Simulation study on short-term load forecasting of electric power parameter optimization based on phase-space reconstruction theory

Huanyu Cai[1]*, Juhui Cao[2], Zhiqiang He[2]

[1]*School of Logistic Engineering University, Chongqing, 401311, China*

[2]*Jiuquan Satellite Launch Center, Lanzhou732750, China*

Received 1 October 2014, www.cmnt.lv

Abstract

With the combination of theoretical achievements related to parameters study on phase-space reconstruction and short-term load forecasting of electric system in physics. The short-term load forecasting of electric power parameter optimization based on phase-space reconstruction theory is put forward, and through theoretical and simulation researches and qualitative analysis, the impact of each parameter on the precision of the power load forecast is obtained and forecasting steps of this method are summarized in this paper. According to the optimized parameters, the prediction is conducted to load, and as a result, the accuracy is improved greatly when compare forecasting results to immediate prediction. A load example is forecasted by using parametric optimization prediction method, and the result shows that parametric optimization prediction method improves the accuracy of load prediction to some extent.

Keywords: load predication, chaos, phase-space reconstruction

1 Introduction

The load predication of the power system is to analyze the influence factor according to the historical data of the load; and then the predication model is set up according to the rules of the factor to reach the purpose [1-3] of predicating the future load. However, the reason that power system guarantee is ineffective is that the supply plan is not accurate and the starting and stopping arrangement of the unit are unreasonable; and the problem can be solved to predicate the future short-term load accurately [4-7]. Thereof, it is very important to improve and enhance the effect of short-term load predication.

We can learn from the phase-space reconstruction theory that two parameters, namely embedded dimension and delay time, should be determined according to Takens theory if we would like to reconstruct time series of a one-dimensional power load into a multi-dimensional phase space, which are of crucial importance in the phase-space reconstruction. Meanwhile, the selection of the number of vectors close to the central point also has certain influence on the precision of the power load forecasting during the power load forecasting.

Based on the above analysis, this paper studies on the three parameters, namely embedded dimension, delay time and vectors close to the central point, hoping that the impact of these three parameters on the forecast precision and their implicit rules could be obtained through study and experiment, then the three parameters are optimized according to these rules, and finally, the load forecasting is carried out based on the optimized parameters.

2 Chaos identification of time sequence and solving of parameter

Before predicating the load, the load data time sequence is judged whether to be chaos sequence or not.

The load data in 2011 of some region is taken as the sample to carry out the simulation test study of the parameter optimization value. 3,000 hours of load data is selected to predicate the short-term load of 24 hours; and the unit of all load data is milliwatt, illustrated in Figure 1.

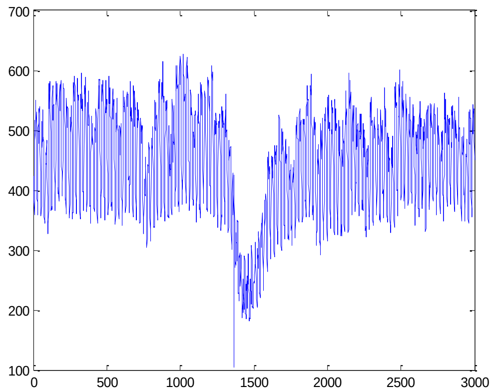

FIGURE 1 Load curve figure of some region in 2011

**Coresponding author's* e-mail: *caihuanyuleu@163.com*

Firstly, the time sequence of the load data is subjected to chaos identification. The simple and common Lyapunov index method is adopted to identify the time sequence of the load. The maximum Lyapunov index is calculated by small-data method [8,9].

