

The power load's signal analysis and short-term prediction based on wavelet decomposition

Huan Wang^{1,2} · Min Ouyang^{2,3} · Zhibing Wang^{2,3} · Ruishi Liang⁴ · Xin Zhou^{2,5}

Received: 19 August 2017 / Revised: 11 October 2017 / Accepted: 31 October 2017 / Published online: 14 November 2017
© Springer Science+Business Media, LLC 2017

Abstract The complex signal represented by power load is affected by many factors, so the signal components are very complicated. So that, it is difficult to obtain satisfactory prediction accuracy by using a single model for the complex signal. In this case, wavelet decomposition is used to decompose the power load into a series of sub signals. The low frequency sub signal is remarkably periodic, and the high frequency sub signals can prove to be chaotic signals. Then the signals of different characteristics are predicted by different models. For the low frequency sub signal, the support vector machine (SVM) is adopted. In SVM model, air temperature and week attributes are included in model inputs. Especially the week attribute is represented by a 3-bit binary encoding, which represents Monday to Sunday. For the chaotic high frequency sub signals, the chaotic local prediction (CLP) model is adopted. In CLP model, the embedding dimension and time delay are key parameters, which determines the prediction accuracy. In order to find the optimal parameters, a segmentation validation algorithm is proposed in this paper.

The algorithm segments the known power load according to the time sequence. Then, based on the segmentation data, the optimal parameters are chosen based on the prediction accuracy. Compared with a single model, the prediction accuracy of the proposed algorithm is improved obviously, which proves the effectiveness.

Keywords Power load · Wavelet transform · Support vector machine · Chaotic local prediction · Segmentation validation

1 Introduction

The prediction of complex time series depends on the development of nonlinear science and machine learning, and many scholars have made valuable exploration in this field [1–4]. Power load is a typical complex time series, which is affected by many factors, such as demand, climate, policy and production condition. Accordingly, the power load data also has complex signal components. The conclusions are as follows: power load data is generated by low dimensional chaotic system, and contains both periodic and chaotic components. Any single prediction model is based on a specific hypothesis, and has a better prediction effect on a particular type of data. For example, the chaotic prediction algorithm is suitable for the prediction of typical chaotic signals (Lorenz signals, Rossler signals). As for the mixed signals represented by power load, it is difficult for a single model to obtain satisfactory prediction results.

In view of the complexity of components, the power load should be decomposed into sub signals of different characteristics by means of mathematical transform method first, then the sub signals are predicted separately. Because of its good time-frequency analysis characteristics, wavelet transform has been used in this paper [5–7]. Using wavelet transform,

✉ Min Ouyang
ouyangmin16@163.com

¹ School of Electron and Information Engineering, University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan 528402, China
² Key Laboratory of Intelligent Information Perception and Processing Technology (Hunan Province), Zhuzhou 412008, China
³ School of Computer, Hunan University of Technology, Zhuzhou 412008, China
⁴ School of Computer Engineering, University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan 528402, China
⁵ School of Sciences, Hunan University of Technology, Zhuzhou 412008, China

we can get the sub signals of the power load. The low frequency sub signal is remarkably periodic, which is affected by temperature. The high frequency sub signals are very similar to typical chaotic signals, in which the Lyapunov exponent [8,9] is positive. Different prediction methods are adopted respectively for different sub signals. The support vector machines (SVM) [10,11] algorithm is adopted for the low frequency sub signals with temperature considered. At the same time, chaotic prediction algorithm [12–14] is applied to other high frequency sub signals.

In SVM model, air temperature and week attributes are included in model inputs. Especially the week attribute is represented by a 3-bit binary encoding, which represents Monday to Sunday. Meanwhile, there are a key problem in chaotic local prediction (CLP) model, the accuracy of phase space reconstruction. Because the parameter of phase space reconstruction, embedding dimension and delay time, can hardly be calculated accurately, which causes the main error of the CLP model. Attempting to solve this problem, the paper proposes a parameter optimization algorithm based on segmentation validation. Based on the initial calculation of phase space parameters, the algorithm can find the optimal reconstruction parameters based on the known data, which can effectively improve the quality of phase space reconstruction and improve the prediction accuracy.

Through the actual example, the prediction result proves: compared with the prediction using a single mathematical model, for the complex signal represented by the power load, the prediction accuracy based on the proposed algorithm is improved obviously. At the same time, the algorithm is also suitable for prediction of other complex signals, such as traffic flow signals and ECG signals. For the prediction of complex signals, granular computing [15–21] is also a promising method. The theory can be used for reference to chaos theory, so as to deepen the cognition of internal dynamics characteristics of complex signals.

2 The wavelet decomposition of power load

2.1 Basic theory of wavelet transform

Wavelet transform is a powerful tool for signal analysis. Wavelet is a wave with finite energy, and is used as base function to analyze transient, unsteady or time-varying signals. Wavelet transform analyzes signals by means of complex expansion, which is similar to the Fourier transform. The difference is that the signal is decomposed into a series of local base functions called wavelets. Therefore, wavelet transform is to expand the signal on a particular wavelet base function. Wavelet transform is divided into two categories: continuous wavelet transform and discrete wavelet transform.

2.1.1 Continuous wavelet transform and discrete wavelet transform

$g(t)$ is a mother wavelet function, and other wavelet basis functions are obtained by stretching and translating the $g(t)$ as follows:

$$g_{a,b}(t) = \frac{1}{\sqrt{a}} g\left(\frac{t-b}{a}\right) \quad (1)$$

In formula (1), a and b are real number ($a \neq 0$), they are scale factor and position factor respectively.

$s(t)$ is a square integrable function, and its continuous wavelet transform is defined as follows:

$$\begin{aligned} W_{g|s}(a,b) &= \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(t) g^*\left(\frac{t-b}{a}\right) dt \\ &= \int_{-\infty}^{\infty} s(t) g_{a,b}^*(t) dt \end{aligned} \quad (2)$$

In formula (2), $g_{a,b}^*(t)$ is conjugate complex of $g_{a,b}(t)$.

Variables a , b , and t are continuous variables. In order to implement the wavelet transform effectively on the computer, a , b and t should take discrete values. In addition, for reducing information redundancy, a and b do not need continuous value. A and B can be discretized in the following manner:

$$a = a_0^j, a_0 > 0, j \in Z \quad (3)$$

$$b = kb_0 a_0^j, b_0 > 0, k \in Z \quad (4)$$

Discrete wavelet basis function $g_{j,k}(t)$ is defined as:

$$g_{j,k}(t) = \frac{1}{\sqrt{a_0^j}} g\left(\frac{t - ka_0^j b_0}{a_0^j}\right) dt, j, k \in Z \quad (5)$$

The discrete wavelet transform is defined as follows:

$$W_{g|s}(j,k) = \int_{-\infty}^{\infty} s(t) \bar{g}_{j,k}(t) dt \quad (6)$$

2.1.2 Wavelet decomposition and reconstruction

In 1989, Mallat [22] unified the construction of wavelet function based on the idea of multi-resolution, and proposed a decomposition and reconstruction algorithm for discrete signals by means of wavelet transform, namely Mallat algorithm. According to the Mallat algorithm, the non-stationary discrete time series can be decomposed into high frequency detail series d_1, d_2, \dots, d_J , and a low frequency series a_J , and J is the maximum decomposition layers. Let the initial time series is S , and the decomposition process is shown in Fig. 1.

