

A new precipitation forecast method based on CEEMD-WTD-GRU

Xianqi Zhang^{a,b,c}, Bingsen Duan^{id a,*}, Shaoyu He^a, Xilong Wu^a and Dong Zhao^a

^a College of Water Resources, North China university of water resources and electric power, Zhengzhou 450046, China

^b Collaborative Innovation Center for Efficient Utilization of Water Resources in the Yellow River Basin, Zhengzhou 450046, China

^c Henan Water Conservancy and Water Transportation Engineering Research Center, Zhengzhou 450046, China

*Corresponding author. E-mail: clcdbs@163.com

 BD, 0000-0002-5599-583X

ABSTRACT

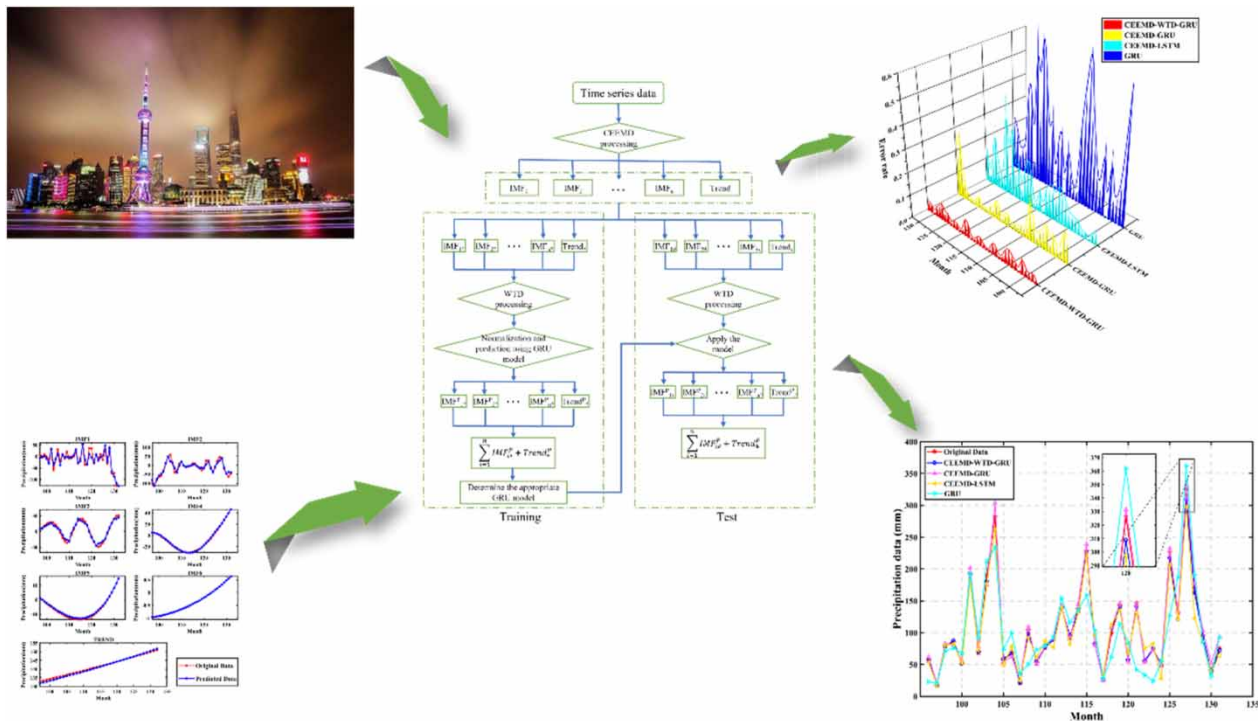
The simulation of precipitation changes can provide references for the prediction and prevention of flood disasters, and has guiding significance for the comprehensive utilization of regional water resources. Precipitation forecasting is difficult due to the randomness and uncertainty of precipitation events. CEEMD can effectively overcome modal aliasing and white noise interference. The WTD process has obvious denoising effects on the original signal. GRU can effectively solve long-term memory and reflection. Based on the advantages of problems such as gradients in propagation, a CEEMD-WTD-GRU precipitation prediction coupling model is constructed. The second decomposition of CEEMD-WTD-GRU can more effectively extract complex time series information. The time series forecasting provided a new method, which effectively improved the accuracy of the forecast and applied it to the forecast of monthly precipitation in Shanghai. The research results show that the average absolute error of the CEEMD-WTD-GRU model is 3.86, the average relative error is 3.30%, and the Nash efficiency coefficient is 0.99. The prediction accuracy is better than the CEEMD-WTD-GRU model without noise reduction, the CEEMD-LSTM model and GRU model, which shows that it has strong nonlinear and complex process learning ability in hydrological factor simulation, and can be used for regional precipitation prediction.

Key words: CEEMD, forecast, GRU neural network, monthly precipitation, Shanghai, wavelet threshold denoising

HIGHLIGHTS

- Complementary ensemble empirical modal decomposition (CEEMD) is a relatively novel data preprocessing method that can effectively reduce the non-smoothness of time series.
- Wavelet threshold noise (WTD) reduction is an excellent noise reduction technology that can effectively reduce the noise in the signal.
- Gated Recurrent Unit (GRU) as a prediction model is more adept at handling long time series.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Precipitation is an important replenishment method for regional water resources, and accurate precipitation prediction can effectively reduce the impact of severe weather. Precipitation is the main climatic factor and an important link in the water cycle. It is of great significance to analyze its change characteristics. Therefore, accurate prediction of precipitation can provide technical support for the sustainable use of regional water resources, flood prevention and disaster reduction, and ecological environment protection (Chen *et al.* 2017). The precipitation time series is a non-stationary and non-linear signal, which can be decomposed and time-frequency analyzed. At the end of the last century, Huang *et al.* (1998) proposed a new method of processing non-stationary signals, EMD (Empirical Mode Decomposition), which has been widely used in various fields of signal processing. Although EMD overcomes the problem of relying on subjective experience when setting basis functions in wavelet analysis, due to its algorithm itself, modal aliasing will occur when IMF decomposition is performed on the original sequence of historical loads. In order to solve this problem, Wu & Huang (2004) proposed a research conclusion on EMD processing white noise, that is, EEMD (Ensemble Empirical Mode Decomposition), but in subsequent studies, it was found that the white noise introduced by EEMD may be mixed into the original signal sequence, causing reconstruction error, therefore, on the basis of EEMD, a new enhanced noise-assisted data analysis method-CEEMD (Complementary Ensemble Empirical Mode Decomposition) (Zhang *et al.* 2021) is proposed. The IMF components decomposed by CEEMD can be further used to denoise WTD (wavelet threshold denoising) (Yue *et al.* 2021) to obtain a more stable component for subsequent prediction work. GRU (Gate Recurrent Unit) is a kind of recurrent neural network (RNN), suitable for processing time series data, through the neural network to learn the changes of each sub-component and further prediction (Chen *et al.* 2021). Like LSTM (Long-Short Term Memory), it is also proposed to solve problems such as long-term memory and gradients in back propagation. Compared with the cumbersome calculation and lower training efficiency of LSTM, GRU can obtain better calculation results with fewer parameters and shorter time (Wang *et al.* 2021a).

