A novel regional annual precipitation predicting model

Xianqi Zhang\textsuperscript{a,b}, Zhijie Zheng\textsuperscript{a,*}, Xiaoyan Wu\textsuperscript{a}

\textsuperscript{a}College of Water Conservancy, North China University of Water Resources and Electric Power, Zhengzhou 450046 China, Tel. +86-156-5583-3573; email: 15655833573@qq.com (Z. Zheng), Tel. +86-158-3719-7937; email: 80324427@qq.com (X. Zhang), Tel. +86-139-3880-4555; email: 940725860@qq.com (X. Wu)

\textsuperscript{b}Collaborative Innovation Center of Water Resources Efficient Utilization and Protection Engineering, Zhengzhou 450046 China

Received 12 September 2019; Accepted 25 February 2020

\section*{Abstract}
Precipitation is a complex system affected by many factors. The evolution of annual precipitation time series is uncertain, non-linear and non-stationary. The accuracy of precipitation prediction is often not high by using a single mathematical method. Based on empirical mode decomposition (EMD) and ensemble empirical mode decomposition, the improved complementary ensemble empirical mode decomposition (CEEMD) can restrain the mode aliasing problem in the process of EMD decomposition, but there is still some noise present in the decomposition sequence, which affects the precision of prediction. The decomposed data of CEEMD are denoised based on wavelet transform, and combining the advantages of Elman neural network in adapting to time-varying and dynamic memory, a novel regional annual precipitation prediction model has been established, this model is applied to the prediction of annual precipitation in Zhengzhou based on the coupled model of CEEMD-wavelet transform-Elman neural network (CWE). The results show that the CWE model has a good prediction effect, and the Nash-Sutcliffe efficiency coefficient of the model is above 0.70. The prediction accuracy of the CWE model is higher than that of the single Elman model and other models, and its relative error and absolute error are lower. In addition, this model can reveal the influencing factors of annual precipitation evolution to a certain extent, and provide a new way for annual precipitation prediction.

\textbf{Keywords:} CEEMD; Wavelet denoising; Elman network; Annual precipitation; Prediction

\section{Introduction}
Scientific and accurate prediction of precipitation has an important guiding significance to regional water resources management, flood control, and disaster reduction, and water environment protection [1–5]. Research on regional precipitation prediction has become a hot topic for scholars at home and abroad. In data-based precipitation prediction, statistical models have traditionally been used [6,7]. Recently, an artificial neural network model has also been introduced into precipitation prediction applications [8–12]. Huang et al. [13] proposed an empirical mode decomposition (EMD) method. It is a new signal analysis method, which has strong local performance and is suitable for the study of non-linear, non-stationary and random signals. However, the intrinsic mode component (IMF) component obtained by the EMD method has the phenomenon of mode aliasing, which makes it difficult to characterize the original signal. Wu et al. [14] proposed an analysis method for noise-assisted data based on EMD, namely the ensemble empirical mode decomposition (CEEMD) method.
decomposition (EEMD) method, which greatly suppresses the modal aliasing of the IMF components in EMD methods. The complementary ensemble empirical mode decomposition (CCEMD) Yeh et al. [15], the model reduces the signal of reconstruction error by adding white noise with opposite values to the original signal and solves the problem of poor EMD completeness. Since Daubechies published an important paper on Wavelet [16], most engineers and scientists have become more and more familiar with wavelet. It has been shown that for many signal denoising methods, the threshold method of wavelet coefficients has near-optimal denoising characteristics. Wavelet denoising not only reduces the complexity of numerical calculation but also produces clearer results. Vidakovic et al. [17] address the shrinkage of wavelet coefficients and induced denoising in the time domain by taking into consideration the “time” characteristics of a noisy signal.

