ISSN: 2414-1895 DOI: 10.6919/ICJE.202009_6(9).0015

RBF Bearing Fault Prediction Model Based on Phase Space Reconstruction

Zhen Song, Jun Song

School of Mechatronic Engineering, Southwest Petroleum University, Chengdu, Sichuan 610500, China.

Abstract

In view of the complex nonlinear characteristics of the temperature time series of the rolling bearing of the tank car, the reconstruction parameters (embedded dimension and delay time) are determined according to the improved C-C algorithm to reconstruct the phase space of the temperature data, and the optimal expansion rate is found by the particle swarm algorithm The factor establishes the Radial Basis Function Neural Network (RBF) for prediction. The test set verification is named PSR-RBF, and compared with the typical algorithms ARMA, LSTM, SVM, the result shows: PSR-RBF neural network model prediction error at about eight out of ten thousand, the PSR-RBF neural network model has higher prediction accuracy compared with the results of other algorithms.

Keywords

Phase space reconstruction, Radial neural network, Particle swarm.

1. Introduction

The current common temperature prediction methods are mainly: continuous method[1], time series analysis method[2] and artificial intelligence algorithm[3]. Although the above prediction methods can predict temperature, they all have their own limitations. The continuous method model, such as the moving average model[4], is the simplest method in the field of temperature prediction. It believes that the temperature prediction value is the moving average of the recent historical temperature time series, but the model has large errors and unstable prediction results. The time series analysis method uses a large amount of historical temperature data to model, and determines the mathematical model of temperature time series through parameter estimation and model testing. The typical prediction models are Auto-Regressive Moving Average ARMA (Auto-Regressive Moving Average ARMA) model[5] and Auto Regressive Integrated Moving Average ARIMA (Auto Regressive Integrated Moving Average ARIMA) model, etc.[6]. The advantages of ARMA and ARIMA models are that they are small in calculation and simple to implement, but they are only suitable for time series with stationary or non-stationary lines. Temperature time series with obvious nonlinear characteristics such as chaos and fractal are difficult to show their complex changes. Make accurate predictions. Artificial intelligence algorithms include genetic algorithms, neural network algorithms, etc., which have the advantages of self-learning, self-organization, and self-adaptation. They are suitable for solving complex problems such as nonlinearities. However, they have slow convergence speed, difficult parameter selection and easy to fall into local optimization. Advantages and disadvantages.

Aiming at the defects of the above model algorithm and the obvious non-linear characteristics of the temperature time series, this paper combines the phase space reconstruction theory to determine the embedding dimension and delay time according to the improved C-C algorithm, and uses the particle swarm algorithm to find the optimal expansion rate factor to establish the path. To the basis function neural network (RBF) prediction model. Make predictions on the test set and perform error analysis.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202009_6(9).0015

2. Introduction to the basic algorithm:

2.1 Normalized

Normalization aims to eliminate the mutual influence between different dimensions and solve the comparability between indicators. Generally, when the amount of data and the data range are large, the convergence speed and accuracy of the model will decrease. In order to solve such problems, we usually scale the data. This article uses the more commonly used (0,1) normalization to map the data to the interval (0,1). The normalization formula is as follows:

$$n(t_i) = \frac{t_i - t_{\min}}{t_{\max} - t_{\min}} \tag{1}$$

The data in the formula is taken from the temperature data of the equipment, t_i is the temperature data to be processed, t_min is the minimum value in the temperature data, and t_max is the maximum value in the temperature data. Through the processing of numerical normalization[7], the original narrow and long elliptical zigzag trend is optimized when the data gradient changes to the trend directly along the shortest path of the center of the circle, which greatly improves the iteration speed. In order to ensure the availability of the prediction algorithm, it is generally necessary to de-normalize the prediction results. The specific calculation formula is as follows:

$$t_i = \frac{(t_{\max} - t_{\min})\,(n(t_i) - n(t_i)_{\max})}{(n(t_i)_{\max} - n(t_i)_{\min})} + t_{\min} \tag{2}$$

2.2 Principal component analysis (PCA)

Principle Component Analysis (PCA for short) is a multivariate statistical method used to test the internal correlation between multiple variables[8]. By finding a set of orthogonal vectors in the sample space, that is, the principal components, to maximize the information in the original data, the main feature components of the multi-dimensional vector data can be extracted, and the internal laws among the multi-dimensional vector data can be revealed[9] . PCA processes the accumulated signal after coherent accumulation and linearly transforms the accumulated signal into a set of principal components arranged in descending order according to the variance. The principal components are not correlated with each other, and then the low-order principal components are extracted for signal reconstruction to remove irrelevant noise and improve the signal-to-noise ratio of the signal[10].