At present, there are many researches on precipitation forecasting. The common precipitation forecasting models can be roughly divided into four categories: time series models, artificial intelligence models, combined forecasting models and hybrid forecasting models. Wang *et al.* (2021b) used the SSVDF model to predict the precipitation in the main flood season of the Yangtze River Basin, and better predicted the spatial distribution of river water anomalies during the main

flood season in the Yangtze River Basin, and [Ge et al. \(2021\)](#) used the time series to improve the two-stage attention mechanism. The precipitation forecast model successfully predicted the two-hour precipitation in Europe. [Sun & Lin \(2003\)](#) used the fuzzy weighted Markov model to take the precipitation data of the Hequ Hydrological Station in Shanxi Province for the past 50 years as an example. The method was specifically applied and received more satisfactory results. [Georgakakos & Hudlow \(1984\)](#) proposed a quantitative precipitation forecasting technique for hydrological forecasting in 1984. In 1998, [Kuligowski & Ba Rros \(1998\)](#) successfully used the artificial neural network numerical weather prediction model to test and forecast four locations in the mid-Atlantic area of the United States. Subsequently, machine learning was widely used in the research and study of precipitation prediction. The relative error of the currently widely used precipitation time series prediction models is generally between 5 and 15%, and there are greater difficulties in predicting the precipitation in a longer period. The use of machine learning to predict precipitation is still in its infancy, and there are few studies on preprocessing the data before prediction. Therefore, the paper combines the advantages of CEEMD and GRU, combined with wavelet threshold denoising technology, establishes a CEEMD-WTD-GRU coupling prediction model, and applies it to the Shanghai monthly precipitation forecast to perform a longer sequence of detailed precipitation changes. Feature analysis and prediction are of great significance.

2. RESEARCH METHODS

2.1. CEEMD (complementary ensemble empirical mode decomposition)

Based on the EMD method and the EEMD method, CEEMD can perfectly solve the modal aliasing phenomenon and has strong adaptability ([Zhang et al. 2021](#)). CEEMD, like EEMD, also assists the analysis by adding white noise. The specific steps are as follows:

- (1) Record the time series as the original signal $p(t)$, the added white noise is marked as $\omega^n(t)$, and the noise coefficient is represented by β_0 , then the time series original signal becomes $p(t) + \beta_0 \omega^n(t)$. The original signal is repeatedly decomposed N times by the EMD decomposition method, and the total average value is calculated according to the EEMD method and defined as the IMF component of the target signal $p(t)$, as shown in Equation (1):

$$IMF_1(t) = \frac{1}{N} \sum_{n=1}^N E_1\{p(t) + \beta_0 \omega^n(t)\} \quad (1)$$

- (2) The remaining component is regarded as the first-order residual $r_1(t)$, as shown in Equation (2):

$$r_1(t) = p(t) - IMF_1(t) \quad (2)$$

- (3) Continue to decompose the signal $r_1(t) + \beta_1 E_1(\omega^n(t))$ repeatedly N times, and define the result after the second decomposition as $IMF_2(t)$, as shown in Equation (3):

$$IMF_2(t) = \frac{1}{N} \sum_{n=1}^N E_1\{r_1(t) + \beta_1 E_1(\omega^n(t))\} \quad (3)$$

- (4) Calculate the k th order residual $r_k(t)$, Among them, $k = 2, \dots, K$, as shown in Equation (4):

$$r_k(t) = r_{k-1}(t) - IMF_k(t) \quad (4)$$

- (5) Decompose the signal $r_k(t) + \beta_k E_k(\omega^n(t))$ after a certain decomposition again, calculate the overall average to get the target signal $IMF_{k+1}(t)$, as shown in Equation (5):

$$IMF_{k+1}(t) = \frac{1}{N} \sum_{n=1}^N E_1\{r_k(t) + \beta_k E_k(\omega^n(t))\} \quad (5)$$

- (6) Repeat the steps (4) and (5) above until a certain residual can no longer be decomposed, stop the decomposition process, and get K IMF components, and the final residual M is shown in Equation (6):

$$R(t) = p(t) - \sum_{k=1}^K IMF_k(t) \quad (6)$$

Therefore, the original time series signal can be expressed by Equation (7):

$$p(t) = \sum_{k=1}^K IMF_k(t) + R(t) \quad (7)$$

From the above process, the basic process of CEEMD decomposition is to perform multiple repeated modal decomposition of the original time series signal. The decomposition process is complete and the original time series signal is accurately reconstructed. The CEEMD method has the same binary filtering characteristics as the EMD method. The IMF components obtained after decomposition are arranged in order from high frequency to low frequency. Usually the first few high frequency components often contain random noise. Therefore, the effect of noise reduction on the obtained IMF component is better.

2.2. WTD (wavelet threshold denoising)

The idea of thresholding wavelet coefficients comes from the theory proposed by [Donoho & Johnstone \(2012\)](#). Donoho first proposed a general threshold denoising formula based on orthogonal wavelet transform, which is a very concise and effective wavelet denoising method. The main idea is to perform wavelet transformation on the signal through the Mallat algorithm and select the generated wavelet coefficients. Since the wavelet coefficients of the noise after wavelet decomposition of the time series signal are smaller than the wavelet coefficients of the original signal, the noise reduction can be achieved by selecting an appropriate threshold and filtering the noise signal ([Yu & Zhen 2021](#)). Wavelet threshold denoising has the advantages of being able to obtain the approximate optimal estimation of the original signal, fast calculation speed and wide adaptability. It is the most widely used wavelet denoising method.

2.3. GRU (gate recurrent unit)

Gated Recurrent Unit (GRU) is a variant of Recurrent Neural Network (RNN), which is very similar to the internal unit of long short-term memory network (LSTM), and was proposed [Cho et al. \(2014\)](#). It is also proposed to solve the problems of long-term memory and gradients in back propagation. The calculation structure is shown in [Figure 1](#).