The spectrum analysis diagram of which the maximum discrete stepping time is equal to 20 and 30 and the maximum Lyapunov index diagram are given below, illustrated in Figure 2 and Figure 3:

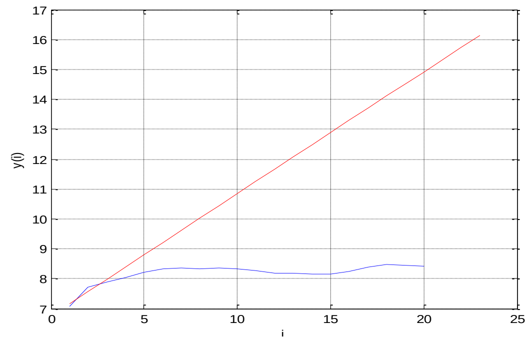

FIGURE 2 The Lyapunov index equal to 0.4086 by the biggest stepping time equal to 20 and fitting step length equal to 3

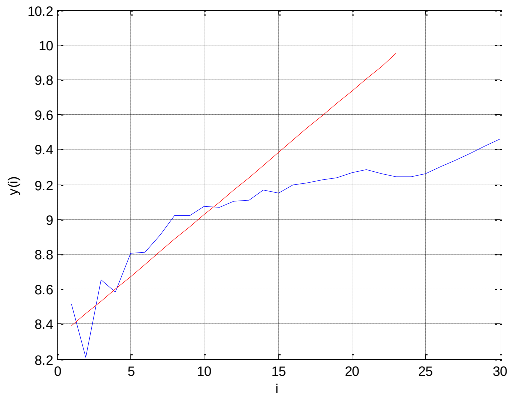

FIGURE 3 The Lyapunov index equal to 0.0709 by the biggest stepping time equal to 30 and fitting step length equal to 3

It can be known from the above experiment that the maximum Lyapunov index of the load time sequence is positive under different stepping time and matching step lengths, namely $\lambda > 0$, so it can draw a conclusion that the time sequence of the power load is chaos; and it can be predicated and researched by the chaos theory.

3 Access of delay time τ

After the chaos attribute of the power load time sequence is tested, the delay time τ of the first parameter affecting the phase space reconstruction is solved. The delay time τ is solved by the mutual information method.

It can be known from the mutual information that it can be used as the time delay τ of phase space reconstruction when the mutual information reaches the minimum delay for the first time; and we can obtain the delay time τ_n =6 from the Figure 4 and Figure 5.

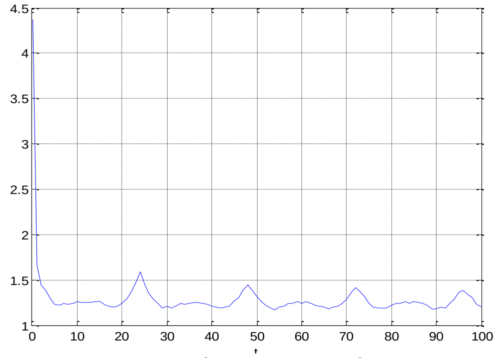

FIGURE 4 Mutual information method for delay time

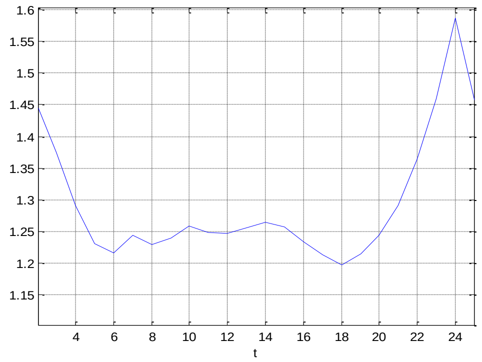

FIGURE 5 The partial enlargement diagram curve use mutual information for delay time

In order to study on the impact of parameters' delay time on the precision of load forecasting, the embedded dimension can be kept as m, and the number of vectors close to the central point can be $m+l$. The optimal delay time τ is searched by changing the delay time τ and based on the changing curve of mean absolute error (MAE) of predict outcomes. In addition, some rules are found through a large number of numerical tests on the impact of changes of the delay time of different parameters on the precision of load forecasting. We will take m=14, $k=1$ as an example to show the result changes of load forecasting as the delay time τ keeps changing. The details are as follows:

TABLE 1 Different delay time and average absolute error

τ	*MAE*	τ	*MAE*	τ	*MAE*
0	0.1343	20	0.1024	40	0.1045
1	0.0349	21	0.1208	41	0.1121
2	0.1006	22	0.0963	42	0.1052
3	0.1158	23	0.1357	43	0.1081
4	0.0989	24	0.0291	44	0.1045
5	0.1118	25	0.0541	45	0.1175
6	0.0991	26	0.0933	46	0.1026
7	0.1160	27	0.1163	47	0.1430
8	0.1078	28	0.1000	48	0.0269
9	0.1075	29	0.1200	49	0.0616
10	0.1018	30	0.0946	50	0.0947
11	0.1008	31	0.1104	51	0.1208
12	0.0635	32	0.1044	52	0.0988
13	0.1048	33	0.1066	53	0.1158
14	0.1081	34	0.1039	54	0.0938
15	0.1076	35	0.0942	55	0.1091
16	0.1088	36	0.0556	56	0.1128
17	0.1059	37	0.0987	57	0.1102
18	0.0991	38	0.0962	58	0.1029
19	0.1105	39	0.1040	59	0.0967

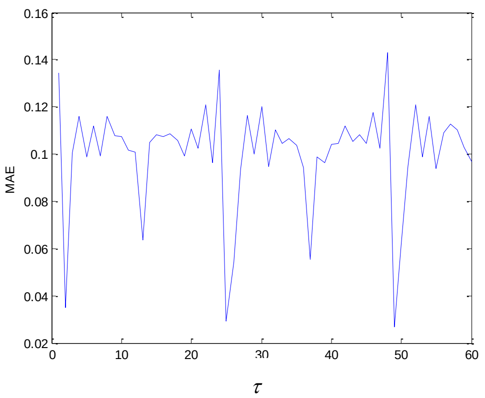

FIGURE 6 The relation curves of delay time τ and MAE

It can be seen from Table 1 and Figure 6 that, the mean absolute error (MAE) fluctuates significantly as the delay time changes and has certain rules to follow. Therefore, the following conclusions can be drawn from the above numerical test results:

1. The selection of delay time τ has a significant impact on the precision of load forecasting, and the predicted data and forecast error fluctuates significantly by changing the value of delay time τ .
2. When $\tau = \tau_n \alpha = 6\alpha, (\alpha = 1,2,3,\ldots)$ and the mean absolute error (MAE) closed to it are significantly lower, the forecast precision is significantly higher.
3. When a double cycle exists in the curve of mean absolute error (MAE), there is cycle $2\tau_n$ (12 herein) and cycle $4\tau_n$ (24 herein).
4. A multiple of τ and values closed to it can be tested constantly during the parameter optimization.
5. The forecast precision doesn't rank the highest when $\tau = \tau_n = 6$, therefore, it is possible and necessary to optimize the value of τ .

3 Determination of embedding dimension *m*

After obtaining the delay time τ , we continue to determine the second parameter and embedding dimension affecting the phase space reconstruction and the predication precision; and G-P method is adopted here.

The specific steps are as follows [10-15]:

1) The parameters known in the previous computational process of phase-space reconstruction: based on the total number of *N*, time series $\{Y_1, Y_2, Y_3 \ldots . Y_N\}$; delay time τ_n and embedded dimension m_n .
2) A new matrix P of $m \times M$ is obtained by carrying out a phase-space reconstruction to time series according to τ_n and m_n, that is, $M = N - (m-1)\tau_n$.

$$P(1) = \{x_{i,1}, x_{i,\tau d+1}, x_{i,2\tau d+1}, \ldots, x_{i,(m-1)\tau d+1}\}$$

$$P(2) = \{x_{i,2}, x_{i,\tau d+2}, x_{i,2\tau d+2}, \ldots, x_{i,(m-1)\tau d+2}\}$$

....

$$P(M) = \{x_{i,M}, x_{i,\tau d+M}, x_{i,2\tau d+M}, \ldots, x_{i,(m-1)\tau d+M}\}$$

3) The maximum difference d_{ij} is calculated by subtracting the corresponding items of components in every two vectors (i, j) in P ;
4) Find the maximum value of d_{ij} and record it as $\max d$; the minimum value and record it as $\min d$;
5) Set the proper step k so that r can change based on k growth rate between $\min d$ and $\max d$.
6) Carry out the following calculations for each $r(i)$:
For every two vectors P , subtract corresponding items of components to find the maximum absolute value and record it as d ;Make statistics of the number of $d \le r$, and record them as $sita$, calculate $C = \frac{2}{M(M-1)} sita$;
Solve the logarithms for C and r and record them as $\ln C(i) = \log_2 C$ and $\ln r(i) = \log_2 r$ respectively.
7) Construct a drawing with $\ln r$ as X-axis and $\ln C$ as Y-axis, and calculate the slope, namely the relevant dimension, in this drawing.

