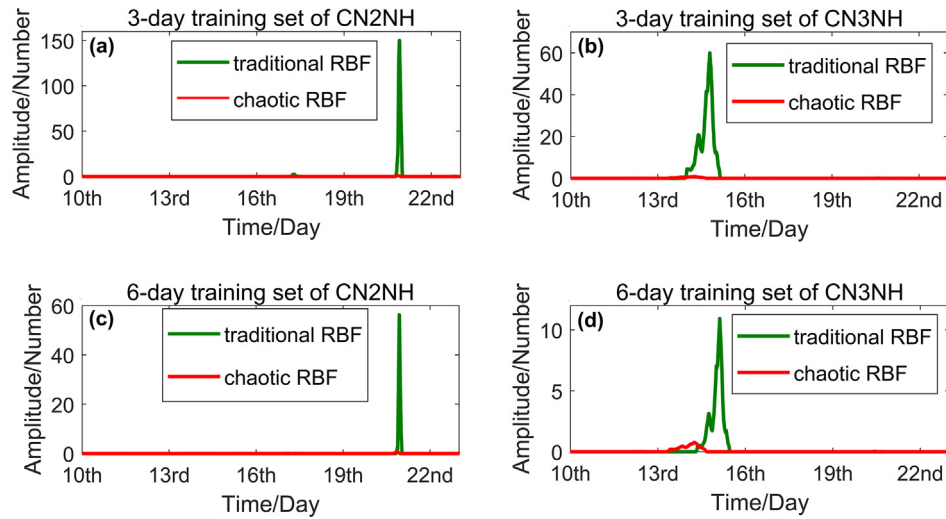


Fig. 8. Absolute error of the proposed chaotic RBF model and the traditional RBF through different training samples and 14-day test samples (a) 3-day training samples of CN2NH; (b) 3-day training samples of CN3NH; (c) 6-day training samples of CN2NH; (d) 6-day training samples of CN3NH.

data. Meanwhile, Fig. 9b and d shows the forecasting result of CN3NH data. From the forecasting results, the forecasted value based on the chaotic RBF neural network forecasting model is consistent with the actual value trend. However, the traditional RBF model cannot track the change of the actual value well in many places, especially in the signal's sharp fluctuation.

Furthermore, we select the absolute error as the evaluation index of the ENPEMF forecasting accuracy, and we utilize it to evaluate the performance of the proposed forecasting model more accurately. The results are presented in Fig. 10. The proposed chaotic RBF model realizes better overall forecasting performance for the ENPEMF data than the traditional RBF model.

To verify the robustness and reliability of the translocation of the forecasted results of the chaotic RBF algorithm, the translocation was measured quantitatively between the forecasted values

of the two algorithms and the actual values using the morphological correlation coefficient as shown in Eq. (19):

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (19)$$

The value of the morphological correlation coefficient obtained is shown in Table 2. As can be seen from Table 2, the correlation values obtained by the chaotic RBF algorithm in this paper are all higher than those obtained by the traditional RBF algorithm, that is, the improved algorithm proposed in this paper has a better forecasting effect.

According to the forecasting results and absolute error results mentioned above, the chaotic RBF forecasting model can fit the

