Short-term data forecasting based on wavelet transformation and chaos theory

Yi Wang 1, a), Cunbin Li 1, b) and Liang Zhang1, c)
1Institute School of Economics and Management, North China Electric Power University, Beijing 102206, China
E-mail: a) wyi@ncepu.edu.cn, b) wyi@ncepu.edu.cn

Abstract. A sketch of wavelet transformation and its application was given. Concerning the characteristics of time sequence, Haar wavelet was used to do data reduction. After processing, the effect of “data nail” on forecasting was reduced. Chaos theory was also introduced, a new chaos time series forecasting flow based on wavelet transformation was proposed. The largest Lyapunov exponent was larger than zero from small data sets, it verified the data change behavior still met chaotic behavior. Based on this, chaos time series to forecast short-term change behavior could be used. At last, the example analysis of the price from a real electricity market showed that the forecasting method increased the precision of the forecasting more effectively and steadily.

1. Introduction
At present, there are many short-term data forecasting methods, such as regression model method and time series method. Not only the wavelet analysis, data mining and neural network are used in the modern forecasting methods, but also the emerging chaos theory provides a new way for the forecasting work. Literature [1-2] used the chaotic model to process the short-term power load forecasting. On the basis of chaos theory, the objective law of the data itself is used to predict the power load. literature [3] gave a short-term load forecasting method according to the phase space reconstruction of load series chaotic features. In literature [4-5], the data sequence was decomposed into low frequency part and high frequency part by wavelet decomposition principle. On this basis, different regression models were established for each sub-sequence and predicted. Literature [6] introduced the big data and data mining technology used in the power system. In literature [7], a short-term price forecasting method based on time series similarity in data mining was proposed.

The usage of data mining technology prediction often does not pay attention to the feature of data sequence itself. In the past, wavelet analysis and chaos theory predicts were mostly lack of the data preprocessing, some only make a simple adjustment of the data, and the "data nail" in the timing data reduced the accuracy of the forecast. Due to the impact of many uncertainties, sometimes data fluctuate with a high degree of non-linear and even chaos.

In this paper, wavelet transform and chaos theory are combined to propose a chaotic time series prediction method based on wavelet transformation. Firstly, according to the characteristics of the time series of data, the data loss method is used to reduce the data and then the maximum Lyapunov exponent is used. Finally, the validity and feasibility of the method is verified by analyzing the data of a real power market.
2. Wavelet Transformation and Time Series Data Preprocessing

2.1 Wavelet Transformation and Haar wavelet
The main feature of wavelet transformation is that the non-stationary time series is transformed into a much more stationary time series than the original sequence. Haar wavelet is a popular wavelet transformation technique. Haar wavelet function is a set of functions that are orthogonal to each other. Haar wavelet is derived from this set of function sets, and it is a single rectangular wave which the support domain in the \( t \in [0, 1] \) range. Because of \( \int \psi(t)dt = 0 \), and \( \int t\psi(t)dt \neq 0 \), the \( \psi(\omega) \) in the point \( \omega = 0 \) only has one-degree zero. \( \psi(t) \) constitutes a family of the simplest orthogonal normalized wavelet family in the multiresolution system of \( a = 2^j \) [8], that is, \( \psi(t) \) is not only orthogonal to \( \psi(2^j t) \), but also orthogonal to its own integer displacement \( \int \psi(t)\psi(t-k)dt = 0 \ (k \in \mathbb{Z}) \). The Haar wavelet basis function is constructed as Eq. 1.

\[
\begin{cases}
1 & 0 \leq t \leq 0.5 \\
-1 & 0.5 \leq t \leq 1 \\
0 & \text{other}
\end{cases}
\]  

(1)

The Haar wavelet basis function value is 1 in the range of \([0, 0.5]\), -1 in the range of \([0.5, 1] \), and 0 for other intervals. The original data affect the Haar wavelet function value, the sequence data value affects the same absolute value of positive and negative function value, these steps can offset the frequent repeated shocks in a very short period time of the original data, while the non-frequent shocks part of the data will not be offset and will be retained. This is the idea and principle of using Haar wavelet for time series data reduction.

2.2 Application of Wavelet Transformation in Time Series Data Preprocessing
The impact of the various risk factors in reality will inevitably be reflected in the historical data. Fig. 1 is the electricity price data from January 1 to August 31 of a power market, it can be seen that the price in some periods of great fluctuations constitute a "data nail." From the waveforms after the decomposition of the wavelet, it is observed that the abrupt changes in these prices lead to violent oscillations of the wavelet waveforms. Some of the changes are largely unpredictable due to the limited amount of information available, and the "data nails" also affect the other periods price forecast. In order to mitigate the adverse effects of these "data nails" in the forecast, it is necessary to properly preprocess the time series data.

First, statistics a period time to obtain the data interval in the probability of a certain probability; and then middle pass filter is used to compress the parts beyond the given range at an appropriate proportion. Compression uses monotonic increasing function, and does not change the size relation of data in time series.