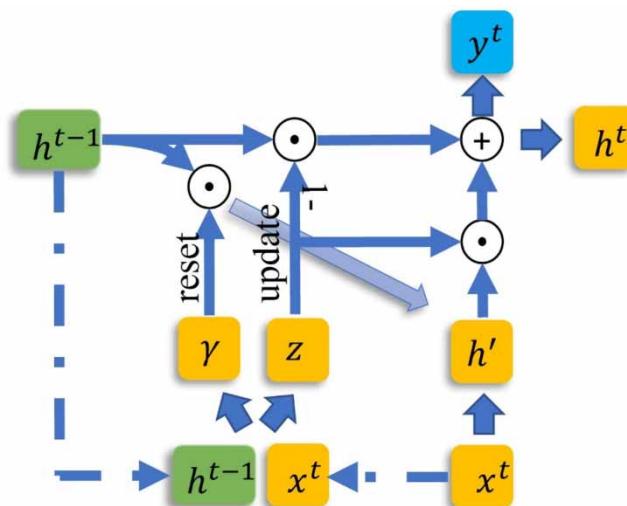


Figure 1 | GRU internal structure.

⊙ in the figure represents the multiplication of the corresponding elements in the matrix, so the two multiplication matrices are required to be of the same type. ⊕ Represents matrix addition. γ is the gate control signal to control reset, z is the gating signal that controls the update. The specific formula is as follows (Zhao *et al.* 2019):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (8)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (9)$$

$$\gamma = \sigma(W^\gamma[h^{t-1}, x^t]) \quad (10)$$

$$z = \sigma(W^z[h^{t-1}, x^t]) \quad (11)$$

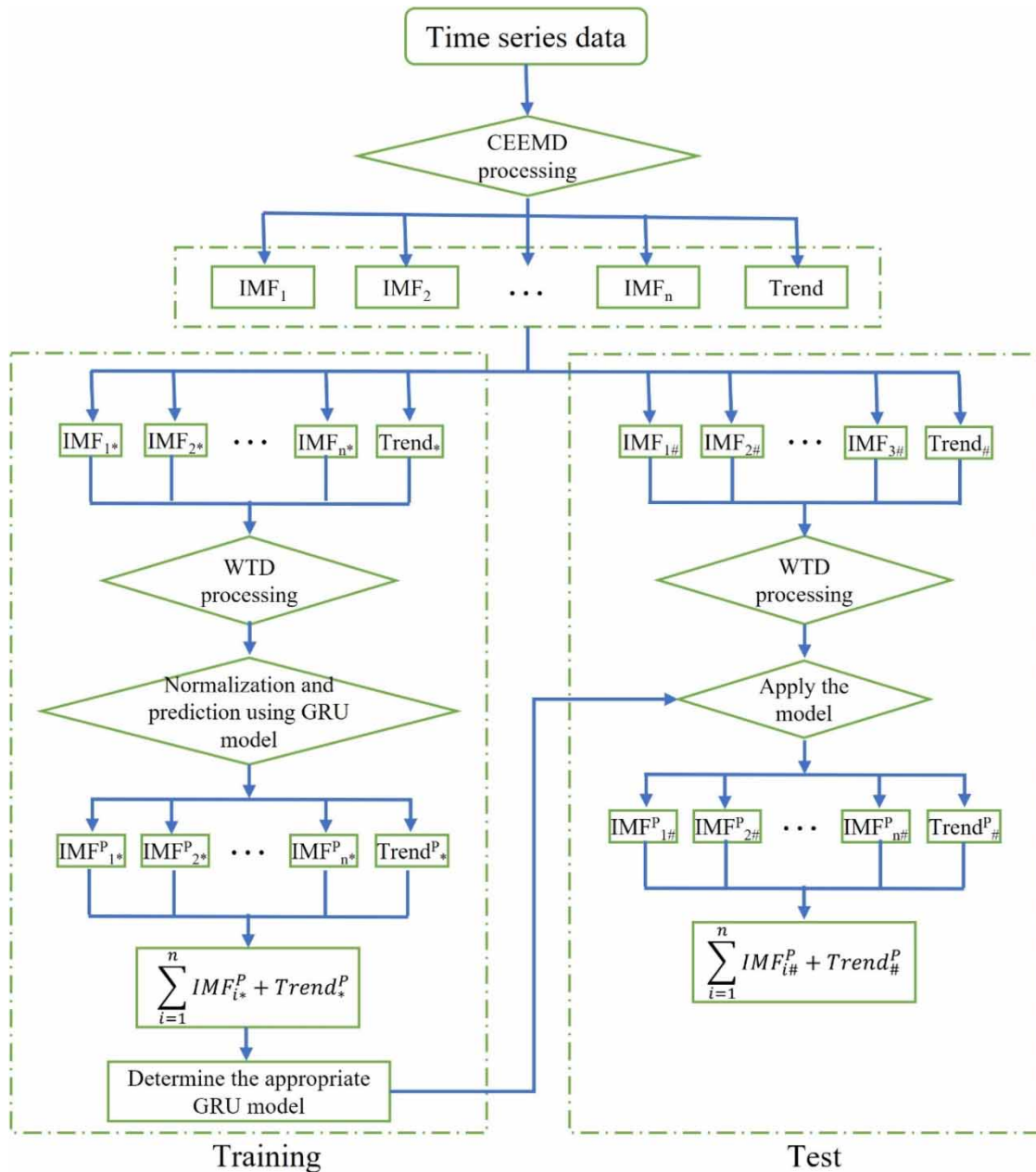


Figure 2 | CEEMD-WTD-GRU coupling model technical route.

$$h' = \tanh(W^h[h^{t-1'}, x^t]) \quad (12)$$

$$h^t = (1 - z) \odot h^{t-1} + z \odot h' \quad (13)$$

Among them: W^y and W^z are the weight matrix of the reset gate and the update gate respectively; W^h is the weight matrix when calculating h' ; \odot represents the connection of the two vectors.

2.4. CEEMD-WTD-GRU coupling model

In order to improve the accuracy of prediction, the CEEMD-WTD-GRU prediction model is proposed. Decompose complex time series data into multiple easy-to-predict IMF components. Each sub-component after decomposition has different feature scales. The components IMF1-IMF3 are optimized for noise reduction. Using the decomposed components for calculation can effectively reduce non-stationarity. GRU neural network, as a cyclic neural network with memory capabilities, builds sub-models for each IMF component and performs machine learning, which can effectively use long time sequence information to make more accurate predictions. The model structure is shown in Figure 2.

In order to measure the prediction accuracy of the CEEMD-WTD-GRU coupling model, the average relative error (MAPE) and Nash efficiency coefficient (NSE) between the original data and the predicted value are used as the evaluation criteria. The specific formula is as follows:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (14)$$

$$\text{NSE} = 1 - \frac{\sum_{t=1}^N (y_t - \bar{y}_t)^2}{\sum_{t=1}^N (y_t - \mu_t)^2} \quad (15)$$

Among them: y_i is the measured value at time i ; \hat{y}_i is the predicted value at time i ; N is the total length of the time series.

