Article

Performance Comparison of Machine Learning Models for Annual Precipitation Prediction Using Different Decomposition Methods

Chao Song [1,2,3] **and Xiaohong Chen** [1,2,3,*]

[1] Center for Water Resources and Environment, Sun Yat-sen University, Guangzhou 510275, China, songch26@mail2.sysu.edu.cn

[2] Guangdong Engineering Technology Research Center of Water Security Regulation and Control for Southern China, Sun Yat-sen University, Guangzhou 510275, China

[3] Key Laboratory of Water Cycle and Water Security in Southern China of Guangdong High Education Institute, Sun Yat-sen University, Guangzhou 510275, China

* Correspondence: author: eescxh@mail.sysu.edu.cn

Abstract: It has become increasingly difficult in recent years to predict precipitation scientifically and accurately due to the dual effects of human activities and climatic conditions. This paper focuses on four aspects to improve precipitation prediction accuracy. Five decomposition methods (time-varying filter-based empirical mode decomposition (TVF-EMD), robust empirical mode decomposition (REMD), complementary ensemble empirical mode decomposition (CEEMD), wavelet transform (WT), and extreme-point symmetric mode decomposition (ESMD) combined with the Elman neural network (ENN)) are used to construct five prediction models, i.e., TVF-EMD-ENN, REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD-ENN. The variance contribution rate (VCR) and Pearson correlation coefficient (PCC) are utilized to compare the performances of the five decomposition methods. The wavelet transform coherence (WTC) is used to determine the reason for the poor prediction performance of machine learning algorithms in individual years and the relationship with climate indicators. A secondary decomposition of the TVF-EMD is used to improve the prediction accuracy of the models. The proposed methods are used to predict the annual precipitation in Guangzhou. The subcomponents obtained from the TVF-EMD are the most stable among the four decomposition methods, and the North Atlantic Oscillation (NAO) index, the Nino 3.4 index, and sunspots have a smaller influence on the first subcomponent (Sc-1) than the other subcomponents. The TVF-EMD-ENN model has the best prediction performance and outperforms traditional machine learning models. The secondary decomposition of the Sc-1 of the TVF-EMD model significantly improves the prediction accuracy.

Keywords: decomposition methods; Elman neural network; difference; precipitation prediction; Guangzhou

1. Introduction

Annual precipitation is significantly influenced by human activities and natural factors, and annual precipitation time-series data show non-stationary characteristics. The accurate prediction of precipitation in a changing environment is a trending topic for hydrologists.

In recent years, linear time-series models have been suggested and widely used for precipitation forecasting. Examples include linear regression models [1], autoregressive integrated moving average (ARIMA) models [2], and multiple linear regression models, which are widely applied because they accurately describe the relationship between multiple influential variables and rainfall [3]. However, linear time-series models require a

close linear relationship between the independent and dependent variables. The precipitation sequence in Guangzhou is influenced by both human activities and natural conditions. Heavy rains are frequent, and the precipitation sequence is variable.

Artificial neural networks (ANNs) are well-suited for analyzing hydrological variables and are widely used for precipitation prediction [4-6]. However, ANNs may not provide good simulation results for some non-stationary and extremely volatile time series data, which may require preprocessing.

Recently, the concept of coupling different models has attracted attention in hydrologic forecasting. These models can be broadly categorized into ensemble models and hybrid models [7]. The concept of ensemble models consists in establishing several different or similar models for the same process and then combining them [8-10]. For example, Kim, et al. [8] integrated five ensemble methods to improve the performance of streamflow prediction. Maryam, et al. [9] integrated wavelet, seasonal autoregressive integrated moving average (SARIMA), and hybrid ANN methods to predict monthly precipitation. Chau and Wu [10] integrated an ANN and support vector regression (SVR) to predict daily precipitation.

With the development of hybrid models, the combination of signal decomposition and machine learning models has been successfully applied to hydrological prediction [11-12]. For example, Tan, et al. [11] used an adaptive EEMD-ANN model to forecast monthly runoff. Chen, et al. [12] combined nonparametric discrete wavelet transform (DWT), EEMD, and parameter weighted least squares (WLS) estimation methods to predict daily precipitation. Complete ensemble empirical mode decomposition (CEEMD), an improvement of the EMD and EEMD, has been used for hydrological prediction [13-14]. For example, Mumtaz, et al. [13] combined CEEMD, random forest (RF), and kernel ridge regression (KRR) to predict monthly precipitation. Ji, et al. [14] integrated CEEMD, least-squares support vector machine, and nearest neighbor bootstrap regression (NNBR) to predict reservoir inflow runoff. Wavelet transforms and neural networks were combined to predict monthly precipitation [15] and daily precipitation [16]. Extreme-point symmetric mode decomposition (ESMD) is a method to deal with non-stationary signals; Qin, et al. [17] used it to predict runoff.

The cited literature indicates a lack of research on the following aspects. First, several researchers used ESMD, CEEMD, and WT for hydrological prediction. However, time-varying filter-based empirical mode decomposition (TVF-EMD) is a method for processing nonlinear signals that has emerged in recent years. To date, few scholars applied it for precipitation prediction, and there is a lack of systematic comparisons of the differences of different decomposition methods regarding the physical mechanisms related to the model. Second, machine learning is not well suited for explaining the physical mechanisms. Few scholars have analyzed the poor simulation performance of machine learning algorithms, which has been attributed to the fact that machine learning algorithms do not provide good results for simulating individual points and are susceptible to noise. In this paper, we use TVF-EMD for precipitation prediction and compare its performance with four other decomposition methods using the variance contribution rate (VCR) and Pearson correlation coefficient (PCC). Wavelet transform coherence (WTC) is used to determine whether the poor simulation performance of the first subcomponent of precipitation (Sc-1) is related to climate indicators, such as the North Atlantic Oscillation (NAO) index, sunspots, and the Nino 3.4 index. Machine learning algorithms often have low simulation performance for the Sc-1. Because of the small proportion of the Sc-1 in the original signal, it has a negligible impact on the overall prediction error of the original sequence. Few scholars have investigated the impact of the Sc-1 on the time-series sequence, especially the effect of the secondary decomposition of the Sc-1 on the overall prediction performance of the model.

The remainder of this paper is organized as follows. The principle of the TVF-EMD, robust empirical mode decomposition (REMD), CEEMD, wavelet transform, ESMD, and Elman neural network (ENN) methods are introduced in Section 2. The five prediction

models (TVF-EMD-ENN, REMD-ENN, CEEMD-ENN, WT-ENN, ESMD-ENN) are described in Section 3. Section 4 provides the discussion, detailing the performances of the models and the advantages of the proposed models over traditional machine learning models. Finally, the conclusions are drawn in Section 5.

2. Method

2.1. Time-Varying Filter-Based EMD

TVF-EMD is a recent method proposed by Li et al [18]. Compared with other decomposition methods, the TVF-EMD has the following advantages: (1) it addresses the separation and intermittence problems [19]. (2) It uses a time-varying filter in the shifting process to address the mode mixing problem and maintains the time-varying features, unlike many other methods, such as EEMD and CEEEMD [19]. (3) The stopping criterion is improved, resulting in a robust performance for low sampling rates.

In the TVF-EMD, the individual components are replaced by local narrow-band signals to improve the performance of EMD. The signals have properties that are similar to those of the intrinsic mode functions (IMFs) but provide a suitable Hilbert spectrum. The implementation steps of the TVF-EMD are shown in Figure 1. More details on the TVF-EMD method can be found in Li et al [18].
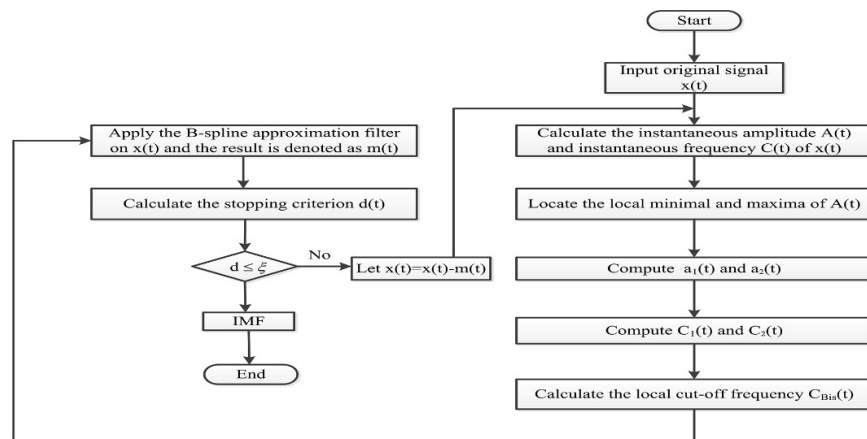

Figure 1. Implementation steps of the time-varying filter-based empirical mode decomposition (TVF-EMD) [20].

