

Chaotic Signal De-Noising Based on Threshold Selection Rules with SNR Evaluations of Wavelet

Sun Hai^{1*}, Gao Huiwang¹ and Ruan Xuejing²

¹Key Laboratory of marine environment and ecology, Ocean University of China,
Qingdao 266100, China

²College of Architecture Engineering, Qingdao Agricultural University,
Qingdao 266109, China
hdsh2008@163.com

Abstract

In nature, the observed Chaos phenomenas were often mixed with noise, the existence of noise made the prediction of chaotic time series generate large errors. Chaotic time series had the characteristic of broadband, which liked noise. So there were some limitations with the traditional method of de-noising. But the wavelet threshold de-noising method had the characteristic of the multi-resolution analysis, and its computational quantity was smaller and the noise filtering effect was better. On the other hand, for different types of signals, with different wavelet base functions and threshold rules, it might have a different effect on the de-noising effect. In order to search for the optimal selection of those parameters, firstly this paper constructed a simulated Lorenz noisy signal, and used this signal to do the de-noising experiment, used the SNR and RMSE as the evaluating indicator, and finally obtained the matching combination of those parameters. At the end of this paper, the de-noising simulation was carried out using China's Shijiao station runoff time series data from 1960 to 1970 in China, and the final results showed the effectiveness of the proposed method in this paper.

Keywords: Wavelet analysis; Threshold de-noising; Chaotic time series; SNR; RMSE

1. Introduction

The research showed that most of the phenomena in nature such as runoff, sunspot number, rainfall and so on had certain chaotic characteristics. The chaotic time series was the result of the chaotic phenomena recorded in fixed interval. The chaotic characteristics of the system could be studied from any single variable time series in the records of the system, then the development trend of the chaos phenomena could be forecasted [1]. But the actual observed time series often contained noise, the existence of noise made the self-similarity characteristic of the chaotic attractors in the system destroyed. As a result, even if the initial conditions were slight different, it would be enlarged in the evolution of time scale. These slight differences of the initial conditions were often caused by noise. So it is important to solve the problem that how to eliminate the noise effectively and to improve the accuracy.

The traditional de-noising method was that, the input data was performed by the method of spectrum analysis or Fourier transform, the noise was regarded as a high frequency component, which could be removed from the data. But the chaotic time series had the characteristic of broadband and been similar to noise, the spectrum of the signal and the spectrum of noise were often overlapped. In the past, chaotic de-noising methods generally aimed at the chaotic system with already known dynamics characteristics. Those de-noising methods were performed by comparing attractors and predicted

*Corresponding Author

trajectory of the system, such as the construction cost function method and the probabilistic method [2-3]. But most of these methods were built on the basis that the noise spectral amplitude was very small, but the spectrum of chaotic time series signal was wide. Because the spectral amplitude of the noise and the spectral amplitude of the signal were often in the same order of magnitude, the traditional de-noising methods were not ideal [4]. Wavelet analysis was a kind of analytical method both in time domain and in frequency domain. At the same time, it also had the characteristics of multi-resolution analysis at different scales [5]. Therefore, this paper started from the characteristics of the chaotic time signal and used different wavelet basis functions and different threshold rules for de-noising simulation experiment. The signal to noise ratio (SNR) and the root mean square error (RMSE) were used as the evaluation criteria to detect the effect of wavelet de-noising on chaotic time series, and finally obtained the best combination of wavelet function and threshold rule.

2. De-Noising Principle and Evaluation Rules

2.1. The Principle of Wavelet Multi-Resolution Analysis and Threshold De-Noising Method

The de-noising method based on wavelet analysis was established on the basis of wavelet multi-resolution analysis, the idea was to construct some orthogonal wavelet basis functions on space $L^2(\mathcal{R})$ with different frequency. Through those orthogonal wavelet basis functions we could decompose the low frequency space, and make the decomposition scale of space $L^2(\mathcal{R})$ higher. The core of the method was to construct a sequence $\{V_j\}_{j \in \mathcal{Z}}$ in the space satisfying the following conditions [6].