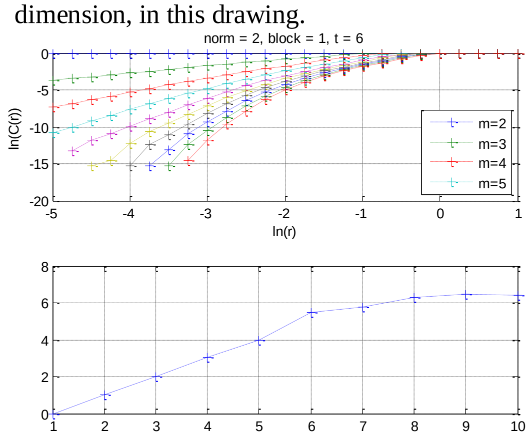

FIGURE 7 Correlation dimension embedding dimension figure

TABLE 2 Corresponding relation between embedding dimension and correlation dimension of attractor

Embedding space dimension (m_i)	Correlation dimension of attractor $d(m_i)$
1	0
2	0.993
3	1.981
4	3.054
5	3.992
6	5.494
7	5.791
8	6.270
9	6.462
10	6.415

It can be known form G-P algorithm that embedding dimension $m_2 > m_1$ is added continuously and the calculation is repeated until the corresponding dimension estimated value $d(m_i)$ is unchanged in a certain error range along the increase of m no longer; at this time, the obtained d is the correlation dimension of attractor. It can be known from the curve Figure 7 and Table 2 that the correlation dimension is steady around 6.47; and d =6.47. Meanwhile, embedding dimension $m \geq 2d+1$ according to the Takens theorem; so the embedding dimension of power load time sequence is m_n =14.

4 Selection of neighboring vector *m+l* of central point

A large number of experiments show that the space distance between each point in the space and the central point is a very important parameter; the predication accuracy depends on the points near the central point. If the distance between the adjacent point and the central point is introduced to the predication process as the matching parameter, the predication precision can be improved to some extent and remove the noise. The selection of the neighboring vectors of the central point has influence on the predication; and to find out the best parameter is of great significance to improve the predication precision.

It can be known from the above section that m meta-regression adopts at least $5m$ data, so that the regression result is reliable. The number of the neighboring vector is above $5m$, namely, in $l = km$, $k = 4$. When the predication is carried out by the weighting one-rank local-region method, $k = 1$.

When delay time and embedded dimension keep unchanged, change the number of vectors closed to the center point, namely changing the value of k, then calculate the mean absolute error (MAE) based on the data for load forecasting, and finally get the following two conclusions independent of examples through a large number of simulation tests and comparative analysis:

1. It can be seen from the mean absolute error (MAE) based on the data for load forecasting that the selection of the number of adjacent points closed to the center point doesn't much affect the precision of load forecasting.
2. In some cases, changing the number of proximal points closed to the center point doesn't affect the predict results.

Take $m = 14$, $\tau = 6$ for example an example to illustrate it. Below are changes of the mean absolute error (MAE) when the value of k keeps changing.

TABLE 3 The average absolute error MAE by *k* value change

k	Average absolute error *MAE*
0, 1, 2, 3	0.0991
4, 5,…,39	0.0990
40, 41,…, 86	0.0985
87, 88,…, 127	0.0981
128	0.0987
129, 130, …, 137	0.0975
138, 139,…., 149	0.0976
150,151,…,159	0.0981
160,161,…183	0.0987
184,…, 187	0.0989
188	0.1016

It can be seen from Table 3 that, when the value of k changes, the mean absolute error (MAE) changes very little and in some certain intervals the change of k value cannot affect the predict result, as shown in Figure 8.