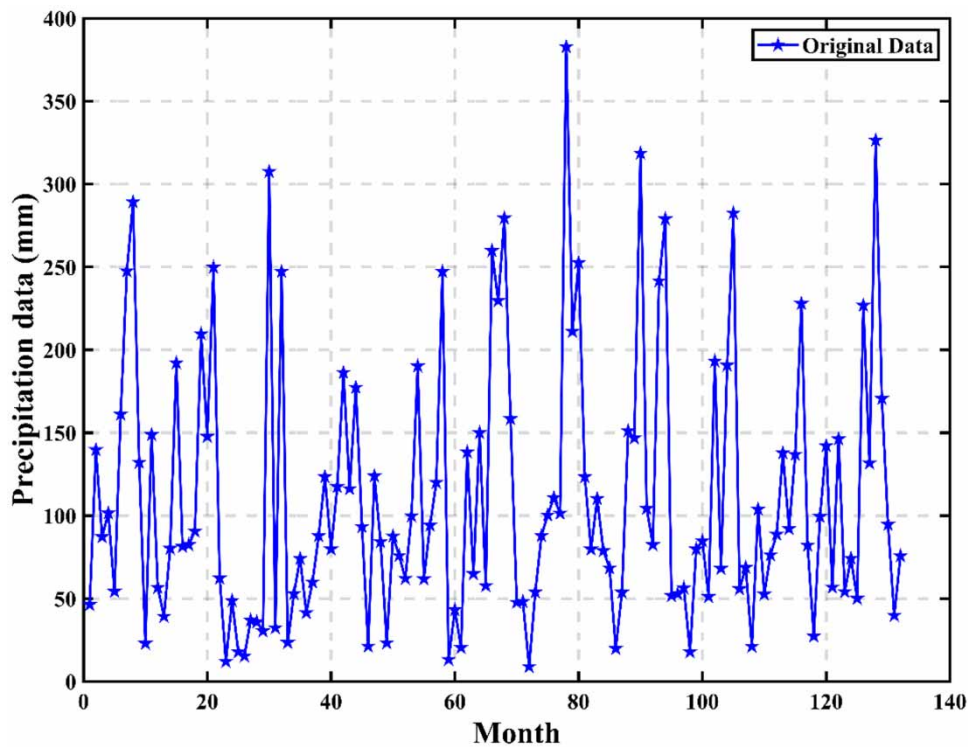


Figure 3 | Monthly precipitation data of Shanghai from 2009 to 2019.

3. EXAMPLE APPLICATION

In order to verify the rationality of the CEEMD-WTD-GRU coupling prediction model, an example application of precipitation data in Shanghai was carried out. The monthly precipitation data of Shanghai from 2009 to 2019 collected through the Shanghai Water Resources Bulletin contains a total of 132 data. As Shanghai is located in the coastal area, the precipitation data is extremely irregular. CEEMD has great advantages in processing non-stationary and non-linear time series data, while the GRU model has a good effect on the learning of longer time series data. Therefore, we apply the CEEMD-WTD-GRU coupling model to Shanghai precipitation simulation work. The monthly precipitation data of Shanghai from 2009 to 2019 is shown in Figure 3.

3.1. CEEMD

Using the CEEMD algorithm to decompose the Shanghai precipitation data, it is found that when the noise amplitude is 0.2 and the noise logarithm is 50, the decomposition effect is ideal. After CEEMD decomposes the time series, six IMF components and one trend component are obtained as shown in Figure 4.

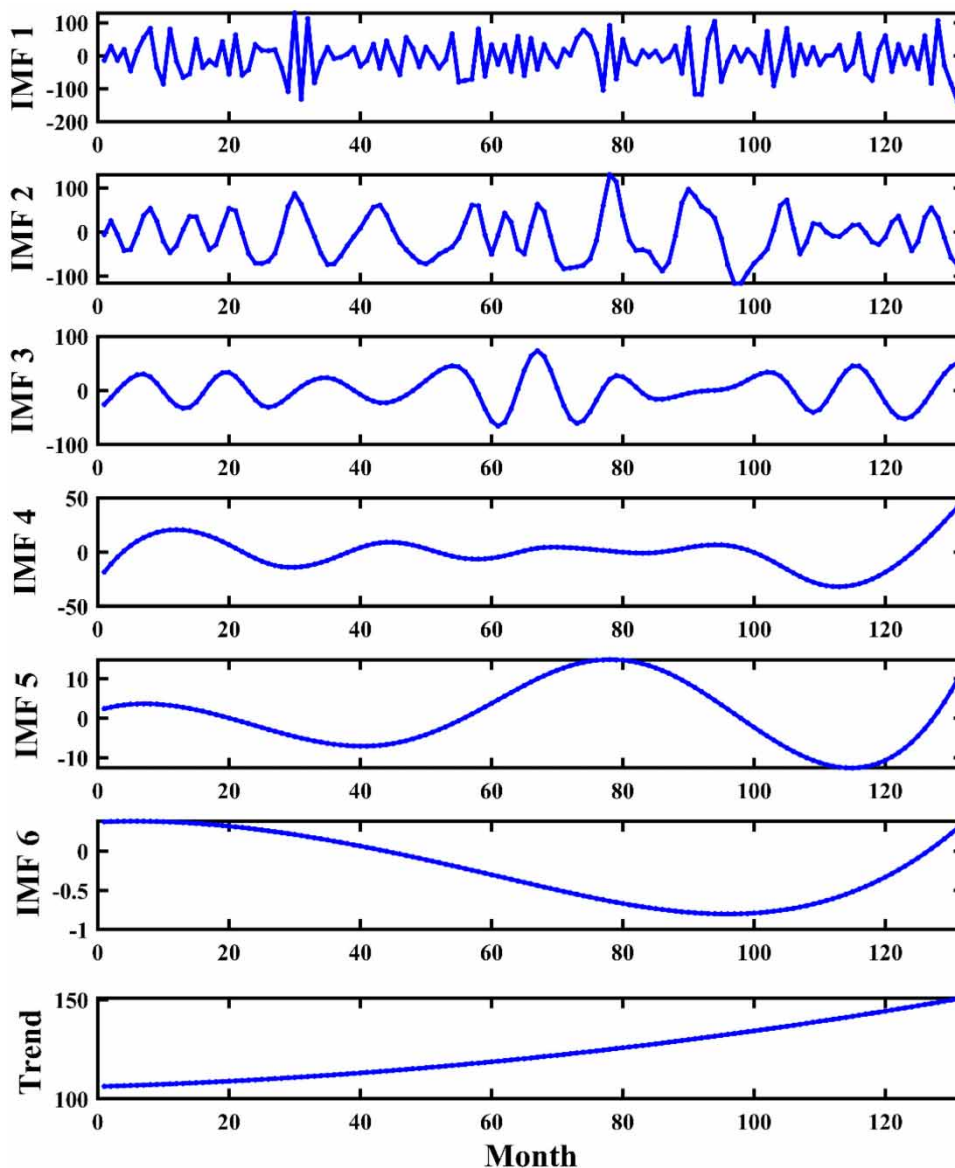


Figure 4 | CEEMD decomposition results of monthly precipitation data in Shanghai.