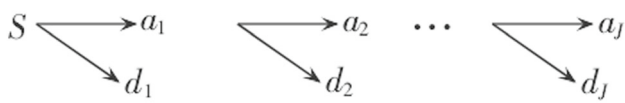


Fig. 1 Process of wavelet decomposition

The above decomposition process can be expressed in formula (7):

$$\begin{cases} a_{j+1} = H(a_j) \\ d_{j+1} = G(a_j) \end{cases} \quad j = 1, 2, \dots, J \quad (7)$$

Among them, $H()$ and $G()$ are low frequency and high frequency decomposition functions, which are similar to the function of low pass filter and high pass filter.

After the time series are decomposed, the decomposition signal of each layer is halved before the relative decomposition, so it is necessary to do the two-interpolation reconstruction for the decomposed signal.

$$\begin{cases} A_j = (H^*)^j a_j \\ D_j = (H^*)^{j-1} G^* d_j \end{cases} \quad j = 1, 2, \dots, J \quad (8)$$

Among them, H^* is a dual operator of H , and G^* is a dual operator of G . After the reconstruction of $d_1, d_2 \dots d_J$ and a_J , the reconstructed series $D_1, D_2 \dots D_J$ and A_J is obtained, whose length is the same as that of the original time series, and:

$$S = D_1 + D_2 + \dots + D_J + A_J \quad (9)$$

There are 2 problems in wavelet decomposition which should be paid attention to:

1. The layer of decomposition can neither be too large nor too small. If the layer is too large, more models should be built to predict the components after decomposition, each model will introduce a certain error, which will lead to a larger prediction error; if the layer is too small, the component with different frequency characteristics in the original signal cannot be separated effectively.
2. Wavelet basis selection. Selecting different wavelet basis will yield different decomposition components, which will affect the final prediction results.

2.2 Example analysis

Power load from Shenzhen is used as an example in this paper. Because of the developed economy and hot climate, the air conditioning load in Shenzhen is relatively high, which causes the power load to fluctuate more. The sampled interval of the power load is 30 min, a total of 48 points in a day. The typical day curve of the power load is shown below:

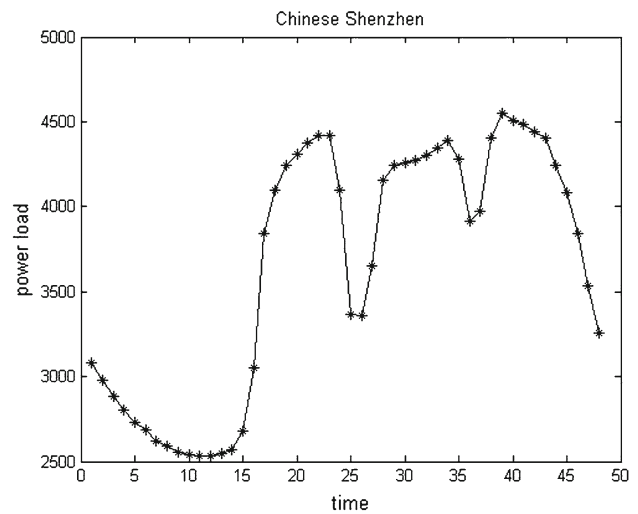


Fig. 2 The typical day power load curve

In the Fig. 2, the power load is minimum at 5–6 a.m. Between 7:30 and 10 in the morning, the power load increases rapidly. Then, at 11, 17 and 20, the power load reaches the three peaks. In a day, the power load fluctuates up and down, which makes the prediction more difficult.

2.2.1 Wavelet decomposition of the power load

One-week power load signal is used as a sample. Firstly, the signal is decomposed by wavelet, and then the characteristics of each sub signal is analyzed. After the experimental comparison, the dmey wavelet base is chosen to decompose and reconstruct the power load in 3 layers. The original signal and the decomposed sub signals are shown in the following Fig. 3.

The first subplot is the original power load signal and other sub plots are low frequency sub signals, high frequency sub signals 1, high-frequency sub signals 2 and high-frequency sub signals 3. The sum of all the sub signals is the original power load signal.

2.2.2 Analysis of the sub signals

1. The low Frequency Sub Signal

Shown in Fig. 3, the low frequency sub signal is obviously periodic, but its fluctuation range also changes. The proportion of air conditioning load to total load is high, so the temperature is much related to the fluctuation of the low frequency sub signals. The low frequency sub signal and the daily maximum temperature are shown in Fig. 4: In Fig. 4, the curve corresponding to the left vertical axis is the low frequency sub signal, and the fold line corresponding to the right vertical axis is the highest daily temperature. Shown in Fig. 4, the daily maximum