Scholars at home and abroad focus on the evolution law of precipitation in large basins, regions and cities. In Adamowski Jan’s research, a method of river flow forecasting in semi-arid watersheds based on a discrete wavelet transform and artificial neural network is proposed [18]. It can be seen that the research methods of the above scholars mainly focus on the traditional statistical model, single neural network and other aspects. Sayemuzzaman and Jha analyzed the time series of 249 precipitation stations in North Carolina from 1950 to 2009 [19]. The Mann–Kendall (MK) test, the Theil–Sen approach (TSA) and the sequential Mann–Kendall (SQMK) test were applied to quantify the significance of trend, the magnitude of trend, and the trend shift, respectively. Regional (mountain, piedmont and coastal) precipitation trends were also analyzed using the above-mentioned tests. Hou et al. [20] used the improved Morlet wavelet neural network to forecast precipitation in western Jilin. The results show that the model has low forecasting error and good performance. Yang et al. [21] based on the precipitation data from 96 weather stations in northwestern China during 1960–2013, the Continuous Wavelet Transform and the Mann–Kendall test were applied to analyze the precipitation spatiotemporal variations at different time scales. The relationships between the original precipitation and different periodic components were investigated. However, the traditional mathematical statistics model cannot learn the high-frequency and mutational data well, and cannot reflect the evolution characteristics of the sequence in the frequency domain. The traditional neural network has the defect of overtraining, which makes the network deviate too far away from the whole. Therefore, in order to reduce the non-stationarity of annual precipitation series, a coupling prediction model is established, which is a new way to improve the accuracy of annual precipitation prediction. CEEMD is a method to effectively deal with the decomposition of nonstationary signals. Compared to common decomposition methods, it effectively reduces the number of iterations and increases the reconstruction accuracy. Wavelet denoising the decomposed components can further reduce the non-stationarity of the sequence. The Elman neural network has a strong nonlinear fitting ability. Based on this, the CEEMD-wavelet transform-Elman neural network (CWE) model for annual precipitation prediction is established.

2. Theories and methods

2.1. Methods of data analysis

2.1.1. Complementary ensemble empirical mode decomposition

The supplement of EEMD Yeh et al. [15] solves the problem of poor completeness of EMD, but the decomposition of the signal depends on the selection of the noise amplitude and the number of integrations, and false components (components without physical meaning) are prone to occur. Based on this, an improved EEMD method is proposed. Complete Ensemble Empirical Mode Decomposition (CEEMD) is improved based on EMD and EEMD. It can reduce modal aliasing in EMD decomposition, and can solve the problem of large error of EEMD reconstruction [22]. The advantage of CEEMD is that it eliminates the pseudo components of the IMF and reduces the non-stationarity of the precipitation sequence.

The calculation steps of CEEMD are as follows:

- $n$ groups of positive and negative white noise sequences were randomly added to the original time series, the mean value of added auxiliary noise sequences is zero, and two sets of collective signals are generated [23].

$$
G_{a} = \begin{pmatrix}
1 & \cdots & S \\
1 & \cdots & -S
\end{pmatrix}
$$

(1)

where $G_{a}$ and $G_{s}$ are time-series signals after adding white noise, $N_{i}$ is auxiliary noise signal and $S$ is the original signal. Finally, there are $2n$ sets of signals were obtained.

- The EMD decomposition method is used to decompose each set signal, and each signal gets a set of $2n$-1 IMF components and a trend term. where the $j$th IMF component of the $i$th component is expressed as $c_{ij}$, $n$ is the number of decompositions for each signal.