2.2. REMD Decomposition Method

Unlike the EMD method, the REMD has strong robustness to noise and outliers; it relies on bilateral filtering to reduce noise.

The steps of the REMD algorithm are as follows [21]:

(1) Find the local maximum and minimum points of the sequence.

(2) Smooth the estimated envelope using Eq. (1) to compute the weights of the points:

$$\omega_{\{m,n\}}[k,j] \propto \exp(-\frac{\Delta I[m,n;k,j]^2}{2\sigma_r^2} - \frac{m^2+n^2}{2\sigma_s^2}) \tag{1}$$

where the function $\Delta I[m,n;k,j]$ represents the difference in intensity between the pixel pairs.

(3) Create the maximum envelope, $E\max$, and the minimum envelope, $E\min$, by solving Eq. (2):

$$E^* = \left[P_e^T P_e + \lambda (D-W)^T (D-W) \right]^{-1} P_e^T P_e I \tag{2}$$

(4) Calculate the average envelope:

$$\overline{E_{(i)}} = (E_{max} + E_{min})/2 \tag{3}$$

(5) Extract the IMF function:

$$h_i^c = I_{(i)} - \overline{E_{(i)}} \tag{4}$$

(6) Repeat the previous steps (2) ~ (5) until the IMF's requirements are met.

(7) Use Eq. (5) to determine if there is a mode-mixing problem of h_i; if so, use Eq. (6) to construct the pseudo-signal M_{cos}. Then, update the obtained IMF as follows:

$$h_i = (h_i^+ + h_i^-)/2 \tag{5}$$

(8) Calculate the residual $R_{es} = I_{(i)} - h_i$; if it does not contain any additional points, stop the EMD process; otherwise, update $i = i + 1$ and $I_{(i)} = R_{es}$ and return to step (1).

(9) Output the final IMF components $\{h_i\}$ and the residual R_{es}.

After the previous nine steps, the precipitation sequence Y is expressed as

$$Y = \sum h_i + R_{es}.$$

More information on REMD can be found in Chen et al [21].

2.3. Complementary Ensemble Empirical Mode Decomposition

The CEEMD is an improvement of EMD and has the following advantages: the residual of the added white noise can be extracted from the noisy data using pairs of complementary ensemble IMFs with added positive and negative white noise; this approach eliminates the residual noise in the IMFs [22].

The steps of the CEEMD algorithm are as follows:

(1) Add white noise in pairs (i.e., positive and negative noise) to the original data:

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} S \\ N \end{bmatrix} \tag{6}$$

where S represents the original data; N is the added white noise; C_1 is the sum of the original data and the positive noise, and C_2 is the sum of the original data and the negative noise.

(2) Obtain the ensemble IMFs from the data with positive noise contributing to a set of IMFs with the residuals of the added positive white noise.

(3) Obtain the ensemble IMFs from the data with negative noise contributing to another set of ensemble IMFs with the residuals of the added negative white noise.

(4) Except for steps (1)–(3), the EMD is used in this method. Therefore, the original signal x(t) can be reconstructed as follows:

$$S(t) = \sum_{i=1}^{n} C_i(t) + r_n(t) \tag{7}$$

where $C_i(t)$ is the ith IMF and $r_n(t)$ is the nth residual.

For more information, please see Huang et al. [23] for the EMD method and Yeh et al. [22] for the CEEMD method.

2.4. Wavelet Transform

Wavelet transform (WT) (figure 2) is a multi-resolution analysis method in the time and frequency domains. The wavelet transform decomposes a time-series signal into different resolutions by performing scaling and shifting. It provides good localization in the time and frequency domains [24].

The continuous wavelet transform (CWT) of signal $x(t)$ is defined as [25]:

$$CWT_x^{\psi}(\tau,s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} x(t)\psi *(\frac{t-\tau}{s})dt \tag{8}$$

Where s is the scale parameter, τ is the translation parameter, $*$ is the complex conjugate, and $\psi(t)$ represents the mother wavelet. The DWT requires less computation time and is simpler to implement than the CWT. The DWT is defined as follows:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{|s_0^j|}} \psi(\frac{t-k\tau_0 s_0^j}{s_0^j}) \tag{9}$$

where j and k are integers that control the wavelet dilation and translation, respectively. s_0 represents a fixed dilation step and τ_0 represents the location parameter. Typical values for these parameters are $s_0 = 2$ and $\tau_0 = 2$ [26].

For a discrete time series, the dyadic wavelet transform is:

$$T_{m,n} = 2^{-j/2} \sum_{i=0}^{N-1} \psi(2^{-j}i-k)s_i \tag{10}$$

where $T_{m,n}$ is the wavelet coefficient for the discrete wavelet with a scale $a = 2^j$ and location $b = 2^j k$. The inverse discrete transform of s_i is given by:

$$s_i = \overline{T} + \sum_{m=1}^{M} \sum_{n=0}^{2^{M-j}-1} T_{m,n} 2^{-j/2} \psi(2^{-j}i-k) \tag{11}$$

The simplified form of Eq. (11) is:

$$s_i = \overline{T}(t) + \sum_{m=1}^{M} W_m(t) \tag{12}$$

where $\overline{T}(t)$ is called the approximation sub-signal at level M, and $W_m(t)$ represents the detailed sub-signals at levels $m = 1, 2, ..., M$.

The Daubechies wavelets exhibit a good trade-off between parsimony and information richness [27]. This paper uses the db5 wavelet to decompose the precipitation data.
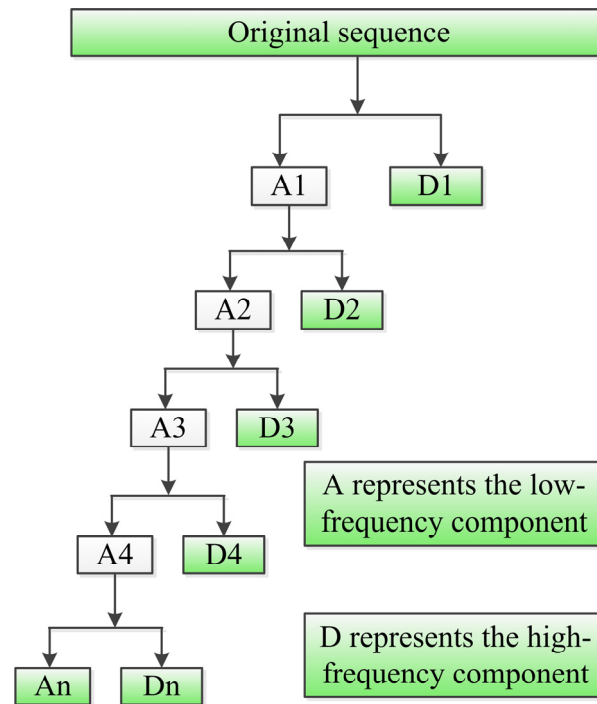
Figure 2. The technical route of WT.

2.5. Extreme-Point Symmetric Mode Decomposition

ESMD is an effective method to deal with nonlinear signals. The ESMD algorithm is divided into the following eight steps [28-29]:

(1) Find all local extreme points (maximum points and minimum points) of the sequence Y and number them as E_i with $1 \leq i \leq n$.

(2) Connect all adjacent E_i with line segments and mark the midpoint as B_i ($1 \leq i \leq n-1$).

(3) Add a left boundary midpoint (B_0) and right boundary midpoint (B_n)

(4) Create p interpolation curves $L_1,...,L_p$ ($p \geq 1$) using these $n+1$ midpoints.

(5) Repeat steps (1)-(4), until $\left| L^* \right| \leq \varepsilon$ (ε is the allowable error) reaches the preset maximum number K. This provides the first mode M_1.

(6) Repeat steps (1)-(5) with the residual $Y - M_1$ to obtain $M_2, M_3...$, until the last residual R has no more than a predetermined number of extreme points.

(7) Change K to a finite integer interval and repeat steps (1)-(6). Next, calculate the variance σ^2 of $Y - R$ and plot σ / σ_0 versus K, where σ^2 is the standard deviation of Y.

(8) Find the minimum σ / σ_0 for (K_{min}, K_{max}) that corresponds to K_0. Then, repeat the previous steps (1) - (6) with K_0 and output all modes. The last residual R is the optimal adaptive global mean curve.