It can be seen from Figure 4 that the trend item of Shanghai precipitation data shows that this time series is increasing month by month. Due to the nonlinearity and non-stationarity of the time series and the characteristics of binary filtering in CEEMD, the waveforms of the components IMF1-IMF3 fluctuate more drastically, which are high-frequency components, and generally random noise will be included.

3.2. WTD

Conventional CEEMD denoising generally directly discards the noise-containing IMF components, but this will cause the problem of high-frequency effective signal loss or incomplete removal of random noise. Therefore, the wavelet threshold noise reduction (WTD) technology is used to further reduce the noise of the three high-frequency components of IMF1-IMF3 to improve the accuracy of the model. The noise reduction results for IMF1-IMF3 are shown in Figure 5.

The red in the figure represents the original data, and the blue represents the noise-reduced component. It can be clearly seen that after the noise reduction process, the fluctuation of the three components of IMF1-IMF3 relative to the original data is significantly reduced, and the stability is significantly improved. Comparing and analyzing the original and denoising coefficients in wavelet decomposition, it can be seen that the number of denoising layers is seven and the soft threshold denoising

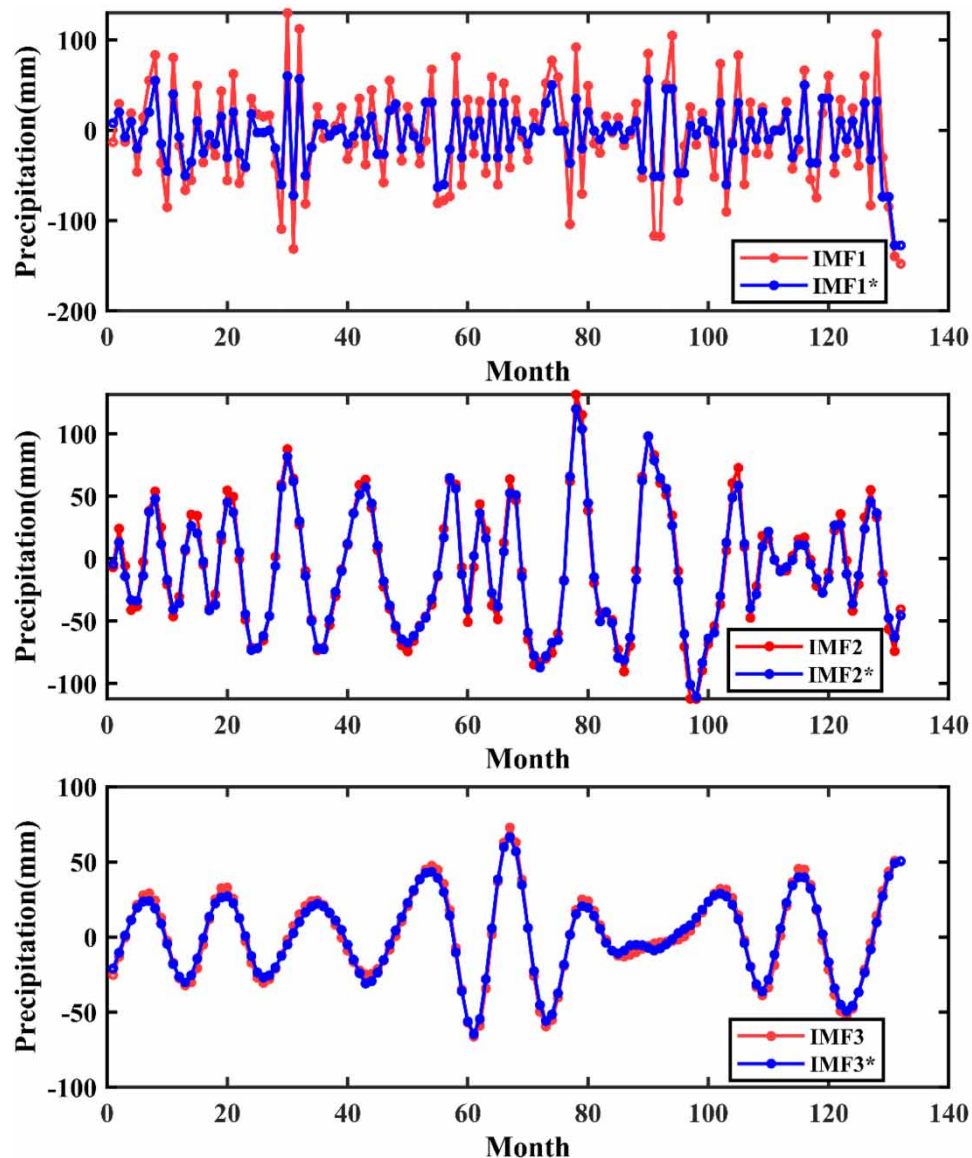


Figure 5 | Wavelet threshold denoising results of IMF1-IMF3 components.

effect is good. It can be seen that the third-order component adopts the soft threshold to reduce the noise, and shows the characteristics of the original signal more completely, and the fluctuation of some details is also more accurate.

3.3. Precipitation forecast

Select Shanghai January 2009–December 2012 precipitation data as the training set, and January 2017–December 2019 data as the prediction sample. After many times of calibration, the selected optimal GRU model parameters are: learning rate, the maximum number of iterations is 421, the gradient threshold is 1, the hidden node is 616, and the initial input and output are both 0. The prediction of the six IMF components and trend items by the GRU model is shown in Figure 6.

It can be seen from Figure 7 that the stability of the Shanghai precipitation time series after CEEMD decomposition and wavelet threshold denoising has been significantly improved, and the volatility has been significantly reduced. The prediction effect of a single component is good. The prediction results of IMF1–IMF6 and trend items are reconstructed and compared with the original precipitation data of Shanghai. The results are shown in Figure 8.

It can be seen from the above results that the CEEMD-WTD-GRU model has good follow-up and volatility in the prediction of precipitation, and the prediction trend is basically consistent with the original data. The maximum relative error is 8.02%, the minimum relative error is 0.30%, and the average relative error is 3.30%. The Nash efficiency coefficient is 0.99, indicating that the model has a small prediction relative error, which further verifies the high accuracy and good stability of the CEEMD-WTD-GRU model. The prediction of the peak and trough positions of the original data shows the excellent learning ability of the CEEMD-WTD-GRU model. The prediction trend is basically the same as the original data, and there is no prediction lag.