Assuming that n signals are coherently accumulated, and each signal has p sampled data, namely X_1,X_2,⋯X_p, the signal coherent accumulation database matrix is obtained:

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n1} & \cdots & x_{np} \end{pmatrix} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_p) \tag{3}$$

In the formula, $X_i = (x_{1i}, x_{2i}, \cdots, x_{ni})^T$, *i*=1, 2…, p represents the first *i* sample data signals containing p points. Perform principal component analysis on the coherent accumulated signal, and reconstruct the signal after denoising. First, it is necessary to standardize the coherent accumulation signal data matrix to eliminate the dimensional influence of variables, that is:

$$\bar{x}_{ij} = x_{ij} - \frac{1}{n}\sum_{j=1}^{n} x_{ij} \tag{4}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202009_6(9).0015

In the formula, x_{ij} represents the i-th sampling point data of the j-th signal. Standardize all data to obtain a new data matrix X_0. Then, the covariance matrix of the data matrix X is calculated, namely

$$\boldsymbol{C}_X = cov(\boldsymbol{X}) = \frac{1}{n-1}\boldsymbol{X}_0\boldsymbol{X}_0^T \tag{5}$$

Then perform eigenvalue decomposition on the covariance matrix C_x to obtain the eigenvector U, namely

$$\boldsymbol{C}_X = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^T \tag{6}$$

In the formula, $D = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n), \lambda_1 > \lambda_2 > \cdots \lambda_n$, represents the diagonal of the eigenvalues in descending order of magnitude Matrix; $U = (u_1, u_2, \cdots, u_n)$ represents the set of eigenvalues$\lambda_i$corresponding to the eigenvector u_i, U is the orthonormal matrix. Use the eigenvector matrix U to linearly transform the coherent accumulation signal data matrix X to obtain each principal component vector Y, namely

$$\boldsymbol{Y} = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{X} \tag{7}$$

In the formula, a feature with a larger value can be calculated to obtain a low-order principal component vector, which usually corresponds to a target signal with a strong correlation. Feature with small value can be calculated to obtain high-order principal component vector, usually corresponding to noise with poor correlation[11]. Then calculate the cumulative contribution rate corresponding to the first k principal components, namely

$$\alpha = \sum_{i=1}^{k}\lambda_i \Big/ \sum_{i=1}^{p}\lambda_i \tag{8}$$

Finally, according to the cumulative contribution rate corresponding to the first k principal components, the selected value of k is determined. The coherent accumulation signal of the first k principal component vectors is retained for signal reconstruction to achieve the purpose of removing part of the noise and improve the signal-to-noise ratio of the reconstructed signal.

2.3 Phase space reconstruction

According to the embedding theorem proposed by Takens, one-dimensional time series can accurately express the dynamic characteristics of the original system after phase space reconstruction. For the time series$\{X_i, i = 1,2, \cdots, n\}$, n is the sequence length, and the reconstructed phase space[12] is $\{X_i, X_i + \tau, \cdots, X_i + (m-1)\tau\}$, i=1, 2, …, M, where X_i is the point in the phase space, m is the embedding dimension, τ is the delay time, and M is The number of phase points in the reconstructed phase space, $M = n - (m-1)\tau$, the reconstructed phase space is expressed as:

$$\begin{bmatrix} y_1^{\mathrm{T}} \\ y_2^{\mathrm{T}} \\ \vdots \\ y_M^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_M \\ x_{1+\tau} & x_{2+\tau} & \cdots & x_{M+\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & \cdots & x_{M+(m-1)\tau} \end{bmatrix} \tag{9}$$

The core of phase space reconstruction is to choose the appropriate embedding dimension m and delay time τ. This paper determines the embedding dimension and delay time according to the improved C-C algorithm.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202009_6(9).0015

2.4 Particle swarm algorithm

Particle Swarm Optimization (PSO) was proposed by Dr. Eberhart and Dr. Kenned in 1995. It originated from the study of bird predation behavior. The basic core of particle swarm optimization algorithm is to use the information sharing of individuals in the group to make the movement of the whole group produce an evolution process from disorder to order in the problem-solving space, so as to obtain the optimal solution of the problem.

$$V_{k+1} = V_k + c_1(P_x - P_z) + c_2(P_x - P_p) \tag{10}$$

Among them,$c_1$and$c_2$are called social learning factors and individual learning factors.