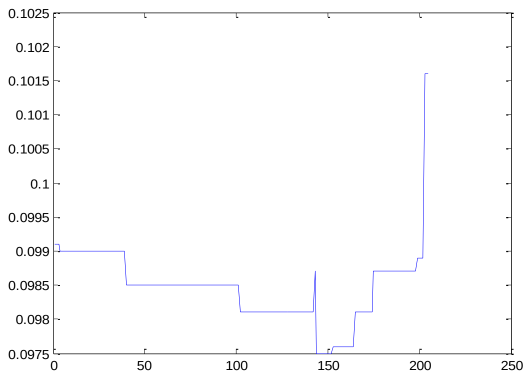

FIGURE8 The curves graph of average absolute error by k value change

5 Simulation test research on optimization value of load predication parameter

Numerical tests are carried out to the three parameters, namely embedded dimension, delay time and number of adjacent points closed to center point, of load forecasting based on chaos theory. And it can be seen from the conclusions obtained from the above numerical tests that, the embedded dimension and delay time have a significant impact on the precision of load forecasting data, while the impact of number of adjacent points closed to center point on the precision of load forecasting data is relatively small. Meanwhile, if the delay time τ keep changing when the embedded dimension and the number of adjacent points closed to center point remain unchanged, the precision of load forecasting near $\tau = \tau_n \alpha = 6\alpha, (\alpha = 1,2,3,\ldots)$ is significantly higher. If the embedded dimension m keep changing when the delay time τ and the number of adjacent points closed to center point remain unchanged, the precision of load forecasting of m =3 is significantly higher than that of other values.

Based on the above numerical tests and analytical conclusions, the combination of two parameters should be firstly tested when optimizing selection of the combination of the three parameters. After the optimal combination of these two parameters is found, the combination of the three parameters can be optimized on that basis; meanwhile, various parameters have different impacts on the precision of load forecasting, therefore, a prior consideration should be given to use the parameter, which has a big effect for the combinatorial test during the selection of parameter combination.

During the numerical test, the number of adjacent points closed to center point remains unchanged. Firstly, test the optimizing combination of delay time τ and embedded dimension m, and when =3 the precision of load forecasting is significantly higher, therefore, the test can be firstly started from the combination of (m, τ)=(3, τ), while the value of τ should be taken as close as possible to $\tau = \tau_n \alpha = 6\alpha, (\alpha = 1,2,3,\ldots)$. Meanwhile, for easy comparison, in addition to the error E_r and mean absolute error MAE, the author sets up the number of a comparative item $E_r \leq 3\%$ and lists the maximum error $E_{\max}$ for the calculation and analysis of the prediction error indicators.

In order to facilitate the analysis and conclusion, above combinations of numerical tests that meet the conditions of $E_{\max} \le 5\%$, $E_r \le 3\%$ and the number $\ge$ 11 and combinations with mean absolute error $MAE \le 0.050$ are selected below, totally 62 pairs, accounting for about 1.69% of the total number of combinations, as shown in Table 4.