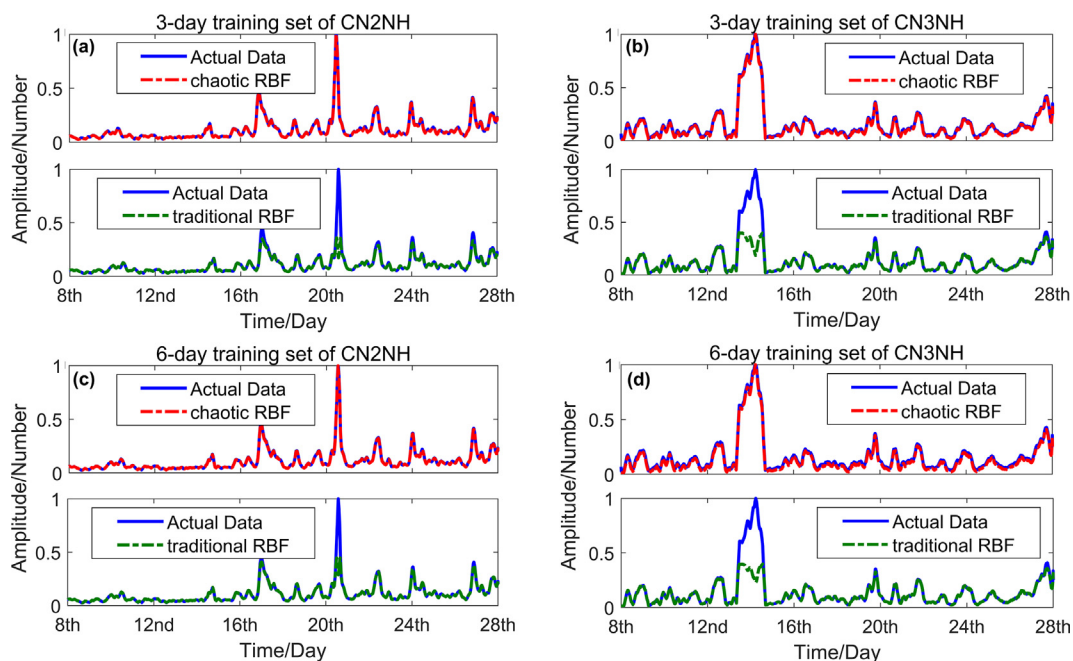


Fig. 9. Results of the proposed chaotic RBF model and the traditional RBF through different training samples and 21-day test samples (a) 3-day training samples of CN2NH; (b) 3-day training samples of CN3NH; (c) 6-day training samples of CN2NH; (d) 6-day training samples of CN3NH.

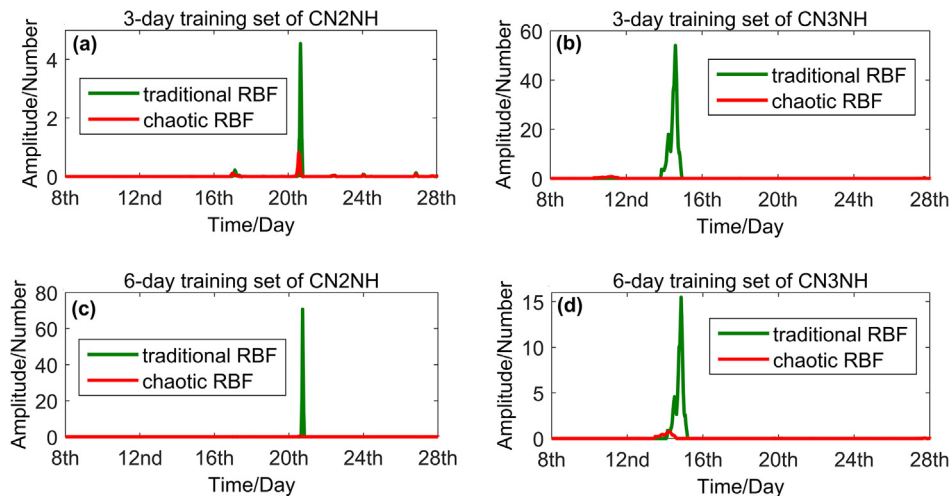


Fig. 10. Absolute error of the proposed chaotic RBF model and the traditional RBF through different training samples and 21-day test samples (a) 3-day training samples of CN2NH; (b) 3-day training samples of CN3NH; (c) 6-day training samples of CN2NH; (d) 6-day training samples of CN3NH.

Table 2
The value of morphological correlation coefficient.

Value of <i>r</i>	CN2NH				CN3NH			
	3&14	6&14	3&21	6&21	3&14	6&14	3&21	6&21
chaotic RBF	0.994	0.995	0.963	0.999	0.993	0.977	0.993	0.979
traditional RBF	0.879	0.877	0.875	0.876	0.856	0.855	0.874	0.867

fluctuation trend of actual ENPEMF signal strength, effectively reduce the forecasting error compared with the traditional RBF model. Therefore, in terms of the overall forecasting effect and accuracy, the algorithm that is proposed in this paper performs better, which is expected to provide support for monitoring geological hazards and magnetic signal anomalies before earthquakes.

5. Discussion

The RBF algorithm has a strong input and output mapping function, and the learning process converges quickly. To solve the problem of low prediction accuracy, we propose the chaotic RBF algorithm based on chaos parameter optimization in this paper. After a lot of experiments and repeated verification, the improved algorithm in this paper can forecast the strength trend of ENPEMF data, and the prediction accuracy is high. Finally, it is proved that the chaotic RBF algorithm in this paper is superior in forecasting the strength trend of the ENPEMF signal compared with the traditional RBF algorithm. At present, there is no exact method and conclusion for earthquake forecasting. We can only focus on the processing and analysis of relevant signals and the research on reliability forecasting methods, to provide support for the monitoring of geological disasters and abnormal magnetic signals before earthquakes. At the same time, as the research on the ENPEMF method and data and the exploration of earthquakes in China are still in the initial stage, the research on other earthquakes needs to be further explored, studied, and verified in the future.