2.5 Radial Neural Network (RBF)

The basic principle of RBF neural network is: RBF is used as the basis of hidden units to form hidden layer space, the input vector is converted from low-dimensional to high-dimensional space, and finally the weighted summation result of hidden units is output[13]. Modeling and forecasting chaotic time series based on RBF network, the grid structure is shown in the figure. It is a three-layer feedforward network: the first layer is the input layer formed by the signal source, the second layer is the hidden layer, and the hidden unit transformation function in the hidden layer is locally distributed, non-negative and nonlinear[14]; third The layer is the output layer. Among them, the first layer to the second layer are nonlinear transformations, and the second layer to the third layer are linear transformations[15].

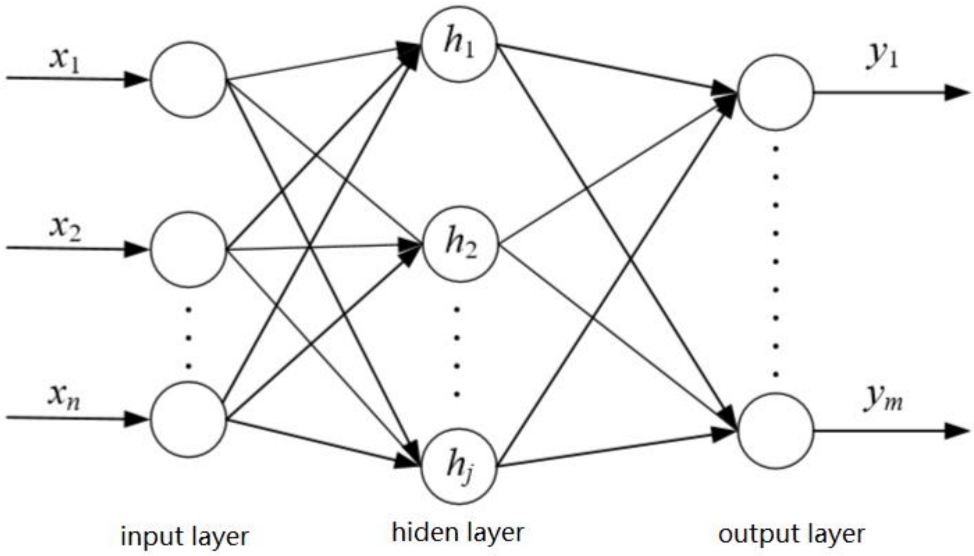

Figure 1-1 Radial basis function neural network

RBF neural network has the following advantages in time series prediction: the network output layer has a linear relationship with the connection weights, has a strong input and output mapping function, has the only best approximation property, and there is no local minimum problem and learning process. It has fast convergence speed and good classification ability.

3. Introduction to the algorithm

3.1 Algorithm flow:

In this article, the data collected by multiple sensors at the same location is listed. In order to eliminate the mutual influence between different dimensions of the data, the entire data is normalized to make

ISSN: 2414-1895 DOI: 10.6919/ICJE.202009_6(9).0015

it have a standardized format for subsequent processing and analysis. When dealing with multidimensional data, it is generally necessary to perform principal component analysis on the data. According to the results of principal component analysis, phase space reconstruction is carried out to identify data features. The radial basis function neural network is constructed using the data features extracted after reconstruction. The specific algorithm flow is as follows.

Based on the RBF neural network to predict the chaotic time series, the first step is to preprocess the data, and then calculate the delay time and embedding dimension based on the improved C-C method, and reconstruct the phase space, on this basis, make the time series non-steady Feature recognition, and then based on RBF neural network to obtain prediction results.