- (1) $V_j \in V_{j-1}$, for any $j \in \mathcal{Z}$
- (2) $\bigcap_{j \in \mathcal{Z}} V_j = \{0\}$, $\bigcup_{j=-\infty}^{+\infty} V_j = L^2(\mathcal{R})$
- (3) $f(t) \in V_j \Leftrightarrow f(2t) \in V_{j-1}$
- (4) For $\forall k \in \mathcal{Z}$, if $\varphi_j(2^{j/2}t) \in V_j \Rightarrow \varphi_j(2^{j/2}t - k) \in V_j$ could be given.
- (5) $\forall \varphi(t) \in V_0$, made $\{\varphi(t - k) | k \in \mathcal{Z}\}$ compose the orthonormal basis of V_0

In Equation (5) the formula $\varphi(t)$ was the scale function. Its function system $\varphi_{j,k}(t) = 2^{-j/2}\varphi(2^{-j}t - k)$, $j, k \in \mathcal{Z}$ was standardized orthogonal, meanwhile it was the standard orthogonal basis of the space V_j . In this way, the space $L^2(\mathcal{R})$ could be decomposed into a set of mutually contained subspaces [7], so the space V_0 could be approximated as follows, that is:

$$V_0 = V_1 + W_1 = V_2 + W_2 + W_1 = \dots \quad (1)$$

The W_j represented the orthogonal complement of V_j in space V_{j+1} . For W_j , we could also get that $\forall \psi(t) \in W_j$ made $\psi_{j,n}(t) = 2^{-j/2}\psi(2^{-j}t - n)$ constitute the standard orthogonal basis functions of space V_j . If $f_j \in V_j$ represented the function system of resolution 2^{-j} used to approximate the space $L^2(\mathcal{R})$, $g_j \in W_j$ represented the approximation error [8]. Then we could get:

$$f_0 = f_1 + g_1 = f_2 + g_2 + g_1 = \dots \quad (2)$$

Any function $f \in L^2(R)$ could be fully reconstructed by the low frequency functions system of resolution 2^{-j} and the high frequency functions of the resolution 2^i ($1 \leq i \leq j$). Due to the different characteristics of the probability distribution of the wavelet decomposition coefficients in different frequency bands of the noise and the chaotic signals, so the frequency distribution of the noise was stronger in some frequency bands and the chaotic signal was stronger in the other part of the frequency bands. Therefore, the wavelet coefficients of noise could be removed according to their distribution characteristics in the frequency bands. Then the wavelet reconstruction of the signal could be carried out according to the processed coefficients, which achieved the purpose of de-noising.

Assuming that the wavelet decomposition coefficient was β_i , the threshold value was λ . There were two kinds of threshold function, as following [9-13]:

$$\beta_i = \begin{cases} \beta_i & |\beta_i| \geq \lambda \\ 0 & \beta_i < \lambda \end{cases} \quad (3)$$

$$\beta_i = \begin{cases} \text{sgn}(\beta_i)(|\beta_i| - \lambda) & |\beta_i| \geq \lambda \\ 0 & |\beta_i| < \lambda \end{cases} \quad (4)$$

The first one was the hard threshold rule, as a result, the wavelet coefficients of this method were not continuous in place λ , and there might be a local vibration after de-noising. Therefore, in this paper, we used the second one called the soft threshold rule, and its de-noising process was mainly divided into 3 steps, that was [14]:

Step1. Obtaining the initial wavelet coefficients $w_{i,j}$ through wavelet multi-scale decomposition of the noisy signals;

Step2. Using the threshold rule above to deal with the initial wavelet decomposed coefficients $w_{i,j}$ and obtaining the new wavelet coefficients $w_{i,j}$;

Step3. The new wavelet coefficients were used to reconstruct the signal to get the pure signal.

2.2. De-Noiseing Effect Evaluation

In the process of simulation experiment, this paper used the SNR and RMSE value as the standard to measure the de-noising effect of the simulated noisy data. The SNR represented the signal to noise ratio, which meant the ratio of the useful signal power to the noise power, its expression was [15-17]:

$$SNR = 10 \cdot \lg \left(\frac{\sum_{k=1}^N s_k^2}{\sum_{k=1}^N (s_k - g_k)^2} \right) \quad (5)$$

As the same, RMSE represented the square root of the sum of the square of values which were the observed value minus the true value divided by of the number of observations N, whose expression was:

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N (s_k - g_k)^2} \quad (6)$$

From the above expressions it could be seen that the greater of the SNR, the smaller of the RMSE, the better of the de-noising effect.