TABLE 4 The parameter combination (m,τ) when prediction precision is higher

(m , τ)	Maximum error $E_{\max}$	Average absolute error MAE	$E_r \le 3\%$
(2, 1)	5.43%	0.0217	19
(2, 25)	9.66%	0.0299	12
(3, 1)	5.34%	0.0209	17
(3, 2)	8.01%	0.0285	14
(3, 5)	8.82%	0.0348	13
(3, 6)	5.02%	0.0210	17
(3, 12)	8.64%	0.0231	17
(3, 24)	9.81%	0.0336	14
(4, 12)	9.61%	0.0277	15
(4, 36)	6.84%	0.0292	14
(5, 1)	9.91%	0.0351	11
(5, 12)	7.87%	0.0332	13
(5, 24)	7.06%	0.0274	13
(5, 48)	8.59%	0.0254	17
(6, 1)	9.50%	0.0301	17
(6, 24)	9.92%	0.0307	14
(6, 48)	7.63%	0.0224	17
(7, 48)	6.23%	0.0202	19
(8, 24)	6.69%	0.0228	18
(8, 36)	4.51%	0.0217	15
(8, 48)	8.74%	0.0312	13
(9, 24)	5.83%	0.0229	17
(9, 36)	6.51%	0.0288	13
(9, 48)	9.22%	0.0355	13
(10, 24)	5.97%	0.0254	15
(10. 48)	9.55%	0.0330	13
(11, 24)	8.25%	0.0245	15
(11, 36)	8.63%	0.0342	12
(11, 48)	9.51%	0.0287	13
(12, 1)	9.33%	0.0403	11
(12, 24)	8.41%	0.0252	15
(12, 48)	9.64%	0.0304	14
(13, 24)	7.83%	0.0250	16
(13, 48)	8.92%	0.0264	16
(14, 1)	9.57%	0.0349	13
(14, 24)	9.10%	0.0291	14
(14, 48)	8.88%	0.0269	14
(15, 24)	7.30%	0.0228	19
(15, 48)	6.58%	0.0263	16
(16, 24)	8.18%	0.0257	16
(16, 48)	8.27%	0.0300	13
(17, 24)	7.19%	0.0283	14
(17, 48)	6.67%	0.0297	14
(18, 24)	9.17%	0.0312	14
(18, 48)	5.93%	0.0292	13
(19, 24)	7.60%	0.0281	16
(19, 48)	7.59%	0.0364	11
(20, 24)	7.85%	0.0281	13
(20, 48)	7.37%	0.0351	11
(21, 24)	7.81%	0.0276	15
(22, 24)	5.79%	0.0261	17
(23, 24)	8.15%	0.0253	16
(23, 48)	6.88%	0.0314	14
(24, 24)	7.70%	0.0237	19
(24, 48)	7.52%	0.0348	11
(25, 24)	7.64%	0.0263	18
(26, 24)	6.82%	0.0255	14
(27, 24)	6.56%	0.0257	16
(28, 24)	6.78%	0.0324	12
(29, 24)	7.83%	0.0298	12
(30, 24)	8.15%	0.0294	13
(31, 24)	6.67%	0.0307	13

The distribution of embedded dimension and mean absolute error, the optimization combination relations between delay time and mean absolute error and between embedded dimension and delay time are analyzed and compared below.