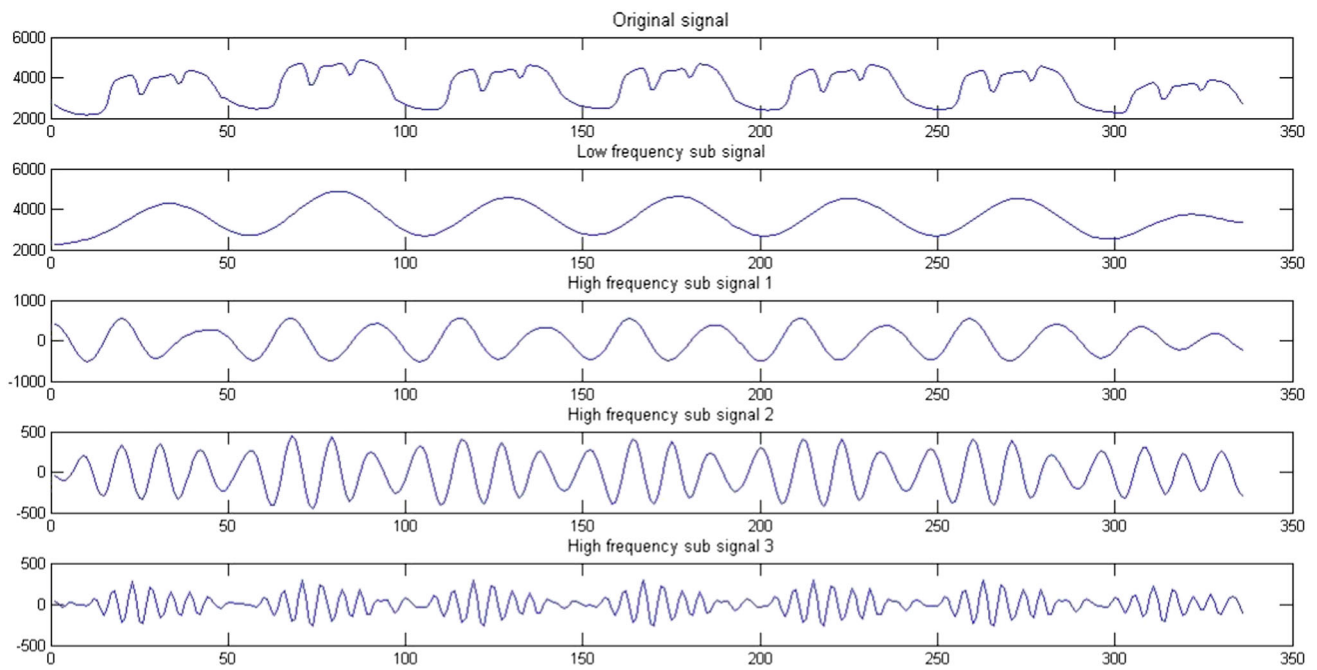


Fig. 3 The original data and each sub data

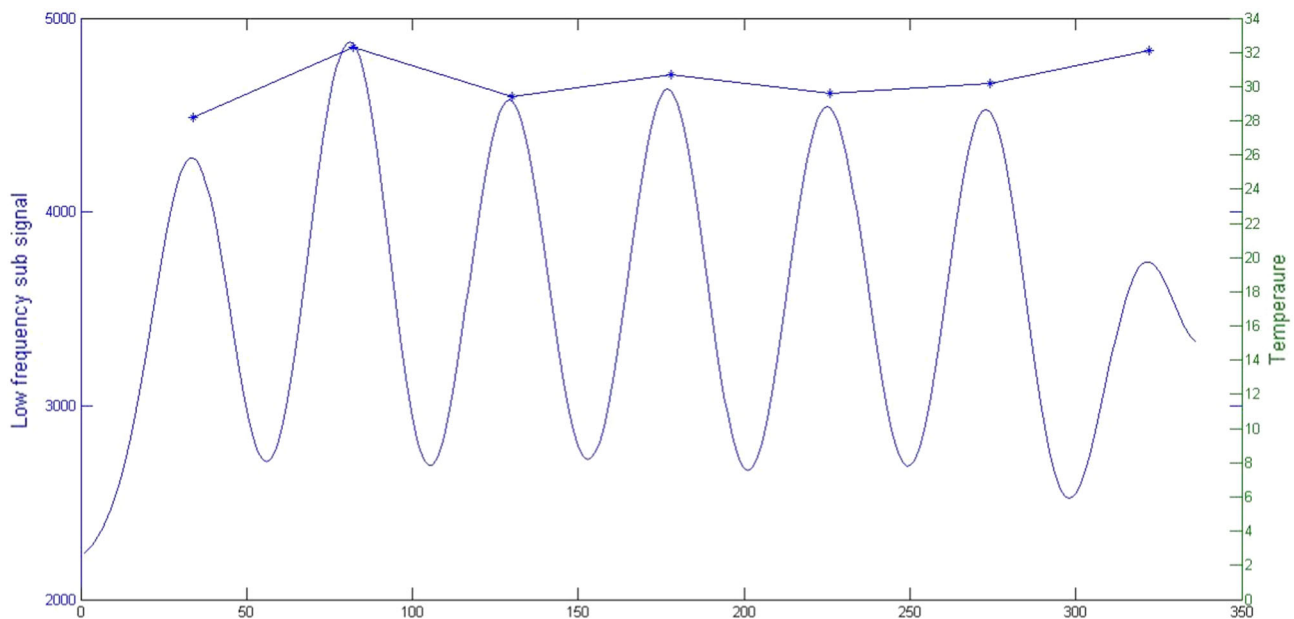


Fig. 4 Low frequency sub signal and daily maximum temperature

temperature is highly correlated with the low frequency sub signal, except for the last day. The reason for the low data correlation on the last day is that the day is Sunday, and Sunday's power load will be much smaller than the workday. In view of the high correlation between the low frequency sub signal and temperature, this paper uses SVM model to describe the relationship between signal fluctuation and air temperature. The specific algorithm will be described in detail later.

2. Other High Frequency Sub Signals

Other high frequency sub signals are highly similar to typical chaotic signals, such as Lorenz signal and Rossler signal. The above high frequency sub signals and typical chaotic signals are drawn as follows:

Shown in Figs. 5 and 6, at different horizontal scales, low frequency sub signal 1, 2 and Rossler chaotic attractor, low

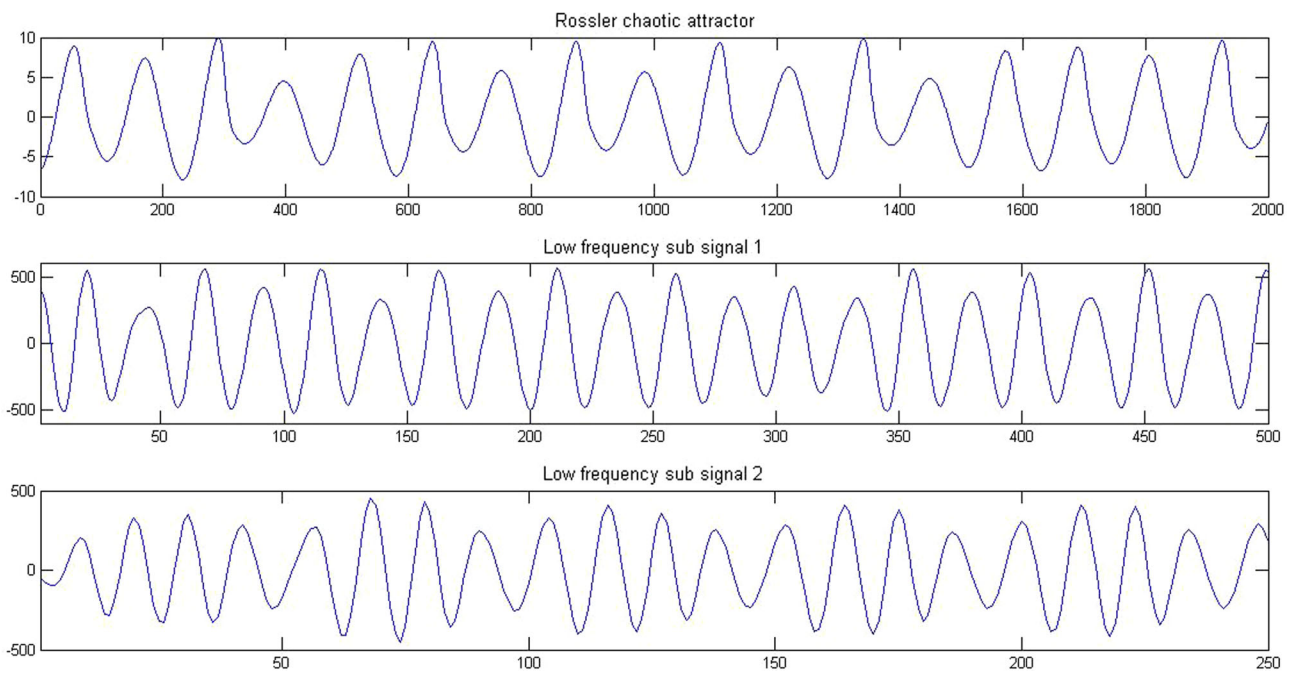


Fig. 5 Rossler chaotic attractor, low frequency sub signal 1 and 2

frequency sub signal 3 and Lorenz chaotic attractor, are very similar to each other. All of these signals show periodicity and instability at the same time. In order to verify the chaotic characteristics of high frequency sub signals, the largest Lyapunov exponents of the above signals are calculated. The basic feature of chaotic systems is the extreme sensitivity to the initial values, and the Lyapunov exponent is a quantitative description of this phenomenon, which characterizes the average exponential rate of convergence or divergence between adjacent orbits in a phase space. The existence of dynamical chaos in the system can be intuitively judged if the largest Lyapunov exponent is greater than zero. Calculating the largest Lyapunov exponent of these high frequency sub signals, the calculating results are 0.64, 0.82 and 1.22, which demonstrates that the high frequency sub signals are chaotic. Based on the chaotic characteristics of these signals, this paper uses CLP algorithm to predict them. The specific content is described in detail later.

3 Power load prediction algorithm based on wavelet decomposition

3.1 The main flow of the algorithm

Based on the wavelet decomposition, the power load is decomposed into low-frequency sub signal (LF), high-frequency sub signal 1 (HF1), high-frequency sub signal 2 (HF2), and high-frequency sub signal 3 (HF3). Combined with the temperature factor (T), SVM model is used to pre-

dict LF; at the same time, CLP algorithm is used to predict HF1, HF2 and HF3. The algorithm flow chart is as follows (Fig. 7).