- The decomposition results are obtained by combining multiple components, the decomposition results are as follows.

$$
c_{j} = \frac{1}{2n} \left( \sum_{i=1}^{n} c_{ij} \right)
$$

(2)

where $c_{j}$ represents the $j$th IMF component obtained by the final decomposition of CEEMD. $n$ is the number of added white noise sequences. Finally, the wave motion or trend of different scales in the signal is decomposed into a series of sequences with different characteristic scales step by step, and each sequence is an eigenmode function component.

$$
U(t) = \sum_{j=1}^{l} c_{j} + \sum_{j=l+1}^{n} c_{j} + r_{n}
$$

(3)

where $U(t)$ represents the sum of IMF components and the remaining components after CEEMD decomposition of the original noisy data, and $r_{n}$ represents the average trend or constant.
2.1.2. Wavelet transform

Wavelet noise reduction is a commonly used modern signal processing method. The research of Donoho et al. [24] shows that wavelet denoising has a wide range of functional adaptability and the optimal adaptive denoising ability. Wavelet denoising can effectively denoise the signal. In order to improve the quality of time series, the signal can be denoised by wavelet transform before the time series is taken as a sample.

In the paper, the Mallat algorithm is considered to implement wavelet decomposition. The process description is shown in Eq. (4).

\[
\begin{align*}
    c^0_n &= f_n \\
    c^j_n &= \sum_{k \in \mathbb{Z}} c^{j-1}_{kn} h^*_{2^n k} \\
    d^j_n &= \sum_{k \in \mathbb{Z}} c^{j-1}_{kn} g^*_{2^n k}
\end{align*}
\]

(4)

where \( f \) is the time domain signal, \( k = 0,1,2, \ldots, n-1 \) (\( n \) is the number of sampling points), \( h^* \) is the pulse response of the conjugate mirror filter, \( j \) is the number of wavelet decomposition layers, this paper considers the optimal decomposition to five layers.

After the signal is decomposed, the high-frequency coefficient threshold value obtained by decomposing is quantified, and a limited threshold value is selected for quantizing the high-frequency coefficient. There are two kinds of traditional threshold functions: hard threshold function and soft threshold function. The soft threshold denoising method is considered in this paper.

Using the characteristic of strong signal reconstruction ability of wavelet analysis, the Mallat reconstruction algorithm is used to reconstruct the decomposed signal. The reconstruction algorithm is actually the inverse process of the decomposition algorithm, the reconstruction algorithm is shown as follows:

\[
\begin{align*}
    c^j_n &= \sum_{k \in \mathbb{Z}} c^{j+1}_{kn} h_{2^n k} + \sum_{k \in \mathbb{Z}} d^{j+1}_{kn} g_{2^n k}
\end{align*}
\]

(5)

The basic flow is to process the high-frequency part of the wavelet results of each layer after decomposition of Eq. (4), and then reconstruct the signals of each layer by using Eq. (5). In other words, the main components reflecting the essential characteristics of signals can be obtained.

2.2. Elman neural network

Elman neural network, which was put forward by Elman [25], is a typical recurrent feedback neural network. The structure of the Elman neural network consists of four layers: an input layer, hidden layer, connection layer and output layer. Different from BP neural network [26], Elman neural network adds a layer of a connective layer in the hidden layer as a time delay operator to realize the dynamic memory of the system, and the output value of hidden layer as the input value of the next time to realize the dynamic feedback to the system. Based on the good memory and stability of the Elman neural network, it has been widely used in various fields. The Elman network structure is shown in Fig. 1.

![Elman neural network structured.](image)

In Fig. 1, \( u \) is the input vector, \( y \) is the output vector, \( x \) is the acceptance layer output vector, and \( x \) is the hidden layer output vector. \( W^3 \) is the connection weight from the input layer to the hidden layer; \( W^2 \) is the connection weight from the input layer to the hidden layer; \( W^3 \) is the connection weight from the hidden layer to the output layer. The expression of the Elman network model is as follows:

\[
y(k) = g(W^3 x(k))
\]

(6)

\[
x(k) = f(W^1 x(k) + W^2 (u(k-1)))
\]

(7)

\[
x_c(k) = x(k-1) + ax_c(k-1) (0 \leq a < 1)
\]

(8)

where \( g \) is the activation function of the output neuron, \( f \) is the activation function of the hidden layer neurons, and \( a \) is the self-joined feedback gain factor. When \( a = 0 \), the network is a standard Elman network; when \( a \neq 0 \), the network is a modified Elman network.