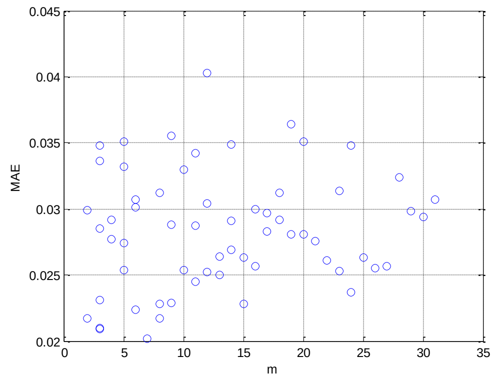

FIGURE 9 The relationship chart m and MAE when forecasting 4

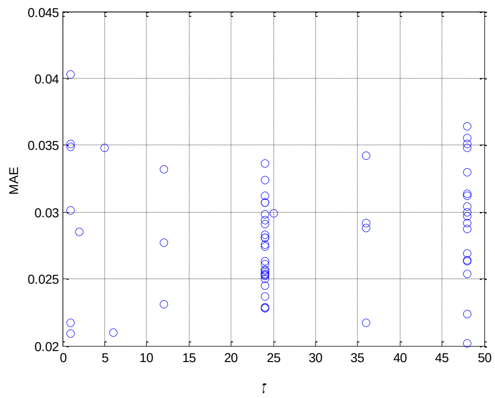

FIGURE 10 The relationship chart τ and MAE when forecasting precision is higher

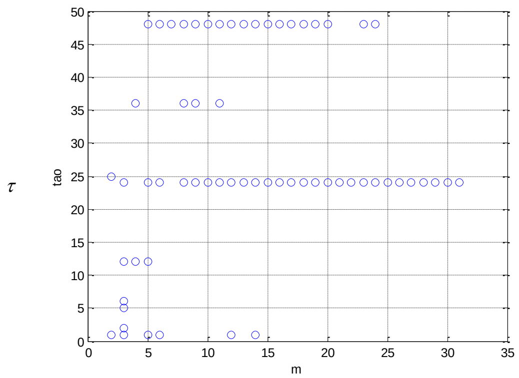

FIGURE 11 The parameter combination (m , τ) when prediction precision is higher

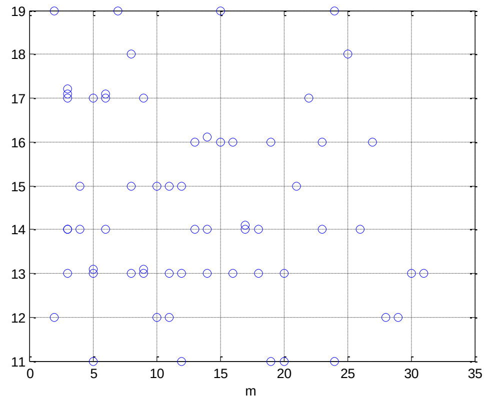

FIGURE 12 The relationships between m and the number of $E \le 3\%$

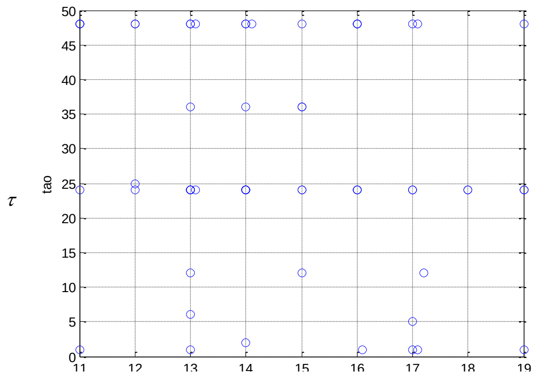

FIGURE 13 The relationships between τ and the number of $E \le 3\%$

It can be seen from Table 4 and Figures 9-13, combinations with delay times of τ =24 and τ =48 are in the majority in the optimized combinations, while for the embedded dimension *m*, combinations with higher forecast precision focus on $m = 1,\ldots,28$. And the following conclusions independent of examples can be drawn through a good deal of numerical calculation and experimental tests:

1. The combination with higher predication precision, namely the embedding dimension *m* in the optimal combination focuses in the range of $\{m \,|\, 0 \le m \le 2m_n, m \in Z\}$; and the parameter combination with high precision is obtained in the range.
2. The higher predication precision can be obtained around the embedding dimension m =3.
3. The optimal combination with higher predication precision is obtained in $\tau = \alpha\tau_n (\alpha = 1,2,3,\ldots)$.
4. Relative to the number of adjacent vectors closed to center point, the selection of embedded dimension and delay time has a significant impact on the forecast precision. Therefore, the optimization of these two parameters must be carried out in order to improve the precision of load forecasting.

It can be obtained from the research on the parameter combination of (m , τ) that: the embedded dimension and delay time gained from a short-term load forecast according to chaos theory aren't the optimal parameters, and the predicted load precision isn't ideal. In addition, it can be learnt from the research on the number of adjacent vectors closed to center point, embedded dimension and delay time above that, the selection of embedded dimension and delay time has a great impact on the precision of load forecasting, while the selection of the impact of the number of adjacent vectors closed to center point is relatively smaller. The parameter combinations are very critical for the precision of load forecasting.

In this example, when the embedded dimension $m \le 14$. The precision is relatively higher, and the proportion of the combination reaches the maximum one, especially when m =3, 5; for the delay time τ , the one calculated by using chaos theory is τ_n =6, the forecast precision is relatively higher when $\tau = 6n, n = 0,1,2$, and the proportion of the combination reaches the maximum one, especially when τ =24, 48. In the previous section, numerical simulation tests are carried out to analyze the impact of the combination between embedded dimension and delay time on the forecast precision, Results of this numerical simulation are given in the following:

Take m =3, τ for an example:

TABLE 5 The predict precision is the highest under combined parameters （3, 12, k） by k equals to 944

	k	Maximum error E_{max}	Average absolute error *MAE*	Number of $E \le 3\%$
（3, 12, k ）	k =1	8.64%	0.0231	17
	k =160	7.90%	0.0256	17
	k =944	7.30%	0.0225	18

TABLE 6 The predict precision is the highest under combined parameters （3, 12, k） when k equals to 2

	k	Maximum error E_{max}	Average absolute error *MAE*	Number of $E \le 3\%$
（20, 24, k ）	k =0	7.85%	0.0282	13
	k =1	7.06%	0.0268	15
	k =2	6.09%	0.0263	16

It can be seen from Table 5 that, when the value of k keeps changing, the mean absolute error changes very little and even remains unchanged in some intervals sometimes, and the same conclusion can be drawn from Table 6. Therefore, changing the value of k can affect the forecast precision slightly; similarly, to optimize the number of adjacent vectors of center point can also improve the forecast precision to some extent.