3.2 PSR-RBF prediction

This article predicts the bearing temperature data. The data used is the temperature data of the rolling bearing of the tank car. In the actual data collection, 4 sensors are used to collect the data at the same location. This article selects more than 40,000 pieces of data collected from August 2018 to March 2020 for prediction processing. The algorithm flow is as follows:

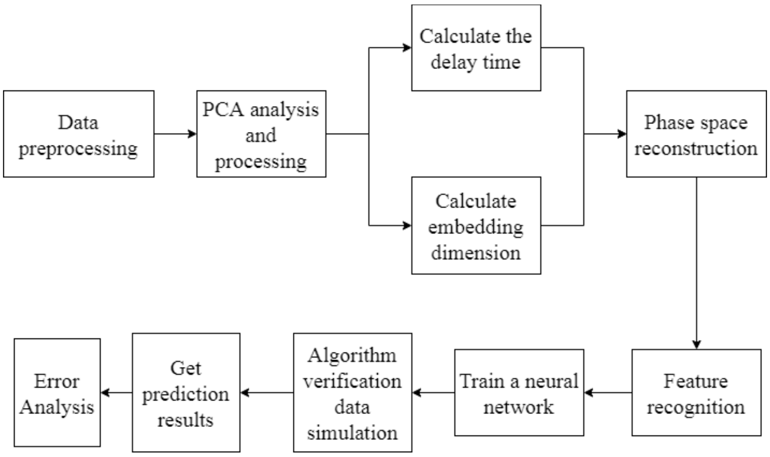

Figure 2-1 Algorithm flow chart

3.3 Data pre-processing

The first step is normalization processing, and the second step is PCA principal component analysis, and the correlation of the four sets of data is shown in the figure:

	Sensor 1	Sensor 2	Sensor 3	Sensor 4
Sensor 1	1	0.998794	0.998563	0.998342
Sensor 2	0.998794	1	0.997857	0.996524
Sensor 3	0.998563	0.997857	1	0.996444
Sensor 4	0.998342	0.996524	0.996444	1

Figure 2-2 Correlation analysis

ISSN: 2414-1895 DOI: 10.6919/ICJE.202009_6(9).0015

As shown in the figure, the four bearing temperature time series collected by the sensors are highly linearly related. In order to simplify the calculation while taking into account the data collected by each sensor, this article will make the average data represent the data at the current moment, namely:

$$X(t) = \frac{x_1(t) + x_2(t) + x_3(t) + x_4(t)}{4} \tag{11}$$

3.4 Phase space reconstruction

After the average data is obtained, the phase space reconstruction is performed on it. The two important parameters of the phase space reconstruction are the delay time \tau and the embedding dimension m, which are directly related to the quality of the phase space reconstruction. This paper adopts the improved C-C method to select the delay parameter τ more accurately, and to select the embedding window τ_w more reliably. The improved C-C algorithm is a simple algorithm based on correlation integral. Its advantages are small calculation amount, easy operation and strong anti-noise ability. Define the associated points as:

$$C(m,N,r,\tau) = \frac{2}{M(M-1)} \sum_{1 \leqslant i < j < M} \theta(r - d_{ij}), r > 0 \tag{12}$$

In the formula, m is the embedding dimension; N is the number of time series; r is the search radius taken in the calculation; τ is the delay parameter;

$$d_{ij} \overset{def}{=} \| x_i - x_j \|_\infty \tag{13}$$

θ is the Heaviside function. Define the test statistics as:

$$S_1(m,N,r,\tau) = C(m,N,r,\tau) - C^m(1,N,r,\tau) \tag{14}$$

When calculating the above formula, use the block average method to calculate, namely

$$S_2(m,N,r,\tau) = \frac{1}{\tau} \sum_{s=1}^{r} [C_s(m,N/\tau,r,\tau) - C_s^m(1,N/\tau,r,\tau)] \tag{15}$$

Define the difference as:

$$\Delta S(m,N,r,\tau) = \max\{S(m,N,r,\tau)\} - \min\{S(m,N,r,\tau)\} \tag{16}$$

When N=1000, m=2, 3, 4, 5, $r = k\sigma/2$, k=1, 2, 3, 4, σ is the standard deviation of the time series $x(n)$, $\tau = 1,2,\cdots,200$, respectively calculate the mean value of $S_1(m,N,r,\tau)$ and $S_2(m,N,r,\tau)$ $\overline{S_1}(\tau)$, $\overline{S_2}(\tau)$ and difference$\Delta\overline{S_1}(\tau)$, $\Delta\overline{S_2}(\tau)$, select the first local minimum point of $\Delta\overline{S_1}(\tau)$ as the delay parameter τ

$$\bar{S}_2(\tau) = \frac{1}{16} \sum_{m=2}^{5} \sum_{k=1}^{4} S_2(m,N,r,\tau) \tag{17}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202009_6(9).0015

$$\Delta\bar{S}_2(\tau)=\frac{1}{4}\sum_{m=2}^{5}\Delta S_2 \tag{18}$$

Take the period point of $|\bar{S}_1(\tau)-\bar{S}_2(\tau)|$ as the optimal embedding window τ_w , $|\bar{S}_1(\tau)-\bar{S}_2(\tau)|$ There is a more significant local peak at the periodic point. The embedding window τ_w is better than the original C-C The global minimum point selected by the method is more reliable. Thus, the embedding dimension m is

$$m=\tau_w/\tau+1 \tag{19}$$

Calculate the embedding dimension m = 3, and the delay time τ=6.