6. Conclusion

We propose a chaotic RBF neural network forecasting model in this paper, which is applied to forecast the ENPEMF time series. First, we utilize the C–C method to calculate the optimal embedded dimension *m* and delay time, which is used to reconstruct the

phase space of the collected ENPEMF data and to determine the chaotic characteristics of the ENPEMF. In addition, the obtained parameters are used as a basis for determining the number of input nodes of the RBF neural network and to optimize the RBF. Finally, we employ the chaotic RBF neural network model trained by different time lengths training samples to forecast the 14-day and 21-day ENPEMF data of CN2NH and CN3NH, and we compare it with the traditional RBF forecasting model. The results demonstrate that the chaotic RBF neural network significantly outperforms the traditional RBF neural network forecasting model in terms of forecasting performance and accuracy. In some existing methods for monitoring anomaly information before some strong earthquakes, the chaotic RBF forecasting model in this paper can assist to a certain extent and improve the monitoring accuracy of anomaly information.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We sincerely acknowledge the reviewers for their constructive amendments in improving the quality of this article. We also thank Min Wu, Haobin Dong, Sidao Ni, Xiangyun Hu, and York Shi for their help and guidance. This research was sponsored by the National Natural Science Foundation of China (61333002), Open Research Foundation of the State Key Laboratory of Geodesy and Earth's Dynamics (SKLGED2018-5-4-E), Foundation of the Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems (ACIA2017002) and 111 projects under Grant (B17040), Open Research Project of the Hubei Key Laboratory of Intelligent Geo-Information Processing (KLIGIP-2017A02). The

work of this paper was also partially supported by the Three Gorges Research Center for geo-hazard, Ministry of Education, and cooperation agreements of Krasnoyarsk Science Center and Technology Bureau, Russian Academy of Sciences. Thanks to the Hubei Provincial Seismological Bureau offering an observation site to place the Earth's natural pulse electromagnetic field device. We also thanks to science and technology workers' seriously reliable routine maintenance in Jiufeng seismic station.