Elman neural network adopts the BP algorithm to modify the weight value, and the learning index function uses the square sum of the error function. The corresponding expression is listed below:

\[
E(w) = \sum_{n=1}^{N} (y_n(w) - \hat{y}_n(w))^2
\]

(9)

where \( \hat{y}_n(w) \) is the target output vector and \( y_n(w) \) is the original output vector.

Elman neural network learning algorithm adopts a momentum gradient descent back-propagation algorithm. The weights and thresholds are optimized by using the difference between the network output values and the output samples, to minimize the sum of squared errors of the network output layer.

3. Predictive coupling model based on CWE

Decomposition–prediction–reconstruction is one of the important ways of nonlinear time series prediction.
According to the “decomposition–prediction–reconstruction” framework, three main steps are involved in the proposed CWE learning paradigm: data decomposition, individual forecast and ensemble forecast.

To verify the effectiveness of the proposed CWE-based prediction model, the annual precipitation sequence was selected as the sample data. Furthermore, some popular forecasting technologies should be also performed as benchmark models for comparison purposes. In this section, the steps of the model are first designed, as described in section 3.1. In the aspect of evaluation criteria, section 3.2 gives the corresponding indicators in detail, the superiority of the proposed model in terms of prediction accuracy can be tested.

3.1. Model steps

The steps of CWE coupling model are as follows:

- Decomposition of precipitation observation data series from 1951 to 2017 at Zhengzhou Meteorological Station by CEEMD, the IMFs and one residue of time series are obtained.
- The IMFs and one residue are denoised by wavelet transform.
- The IMFs and one residue after wavelet de-noising are standardized. If the range of input or output data of the network varies greatly, the prediction model of the network will have a big error, so we must standardize the data and make the range of the data in [0,1]. The corresponding expression is listed below:

\[ y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \]  

(10)

where \( x \) is the original value of the \( t \) moment; \( x_{\min} \) is the minimum value of the sequence; \( x_{\max} \) is the maximum value of the sequence; \( y \) is the normalized value of the \( t \) moment.
- By using Elman neural network, the training data of IMF components and trend items are repeatedly adjusted to make the prediction of IMF components and trend items reach the best effect.
- Finally, the predicted IMF components and one residue are accumulatively restored and compared with the original data.

The calculation process of the CWE coupling prediction model is shown in Fig. 2.

3.2. Evaluation criteria of predicting performance

In order to better reflect the prediction effect of the CWE coupling model, four main criteria are used for evaluation of prediction, relative error (RE), mean relative error (MRE), mean absolute error (MAE), root mean square error (RMSE), and Nash–Sutcliffe efficiency coefficient (NSE) are taken as evaluation indicators.

\[ \text{RE} = \frac{|y_t - \bar{y}_t|}{y_t} \times 100\% \]  

(11)

\[ \text{MRE} = \frac{1}{N} \sum_{t=1}^{N} \left| \frac{y_t - \bar{y}_t}{y_t} \right| \times 100\% \]  

(12)

\[ \text{MAE} = \frac{1}{N} \sum_{t=1}^{N} |y_t - \bar{y}_t| \]  

(13)

\[ \text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (y_t - \bar{y}_t)^2} \]  

(14)

\[ \text{NSE} = 1 - \frac{\sum_{t=1}^{N} (y_t - \bar{y}_t)^2}{\sum_{t=1}^{N} (y_t - \bar{y})^2} \]  

(15)

where \( y_t \) is the actual value of the \( t \) moment; \( \bar{y}_t \) is the prediction value of the \( t \) moment, \( \mu_t \) is the total average of observations, \( N \) is the number of time series.
4. Case study

4.1. Data source

The data source of this study is the China ground international exchange station. Precipitation data of Zhengzhou meteorological stations from 1951 to 2017 were selected, and preliminary quality control was carried out to make the experimental data more representative. The location of the Zhengzhou meteorological bureau is shown in Fig. 3.