3.5 PSR-RBF neural network prediction model

In a radial basis function neural network, the expansion rate will greatly affect the accuracy of the model. Normally, this parameter is selected by manual testing. In this article, in order to obtain more suitable parameters as soon as possible, this article will use particle swarm algorithm for parameter selection, the number of particles is set to 100, the learning factors c_1 and c_2 are both set to 2, and the number of iterations is set to 1000 Configure this parameter adaptively.

4. Experimental comparison

4.1 Lab environment

For the sake of generality, this article will compare with the following typical algorithms, like PSR-RBF, select 50 for training and 20 for prediction. Use MATLAB as the programming language. The system configuration is Intel(R) Core (TM) i5-8265U CPU @1.60GHz 1.80 GHz 8.00 GB.

4.2 Comparative Results

Autoregressive moving average (ARMA), long short-term memory network (LSTM), support vector machine (SVM) and the prediction model PSR-RBF in this paper are compared, and the prediction results are shown in the figure.

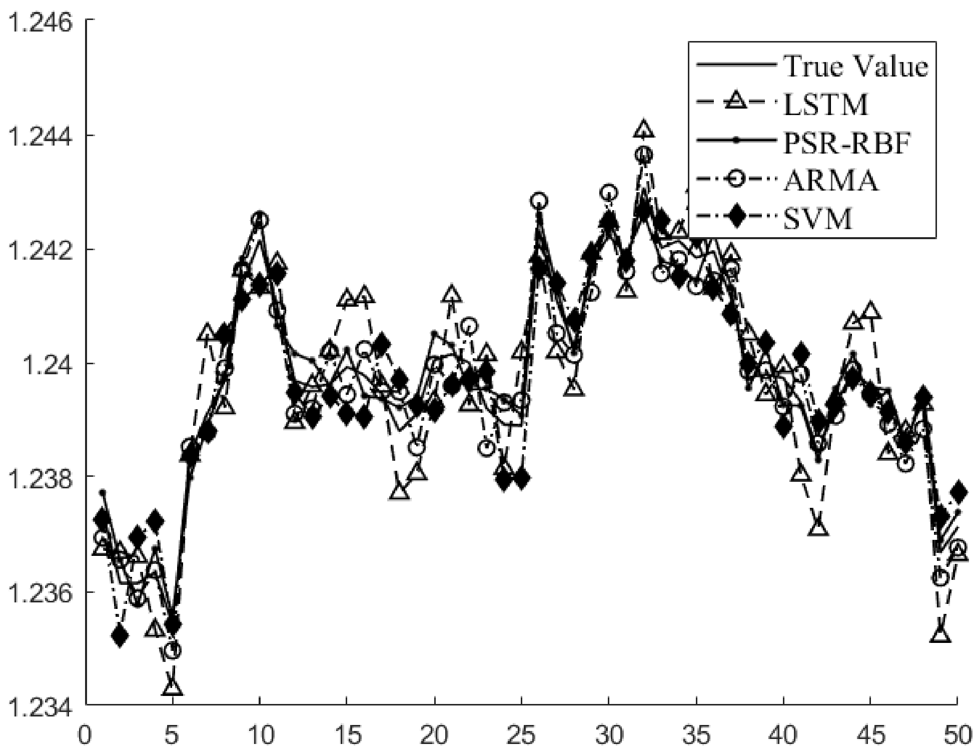

Figure 3-1 Comparison of experimental results

ISSN: 2414-1895
DOI: 10.6919/ICJE.202009_6(9).0015

4.3 Evaluation method

There are many evaluation indicators for chaotic time series prediction models. The commonly used indicators are as follows: absolute value of average relative error (MAPE), absolute value of maximum relative error (M_RE), root mean square error (RMSE)[16].