References

- Bagheri, V., Uromeihy, A., Fatemi Aghda, S.M., 2019. Predicting the probability of rockfalls occurrence caused by the earthquake of Changureh-Avaj in 2002 using LR, MLP, and RBF methods. *B. Eng. Geol. Environ.* 78 (5), 3119–3141.
- Billings, S.A., Zheng, G.L., 1995. Radial basis function network configuration using genetic algorithms. *Neural Networks* 8 (6), 877–890.
- Casdagli, M., 1989. Nonlinear prediction of chaotic time series. *Physica D* 35 (3), 335–356.
- Chen, D.Y., Liu, Y., Ma, X.Y., 2012. Two-parameter joint estimation of phase space reconstruction in chaotic time series based on radial basis function neural network. *Acta Phys. Sin.* 61 (10), 1–10 (in Chinese with English abstract).
- Cui, W.Z., Zhu, C.C., Bao, W.X., Liu, J.H., 2004. Prediction of the chaotic time series using support vector machines. *Acta Phys. Sin.* 53 (10), 3303–3310 (in Chinese with English abstract).
- Dai, B., Chen, B., 2016. Chaos-based dam monitoring sequence wavelet RBF neural network prediction model. *Water Resour. Hydr. Eng.* 47 (2), 80–85.
- Gan, M., Peng, X.Y., Peng, H., 2009. Two hybrid parameter optimization algorithms for RBF neural networks. *Control & Decision* 24 (8), 1172–1176.
- Han, P., Hattori, K., Hirokawa, M., Zhuang, J., Chen, C.-H., Febriani, F., Yamaguchi, H., Yoshino, C., Liu, J.-Y., Yoshida, S., 2014. Statistical analysis of ULF seismomagnetic phenomena at Kakioka, Japan, during 2001–2010. *J. Geophys. Res.-Space* 119 (6), 4998–5011. <https://doi.org/10.1002/2014JA019789>.
- Han, P., Hattori, K., Zhuang, J., Chen, C.H., Liu, J.Y., Yoshida, S., 2017. Evaluation of ULF seismo-magnetic phenomena in Kakioka, Japan by using Molchan's error diagram. *Geophys. J. Int.* 208, 482–490. doi: 10.1093/gji/ggw404.
- Hao, G.-C., Bai, Y.-X., Liu, H., Zhao, J., Zeng, Z.-X., 2018a. The Earth's natural pulse electromagnetic fields for earthquake time-frequency characteristics: Insights from the EEMD-WVD method. *Isl. Arc.* 27 (4), e12256. <https://doi.org/10.1111/iar.2018.27.issue-410.1111/iar.12256>.
- Hao, G.C., Bai, Y.X., Wu, M., Wang, W., Liu, H., 2018b. Time-frequency analysis of the Earth's natural pulse electromagnetic field before earthquake based on BSWT-DDTFA method. *Chin. J. Geophys.* 61 (10), 4063–4074.
- Hao, G.C., Chen, Z.C., Zhao, J., Zeng, Z.X., Liu, H., Sibagatulin, V.G., Kang, F., 2016. Time-frequency analysis of the Earth's natural pulse electromagnetic field signal before and after the Lushan $M_s7.0$ earthquake based on NSFT-WVD transform. *Earth Sci. Front.* 23 (1), 276–286 (in Chinese with English abstract).
- Hattori, K., Han, P., Yoshino, C., Febriani, F., Yamaguchi, H., Chen, C.-H., 2013. Investigation of ULF seismo-magnetic phenomena in Kanto, Japan during 2000–2010: case studies and statistical studies. *Surv. Geophys.* 34 (3), 293–316. <https://doi.org/10.1007/s10712-012-9215-x>.
- Hong, M., Shao, D., Wu, T., Zhang, S., Zhang, Y., Wang, L., Qian, X., 2018. Short-impending earthquake anomaly index extraction of GNSS continuous observation data in Yunnan, Southwestern China. *J. Earth Sci-China* 29 (1), 230–236.
- Hou, Q., 2016. Small Target Detection Algorithm Research in Sea Clutter Based on Spatio-Temporal Chaos. M.S. thesis, Chongqing University, 64 pp (in Chinese).
- Kim, H.S., Eykholt, R., Salas, J.D., 1999. Nonlinear dynamics, delay times, and embedded windows. *Physica D* 127 (1), 48–60.
- Kugiumtzis, D., 1996. State space reconstruction parameters in the analysis of chaotic time series—the role of the time window length. *Physica D* 95 (1), 13–28.
- Leung, H., Lo, T., Wang, S., 2001. Prediction of noisy chaotic time series using an optimal radial basis function neural network. *IEEE T. Neural Networ.* 12 (5), 1163–1172.
- Li, G., Lv, G.F., 2014. Prediction of concrete strength based on regular RBF neural network. *Electron Eng. Des.* 22(13), 52–54+57. (in Chinese with English abstract).
- Liu, Q., Sun, P., Fu, X., Zhang, J., Yang, H., Gao, H., Li, Y., 2020. Comparative analysis of BP neural network and RBF neural network in seismic performance evaluation of pier columns. *Mech. Syst. Signal Pr.* 141, 106707. <https://doi.org/10.1016/j.ymssp.2020.106707>.
- Lorenz, E.N., 1969. Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.* 26 (4), 636–646.
- Lv, J.H., Lu, J.A., Chen, S.H., 2005. Chaotic time series analysis and its application. Wuhan University Press, Wuhan, pp. 63–66 (in Chinese with English abstract).
- Malyshkov, Y.P., Dzhumabaev, K.B., 1987. Earthquake prediction from parameters of Earth's pulse EM field. *Vulkanologiya i Seismologiya* 1, 97–103.
- Malyshkov, Y.P., Malyshkov, S.Y., 2009. Periodicity of geophysical fields and seismicity: possible links with core motion. *Russ. Geol. Geophys.* 50 (2), 115–130.
- Malyshkov, Y.P., Malyshkov, S.Y., 2011. Eccentric rotation of the earth's core and lithosphere: Origin of deformation waves and their practical application. In: Phillips, J.M. (Ed.), *The Earth's Core: Structure, Properties and Dynamics* chapter 6, 1–95.
- Packard, N.H., Crutchfield, J.P., Farmer, J.D., Shaw, R.S., 1980. Geometry from a time series. *Phys. Rev. Lett.* 45 (9), 712–716.
- Rosenstein, M.T., Collins, J.J., De Luca, C.J., 1993. A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D* 65 (1–2), 117–134.
- Ueda, S., Zhang, F., Zhang, S.W., 2014. Earthquake prediction in Japan. *Recent Dev. World Seismol.* 2, 1–10.
- Vorobyov, A.A., 1979. On probability of electrical discharges in the Earth's interior. *Geol. Geoph.* 12, 3–13.
- Wu, Y., Zhong, B.o., Luo, Z., 2018. Investigation of the Tikhonov regularization method in regional gravity field modeling by Poisson wavelets radial basis functions. *J. Earth Sci-China* 29 (6), 1349–1358.
- Xu, P.P., 2017. Weak target detection based on the chaotic characteristics of sea clutter M.S. thesis. Dalian Maritime University, p. 81 (in Chinese with English abstract).
- Yuan, J.H., Gu, Z.W., Chen, B., Wang, C., Di, C.Z., Gao, J.T., 2014. Seismomagnetic observation and research in USA. *J. Seismol. Res.* 37 (1), 163–169.
- Zhang, J.S., Xiao, X.C., 2000. Predicting low-dimensional chaotic time series using Volterra adaptive filters. *Acta Phys. Sin.* 49 (3), 403–408 (in Chinese with English abstract).
- Zhang, Y.M., Qu, S.R., Wen, K.G., 2007. A short-term traffic flow forecasting method based on chaos and RBF neural network. *Syst. Eng.* 25 (11), 26–30.