3. Simulation Experiment and Analysis

In this paper, we used different wavelet basis functions and threshold functions to deal with the noise of chaotic time series generated by Lorenz system and China's Shijiao station runoff time series data in China, comparing with the SNR and RMSE of the de-noised signal and the original signal to determine the effectiveness of de-noising process, the analysis process was as follows.

3.1. The Original Signal and Noisy Signal of Lorenz System

The dynamical equation of Lorenz chaotic system was shown in the Formula (7):

$$\begin{cases} \frac{dx}{dt} = -\sigma x + \sigma y \\ \frac{dy}{dt} = \gamma x - y - xz \\ \frac{dz}{dt} = xy - bz \end{cases} \quad (7)$$

In the above equation, $\sigma = 10$, $\gamma = 28$, $b = \frac{8}{3}$ at this time, the system was in the form of chaos. The 4 order Runge-Kutta method could be used to solve the problem. And took the first 3000 points of the X axis of the Lorenz system as the simulation signals of the chaotic signal, as shown in Figure 1. Took the noisy Lorenz sequence signal who's SNR equaled to 15 as contrast sequence, as shown in Figure 2.

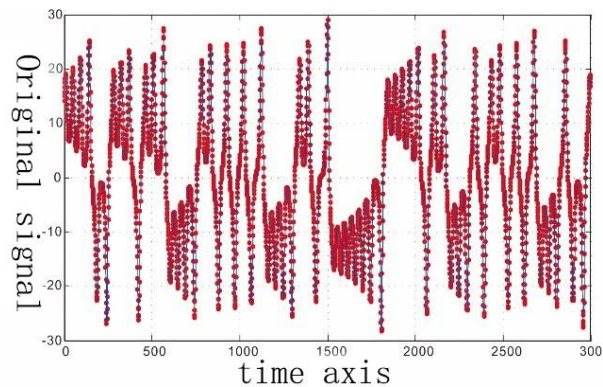


Figure 1. The Original $x(t)$ Sequence of Lorenz System

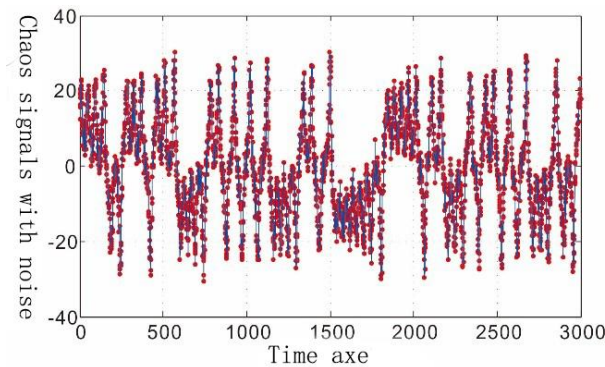


Figure 2. Lorenz Sequence $x(t)$ with White Noise

For Lorenz system, when the embedding dimension m was 5 and the delay time τ was 1, the sequence could be reconstructed well in phase space. At the same time, the wavelet decomposition scale J should be less than or equal to m , and less than the average period of data track. Therefore, this paper took the wavelet decomposition scale $J=4$, and carried on the 4 layer wavelet decomposition.