After the previous eight steps, the precipitation sequence Y is expressed as $Y = \sum M_i + R$. In other words, after using the ESMD method, the precipitation sequence Y is decomposed into a series of the IMF and a trend.

2.6. Elman Neural Network

The Elman neural network (ENN) (figure 3) [30] is a feedback neural network composed of four layers, including the input layer, hidden layer, context layer, and output layer [31]. The addition of internal feedback of the receiver layer increases the ENN's sensitivity to historical data and its capacity to forecast the time-series data and handle dynamic information; thus achieving dynamic modeling [32].

The ENN is defined as follows:

$$x_{jt}^0(k) = f_h[\sum_{i=1}^m (\omega_{ij}^i x_{it}(k) + b_i) + \sum_{j=1}^n \omega_j^r x_{jt}^0(k-1)] \tag{13}$$

$$y_{t+1}(k) = f_0[\sum_{j=1}^m \omega_j^0 x_{jt}(k) + b_j] \tag{14}$$

where x_{jt} is the input to the network at time t, $x_{jt}^0 (j = 1,2,3,...,n)$ is the output of the hidden layer at time t, y_{t+1} is the output of the network at time $t+1$, ω_{ij}^j is the connection weight of the input layer to the hidden layer, ω_j^r is the connection weight of the context layer to the hidden layer, ω_j^0 is the connection weight of the hidden layer to the output layer. f_h and f_0 are the transfer functions of the hidden layer and the output layer, respectively.

More information on the ENN method is available in Elman [30] and Ardalani-Farsa and Zolfaghari [33].
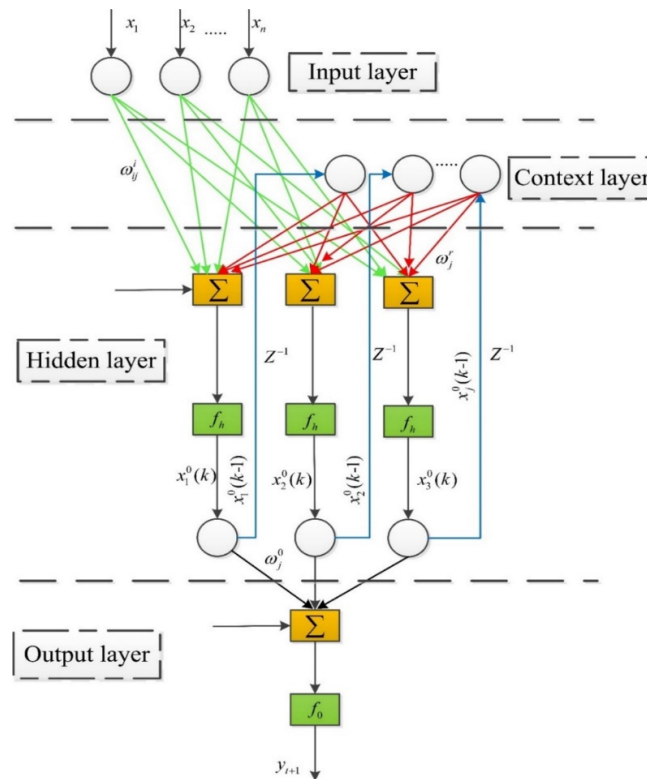

Figure 3. The schematic diagram of the Elman neural network (ENN).

2.7. Performance Evaluation Criteria

Several commonly used evaluation criteria are used to evaluate the prediction performance of the model.

The root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_i')^2}$$

(15)

where x_i represents the true value, x_i' represents the predicted value, and N is the number of predicted values; x_i and x_i' in the following equations have the same meaning.

The mean absolute error (MAE):

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|x_i - x_i'\right|$$

(16)

The mean absolute percentage error (MAPE):

$$MAPE(\%) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - x_i'}{x_i}\right| * 100$$

(17)

The coefficient of determination (R^2):

$$R^2 = 1 - (\sum_{i=1}^{N}(x_i - x_i')^2 / \sum_{i=1}^{N}(x_i - \bar{x})^2)$$

(18)

where $\bar{x}$ represents the average value, and the range of R^2 is [0-1]; the larger the R^2, the better the fitting performance.

3. Empirical Study

3.1. Study Area and Data Description

Located in southern China and the lower reaches of the Pearl River, Guangzhou is the capital city of Guangdong province. It is a first-tier city in China. Guangzhou is located at longitude 112°57″—114°3″E and latitude 22°26″—23°56″N, as shown in Figure 4.
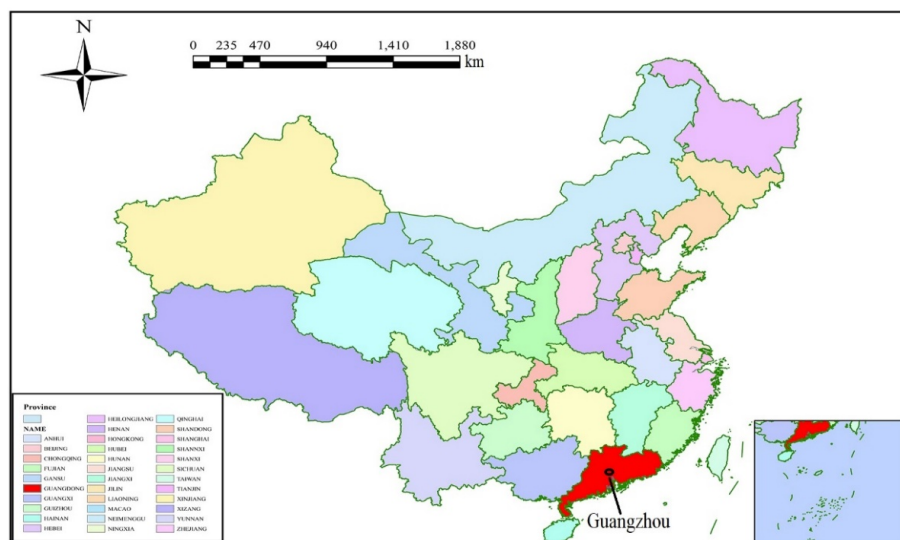

Figure 4. The region of Guangzhou.

The annual precipitation data were obtained from the continuously measured data of the China Meteorological Data Network (http://data.cma.cn/data/cdcdetail/dataCode/SURF_CLI_CHN_MUL_YER_CES.html) and ranged from 1951 to 2019. The model was trained with the data sets from 1951 to 2014 and validated with datasets from 2015 to 2019.
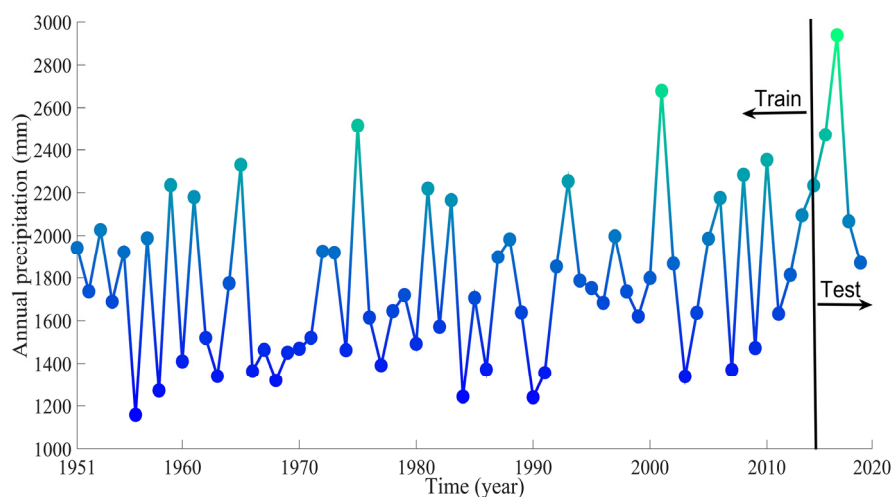
Figure 5. Annual precipitation in Guangzhou.

As shown in Figure 5, the annual precipitation of Guangzhou exhibits an increasing trend and high variability, indicating non-stationarity. It is observed that 1980 was a demarcation point, with a slow increase in precipitation before 1980 and a faster increase after 1980.

3.2. Model Forecasting Results

This section describes the results of the five prediction models (TVF-EMD-ENN, REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD-ENN). Each prediction model decomposed the precipitation sequence into several subcomponents using signal decomposition. The ENN was used to predict these subcomponents. Subsequently, the predicted values of the subcomponents were converted into the predicted precipitation values through summation.

The same parameters were used for the ENN to conduct a fair comparison of the prediction accuracy of the different models. We used the data of the past three days as input to predict the output on the fourth day. The number of neurons in the hidden layer was 3, the number of iterations was 1000, and the training function was traingdx.