4. DISCUSSION

The CEEMD-WTD-GRU model has shown good results in the precipitation prediction test in Shanghai. In order to reflect the accuracy improvement effect of the research model in this paper, the CEEMD-GRU model, CEEMD-LSTM model (Zhang *et al.* 2020) and GRU (Zhang *et al.* 2017) model without noise reduction are used to compare with the prediction results of this paper. Calculate the errors of the predicted and actual values of different models and their Nash efficiency coefficient (NSE), and the results are shown in Figures 9 and 10.

It can be seen from Figure 9 that the prediction accuracy of the GRU model is poor, and the prediction results of the other several prediction models are roughly the same as the original data. Among them, the CEEMD-WTD-GRU model has the best effect. It can be seen more clearly from Figure 10 that the relative error between the prediction result of the CEEMD-WTD-

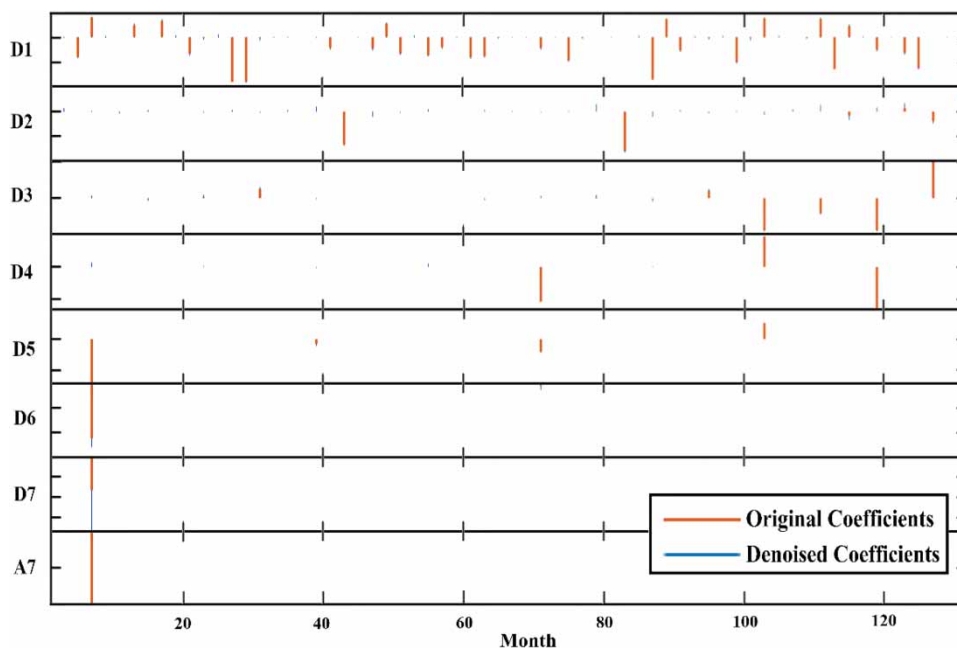


Figure 6 | Wavelet threshold denoising decomposition coefficient graph.

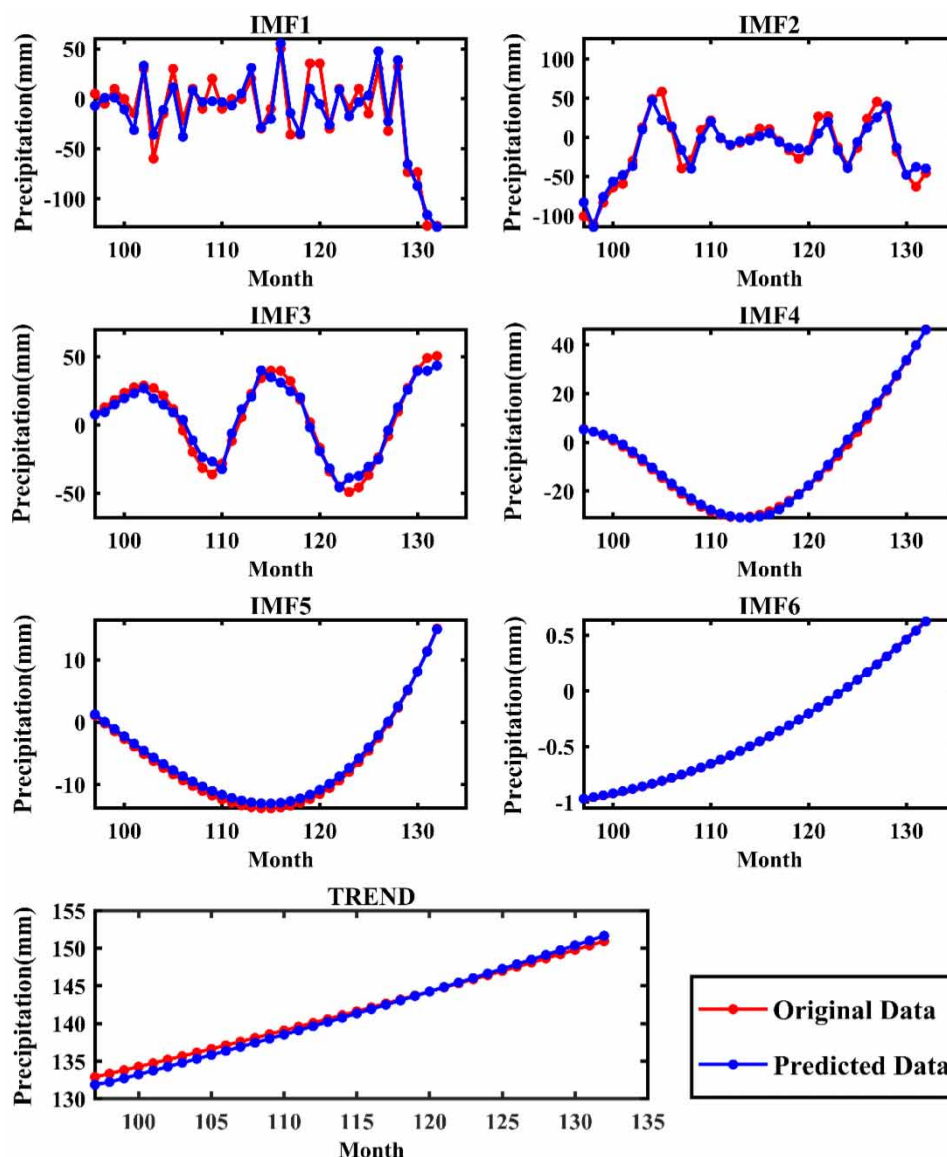


Figure 7 | Forecast results of each component and trend item.

GRU model and the original data is the smallest. It shows that after the CEEMD decomposition is converted into multiple IMF components, the noise reduction is performed first, and then the CEEMD-WTD-GRU coupling model is established for prediction, which can effectively improve the prediction accuracy on the original basis. The average relative error and Nash efficiency coefficient (NSE) of various algorithms are shown in Table 1.