3.2 LF prediction based on SVM

3.2.1 Basic theory of SVM

SVM is a machine learning method, which is based on the statistical learning theory and the principle of structural risk minimization. For the limited sample set, SVM can find the best compromise between model complexity and learning ability to obtain the best generalization ability. The algorithm is essentially a convex quadratic programming problem with no local optimum, which can effectively solve the small sample and nonlinear regression problems. Through the inner product kernel function, SVM can map a low dimensional nonlinear regression problem into a high dimensional space, and then become a linear regression problem, which reduces the difficulty of solving complex regression problems.

The training of SVM can be reduced to solve a linear constrained a quadratic programming (QP) problem. At present, the most typical training algorithm is sequential minimal optimization (SMO) [23,24]. The algorithm decomposes a large QP problem into a series of small scale QP sub-problems. To avoid the use of numerical methods to optimize the QP in the inner loop, the convergence speed of the algorithm is increased by thousands of times compared with the traditional method.

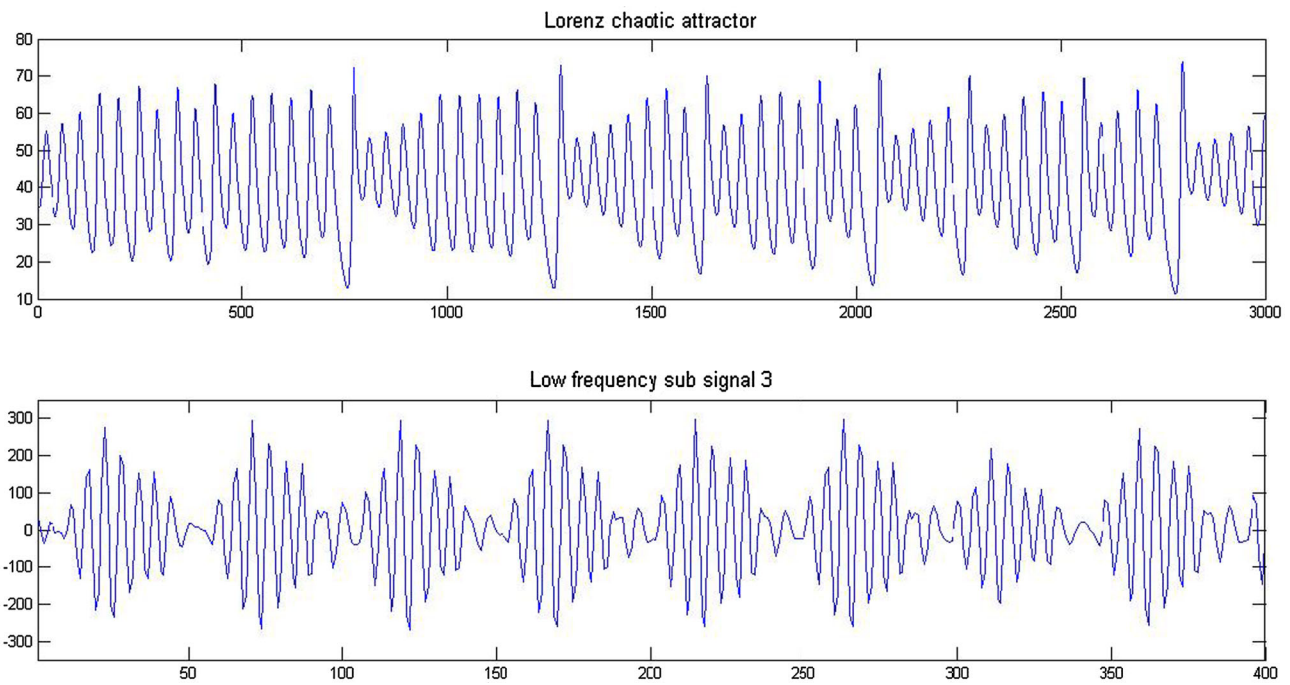


Fig. 6 Lorenz chaotic attractor and low frequency sub signal 3

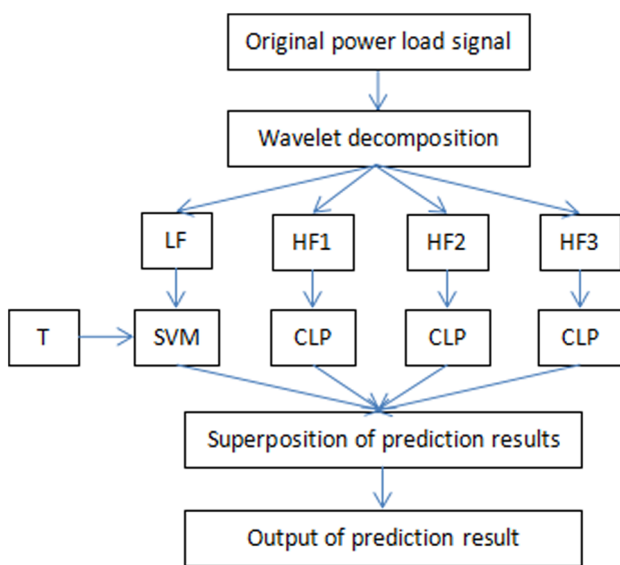


Fig. 7 The algorithm flow chart

3.2.2 Generation and pre-processing of training samples

LF is associated with temperature and week attributes, so these factors should be taken into account when predicting. Before prediction using SVM, the input and output of the model must be defined. The SVM model is trained using a rolling training method, that is, the LF value at the same time of the first seven days is used to predict the LF value at the same time of the next day.

The input sample consists of 3 feature vectors, which are:

1. $L = [L_1, L_2, L_3, L_4, L_5, L_6, L_7]$, LF value for the corresponding time of 7 days prior to the forecast date;
2. T , the temperature of the forecast day;
3. $W = [001, 010, 011, 100, 101, 110, 111]$, the feature vector is a week attribute, represented by a 3 bit binary encoding, which represents Monday to Sunday.

The output sample is the LF value for the corresponding time of the forecast day.

In order to improve the quality of training, the input set and output set should be normalized before training. The normalized formulas corresponding to the load data and the air temperature data are:

$$\bar{L}_i = \frac{L_i - L_{min}}{1.5L_{max} - L_{min}} \tag{10}$$

$$\bar{T}_i = \frac{T_i - T_{min}}{1.2T_{max} - T_{min}} \tag{11}$$

L_{max}, L_{min} are the maximum and minimum values in the power load data set, T_{max}, T_{min} are the maximum and minimum values in the temperature data set.

After the prediction results are obtained, the correlation coefficient is used to evaluate the accuracy of the prediction results. The correlation coefficient r is defined as below:

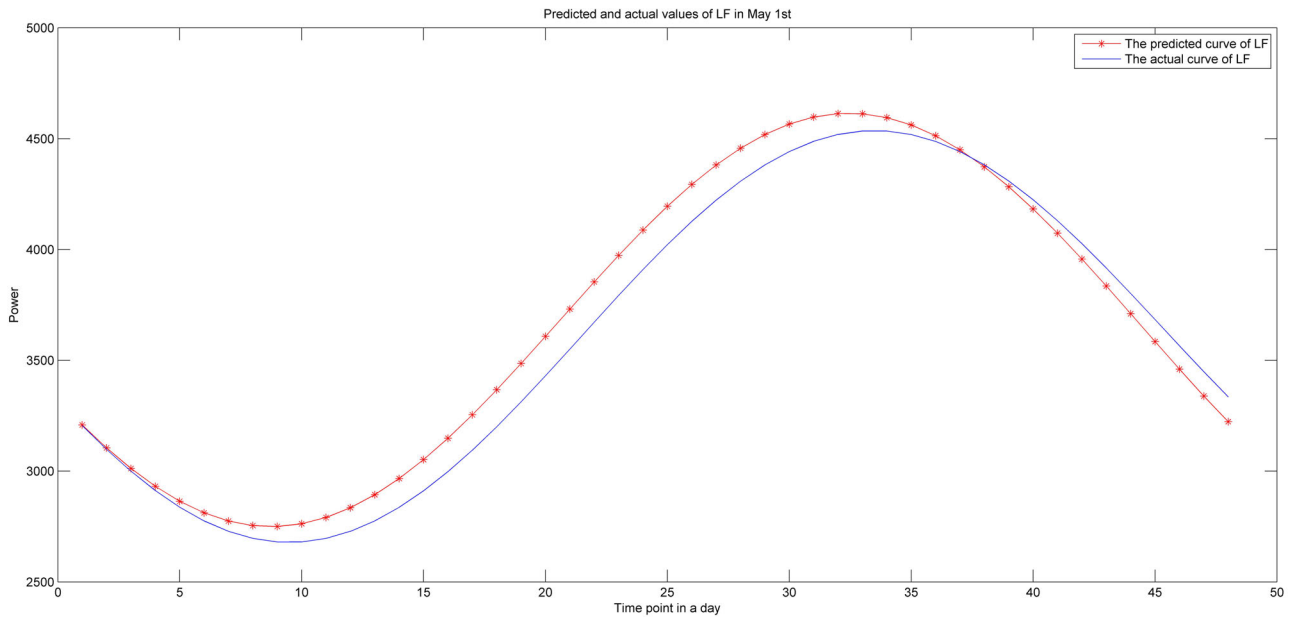


Fig. 8 Predicted and actual values of LF in May 1st

$$r = \frac{\sum_{i=1}^n [(P_{mi} - \bar{P}_m) \cdots (P_{pi} - \bar{P}_p)]}{\sqrt{\sum_{i=1}^n (P_{mi} - \bar{P}_m)^2 \cdots \sum_{i=1}^n (P_{pi} - \bar{P}_p)^2}} \quad (12)$$

In formula (12), P_{mi} is the actual value of a certain moment, \bar{P}_m is the average of all actual value samples; P_{pi} is the predicted value of a certain moment, \bar{P}_p is the average of all predicted value samples.

3.2.3 Example analysis

This paper uses the power load of Shenzhen city to verify the above algorithm, predicts 48 data points in 1 May based on 30 days power load and corresponding air temperature data in April. The prediction results are shown in the following figure:

In Fig. 8, the solid line with an asterisk is the predicted value, and the ordinary solid line is the actual value. It can be seen that the predicted values are in good agreement with the actual values. The correlation coefficient r between the predicted and actual values is 0.9904. Because the prediction step length is 48, the prediction accuracy is satisfactory.

3.3 HF prediction based on chaotic local prediction (CLP)

3.3.1 Traditional chaotic local prediction algorithm

For a n -dimensional chaotic system, $Y(t + 1) = F(Y(t))$, $Y(t), Y(t + 1) \in \mathbb{R}^n$, Where t represents timing sequence, F is a smooth continuous function. For this system, a one-dimensional time series $\{x(t), t = 1, 2, \dots, N\}$ can be

observed. Based on the phase space reconstruction theory, the phase points sequence in the m -dimensional phase space can be obtained:

$$X(t) = \{x(t), x(t + \tau), \dots, x(t + (m - 1)\tau)\} \quad t = 1, 2, \dots, L \quad (13)$$

In formula (13), m is the embedding dimension, and τ is the delay time, $L = N - (m - 1)\tau$.

According to the Takens embedding theorem [25], when τ is appropriate and $m \geq 2n + 1$, the deterministic mapping F^m exists:

$$X(t + 1) = F^m(X(t)) \quad (14)$$

Formula (14) is the reconstruction system, which has the same dynamic characteristics as the original system.

For the calculation of m and τ , Gautama introduces the differential entropy [26] in information theory into phase space reconstruction. Differential entropy can effectively evaluate the degree of disorder of time series under different embedding conditions, and then determine the embedding dimension and delay time at the same time.

For discrete time series, the differential entropy can be estimated using the Kozachenko–Leonenko (K–L) method:

$$H(x, m, \tau) = \sum_{j=1}^N \ln [N\rho(j)] + \ln 2 + C_E \quad (15)$$

In formula (15): N is the data length, C_E is Euler constant, $\rho(j)$ is the Euclidean distance between phase point j and its nearest neighbor.

Since the upper formula is not robust to the embedding dimension, it is necessary to standardize the $H(x_i, m, \tau)$ via the surrogate data $x_{s,i}$ of x . The surrogate data $x_{s,i}$ of x is generated by iterative amplitude matched Fourier transform, $H(x_i, m, \tau)$ is the differential entropy of the no.i surrogate data of original data. After using surrogate data, the distribution characteristics of the signals are not affected. Use formula (15) can compute the original sequences $H(x_i, m, \tau)$ and surrogate data $H(x_{s,i}, m, \tau)$, thus the differential entropy rate can be expressed as:

$$I(m, \tau) = \frac{H(x, m, \tau)}{H(x_{s,i}, m, \tau)_i} \quad (16)$$

In formula (16), $\langle \bullet \rangle_i$ means the average value of i times calculation. After adding the penalty factor to the embedding dimension, the differential entropy rate is:

$$R_{ent}(m, \tau) = I(m, \tau) \left(1 + \frac{m \ln N}{N} \right) \quad (17)$$

The smaller the differential entropy rate is, the more orderly the chaotic attractor in the phase space is, the better the dynamic characteristics of the system can be reflected. Therefore, the m and τ when the differential entropy rate is minimum are the calculation value of embedding dimension and the delay time.

After the reconstruction of phase space, takes $X(L)$ as the reference phase point, and calculates the Euclidean distance between other phase points and $X(L)$. According to the Euclidean distance values, q phase points are chosen as the reference neighborhood $X(L_i)$, $i = 1, 2, \dots, q$. According to the correlation between the points in the reference neighborhood and $X(L)$, computes weights for each phase point in the neighborhood. The Euclidean distance of each point in the neighborhood to $X(L)$ is d_i , $i = 1, 2, \dots, q$, d_{\min} is the minimum. Then the weight of the point i in the neighborhood is:

$$P_i = \frac{\exp[-(Cor_i/Cor_{min})]}{\sum_{i=1}^q \exp[-(Cor_i/Cor_{min})]} \quad (18)$$

For the phase points in the reference neighborhood, the evolution in the phase space is described by a first order function. Taking into account the weight of each phase point, the least squares method is used to minimize the formula (18), and the suitable first order parameters a and b can be obtained:

$$\min \sum_{i=1}^q P_i [X(L_i + h) - aR - bX(L_i)]^2 \quad (19)$$

$R = [1, 1, \dots, 1]^T$, h is the prediction step.

The first order parameters, a and b , are substituted into the formula (20):

$$\ddot{X}(L + h) = aR + bX(L) \quad (20)$$

The last element $\ddot{X}(L + h)_m$ in $\ddot{X}(L + h)$ is the h -step prediction value of the original time series.

3.3.2 Segmentation verification of traditional chaotic local algorithm

The main problem of traditional CLP algorithm is the calculation of the parameters for phase space reconstruction. For a real time series, the calculated embedding dimension and time delay are usually not optimal, so the method of segmentation verification can be used to find the optimal parameters according to the calculated values. The idea of segmentation verification comes from cross validation [27,28]. Cross validation is an efficient method for parameter optimization in pattern recognition. The data set DS is divided into k mutually exclusive subsets:

$$DS = DS_1 \cup DS_2 \cdots \cup DS_k, DS_i \cap DS_j = \emptyset (i \neq j) \quad (21)$$

The union of $k - 1$ subsets is regard as a training set, and the remaining subset is regard as a test set. In this way, K times training and testing can be carried out, and finally the average of K times test classification results can be obtained. The model parameters with the highest classification accuracy are the optimal parameters, which can be used to classify the unknown data.