Given \([B,T]\) is the data interval in the probability of a certain probability, \( P \) is the predicted value, then the middle pass filter function is listed as Eq. 2.

\[
F(P) = \begin{cases}
T + a \ln(P - T + 1) & P > T \\
P & P \in [B,T] \\
B - a \ln(B - P + 1) & P < B
\end{cases}
\]

(2)

The filter function for \( P \) is continuous and monotonous, \( a \) is the experimental tunable parameter, and used to adjust the compression ratio of the data beyond the statistical interval part. Dealt with this filter function, the original size order relationship of the data sequence will not be destroyed, and the greater the data deviation from the statistical interval, the greater the compression ratio. With such preprocessing, the data is smoothed while maintaining the original size data relationship.

Wavelet decomposition techniques can reflect the regularity of sequences on different scales, and short-term predictions are mainly found in the history data. There are a lot of concrete and
unpredictable factors in reality that work to give data rich in details. Such as sudden peaks or troughs, and these details create obstacles to reveal the general laws of data changes, and Fig. 2 is the data time series after wavelet reduction.

3. Forecast Model of Chaotic Time Series Based on Wavelet Transformation

3.1 Phase Space Reconstruction Theory and Phase Space Reconstruction Parameters Selection

Packard et al originally proposed the theory of phase space reconstruction is to restore the chaotic attractors in high-dimensional phase space. The chaotic attractors reflect the regularity of the chaotic system, which means that the chaotic system will eventually fall into a particular trajectory. The evolution of any component of the system is determined by the other components that interact with it, so that the system itself can be predicted from the time series data of a component. The most commonly used reconstruction method is the delay coordinates vector method proposed by Packard in 1950. Takens proves that a suitable embedding dimension can be found, that is, if the dimension of the delay coordinate \( m \geq 2d + 1 \) (\( d \) is the dimension of the power system), in this embedding dimension space, the regular trajectory can be recovered, that is, the trajectory in the reconstructed space \( R^m \) of the original power system keeps the diffeomorphism, thus laying a solid theoretical foundation for the prediction of chaotic time series [9].

A time series \( x(i), i = 1, 2, \cdots, N \), the appropriate time delay \( \tau \) and embedding dimension \( m \) can be used to extend the time series into a phase distribution of a \( M \) - dimensional phase space:

\[
Y = \{x(i), x(i + \tau), \cdots, x[i + (m - 1)\tau]\} \tag{3}
\]
In Eq. 3, \( i = 1, 2, \cdots, M \), \( M = N - (m - 1) \tau \) and \( M \) is the number of phase points in the phase space.

In phase space reconstruction, the time delay \( \tau \) and embedding dimension \( m \) have great significance, their values are also difficult to select. In practice, the methods of time delay are mainly autocorrelation method, complex autocorrelation method, mutual information method and C-C method. The embedded dimension is mainly obtained by GP algorithm, and the C-C method can be used to determine the time delay get the appropriate embedded dimension.

### 3.2 Chaotic Time Series Forecasting Method Based on Lyapunov Exponent

Chaotic time series forecasting is carried out directly according to the objective law calculated by the data sequence itself. The commonly used methods include the global forecast method, the local forecast method and the Lyapunov exponent method.

"Butterfly effect" is a classic example of chaos theory, as it describes, the orbits generated by two very close initial values are exponentially separated with time, and the Lyapunov exponent is a feature that quantitatively describes the phenomenon. It is a good predictor of the system as a feature to quantify the divergence of the initial orbit and estimate the chaotic level of the system. \( X_M \) can be set as the prediction central point, the nearest point of \( X_M \) in the phase space is \( X_k \), the distance is \( d_M(0) \), and the maximum Lyapunov exponent is \( \lambda_1 \), that is:

\[
d_M(0) = \min_j \|X_M - X_j\| = \|X_M - X_k\| \tag{4}
\]

\[
\|X_M - X_{M+1}\| = \|X_k - X_{k+1}\| e^{\lambda_1} \tag{5}
\]

In the above equations, the point \( X_{M+1} \) only has one last component, so \( x(t_{n+1}) \) is predictable.

### 3.3 Forecasting of Chaotic Time Series Based on Wavelet Transformation

Before making short-term forecasting, the data is preprocessed, and the data is reduced by wavelet transformation, and it is integrated into the chaotic time series prediction process as shown in Fig. 3. The forecasting steps are as follows:

1. Data preprocessing mainly includes data selection and data transformation. Choosing the appropriate data interval is an important step before wavelet transformation and chaos prediction.

2. The prediction of chaotic time series is carried out after data reduction, and the time delay and embedding dimension are calculated in turn, then the correlation dimension and the largest Lyapunov exponent are used to complete the prediction.

3. Analyze the results of the forecast and provide users with knowledge.

![FIGURE3 Forecasting of chaotic time series based on wavelet transformation](attachment:fig3.png)

### 4. Case Study

Using the data from January 1st to August 31st in a power market as a predictive data source, Fig. 4
shows the self-correlation method to determine the delay time $\tau = 12$, which is consistent with the results of the mutual information method. As shown in Fig. 5, the embedding dimension $m=6$ is obtained from the pseudo neighborhood point.