The study found that the CEEMD-WTD-GRU algorithm is more accurate in predicting the peaks and troughs, there is no hysteresis, and it can effectively reduce the adverse effects caused by extreme weather. Analyze model calculation results and error comparison. The main error of this algorithm is that the short time series is the main reason for the error of this algorithm. Because the GRU model used in this paper has good learning ability for nonlinear and non-stationary long-term series data, it uses longer time series data that can effectively reduce the average relative error in the forecast and improve the Nash efficiency coefficient (NSE).

5. CONCLUSION

- (1) The simulation prediction of Shanghai's precipitation data from 2017 to 2019 shows that the model has good follow-up and consistency. Compared with the currently more widely used CEEMD-LSTM model and GRU model, the noise-

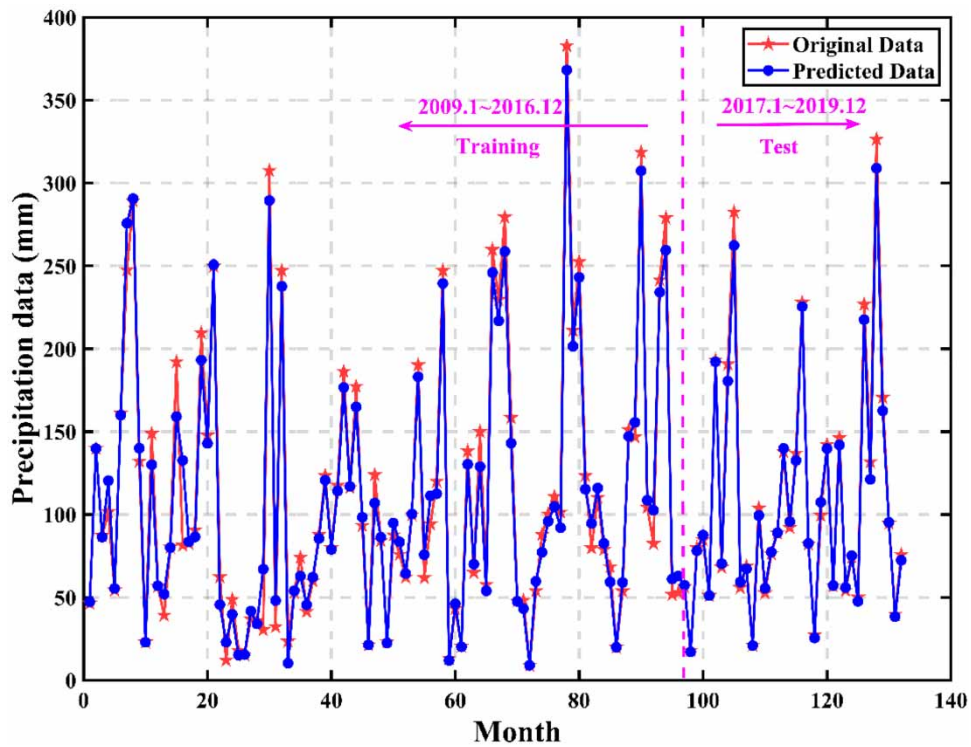


Figure 8 | Comparison of the prediction results of the CEEMD-WTD-GRU model with the original data.

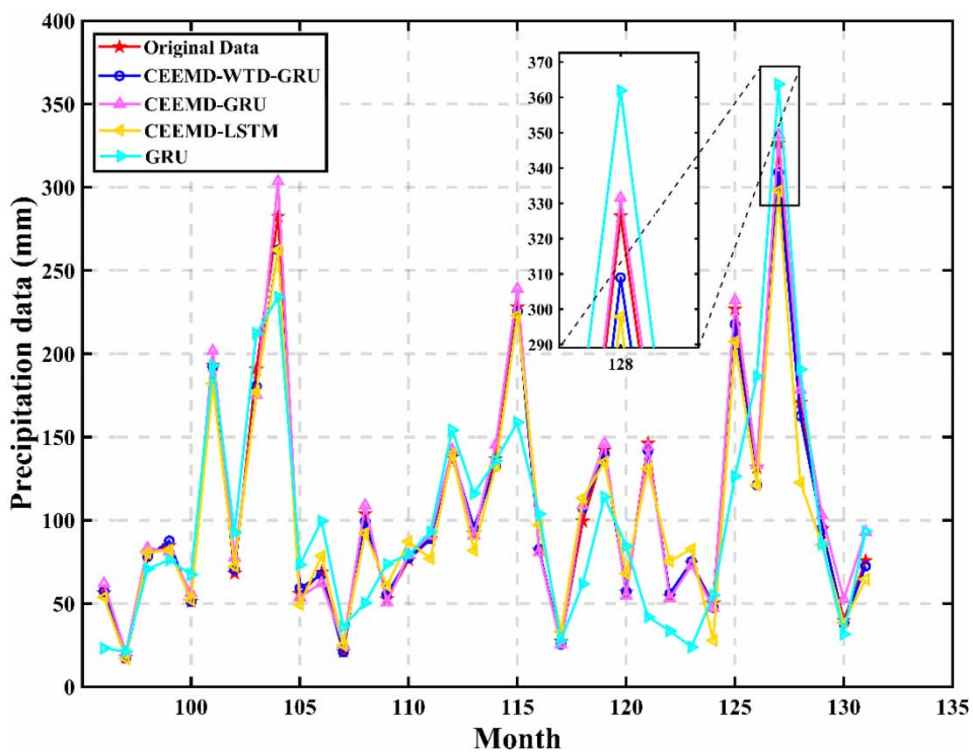


Figure 9 | Comparison of prediction results of multiple algorithms with original data.