The time series prediction problem is different from the classification problem. The data sequence in the time series cannot be arbitrarily disrupted. So the order of the data sets cannot be exchanged when the model parameters are identified. For this problem, the known data set DS can be divided into n subsets:

$$DS = DS_1 \cup DS_2 \cdots \cup DS_n, DS_i \cap DS_j = \emptyset (i \neq j) \quad (22)$$

$D_1 \dots D_{m-1}$ can be used as the training set, and D_m as the test set, $k < n$; Then $D_2 \dots D_m$ can be used as the training set, and D_{m+1} as the test set. And so on, so that $n - m$ times training and testing can be carried out, and finally the average of $n - m$ times test results are obtained. The model parameters with the highest prediction accuracy are the optimal parameters, which can be used to predict the unknown data. In the

Table 1 The parameters and optimization range of phase space reconstruction

	Embedding dimension	Optimization range	Delay time	Optimization range
HF1	4	[2–8]	5	[1–10]
HF2	6	[2–12]	3	[1–7]
HF3	5	[2–10]	2	[1–5]

Table 2 The best reconstruction parameters

	Best embedding dimension	Best delay time
HF1	3	1
HF2	4	4
HF3	6	2

process of time series prediction, new data will be added, and the training set and test set will be updated according to the timing of the data. Based on this method, the optimal model parameters can be adaptively identified according to the known data, and the prediction accuracy can be guaranteed.

For the prediction of chaotic time series, calculate the embedding dimension m and the delay time τ at first, then select the corresponding threshold tm and τ . In the range of $[m - tm, m + tm]$ and $[\tau - \tau, \tau + \tau]$, select the optimal parameters using the above method, can guarantee the quality of the reconstructed phase space, so as to improve the accuracy of chaotic prediction algorithm.

3.3.3 Example analysis

The above algorithm is used to predict the high frequency sub signals HF1, HF2 and HF3. First of all, the differential entropy algorithm is used to calculate the embedding dimension and delay time of HF1, HF2 and HF3. The result of calculation and the range of parameter optimization are shown in Table 1.

Aiming at the above 3 signals, the segmentation validation method is used to find the best phase space reconstruction parameters in the optimization range. After calculation, the best reconstruction parameters corresponding to the 3 groups of signals are shown in Table 2.

Using the 3 sets of parameters to predict HF1, HF2 and HF3 based on CLP algorithm, the prediction results are shown in Figs. 9, 10 and 11.

In these figures, the solid line with an asterisk is the predicted value, and the ordinary solid line is the actual value. The correlation coefficient r between the predicted and actual values of the above three signals are 0.9990, 0.9401 and 0.8720. From the above prediction results, it can be seen that the higher the signal frequency, the greater the difficulty