3.2. Basis Functions and Threshold Rules

The wavelet function we decided to select the symlets wavelet basis functions system which had the characteristic of better orthogonal symmetry and compact support. Symlets wavelet system was usually expressed as symN (N=2, 3... 8). And the threshold rule was selected by 4 kinds of classical threshold rules, as follows:

(1) Soft threshold rule based on Stein's unbiased risk estimation. The method was to rearrange the decomposition coefficients of each layer according to the square of those coefficients from small to large order, though the process we could get the following vector:

$$W = \{w_1, w_2, \dots, w_n\} \quad (8)$$

N was the number of wavelet coefficients after decomposition, and then constructed another vector R , the element was:

$$r_i = \left[n - 2i + (n - i)w_i + \sum_{k=1}^i w_k \right] / n, \quad (9)$$

$$i = 1, 2, 3, \dots, i$$

The minimum value in the sequence r_i was the risk vector r_i , and the corresponding wavelet coefficients were w_{mn} . At the same time, σ represented the mean square deviation of noise $x(t)$. The Stein's unbiased risk estimation of the threshold value could be calculated (10):

$$T_1 = \sigma \sqrt{w_{mn}} \quad (10)$$

(2) Heuristic threshold rule. The heuristic threshold combined the results of the unbiased risk estimation threshold rule and the fixed threshold de-noising rule. The basic idea was using the fixed threshold de-noising rule if the signal to noise was relatively large, otherwise, using unbiased risk estimation threshold de-noising rule. Making

$S = \sum_{i=1}^n w_i^2, \alpha = (S - n) / n, \beta = (\log_2 n)^{\frac{3}{2}} \sqrt{n}$, then the heuristic threshold value could be calculated by (11):

$$T_2 = \begin{cases} T_1 & \alpha < \beta \\ \min(T_1, T_3) & \alpha \geq \beta \end{cases} \quad (11)$$

(3) Fixed threshold rule. If the length of the noise signal $x(t)$ was N , the mean square deviation of the noise $x(t)$ was expressed by σ , and the fixed threshold value could be calculated by the following method:

$$T_3 = \sigma \sqrt{2 \cdot \ln N} \quad (12)$$

(4) The extreme value threshold rule. The threshold value was calculated by using the minimax principle, and the minimum mean square error could be obtained by using this method, its expression was:

$$T_4 = \begin{cases} 0 & n \geq 32 \\ 0.3936 + 0.1829 \log_2 n & n < 32 \end{cases} \quad (13)$$

3.3. Simulation of Noisy Chaotic Signal De-Noiseing Process

In the simulation experiment, the simulation experiment were processed by different wavelet basis functions and threshold functions. And finally getting the signal after de-noising, according to the evaluation rules to select the relatively optimal parameters. First we selected the sym2 wavelet function, and then selected the above four threshold rules for signal de-noising, the de-noising effect was shown in Figure 3.

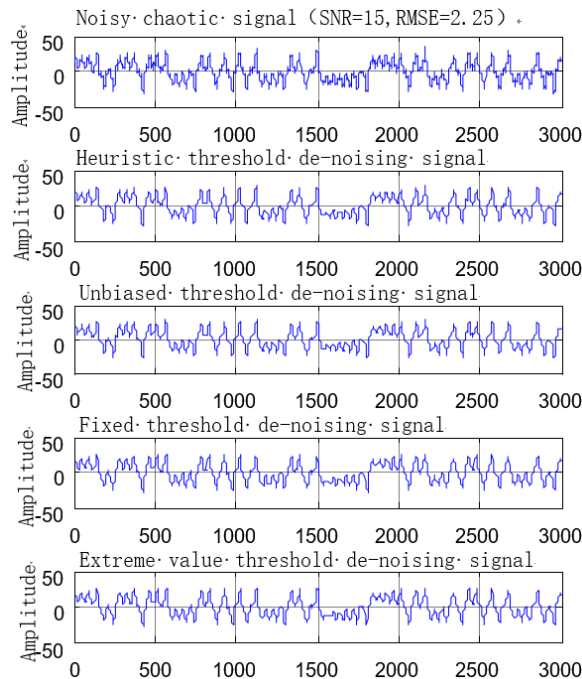


Figure 3. Comparison of the De-Noiseing Effect of the Four Threshold Rules Based on Sym2 Wavelet Function

The SNR of the original chaotic noisy signal was 15, and the RMSE value was 2.25. The de-noising effect of the four threshold rules were shown in Table 1. It could be seen that the heuristic de-noising effect was the best, followed by the unbiased risk, the fixed de-noising effect was the worst.