The following conclusions for the three parameter combinations independent of examples can be drawn through a good deal of numerical calculation and experimental tests:

1. In the combination of parameter (m , τ , k), m and τ have great influence on the precision of the predication result; k has little influence on the precision of the predication result; and in the value test with changed k value, the predication data change is gentle.
2. When the parameter combination (m , τ) reaches the optimization, the predication precision can be improved by changing k value.

6 Conclusion

Since the adding-weight one-rank local-region method is a forecasting method with the best predictive effect in the present chaos theory, and in view of current situation that there are many researches on the load forecasting method but few researches on the parameter optimization, this paper puts forward a short-term load forecasting of electric power parameter optimization based on the phase-space reconstruction theory, and through theoretical and simulation researches and qualitative analysis, the impact of each parameter on the precision of load forecasting is obtained and short-term load forecasting methods for power systems are studied and improved.

References

Fu Y, Zhang L 2003 Forecasting and decision-making theory and method *Harbin Harbin Engineering University press* 4-11

[1] Zhao X 2002 Analysis and predication of load predication features of power in China *Beijing: China Electric Power Press* 3-6

[2] Xiao G, Wang C, Zhang F 2001 Power load predication *Beijing: China Electric Power Press* 2-10

[3] Xiu C, Liu X, Zhang Y 2003 Selection of delay time and embedding dimension of phase space reconstruction *Journal of Beijing institute of technology* **23**(2) 219-24

[4] Xiao F, Yan G, Han Y 2005 Information theory method for confirming chaos time sequence phase space *Acta Physica Sinica* **52**(2) 550-6 (*in Chinese*)

[5] Lv X, Cao B, Zeng M, Huang S, Liu X 2006 Algorithm for confirming information of delay time *Computational physics* **23**(2) 184-8

[6] Yang Z, Zhang G, Chen H, Lin H 2005 Parameter optimization for load tendency and chaos of short-term load *Power System Technology* **29**(4) 27-30

[7] Ding S 1981 Multivariate analysis method and application *Changchun: Jilin People's Press* (*in Chinese*)

[8] Yang S, Wu Y, Wang Z 1996 Engineering application of time sequence analysis *Wuhan: Elsevier Science Publishers LTD*

[9] Yang Z, Lin H 2012 Discussion on reason of chaos features of load record of power system *Power system automation* **26**(10) 18-23

[10] Yang Z, Lin H 2013 Value test and analysis of parameter optimization of short-term load predication phase space reconstruction *Power system automation* **27**(16) 40-4

[11] Yang W, Gong J 2010 Traffic flow predication based on phase space reconstruction and LSSVM *Traffic technology* **242**(5) 78-80

[12] Yang Shaoqing, Jia Chuanying, two practical phase space reconstruction method 2002 **51**(11) 2452-8

[13] Guzi T 2006 Parameter optimization method of power short-term load time sequence chaos phase space reconstruction *Proceedings of the CSEE* **26**(14) 18-23

[14] Liu S, Zhu S, Yu X 2008 Optimization research on phase space reconstruction *Data acquisition and processing* **23**(1) 65-9

Authors

Cai Huanyu, March 16,1985, Shichuan, China.

Current position, grades: Dr student of School of Logistic Engineering University. Master's degree.
University studies: School of Logistic Engineering University (2002-2006).
Scientific interest: power system operation analysis and control, operation and management of power grid.
Publications: 4 papers.

Cao Juhui, December 10, 1964, Hunan, China.

Current position, grades: School of Logistic Engineering University .Doctor degree,professor,and doctorial supervisor.
Scientific interest: military logistic service, military engineering materials, barracks service.
Publications: 43 papers, 2 monograph.

He Zhiqiang, October 26, 1985, Hunan, China.

Current position, grades: Master's degree. Seniore engineer and expert on electrical engineering of Jiuquan Satellite Launch Center.
Scientific interest: power system operation analysis and control, power reliability and security.
Publications: 4 papers.