$$\mathrm{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{\hat{y}(t)-y(t)}{y(t)}\right| \tag{20}$$

$$\mathrm{M\_RE} = \max\left|\frac{\hat{y}(t)-y(t)}{y(t)}\right| \tag{21}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{y}(t)-y(t)\right)^2} \tag{22}$$

Table 3-1 Error analysis and comparison

Evaluation index	ARMA	LSTM	SVM	PSR-RBF
MAPE($\times 10^{-4}$)	2.435	5.209	3.861	1.7975
M_RE($\times 10^{-4}$)	5.575	12.1	7.884	3.981
RMSE($\times 10^{-4}$)	3.638	7.496	5.471	2.624

Analysis can find that PSR-RBF greatly improves accuracy

5. Summary and Outlook

In this paper, the phase space is reconstructed from the temperature data of the rolling bearing of the tank car, and the particle swarm algorithm is used to find the optimal diffusion rate factor to establish the RBF neural network prediction model. The prediction results and the results of the classic algorithms ARMA, LSTM, and SVM are evaluated by MAPE, M_RE and RMSE. Compared with the methods, the experimental results show that the PSR-RBF model greatly improves the prediction accuracy. It is of great significance to the operation and maintenance of tank trucks, which improves production efficiency and reduces the risk of safety accidents and enterprise operating costs.

References

[1] ALEXIADIS M C, DOKOPOULOS P S, SAHSAMANOGLOU H S, et al. Short-term forecasting of wind speed and related electrical power[J]. Solar Energy, 1998, 63(1) :61 -68.

[2] KAMAL L, JAFRI Y Z. Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. [J]. Solar Energy, 1997, 61(1) :23 -32.

[3] RONG F U, WANG W Q, GUI-XIONG H E. The forecasting of wind speed in wind farm based on the meteorological factors with BP neural network[J]. Renewable Energy Resources, 2009, 27 (2): 86 -89.

[4] CHAN J C C. Moving average stochastic volatility models with application to inflation forecast [J]. Journal of Econometrics, 2013, 176(2) :162 -172.

[5] CHE J, WANG J. Short-term electricity prices forecasting based on support vector regression and Auto-regressive integrated moving average modeling[J]. Energy Conversion & Management, 2010, 51(10): 1911 -1917.

[6] HAN Shuang, YANG Yong-ping, LIU Yong-qian. Application study of three methods in wind speed prediction[J]. Journal of North China Electric Power University, 2008, 35(3): 57 -61.

[7] Xiao C W, Li P, Feng Z M. Monitoring annual dynamics of mature rubber plantations in Xishuangbanna during 1987 -2018 using Landsat time series data: A multiple normalization approach[J]. International Journal of Applied Earth Observation and Geoinformation, 2019, 77:30 -41.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202009_6(9).0015

[8] ZHANG Q, CONG P, LU Y, et al. Airborne electromagnetic data levelling using principal component analysis based on flight line difference[J]. Journal of Applied Geophysics, 2018, 151:

[9] ABDI H, WILLIAMS L J. Principal component analysis[J]. Wiley Interdisciplinary Reviews: Computational Statistics, 2010, 2(4): 433- -459.

[10] QIN L, LIU S, LONG T, et al. Wind field reconstruction using dimension reduction of CFD data with experimental vali- dation[J]. Energy, 2018, 151:272 -288.

[11] XIE B, WANG L, LINJ, et al. An improved principal component analysis approach to noise removal for airborne elect romagnetic data [C]// Proc. of the SEG Annual Meeting,2014.

[12] BO Lin, LU Chao, ZHAO Xin. A method of enhancing the fault characteristics in early stage of rolling bearing based on phase space reconstruction and ICA with reference [J] Journal of Vibration, Measurement & Diagnosis, 2016, 36 (6): 10971102.

[13] HUANG Ji. Prediction of Shanghai composite index based on RBF neural network[D]. Hefei: University of Science and Technology of China, 2009.

[14] CHEN Guang. Dynamic simulation of energy model and oxygen system in iron and steel enterprises[D]. Shenyang: Northeastern University, 2002.

[15] WU Hua-wen, WANG Fu-zhang. Chaotic time series prediction of railway passenger and cargo traffic based on RBF neural network [C]/ /National doctoral academic conference, Qingdao, 2012.

[16] LU Z B, CAI Z M, JIANG K Y. Determination of embedding parameters for phase space reconstruction based on improved C 一C method [J]. Journal of System Simulation, 2007, 19(11): 2527- 2529.