Table 1. Comparison of the De-Noiseing Effect of Sym2 Wavelet Function Matching 4 Kinds of Threshold Function According to the SNR and RMSE Value

| Threshold rule | SNR | RMSE | De-noising effect order |
|----------------|-------|------|-------------------------|
| Heuristic | 21.6 | 1.06 | 1 |
| unbiased | 21.21 | 1.10 | 2 |
| fixed | 15.16 | 2.21 | 4 |
| Extreme Value | 17.04 | 1.78 | 3 |

For symN (N=3, 4,... 8) wavelet basis functions, the final SNR and RMSE results could also been calculated with different Threshold rules , as shown in Table 2- Table 7:

Table 2. Comparison of the De-Noising Effect of Sym3 Wavelet Function Matching 4 Kinds of Threshold Function According to the SNR and RMSE Value

| Threshold rule | SNR | RMSE | De-noising effect order |
|----------------|-------|------|-------------------------|
| Heuristic | 20.87 | 1.14 | 2 |
| unbiased | 21.78 | 1.03 | 1 |
| fixed | 15.87 | 2.04 | 4 |
| Extreme Value | 18.48 | 1.51 | 3 |

Table 3. Comparison of the De-Noising Effect of Sym4 Wavelet Function Matching 4 Kinds of Threshold Function According to the SNR and RMSE Value

| Threshold rule | SNR | RMSE | De-noising effect order |
|----------------|-------|------|-------------------------|
| Heuristic | 21.47 | 1.07 | 2 |
| unbiased | 22.09 | 0.99 | 1 |
| fixed | 15.92 | 2.02 | 4 |
| Extreme Value | 18.3 | 1.54 | 3 |

Table 4. Comparison of the De-Noising Effect of Sym5 Wavelet Function Matching 4 Kinds of Threshold Function According to the SNR and RMSE Value

| Threshold rule | SNR | RMSE | De-noising effect order |
|----------------|-------|------|-------------------------|
| Heuristic | 22.03 | 1.00 | 2 |
| unbiased | 22.50 | 0.95 | 1 |
| fixed | 16.88 | 1.81 | 4 |
| Extreme Value | 18.90 | 1.43 | 3 |

Table 5. Comparison of the De-Noising Effect of Sym6 Wavelet Function Matching 4 Kinds of Threshold Function According to the SNR and RMSE Value

| Threshold rule | SNR | RMSE | De-noising effect order |
|----------------|-------|------|-------------------------|
| Heuristic | 22.56 | 0.94 | 2 |
| unbiased | 22.67 | 0.93 | 1 |
| fixed | 16.32 | 1.93 | 4 |
| Extreme Value | 18.88 | 1.44 | 3 |

Table 6. Comparison of the De-Noising Effect of Sym7 Wavelet Function Matching 4 Kinds of Threshold Function According to the SNR and RMSE Value

| Threshold rule | SNR | RMSE | De-noising effect order |
|----------------|-------|------|-------------------------|
| Heuristic | 22.52 | 0.94 | 2 |
| unbiased | 22.77 | 0.92 | 1 |
| fixed | 16.30 | 1.94 | 4 |
| Extreme Value | 18.98 | 1.42 | 3 |

Table 7. Comparison of the De-Noiseing Effect of Sym8 Wavelet Function Matching 4 Kinds of Threshold Function According to the SNR and RMSE Value

| Threshold rule | SNR | RMSE | De-noising effect order |
|----------------|-------|------|-------------------------|
| Heuristic | 22.88 | 0.91 | 1 |
| unbiased | 22.52 | 0.94 | 2 |
| fixed | 16.15 | 1.97 | 4 |
| Extreme Value | 18.52 | 1.50 | 3 |

It could be seen from the results that the use of sym6 wavelet basis functions, meanwhile selecting the heuristic and unbiased threshold function for chaotic noisy signals the de-noising effect was better. The fixed and extreme threshold de-noising effect was relatively poor.