3.2.1. TVF-EMD-ENN Forecasting Results

The parameters of the max output of the IMFs, the b-spline order, and the stopping criterion were 6, 26, and 0.1, respectively, for the TVF-EMD. The decomposition results of annual precipitation in Guangzhou are shown in Figure 6.
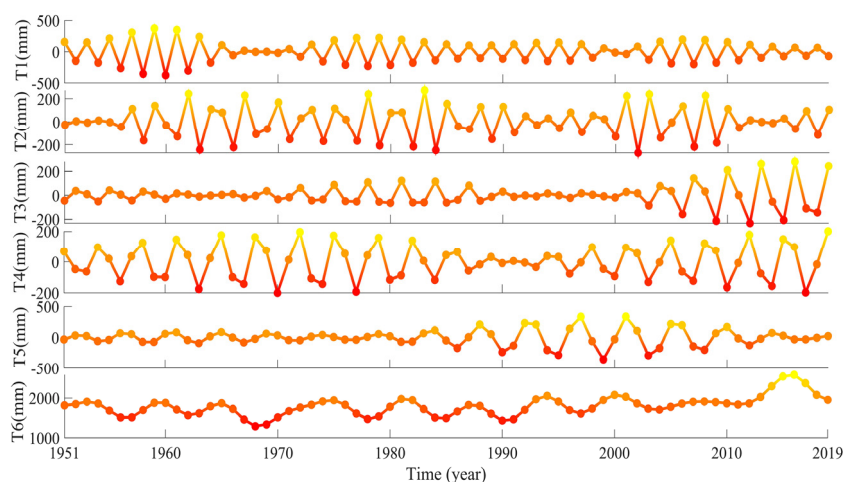
Figure 6. The precipitation decomposition result of TVF-EMD.

Note. Tn represents the nth subcomponent obtained from TVF-EMD.

The annual precipitation sequence was decomposed into six subcomponents; the subsequence shows regular periodic changes. However, with an increase in the number of decomposition layers, the frequency and amplitude of the subsequence decrease, and the non-stationarity of the sub-sequences weakens, providing a suitable data input for the ENN for the prediction.

After the precipitation sequence was decomposed, the six subcomponents obtained from the TVF-EMD decomposition were predicted by the ENN. The absolute error of the prediction is shown in Figure 7.
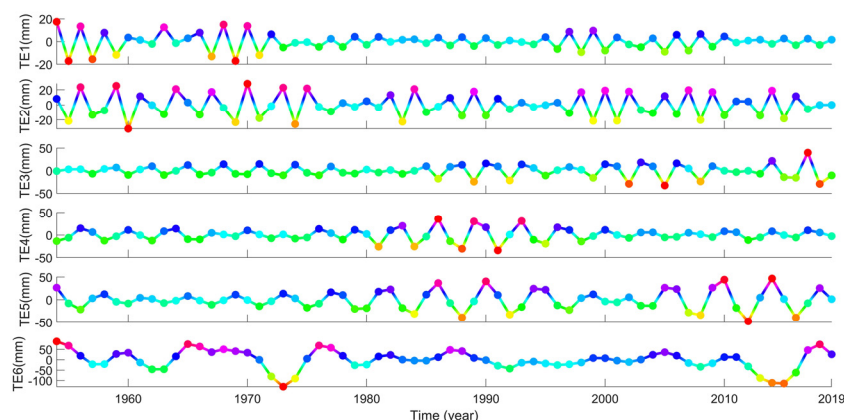

Figure 7. The absolute prediction error of TVF-EMD-ENN.

Note. TEn represents the absolute prediction error of the nth subcomponent obtained from TVF-EMD-ENN.

It is observed in Figure 7 that T1 and T2 had a better prediction performance than the other components, with an absolute error of about 10 mm. However, the prediction result of the test set showed that T6 provided slightly worse results, with an absolute prediction error of around 50 mm, which was the opposite of the results of the training period. T6 showed poor prediction performance on the test set. The effect on the prediction accuracy of the total precipitation will be presented in the discussion section.

3.2.2. REMD-ENN Forecasting Results

We set the maximum number of iterations to 30, the maximum number of IMF to 6, and the end extension rate of the original data to 0.2 in the REMD. The decomposition results are shown in Figure 8.
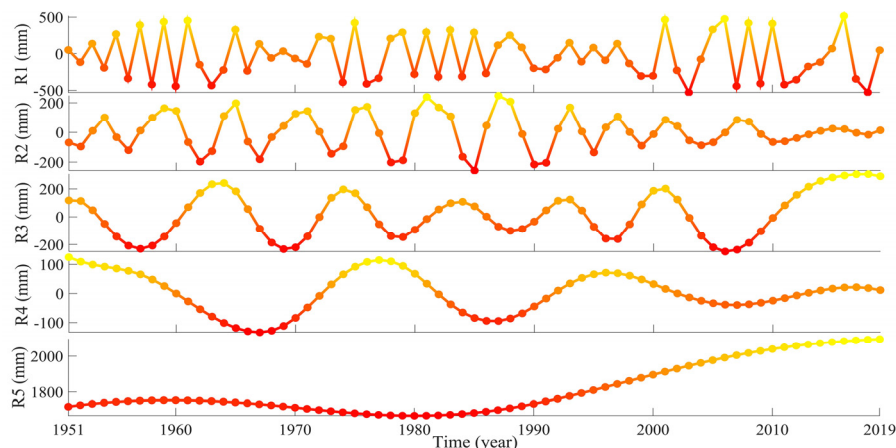

Figure 8. The precipitation decomposition result of robust empirical mode decomposition (REMD).

Note. Rn represents the nth subcomponent obtained from REMD.

The first subcomponent R1 obtained from the REMD decomposition exhibits some non-stationarity and uncertainty, and its decomposition performance is worse than that of the remaining four subcomponents.

Figure 9. The absolute prediction error of REMD-ENN.

Note. REn represents the absolute prediction error of the nth subcomponent obtained from REMD-ENN.

As shown in Figure 9, the absolute prediction errors and fluctuations of R1 and R2 were relatively large, and the prediction performance was inferior to that of the T1 and T2 of the TVF-EMD-ENN model. The last sub-component, R5, provided the best results, with an absolute error of less than 4 mm.

3.2.3. CEEMD-ENN Forecasting Results

We set the standard deviation of the noise to 0.2, the ensemble number to 100, and the decomposition level to 6. The decomposition results of the CEEMD are shown in Figure 10.
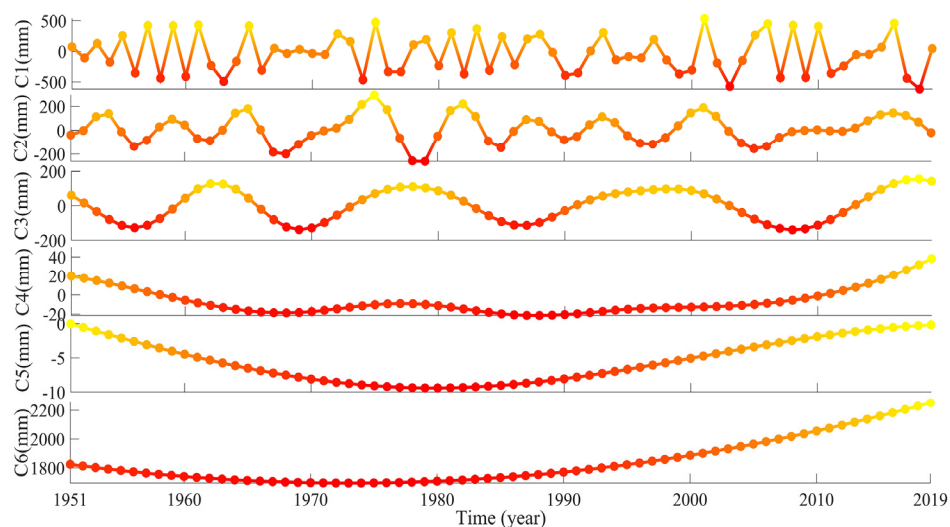
Figure 10. The precipitation decomposition result of ensemble empirical mode decomposition (CEEMD).

Note. Cn represents the nth subcomponent obtained from CEEMD.