As can be seen from Fig. 4, the annual precipitation in Zhengzhou from 1951 to 2017 is influenced by the complexity, diversity and variability of meteorological conditions, and there is a lot of randomness and uncertainty in the precipitation process. The maximum annual precipitation is 1,040.7 mm, the minimum annual precipitation is 353.2 mm, and the minimum annual precipitation is in 2013, which brings some difficulties to the prediction of annual precipitation.

4.2. Complementary ensemble empirical mode decomposition

The CEEMD decomposition model was used to decompose the annual precipitation data of Zhengzhou city from 1951 to 2017. After repeated testing, when the maximum decomposition is 5; the noise frequency is 100, the noise amplitude is 0.2, the CEEMD has the best decomposition effect on precipitation. The result of decomposition is shown in Fig. 5.

The annual precipitation sequence of Zhengzhou city was decomposed into 4 IMF components and a corresponding residue by using the CEEMD method. Among them, the IMF1 component has the highest volatility, the highest frequency, and the shortest wavelength; the amplitude of the IMF2 – Residue gradually decreases, the frequency gradually decreases, and the wavelength gradually becomes larger. Due to the large non-stationary nature of the original data and strong volatility, the proportion of IMF1–IMF4 components is gradually reduced. At the same time, it can be seen from the residual term that the annual precipitation in Zhengzhou city is decreasing. It can be seen that the volatility and tendency of the series are greatly reduced by CEEMD decomposition. Therefore, this decomposition method is helpful to transform nonlinear and non-stationary sequences into a stationary time series, which can improve the performance of the sequences.

4.3. Wavelet denoising

Wavelet denoising can effectively denoise non-stationary signals, the IMF’s and trend terms generated by CEEMD decomposition are denoised by wavelet denoising to improve the quality of the sequence.
Due to the use of wavelet analysis for rotor fault signal noise reduction, it is difficult to determine the number of wavelet decomposition layers, and the effect of noise reduction is closely related to the speed of the fault rotor and signal sampling frequency, so the noise reduction process is difficult to complete automatically. This paper aims at this problem by means of a large number of experiments.

First, the original data are resampled and then decomposed to the specified number of layers by wavelet transform. Finally, the Donoho soft threshold method is used to achieve automatic noise reduction. The continuity of the wavelet coefficients obtained by the soft threshold is good, so the signal does not generate additional oscillations. The data after wavelet denoising is shown in Fig. 6.

As can be seen from Fig. 6, the wavelet denoising can denoise the non-stationary signal, and the effect is very good, which improves the quality of the sequence very well.

### 4.4. Annual precipitation prediction

The IMF components and trend items from 1951 to 2012 were used as training samples, and the IMF components and trend items from 1951 to 2012 are denoised by wavelet denoising.
trend items from 2013 to 2017 (5 y) were used as test samples. The activation function of the hidden layer is traingdx, the learning function is pure in, the number of iterations is 1,000, the number of hidden layer nodes is 20, and the number of input layer neurons is 4. The training target error tolerance is $10^{-5}$. The network prediction effect and absolute error are shown in Fig. 7.

It can be seen from Fig. 7, after CEEMD decomposition and wavelet denoising, the fluctuations and non-stationarity of the annual precipitation time series in Zhengzhou have been greatly reduced, and the fitting effect of the true and predicted values of the IMF 1 - residual is getting better and better, except for a few years of relative error is bigger, the relative error of the other component index

Fig. 7. Prediction effects and absolute errors of (a) IMF1, (b) IMF2, (c) IMF3, (d) IMF4, and residue.
showed a trend of a gradual decline, which enabling the Elman network to better predict its components and trend terms.

The prediction results of IMF1–IMF4 and residue are reconstructed into the prediction values of annual precipitation and compared with the original values of annual precipitation. The calculation prediction error is shown in Table 1.