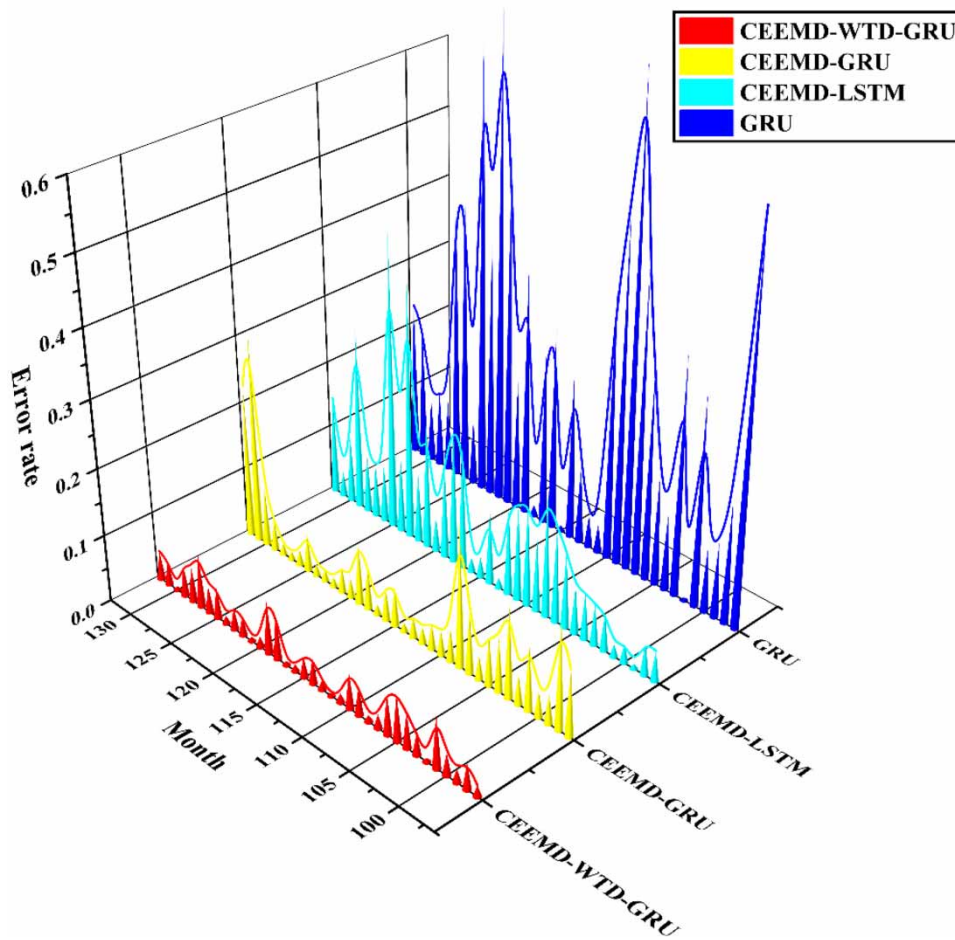


Figure 10 | Errors between the prediction results of multiple algorithms and the original data.

Table 1 | Comparison of prediction errors

Predictive model	Average relative error	NSE
CEEMD-WTD-GRU	3.30%	0.99
CEEMD-GRU	6.96%	0.94
CEEMD-LSTM	11.72%	0.89
GRU	27.66%	0.78

reduced CEEMD-WTD-GRU model is 52.58% higher than the CEEMD-GRU model without noise reduction, reducing the average relative error to 3.30%. The Nash efficiency coefficient reaches 0.99, and the performance of various indicators shows that the model is effective and the results are accurate.

- (2) The currently widely used CEEMD model generally lacks a noise reduction process, and the IMF components obtained by long-term sequence decomposition usually have many peaks and valleys. The data without noise reduction will have a greater impact on subsequent predictions. Compared with the CEEMD-GRU model without noise reduction, the accuracy of the proposed CEEMD-WTD-GRU model is significantly improved.
- (3) It should be pointed out that this method is mainly used for the prediction of long-term series. Although the application effect is good in scenarios with sufficient original data, it has certain limitations in the context of lack of data support. In future predictions, algorithms with stronger learning capabilities can be used to achieve further results.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

FUNDING

The authors wish to thank the Key Scientific Research Project of Colleges and Universities in Henan Province (CN) [grant numbers 17A570004].

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

- Chen, L., Gu, W., Bo, Z. & Liu, X. 2017 The statistical downscaling method of summer rainfall prediction over the Huang-Huai Valley. *Journal of Applied Meteorology* **28** (02), 129–141.
- Chen, K., Zhang, X. & Li, N. 2021 Application of CEEMD and adaptive wavelet threshold combined noise reduction in the OPAX method. *Journal of Vibration and Shock* **40** (16), 192–198.
- Cho, K., van Merriënboer, B., Gülçehre, C., Bougares, F., Schwenk, H. & Bengio, Y. 2014 Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Computer Science* arXiv:1406.1078, 2014.
- Donoho, D. L. & Johnstone, I. M. 2012 Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association* **90** (432), 1200–1204.
- Ge, M., Lu, Z., Liang, S. & Xia, Y. 2021 Intelligent precipitation forecast based on improved dual-stage attention mechanism. *Journal of Nanjing University of Information Technology (Natural Science Edition)* 1–10. [2021-08-31]. Available from: <http://kns.cnki.net/kcms/detail/32.1801.N.20210630.0844.004.html>.
- Georgakakos, K. P. & Hudlow, M. D. 1984 Quantitative precipitation forecast techniques for use in hydrologic forecasting. *Bulletin of the American Meteorological Society* **65** (11), 1186–1200.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C. & Liu, H. H. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **454** (1971), 903–995.
- Kuligowski, R. J. & Ba Rros, A. P. 1998 Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks. *Weather & Forecasting* **13** (4), 1194–1204.
- Sun, C. & Lin, X. 2003 Research on fuzzy Markov chain model with weights and its application in predicting the precipitation state. *Journal of System Engineering* **04**, 294–299.
- Wang, J., Cui, Q. & Sun, X. 2021a A novel framework for carbon price prediction using comprehensive feature screening, bidirectional gate recurrent unit and Gaussian process regression. *Journal of Cleaner Production* **314**, 128024.
- Wang, L., Zhang, F., Min, Y., Qiu, H., Zhang, X. & Zi, L. 2021b Long-term precipitation prediction based on multiple climatic factors in Changjiang River Basin. *Yangtze River* **52** (07), 81–87.
- Wu, Z. & Huang, N. E. 2004 A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **460** (2046), 1597–1611.
- Yu, H. & Zhen, T. 2021 Research on optimal wavelet base selection based on wavelet threshold denoising. *Modern Electronic Technology* **44** (17), 86–89.
- Yue, L., Yu, X. & Li, Z. 2021 EMD-GRU shield vertical attitude prediction based on optimal combination weighting. *Tunnel Construction (Chinese and English)* **41** (S1), 167–173.
- Zhang, B., Xiong, D. & Su, J. 2017 A GRU-Gated attention model for neural machine translation. *IEEE Transactions on Neural Networks and Learning Systems* PP(99). doi:10.1109/TNNLS.2019.2957276
- Zhang, Y., Yan, B. & Aasma, M. 2020 A novel deep learning framework: prediction and analysis of financial time series using CEEMD and LSTM. *Expert Systems with Applications* **159**, 113609.
- Zhang, J., Liu, Z. & Wang, M. 2021 Research on natural gas price forecasting model based on CEEMD-ELM-ARIMA. *Natural gas and oil* **39** (04), 129–136.
- Zhao, B., Wang, Z., Ji, W., Gao, X. & Li, X. 2019 A short-term power load forecasting method based on attention mechanism of CNN-GRU. *Power Grid Technology* **43** (12), 4370–4376.

First received 11 November 2021; accepted in revised form 19 January 2022. Available online 2 February 2022