The C1 component had relatively large frequency, amplitude, and variability. The C6 component provided the best results, with the smoothest curve and the least amount of non-stationarity.
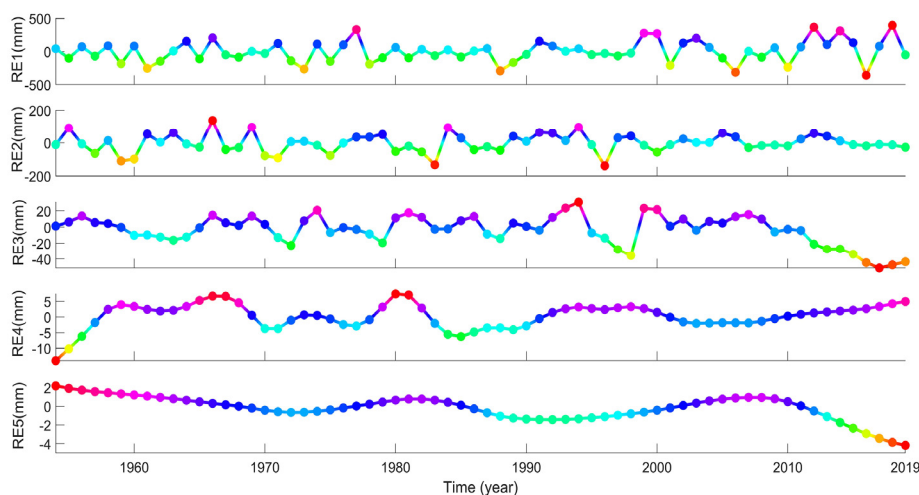

Figure 11. The absolute prediction error of CEEMD-ENN.

Note. CEn represents the absolute prediction error of the nth subcomponent obtained from CEEMD-ENN.

As shown in Figure 11, the prediction performance of the first subcomponent was relatively poor for the test data. The absolute value of the prediction error and the

fluctuation were large. The prediction performance of the remaining components is relatively good.

3.2.4. WT-ENN Forecasting Results

The wavelet basis function was db5, and the number of decomposition layers was 6. The decomposition results of the WT are shown in Figure 12.

Figure 12. The precipitation decomposition result of the WT.

Note. Xn represents the nth subcomponent obtained from the WT.

The WT of the X1 and X2 components was slightly worse than that of the other components, exhibiting some variability.
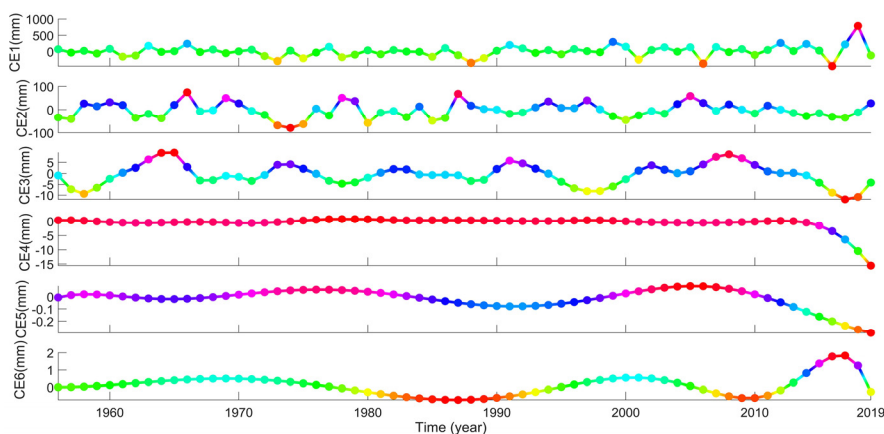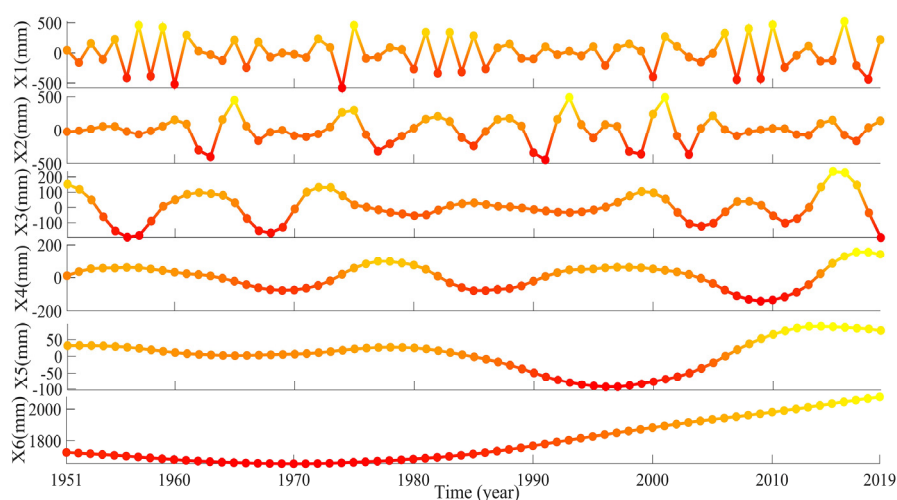

Figure 13. The absolute prediction error of WT-ENN.

Note. XEn represents the absolute prediction error of the nth subcomponent obtained from WT-ENN.

As shown in Figure 13, the first and second subcomponents of the WT showed a poor prediction performance in the test set, as well as a large absolute error and volatility. The prediction performance of the first and second subcomponents in the training set was poor, around 1973, which was similar to the results of the REMD-ENN model.

3.2.5. ESMD-ENN Forecasting Results

We set the maximum number of iterations to 40, the minimum number of the remaining mode extremum points to 4, and the best screening number to 34. The decomposition results of the ESMD are shown in Figure 14.

Figure 14. The precipitation decomposition result of ESMD.

Note. En represents the nth subcomponent obtained from ESMD.

The E1 component performed slightly worse than the other component, exhibiting some randomness and volatility. The prediction performance of the ESMD-ENN model is shown in Figure 15.
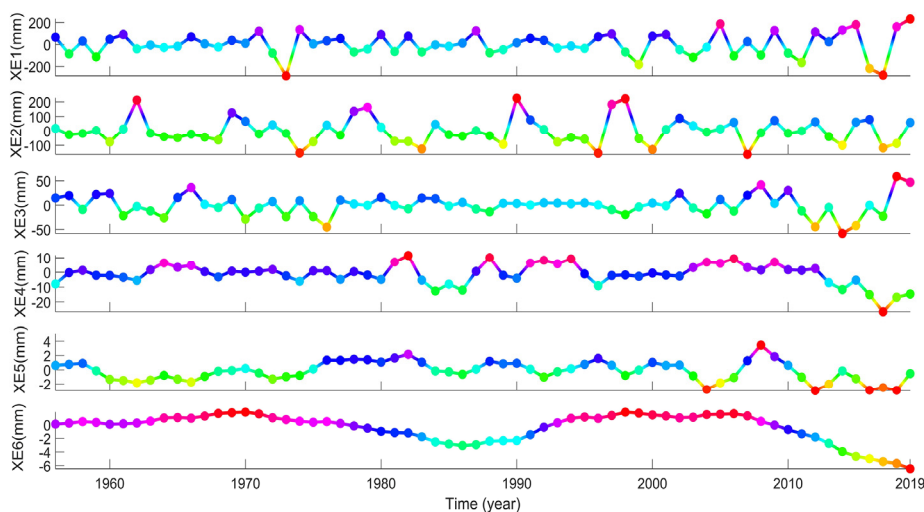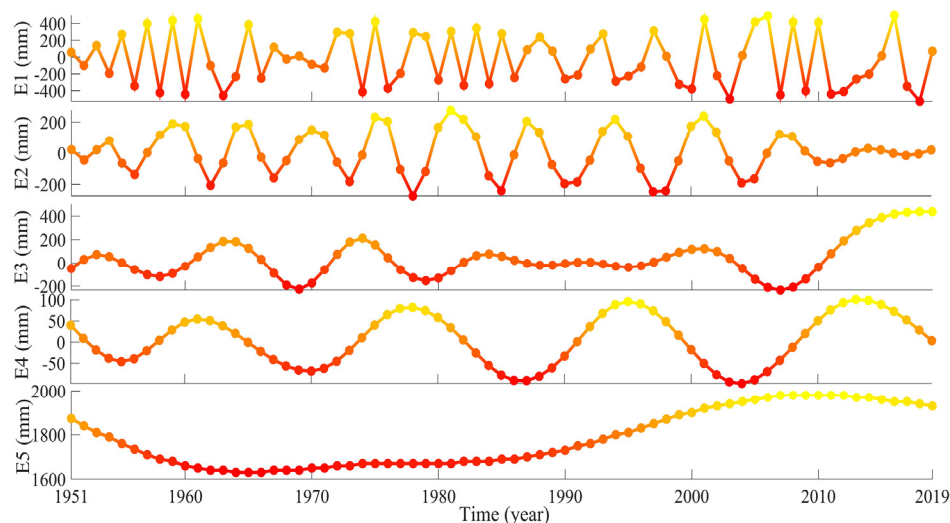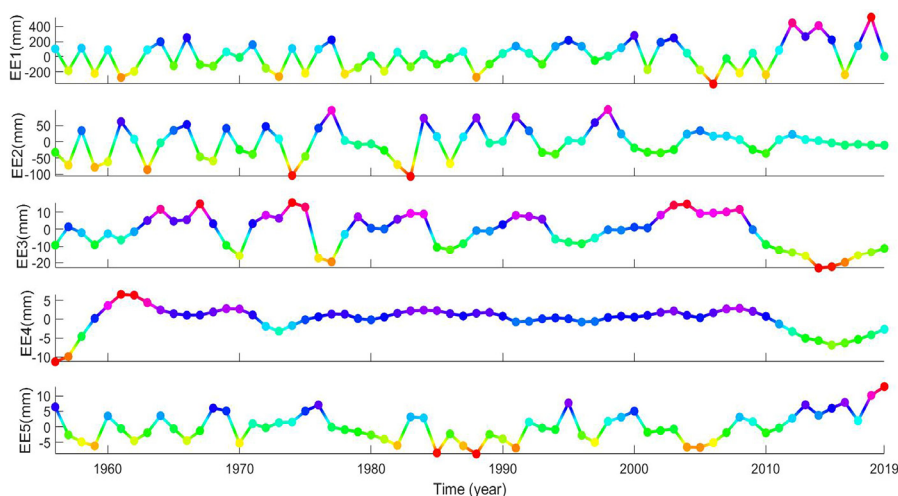