4. Real Chaotic Noisy Signal De-Noiseing

This paper used China's Shijiao station runoff time series data as the actual chaotic noisy signal. The runoff data seemed non-stability in time, but it also had the orderly regularity and chaos characteristics. At present, many researches had indicated that the hydrological time series had some chaotic characteristics. The chaos prediction method could be used to predict hydrological series. This paper was intended to evaluate whether or not it could significantly improve the chaotic characteristics by using the method of this paper, and improve the accuracy of the prediction. Therefore, this paper analyzed the statistical data of the 132 groups of monthly average runoff data in the hydrological station from 1960 to 1970. The hydrologic runoff data recorded during the period are shown in Figure 4.

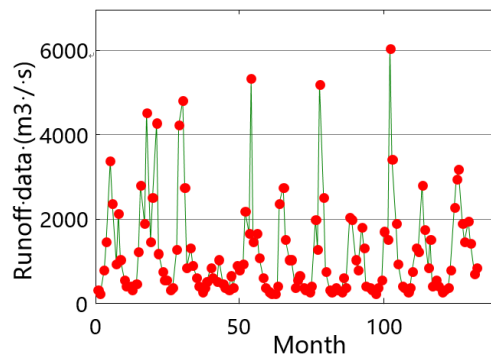


Figure 4. Monthly Mean Runoff Data of the Hydrologic Station in 132 Months

4.1. Determining the Time Delay and Embedding Dimension of the Reconstructed Phase Space

To predict the chaotic time series, we must first determine the spatial time delay. Self-correlation function method was often used to calculate spatial time delay. The method was to take the log time that made the self-correlation function through or near the zero point at the first time. At this time, the time delay was the minimum value to make the new sequence to achieve linearly independent.

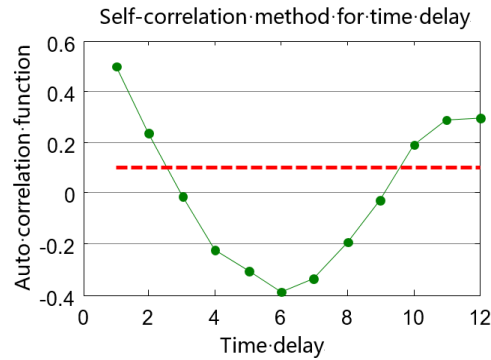


Figure 5. Self-Correlation Function of the Runoff Data

As could be seen from Figure 5, when the time delay was 3, its autocorrelation function was reduced near to 0, so the time delay of the sequence was 3.

Embedding dimension was usually implemented by the G-P algorithm. The embedding dimension was assigned 2, 3, 4... respectively. For each of the embedding dimensions, calculated the value of $\ln(r)$ and $\ln(Q)$, and fitted the slope of the data, which was the correlation dimension of the chaotic system. The following figure was the $\ln(r) - \ln(Q)$ correlation diagram obtained by the G-P algorithm in the case of different embedding dimensions.

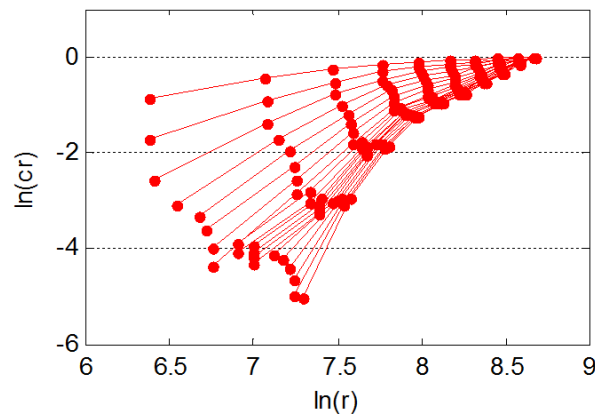


Figure 6. $\ln(r) - \ln(Q)$ Diagram of the Original Data When the Time Delay was 3 in the Case of Different Embedding Dimensions.

With the increase of the embedding dimension, when $m=12$, the correlation dimension was no longer increased and reached saturation, as shown in Figure 7.