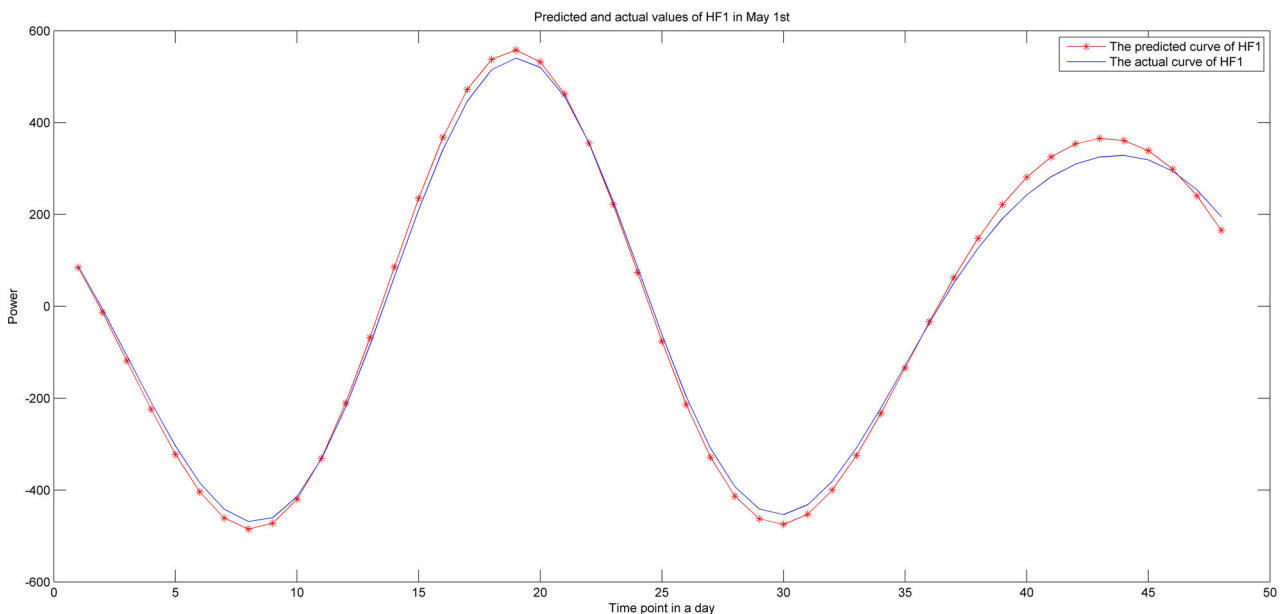


Fig. 9 Predicted and actual values of HF1 in May 1st

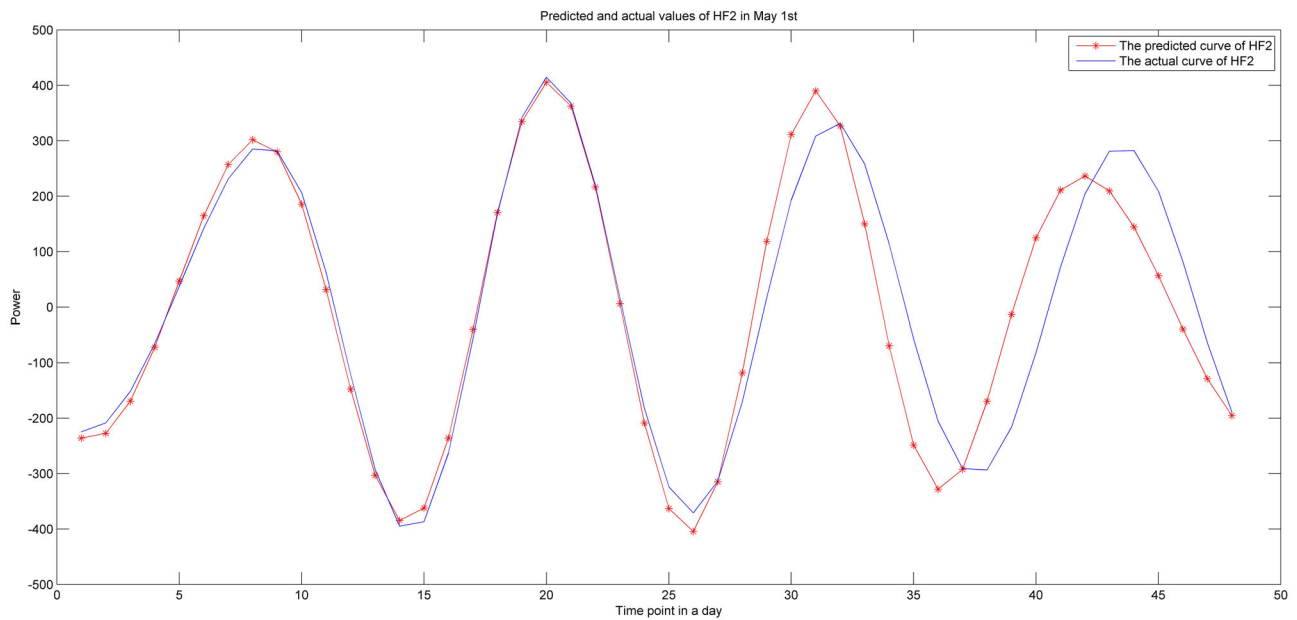


Fig. 10 Predicted and actual values of HF2 in May 1st

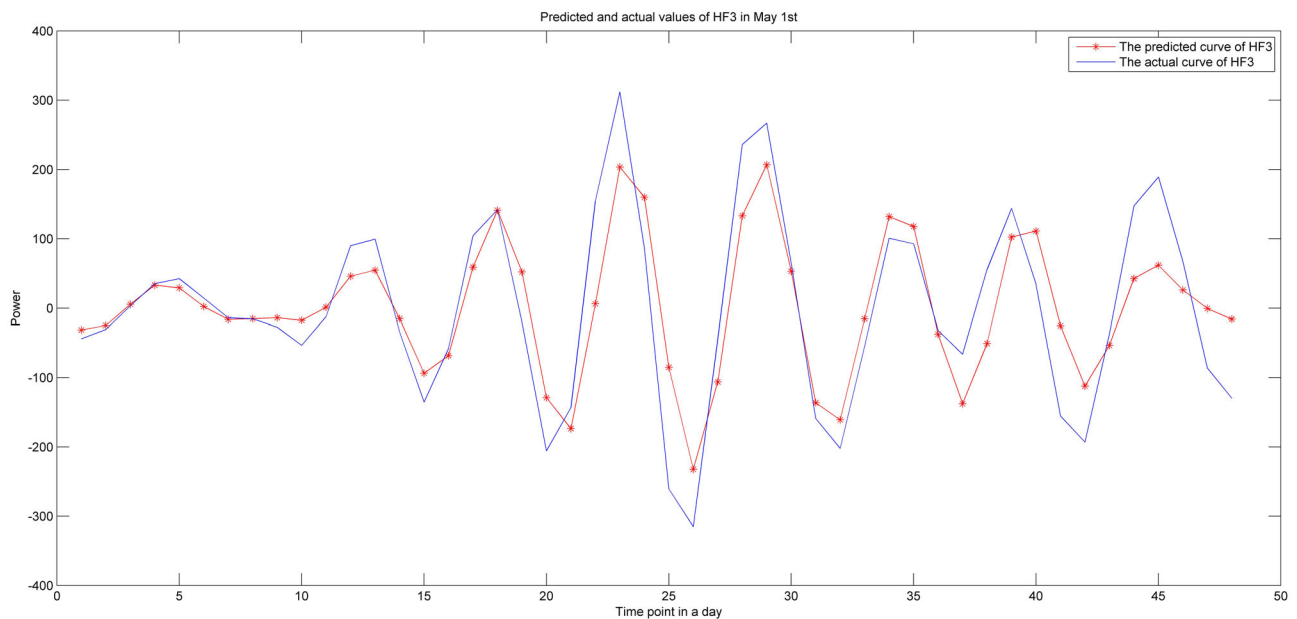


Fig. 11 Predicted and actual values of HF3 in May 1st

of prediction, and the prediction accuracy is also reduced. But the overall prediction accuracy is satisfactory.

4 Superposition of prediction results

In Sect. 3, the sub signals of power load are predicted respectively. In this section, the prediction results of each of the sub signals are superimposed together. Sum up the prediction results of LF, HF1, HF2, HF3, and get the total prediction

results. In order to verify the effectiveness of the proposed algorithm, a single CLP model is used to predict the power load at the same time. All these prediction values are shown in Fig. 12.

In Fig. 12, the black solid line represents the actual power load, the star line represents the prediction value of the algorithm in this paper, and the diamond dotted line represents the prediction value of a single chaotic model. Shown in Fig. 12, the prediction value of the algorithm in this paper is obviously more consistent with the actual value. In order

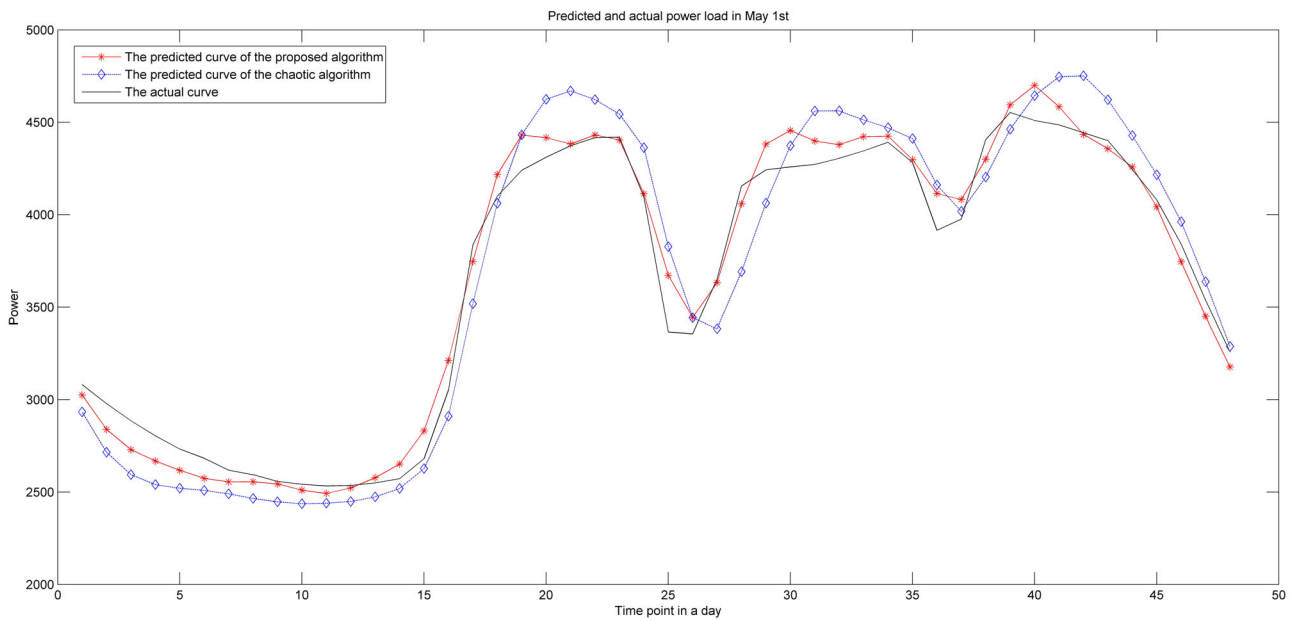


Fig. 12 Predicted and actual power load in May 1st

Table 3 The prediction accuracy

	Correlation coefficient	Average relative error (%)
The proposed algorithm	0.9932	2.00
The single chaotic algorithm	0.9811	3.87

to quantify the prediction accuracy, the prediction accuracy measurements of the two sets of predicted values are shown in Table 3.

According to the data in Table 3, the prediction accuracy of this algorithm is improved obviously, which verifies the effectiveness.

5 Conclusion

In this paper, the dmey wavelet is used to decompose the power load into 3 layers, and the original power load signal is decomposed into low frequency sub signal (LF), high frequency sub signal1 (HF1), high frequency sub signal2 (HF2), high frequency sub signal3 (HF3). It is found that LF has obvious periodicity and is affected by the temperature and the properties of the week. And HF1, HF2 and HF3 are proved to be typical chaotic signals.