Figure 15. The absolute prediction error of ESMD-ENN.

Note. EEn represents the absolute prediction error of the nth subcomponent obtained by ESMD-ENN.

E1 had a poor predictive performance with a large absolute error and considerable instability. E1 also performed poorly in the test set, with large absolute errors. In addition, around 1975, the prediction performance of E2 and E3 of the training set was poor, with strong variability.

The results of the five models indicated that the TVF-EMD model has the best prediction performance for the first subcomponent, whereas the REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD-ENN models had poor prediction performance for the first subcomponent. Although some of the models did not perform well in the prediction of individual subcomponents, the subcomponents account for a small proportion of the precipitation and have a negligible influence on the overall prediction error. Thus, we do not discuss their influence here. Since the first subcomponent obtained from the signal decomposition is the one with the largest amount of non-stationarity, we discuss the influence of the first sub-component on the overall precipitation prediction error below.

4. Results and Discussion

The results in Section 3.2.2~3.2.5 indicated that the REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD-ENN models showed poor simulation performance for predicting the Sc-1. However, it is unclear if climate indicators are the reason for poor prediction performance. Thus, a quantitative analysis of the differences between the five decomposition methods is required to determine if the subcomponents with poor prediction performance influence the precipitation prediction results. A comparison of the proposed methods with traditional machine learning models was performed to assess the advantages and disadvantages of the methods.

4.1. Reason for the Poor Prediction Performance of Sc-1

It is well known that the simulation results of machine learning algorithms have uncertainties. Few studies investigated whether the poor prediction performances of the subcomponents of precipitation obtained by different decomposition methods are related to climate indicators.

We found that Sc-1 provided unsatisfactory results in the test set. However, it is unknown whether these anomalies are related to sunspots or the El Nino Southern Oscillation. Here, we use Wavelet transform coherence (WTC) to describe the degree of correlation between the precipitation subcomponents and the climate series to ascertain whether the poor simulation performance of Sc-1 at some points is related to the climate indicators.

The WTC of two time series X and Y is defined as:

$$R_n^2(s) = \frac{\left|S(s^{-1}W_n^{XY}(s))\right|^2}{S(\left|s^{-1}W_n^X(s)\right|^2)*S(\left|s^{-1}W_n^Y(s)\right|^2)} \qquad (19)$$

where S is a smoothing operator that depends on the wavelet type, and

$$W_n^{XY}(s) = W_n^X(s).W_n^Y*(s) \qquad (20)$$

where $*$ is the complex conjugate, $R_n^2(s)$ is between 0 and 1; 0 indicates no correlation between the sequences, and 1 indicates a strong correlation between the sequences.

The confidence levels of the red-noise background were assessed. Monte Carlo methods were adopted to estimate the statistical significance of the wavelet coherence [34]. The significance level at each scale was calculated solely from values outside the cone of influence (COI). A detailed description of the calculation of WTC is provided by Grinsted et al [35].

The NAO index and the Nino 3.4 index were obtained from https://www.psl.noaa.gov/data/climateindices/; both are monthly averages. The sunspot

data representing the annual mean sunspot number were obtained from http://www.sidc.be/silso/infosnytot.

The results of the WTC analysis of the four decomposition methods (REMD, CEEMD, Wavelet, ESMD) to determine the relationships between the SC-1 and the NAO index, the Nino 3.4 index, and the sunspots are shown in Figure 16.
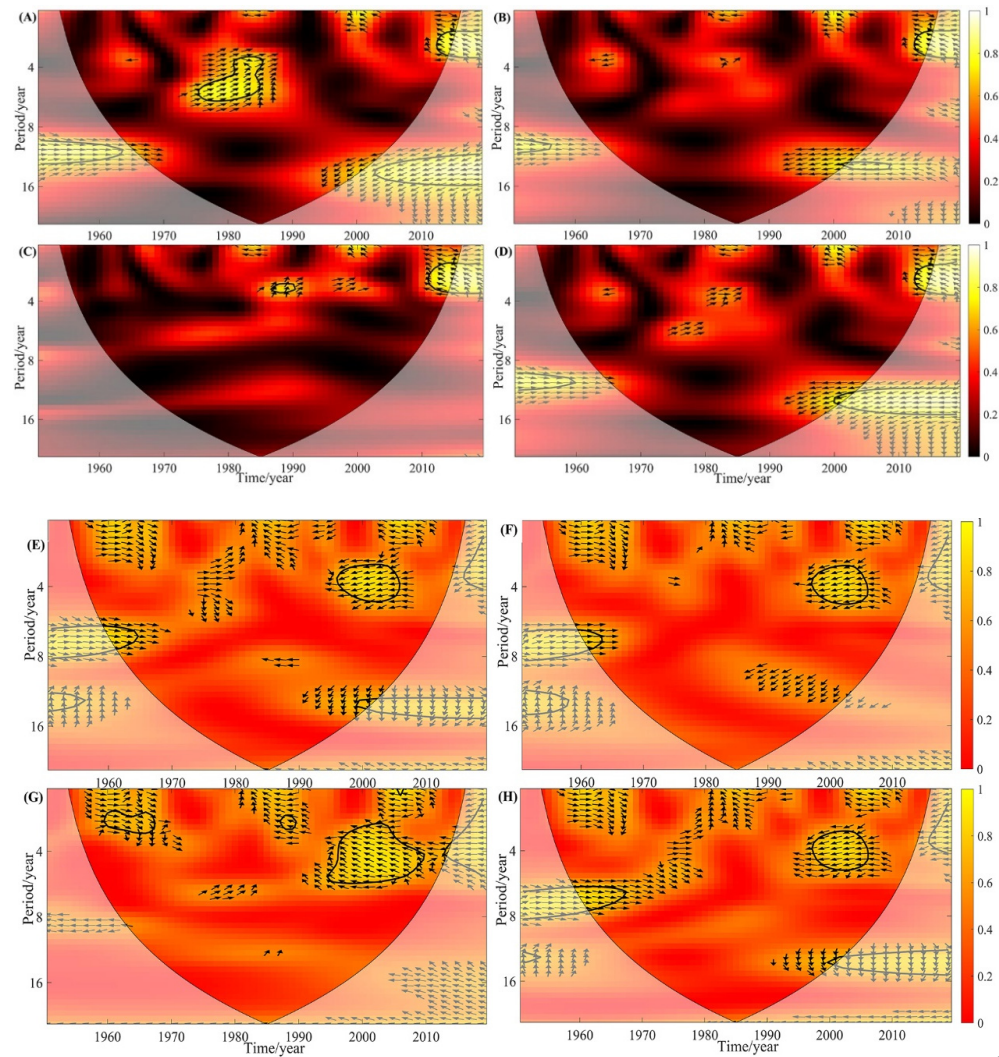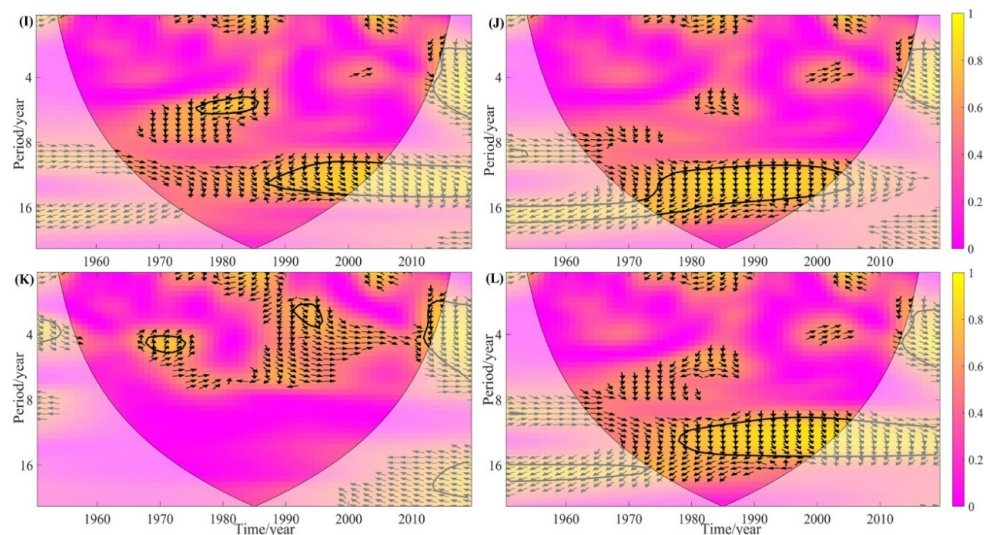
Figure 16. The wavelet transform coherence (WTC) between Sc-1 and the sunspot data (A-D), North Atlantic oscillation (NAO) (E-H), and the Nino 3.4 index (I-L).