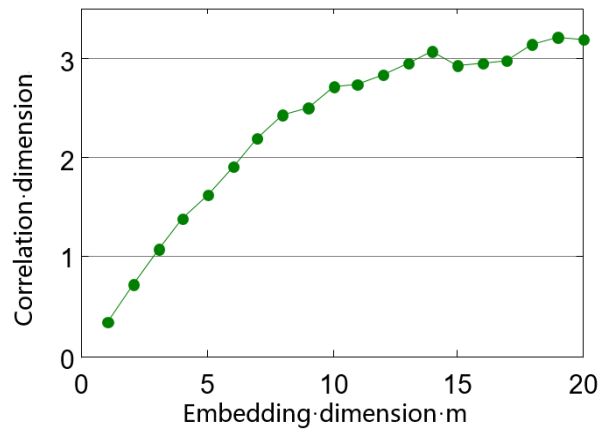


Figure 7. Relation between System Correlation Dimension and Embedding Dimension

The above research showed that the time series had chaotic characteristics, at the same time, $m=12$ was the minimum embedding dimension of the phase space of the data. On the other hand, the Wolf method was also used to calculate the maximum Lyapunov exponent of the monthly runoff series data, which was $0.3329 > 0$, and this also showed that the system had some chaotic characteristics.

4.2. Prediction Effect Evaluation

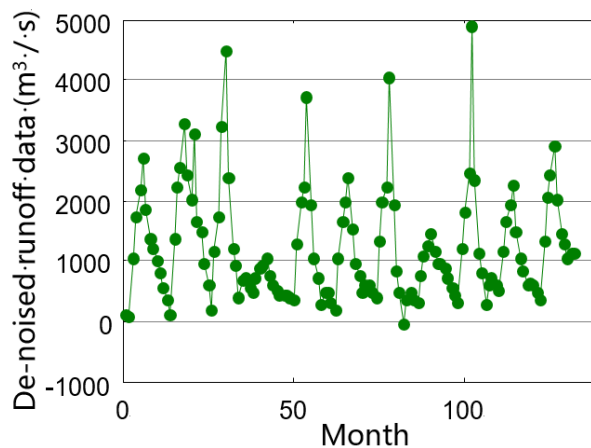


Figure 8. De-Noised Monthly Mean Runoff Data of Hydrologic Station during 132 Months

Using RBF neural network to predict the original data, at the same time, took the first 66 numbers as the sample sequence, and the last 66 number as the prediction sequence. Compared the RMSE value between the predicted data and the original data, and get its RMSE value was 2.29, the result is shown in Figure 9(1). Meanwhile we used the method of this paper to deal with the same de-noised original data, the de-noising results as shown in Figure 9(2).

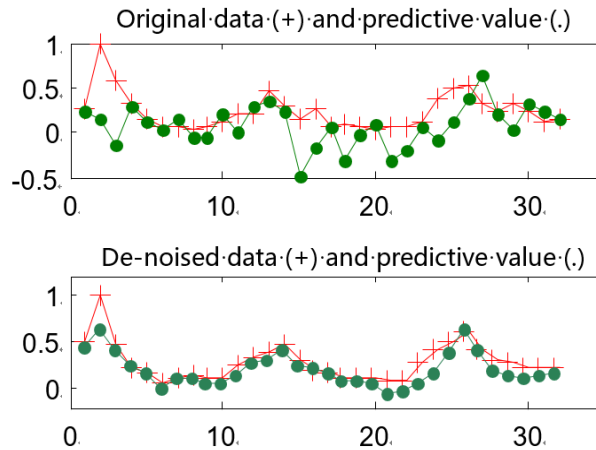


Figure 9. Comparison of the Predicted Effect of the Runoff Before and After De-Noising

Recalculated the RMSE value of the predicted value, and its value was 0.31, the prediction effect is better.

5. Conclusion

A lot of experimental data had to be de-noising processed. For different characteristics of the noisy signal, if we chose different wavelet basis function and threshold rules, then the de-noising effect was not the same. In this paper, wavelet theory was combined with the de-noising process for the single time variable of chaotic system, and solved the problem of the parameter selection of wavelet de-noising in chaotic system. Used SNR and RMSE to evaluate the de-noising effect. And used the chaotic time series generated by the Lorenz system and the chaotic time series of the hydrologic runoff data to perform the de-noising simulation. The experimental results showed that this method could improve the chaotic characteristics of time series and improve the accuracy of prediction.