As shown in Table 1, the absolute error and relative error of the CWE coupling model in predicting annual precipitation in Zhengzhou are within a reasonable range; the maximum and minimum of relative errors are 17.14% and 4.96%, and the average relative error is 11.77%, the model prediction relative error is relatively small.

In order to understand the training and prediction of the annual precipitation time series in Zhengzhou more intuitively and in more detail, the effect diagram of precipitation forecast from 2013 to 2017 is drawn, and the absolute error map of the real value and forecast value from 2013 to 2017 is made, as shown in Fig. 8. Finally, the MAE, RMSE, MRE, and NSE of the true and predicted annual precipitation in Zhengzhou from 2013 to 2017 were calculated. The results are shown in Table 2.

It can be seen from Table 2 and Table 3, the CWE coupling prediction model has a good prediction effect, the fitting effect between the real value and the prediction value is good, the prediction effect in 2016 is slightly poor, and the overall error is small. The NSE is 0.71, which is close to 1, indicating that the hydrological model has good quality and high model credibility [27].

4.5. Discussion

In order to verify the superiority of the CWE model, the Elman model, EMD-Elman model, EEMD-Elman model and CEEMD-Elman model were used to make predictions, and the prediction results of all models were compared with the real values. The comparison results are shown in Table 3.

From Fig. 9, prediction accuracy of the “decomposition–prediction–reconstruction” model is higher than that of the single neural network prediction model. It can be seen that the CWE coupling model is more obvious than other models when predicting the annual precipitation in Zhengzhou, and the predicted value and the original value have the highest fitting degree.

It can be seen from Table 3 that the average absolute error, root mean square error and average relative error of the CWE coupling model are smaller than the other four models when predicting annual precipitation. The CWE coupling model overcomes the disadvantages of white noise and other noises on other network models, and overcome their shortcomings in learning high frequency and non-stationary data. At the same time, the errors of the four “decomposition–prediction–reconstruction” models are significantly smaller than that of a single Elman network. The “decomposition–prediction–reconstruction” model has obvious advantages in predicting precipitation over a single neural network. The reason is that the smoothness of the original time series is greatly improved after the decomposition, and the prediction accuracy is improved.

5. Conclusion

In this paper, a new coupling model of regional annual precipitation prediction is constructed, which is based on the CEEMD decomposition and wavelet de-noising of
non-stationary annual precipitation time series and the adaptive ability of Elman neural network. Through the application of the annual precipitation prediction in Zhengzhou city, the average absolute error, root mean square error and average relative error of the model are all lower. The NSE of the model reaches 0.71, which indicates that the prediction effect is better. Empirical research shows that the proposed “decomposition–prediction–integration” model can significantly improve prediction performance because it can be statistically superior to other popular prediction methods, including single popular prediction tools (such as ARIMA, ANN, ENN, LSSVR), as well as other data combination tools and integration methods in terms of accuracy. This further shows that, the proposed CWE-based coupling model has effective decomposition algorithms and powerful, fast and stable predictive tools. It can deal with complex time series prediction tasks, especially high volatility and irregular annual precipitation prediction tasks, which have broad application prospects. The prediction algorithm based on the CWE model proposed in this paper can be used not only for forecasting annual precipitation but also for other time series forecast situations, such as sediment, runoff, stock, and crude oil prices. In addition, this study is based on time series analysis and does not consider the physical mechanism and long-term forecast of precipitation. How to make a more comprehensive analysis and improve the prediction accuracy will be the focus of the next research.

Acknowledgments

This work is financially supported by Technology Research Center of Water Conservancy and marine traffic engineering, Henan Province, Water Environment Governance and Ecological Restoration Academician Workstation of Henan Province, Program for Science & Technology Innovation Talents in Universities of Henan Province (No.15HASTIT049). Our gratitude is also extended to reviewers for their efforts in reviewing the manuscript and their very encouraging, insightful and constructive comments.

References