Note. (A) represents the relationship between Sc-1 of REMD and the sunspots, (B) represents the relationship between SC-1 of CEEMD and the sunspots, (C) represents the relationship between Sc-1 of the WT and the sunspots, and (D) represents the relationship between Sc-1 of ESMD and the sunspots. (E)-(H) and (I)-(L) reflect the corresponding relationships between the Sc-1 of the same models with the NAO index and the Nino 3.4 index, respectively. The period is measured in years. Thick contours denote the 5% significance levels. The pale regions denote the cone of influence (COI), where edge effects might distort the results. The colors denote the strength of the wavelet transform.

The correlation between Sc-1 and the sunspots was high for the REMD, CEEMD, and ESMD methods and low for the WT method, which was reflected in the period of 11 years. However, there were similarities between the four methods. In the period of 2-4 years, the Sc-1 and the sunspots showed a strong significant correlation in 2010-2015, with a correlation coefficient of 0.9. In addition, due to the boundary effect, if the length of the precipitation and sunspot sequence was relatively long, the Sc-1 and sunspots showed a strong correlation in the period of 2-4 years after 2015. Therefore, we can infer that the poor simulation results of the Sc-1 in the test sets of the REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD-ENN may be related to sunspots.

The correlation between Sc-1 and the NAO was high for the REMD, CEEMD, and ESMD methods and low for the WT method, which was reflected in the period of 6–8 years from 1960 to 1965, and there were no arrows before 1960 to 1965. For the REMD, CEEMD, and ESMD methods, there was a significant positive correlation between Sc-1 and the NAO index in the period of 6–8 years from 1960 to 1965, with a correlation coefficient of 0.9. The prediction result of Sc-1 for the REMD-ENN, CEEMD-ENN, and ESMD-ENN models in 1960–1964 was unsatisfactory. It can be inferred that this poor performance was strongly related to the NAO. In addition, in the period of 3–6 years, the Sc-1 and the NAO index showed a strong correlation between 1995 and 2005, with a correlation coefficient of 0.9. The Sc-1 of the REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD-ENN models had poor prediction performance from 1995 to 2005, which may be related to the NAO.

The correlation between Sc-1 and the Nino 3.4 was high for the REMD, CEEMD, and ESMD methods and low for the WT method, which was reflected in the period of 8–16 years. There were few arrows in the relationship between Sc-1 and the Nino 3.4 index between 1951 and 2019. However, the Sc-1 of the WT-ENN showed poor simulation performance in 1973, in the period of 4–5 years, the Sc-1 and the Nino 3.4 index had a strong

correlation between 1969 and 1974 (correlation coefficient of 0.9). It can be inferred that the poor performance of the Sc-1 of the WT-ENN model in 1973 was related to the Nino 3.4 index. For the REMD, CEEMD, and ESMD methods, in the period of 10–16 years, the Sc-1 and Nino 3.4 had a strong correlation between 1990 and 2005 (correlation coefficient of 0.9). The Sc-1 of the REMD-ENN, CEEMD-ENN, and ESMD-ENN models in 1990–2005 showed poor prediction performance, which was likely related to the Nino 3.4 index. In addition, due to the boundary effect, if the length of precipitation and Nino 3.4 was longer, the Sc-1 and Nino 3.4 index showed a strong correlation in the period of 3–7 years after 2015. Therefore, we can infer that the poor simulation performance of the Sc-1 in the test set of the REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD-ENN models was likely related to the Nino 3.4 index.

The relationship between the Sc-1 of the TVF-EMD and the sunspots, NAO index, and Nino 3.4 index is shown in Figure 17.
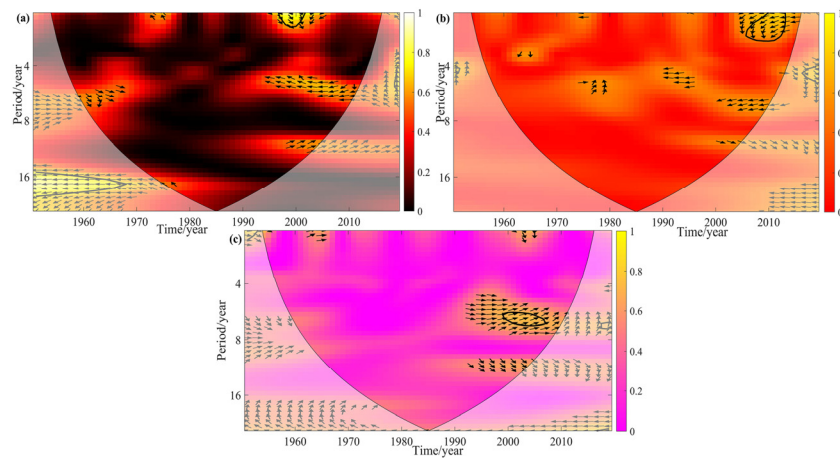
Figure 17. The WTC between Sc-1 of TVF-EMD and the sunspots, NAO index, and Nino 3.4 index.

Note. (a) represents the relationship between the SC-1 and the sunspots, (b) represents the relationship between the Sc-1 and the NAO index, and (c) represents the relationship between the Sc-1 and the Nino 3.4 index.

The results show a relatively low correlation between the Sc-1 of TVF-EMD and the sunspots, NAO index, and the Nino 3.4 index. The Sc-1 decomposed by the TVF-EMD is more stable and less affected by the sunspot, NAO index, and Nino 3.4 index than those of the other methods. Therefore, the Sc-1 prediction performance of TVF-EMD provides the best prediction performance and most stable results among the methods.

4.2. The Difference between the Five Decomposition Methods

The VCR was used to determine the proportion of the subcomponents in the original signal. The VCR is the percentage of the sum of the variance of each IMF component and the variance of each IMF component obtained by decomposition. The variance contribution of the IMF component reflects the contribution of the decomposed signal of this frequency to the entire sequence. The greater the VCR, the greater the contribution of the signal to the sequence. The variance is defined as follows:

$$\sigma_x^2 = E^2(x) - (E(x))^2 \tag{21}$$

where $E(x)$ is the expected value of x.

The PCC was used to determine the correlation between the subcomponent sequences obtained from each decomposition method and the precipitation sequence. The larger the PCC, the stronger the correlation between the sequences. The PCC is defined as follows:

$$r = (\sum XY - \frac{\sum X \sum Y}{N}) / \sqrt{(\sum X^2 - \frac{(\sum X)^2}{N})(\sum Y^2 - \frac{(\sum Y)^2}{N})} \qquad (22)$$

where N is the length of the time sequence, X represents the subcomponent sequence obtained from the decomposition method, and Y represents the original precipitation sequence.