Acknowledgements

The authors are grateful to the director and workmate, Gao Huiwang and Ruan Xuejing, who gave us active support and lots of advice. What's more, the authors are grateful "Key Laboratory of marine environment and ecology, Ocean University of China" and "Application research project of post-doctoral personnel in Qingdao", because it provides additional room and financial support.

References

- [1] M. Xiang, "Study on Nonlinear De-noising Method of Chaotic Time Series", Dalian University of Technology. Dalian, (2007).
- [2] Z. Jako, "Application of Noise Reduction to Chaotic Communications: a Case Study", IEEE Transactions on Circuits and Systems: Fundamental Theory and Applications, Washington D C, (2000), October 9-12.
- [3] H. Dediou and A. Kisel, "Communications with Chaotic Time Series: Probabilistic Methods for Noise reduction", International Journal of Circuit Theory and Applications, vol. 27, no. 1, (1999), pp. 577-587.
- [4] H. Min, L. Yuhua, X. Jianhui and S. Zhiwei, "Chaotic Signal De-noising Based on Threshold Selection of Wavelet Transform", Information and Control, vol. 34, no. 5, (2005), pp. 543-547.
- [5] Y. Yong, S. Chao and D.Hao, "Application of Wavelet De-noising Method in the Magnetic Signal Detection", Information & Communications, vol. 2, no. 3, (2015), pp. 6-8.
- [6] L. Cheng, "Wavelet and spectrum analysis", Shanghai Jiao Tong University, Shanghai, (2009).
- [7] F. Qibin, "Wavelet Analysis", Wuhan University Press, Wuhan, (2008).

- [8] H. Daihai, X. Jianxue and C. Yonghong, "Research on the Wavelet Transform Methods for the Reconstruction of the Nonlinear Dynamics System", *Journal of Vibration Engineering*, vol. 12, no. 1, (1999), pp. 27-32.
- [9] T. Schreiber, "Determination of the noise level of chaotic time series", *Physical Review E*, vol. 48, no. 1, (1993), pp. 13-16.
- [10] A. Leontitsis, J. Pange and T. Bountis, "Large noise level estimation", *International Journal of Bifurcation and Chaos*, vol. 13, no. 8, (2003), pp. 2309-2313.
- [11] K. Urbanowicz and J. A. Holyst, "Noise-level estimation of time series using coarse-grained entropy", *Physical Review E*, vol. 67, no. 4, (2003), pp. 46-48.
- [12] D. L. Donoho, "De-noising by soft-thresholding", *IEEE Transactions on Information Theory*, vol. 41, no. 3, (1995), pp. 613-627.
- [13] Z. Ruizhen and S. Guoxiang, "The self-noise improved de-noising method based on the wavelet transform", *Xi'an University of Electronic Science and technology journal*, vol. 27, no. 5, (2000), pp. 619-622.
- [14] H. Tengfei, L. Bangyi and X. Jixia, "Noise estimation and reduction for chaotic time series by wavelet variance decomposition", *Journal of Computer Applications*, vol. 33, no. 3, (2013), pp. 890-895.
- [15] Fan Qibin, "Wavelet Analysis", Wuhan University Press, Wuhan, (2008).
- [16] H. Daihai, X. Jianxue and C. Yonghong, "Research on the Wavelet Transform Methods for the Reconstruction of the Nonlinear Dynamics System", *Journal of Vibration Engineering*, vol. 12, no. 1, (1999), pp. 27-32.
- [17] T. Xiaorui, L. Hong and S. Lifu, "Wavelet analysis of forest fire in Heilongjiang", *Journal of natural disasters*, vol. 2, no. 4, (2007), pp. 15-19.

Author



Sun Hai, Sun Hai received his doctor's degree in Harbor, coastal and offshore engineering from Ocean University of China in Qingdao, China. He is currently a lecturer in Ocean University of China. His research interest is mainly in the area of intelligent algorithm in mixed variables optimal design of structure, Port management information system, Computer Software. He has published several research papers in scholarly journals in the above research areas and has participated in several books and many practical projects.