According to the different characteristics of LF and HF, different prediction methods are adopted. SVM is used to predict the LF, and temperature and week attributes are used as influence factors in the prediction. For HF, an improved CLP algorithm is adopted. In traditional local prediction algorithm, the parameters of phase space reconstruction cannot be accurately calculated. The accuracy of phase space recon-

struction limits the final prediction accuracy. In this paper, a method of parameter optimization based on segmentation verification is proposed, which can find the best parameters of phase space reconstruction.

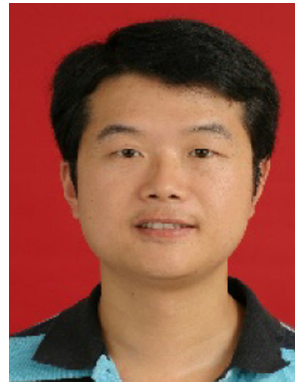
The above algorithm is applied to the prediction of power load in Shenzhen, China, and the prediction results show its effectiveness.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61300095 and 11604094), Natural Science Foundation of Hunan Province, China (Grant No. 2017JJ2068) and the Science Research Foundation of Hunan Provincial Education Department (Grant Nos. 16C0479 and 14A037).

References

1. Hu, D.G., Shu, H., Hu, H.D.: Spatiotemporal regression Kriging to predict precipitation using time-series MODIS data. *Cluster Comput.* **20**(1), 347–357 (2017). SI
2. Zhang, G.W., Xu, L.Y., Xue, Y.L.: Model and forecast stock market behavior integrating investor sentiment analysis and transaction data. *Cluster Comput.* **20**(1), 789–803 (2017)
3. Hirata, Y., Aihara, K.: Improving time series prediction of solar irradiance after sunrise: comparison among three methods for time series prediction. *Solar Energy* **149**, 294–301 (2017)
4. Yang, G.L., Cao, S.Q., Wu, Y.: Recent advancements in signal processing and machine learning. *Math. Probl. Eng.* **2014** Article ID 549024 (2014)

5. Moreau, F., Gibert, D., Holschneider, M., et al.: Identification of sources of potential fields with the continuous wavelet transform: basic theory. *J. Geophys. Res.* **104**(B3), 5003–5013 (1999)
6. Mallat, S.: *A Wavelet Tour of Signal Processing: The Sparse Way*, vol. 3. Elsevier, Amsterdam (2009)
7. Kumari, G.S., Kumar, S.k.: Electrocardio graphic signal analysis using wavelet transforms. In: 2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), pp. 1–6 (2015)
8. Rosenstein, M.T., Collins, J.J., Luca, C.J.D.: A practical method for calculating largest Lyapunov exponents from small data sets. *Phys. D* **65**(1–2), 117–134 (1993)
9. Yong, Z.: New prediction of chaotic time series based on local Lyapunov exponent. *Chin. Phys. Lett.* **22**(5) Article ID 020503 (2013)
10. Corts, C., Vapnik, V.: Support vector networks. *Mach. Learn.* **20**(3), 273–297 (1995)
11. Petra, V., Anna, B.E., Viera, R., Slavomír, Š., et al.: Smart grid load forecasting using online support vector regression. *Comput. Electr. Eng.* **000**, 1–16 (2017)
12. Su, L.Y., Li, C.L.: Local prediction of chaotic time series based on polynomial coefficient autoregressive model. *Math. Probl. Eng.* **2015**, Article ID 901807
13. Qu, J.L., Wang, X.F., Qiao, Y.C. et al.: An improved local weighted linear prediction model for chaotic time series. *Chin. Phys. Lett.* **31**(2) Article ID 020503 (2014)
14. Frandes, M., Timar, B., Timar, R., Lungeanu, D.: Chaotic time series prediction for glucose dynamics in type 1 diabetes mellitus using regime-switching models. *Sci. Rep.* **7** Article ID 6232 (2017)
15. Livi, L., Sadeghian, A.: Granular computing, computational intelligence, and the analysis of non-geometric input spaces. *Granul. Comput.* **1**(1), 13–20 (2016)
16. Antonelli, M., Ducange, P., Lazzarini, B., Marcelloni, F.: Multi-objective evolutionary design of granular rule-based classifiers. *Granul. Comput.* **1**(1), 37–58 (2016)
17. Lingras, P., Haider, F., Triff, M.: Granular meta-clustering based on hierarchical, network, and temporal connections. *Granul. Comput.* **1**(1), 71–92 (2016)
18. Skowron, A., Jankowski, A., Dutta, S.: Interactive granular computing. *Granul. Comput.* **1**(2), 95–113 (2016)
19. Dubois, D., Prade, H.: Bridging gaps between several forms of granular computing. *Granul. Comput.* **1**(2), 115–126 (2016)
20. Yao, Y.: A triarchic theory of granular computing. *Granul. Comput.* **1**(2), 145–157 (2016)
21. Ciucci, D.: Orthopairs and granular computing. *Granul. Comput.* **1**(3), 159–170 (2016)
22. Mallat, S.: Multi-resolution approximations and wavelet orthogonal bases of $L_2(\mathbb{R})$. *Trans. Am. Math. Soc.* **315**, 67–87 (1989)
23. Smola, A., Scholkopf, B.: *A Tutorial on Support Vector Regression*. Royal Holloway College, London (1998)
24. Zhang, H.R., Han, Z.Z.: An improved sequential minimal optimization learning algorithm for regression support vector machine. *J. Softw.* **12**(3), 2006–2013 (2003)
25. Takens, F.: *Detecting Strange Attractors in Fluid Turbulence*. Springer, Berlin (1981)
26. Gautama, T., Mandic, D.P., Van Hulle, M.M.: A differential entropy based method for determining the optimal embedding parameters of a signal. In: 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing. Hong Kong, China: IEEE, pp. 29–32 (2003)
27. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. *Int. Joint Conf. Artif. Intell.* **14**, 1137–1143 (1995)
28. Rao, C.R., Wu, Y.: Linear model selection by cross-validation. *J. Stat. Plan. Inference* **128**(1), 231–240 (2005)



Huan Wang received his Ph.D. degree in electrical engineering from Hunan University, China in 2010. From 2010 to 2012, he was a electrical design engineer of Guangdong electric design institute, China. From 2012 till now, he is a lecturer of University of electronic science and technology of China, Zhongshan Institute. His research interests include time series analysis and machine learning for nonlinear systems.



Min Ouyang received his M.S. degrees in information engineering from Central South University, China in 2009. Currently, he is a research staff at Key Laboratory of Intelligent Information Perception and Processing Technology of Hunan University of Technology, China. His research interests include Complex signal analysis and machine learning for nonlinear systems.



Zhibing Wang received the M.S. degree from Hunan University of Technology in 2010. He is currently an associate professor with the School of Computer, Hunan University of Technology. His current research interests include software engineering, trusted software and IoT networking.



Ruishi Liang received his Ph.D. degree in computer software and theory from Sun Yat-sen University in 2010. From 2010 to 2014, he was an assistant professor of school of computer engineering at University of Electronic Science and Technology of China, Zhongshan Institute. From 2014 till now, he is an associated professor of school of computer engineering at University of Electronic Science and Technology of China, Zhongshan Institute. His research inter-

ests include artificial intelligence, automated planning and its application, machine learning and big data.



Xin Zhou received the M.S. and Ph.D. degrees from Central South University in 2006 and 2012, respectively. From 2014 to 2016, she was a postdoctoral researcher at the School of Physics and Electronics, Hunan University. She is currently an associate professor with the School of Sciences, Hunan University of Technology. Her current research interests include surface plasmon polaritons, metamaterials, and material studies.