Under the known delay time and embedding dimension, the small data method [10] is used to find the maximum Lyapunov exponent $\lambda = 0.0921 > 0$, and the chaotic characteristics of the data sequence are also proved. Generally, the longest forecast time is defined as $T_m = 1/\lambda$, it represents the longest time required to double the system state error. It can be used as one of the reliability indicators of short-term forecasts. By using of Eq. 5, the electricity price of a power market in September 1st is predicted in 10 steps. The forecast results are shown in Table 1.

![FIGURE 4](image1.png)  
**FIGURE 4** The delay time $\tau$ calculated with autocorrelation method

![FIGURE 5](image2.png)  
**FIGURE 5** Embedding dimension calculated with pseudo-proximity

| Time | Real value | Forecasting value | Relative error(%) |
|------|------------|-------------------|-------------------|
| 1    | 25.4453    | 24.4364           | -3.96             |
| 10   | 32.999     | 36.062            | 9.28              |
As can be seen from Table 1 and Fig. 6, the forecasting errors of the 10 periods are within 10%. The average error is 5.65% by calculation. Compared with the general model predicted by time series, the accuracy is very high.

5. Conclusions
In order to reduce the influence of "data nail" on the short-term forecasting, the Haar wavelet transform is used in this paper, and the data are preprocessed before forecasting. Wavelet analysis can highlight the characteristics of sequence in the frequency domain, but if only the removal of the reconstruction after details of the data as the forecasting data source, some of the missing information leads to the forecast error increases. In the process of wavelet reconstruction, the threshold of each layer of high frequency coefficients is set. The part that exceeds the threshold is defined as the singularity. The middle pass filter method is used to compress the data value at the singular point. It is found that this can effectively improve the forecasting accuracy.

In the forecasting of chaotic time series, the prediction method based on the maximum Lyapunov exponent method is adopted. The maximum Lyapunov exponent forecasting method is a very practical prediction method. It avoids the complicated calculation and error of advanced mathematics. The simple method is often successfully avoid the calculation of defects caused. At present, many scholars have put forward a lot of improved short-term forecasting methods for the maximum Lyapunov exponent, which is an incentive for the author of this paper.

Acknowledgment
This work is supported by National Natural Science Foundation of China (No.71671065) and the “Research on cascading failures and control strategies of physical integration system for power grid information” (No. 5204BB1600CP) of National Power Grid Corp science and technology project, all supports are gratefully acknowledged.

References
[1] Zhang Shuqing, Shi Rongyan, Dong Yulan, Li Pan, Ren Shuang and Jian Wanlu. 2015. Bivariate Wavelet Threshold Denoising and Improved Chaotic Forecasting Model and Its Application in Short-term Power Load Forecasting[J], Proceedings of the CSEE, 35(22) pp. 5723-5730.
[2] Li Dongdong, Qin Zishan, Lin Shunfu, Zheng Xiaoxia and Wang Tianxiang. 2015. Short-term Load Forecasting for Microgrid Based on Method of Chaotic Time Series[J], Proceedings of the CSU-EPSA, 27(5) pp. 14-18
[3] Gu Xi and Liao Zhiwei. 2017. Short-term Load Forecasting Based on Phase Space Reconstruction and Gaussian Process Regression [J], Power System Protection and Control, 45(5) pp. 73-79.
[4] Niu Lixiao, Wang Zhengfang, Cang Chuanzhi, Shang Wenli and Zhang Shengshan. 2014. Hybrid
model based on wavelet transform and ARIMA for short-term electricity price forecasting[J], Application Research of Computers, 31(3) pp.688-691.

[5] Xie Pinjie, Tan Zhongfu, Shang Jincheng, Hou Jianchao and Wang Mianbin. 2008. Short-Term Electricity Price Forecasting Based on Wavelet Transform and Generalized Autoregressive Conditional Heteroskedasticity Model Research on Short-term Electricity Price Forecasting[J], Power System Technology, 32(16) pp.96-100.

[6] Wu Keqi and Cheng Fan. 2016. Data Mining Technology and Its Application in Power System[J], Electrotechnics Electric, 6 pp.28-32.

[7] Yuan Guichuan, Cheng Li and Wang Jianquan. 2003. Electric Price Forecasting Using Data Mining, Proceedings of the EPSA,15(2) pp.19-22.

[8] Chai Yanling and Tang Zhita. 2016. Application of Haar Wavelet Transform in Well-Logging Curve Similarity [J], Computer Technology and Development, 26(1) pp.119-123.

[9] Wu Huawen. 2014. Chaotic time series analysis and the research of its application in railway freight volume prediction[D], China Academy of Railway Sciences.

[10] Rosenstein M T, Collins J J, and De Luca C J.A. 1993. Practical method for calculating largest Lyapunov exponents from small data sets[J]. Physica: D, 65 pp.117-134.