The VCR and PCC values of the subcomponents obtained from the five decomposition methods are shown in Figure 18.

VCR values (%)					
CEEEMD	67.3	9.5	5.7	0.1	0	17.4
ESMD	60	11.9	15.9	2.2	10	■
REMD	58	9.2	16.6	3.1	13.2	■
TVF-EMD	20	13	6	8.1	12	40.9
Wavelet	48.2	27	6.3	3.6	1.8	13.1
	Sc-1	Sc-2	Sc-3	Sc-4	Sc-5	Sc-6

PCC values					
CEEEMD	0.75	0.44	0.27	0.38	0.34	0.44
ESMD	0.69	0.3	0.42	0.25	0.37	■
REMD	0.71	0.33	0.39	0.17	0.4	■
TVF-EMD	0.44	0.36	0.25	0.29	0.35	0.65
Wavelet	0.67	0.51	0.3	0.23	0.22	0.41
	Sc-1	Sc-2	Sc-3	Sc-4	Sc-5	Sc-6

Figure 18. The variance contribution rate (VCR) and Pearson correlation coefficient (PCC) values of the five decomposition methods.

Note. The black area indicates that this item does not exist

The VCR values of the Sc-1 of the REMD, CEEMD, Wavelet, and ESMD were relatively large. The average value was 58%, indicating that the Sc-1 of the four methods contributed considerably to the precipitation sequence. However, the VCR of the Sc-1 of the TVF-EMD was relatively small (20%), demonstrating that the Sc-1 of the TVF-EMD contributed relatively little to the precipitation sequence. In addition, the VCR value of the last subcomponent of TVF-EMD was the largest among the four methods, indicating that the last subcomponent of the TVF-EMD contributed most to the precipitation sequence. The non-stationarity of the subcomponent weakened with an increase in the number of signal decomposition layers, and the prediction performance of the subcomponents of the ENN improved. Since the Sc-1 of CEEMD, ESMD, REMD, and WT contributed more to the precipitation sequence, the prediction performance was poor. In contrast, the Sc-1 of TVF-EMD contributed little to the precipitation sequence. Thus, the prediction performance of the TVF-EMD-ENN model was better than that of the REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD models. A comparison of the model prediction results is described in Section 4.3.

As shown in Figure 18, the correlation between the Sc-1 obtained from the TVF-EMD and precipitation was the lowest among the models, with a PCC value of 0.44. The PCC values of the other four methods were above 0.67, indicating a moderate correlation between the Sc-1 and precipitation. In contrast, the other components had a weak correlation with precipitation. If the Sc-1 has a poor prediction performance, it influences the prediction performance of precipitation, even if the effect is not significant. Figure 19 illustrates

the interaction between the Sc-1 and the precipitation prediction error. If the Sc-1 prediction error is large, the relative error of precipitation prediction is also large.

4.3. Advantages of the Proposed Models over Traditional Machine Learning Models

The prediction models established in this paper are based on decomposition and reconstruction. First, the precipitation series was decomposed into several subcomponent sequences. Then, the ENN model was used to predict these components, and the precipitation prediction was obtained by summing the prediction values of the subcomponents. As the number of signal decomposition layers increased, the non-stationarity and volatility of the precipitation subcomponents weakened, improving the performance of the neural network for predicting the subcomponents.

The advantages of the decomposition-reconstruction model are two-fold. First, the prediction error of precipitation is determined by the subcomponents. Even if the prediction error of the individual subcomponent is large, the subcomponent may account for a small proportion in the precipitation sequence, and the error does not significantly influence the precipitation prediction performance. Second, the precipitation subcomponents obtained from the decomposition method have positive and negative values, canceling the error. In traditional machine learning models without decomposition, a large prediction error of certain points significantly affects the overall prediction performance. These advantages are reflected in the results in Figure 19, showing the relative prediction errors of the five models.
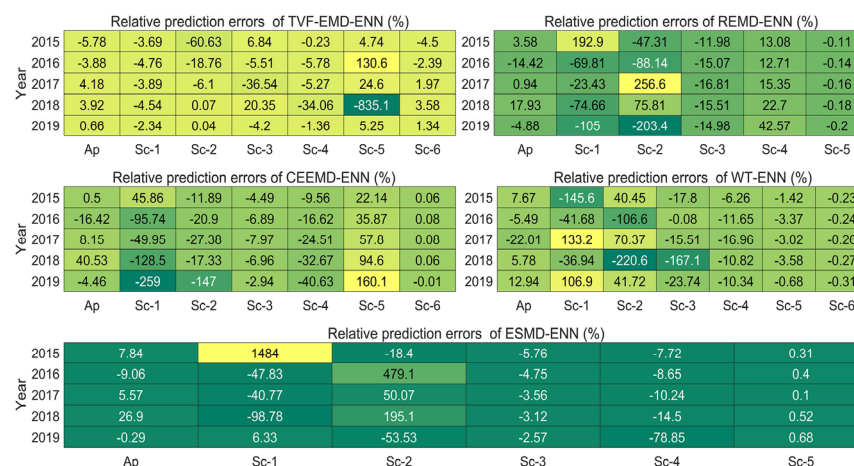
Relative prediction errors of TVF-EMD-ENN (%)

Year	Ap	Sc-1	Sc-2	Sc-3	Sc-4	Sc-5	Sc-6
2015	-5.78	-3.69	-60.63	6.84	-0.23	4.74	-4.5
2016	-3.88	-4.76	-18.76	-5.51	-5.78	130.6	-2.39
2017	4.18	-3.89	-6.1	-36.54	-5.27	24.6	1.97
2018	3.92	-4.54	0.07	20.35	-34.06	-835.1	3.58
2019	0.66	-2.34	0.04	-4.2	-1.36	5.25	1.34

Relative prediction errors of REMD-ENN (%)

Year	Ap	Sc-1	Sc-2	Sc-3	Sc-4	Sc-5
2015	3.58	192.9	-47.31	-11.98	13.08	-0.11
2016	-14.42	-69.81	-88.14	-15.07	12.71	-0.14
2017	0.94	-23.43	256.6	-16.81	15.35	-0.16
2018	17.93	-74.66	75.81	-15.51	22.7	-0.18
2019	-4.88	-105	-203.4	-14.98	42.57	-0.2

Relative prediction errors of CEEMD-ENN (%)

Year	Ap	Sc-1	Sc-2	Sc-3	Sc-4	Sc-5	Sc-6
2015	0.5	45.86	-11.89	-4.49	-9.56	22.14	0.06
2016	-16.42	-95.74	-20.9	-6.89	-16.62	35.87	0.08
2017	0.15	-49.95	-27.30	-7.97	-24.51	57.8	0.00
2018	40.53	-128.5	-17.33	-6.96	-32.67	94.6	0.06
2019	-4.46	-259	-147	-2.94	-40.63	160.1	-0.01

Relative prediction errors of WT-ENN (%)

Year	Ap	Sc-1	Sc-2	Sc-3	Sc-4	Sc-5	Sc-6
2015	7.67	-145.6	40.45	-17.8	-6.26	-1.42	-0.23
2016	-5.49	-41.68	-106.6	-0.08	-11.65	-3.37	-0.24
2017	-22.01	133.2	70.37	-15.51	-16.96	-3.02	-0.26
2018	5.78	-36.94	-220.6	-167.1	-10.82	-3.58	-0.27
2019	12.94	106.9	41.72	-23.74	-10.34	-0.68	-0.31

Relative prediction errors of ESMD-ENN (%)

Year	Ap	Sc-1	Sc-2	Sc-3	Sc-4	Sc-5
2015	7.84	1484	-18.4	-5.76	-7.72	0.31
2016	-9.06	-47.83	479.1	-4.75	-8.65	0.4
2017	5.57	-40.77	50.07	-3.56	-10.24	0.1
2018	26.9	-98.78	195.1	-3.12	-14.5	0.52
2019	-0.29	6.33	-53.53	-2.57	-78.85	0.68

Figure 19. The relative prediction errors of the five models.

Note. Sc-n represents the nth subcomponent obtained from the decomposition method. Ap=Annual precipitation

Except for the TVF-EMD, the relative prediction errors of the Sc-1 of the other four methods were relatively large. The performance of the ESMD-ENN model was the worst, with the highest relative errors of Sc-1 in 2015 and 2018 (1484% and –98.78%, respectively). The prediction performance of Sc-2 was slightly better than that of Sc-1. The prediction performance of the REMD-ENN model was unsatisfactory, especially in 2017–2019. The non-stationarity of the subcomponents weakened with an increase in the number of signal decomposition layers. The prediction performance of the subcomponents increased from Sc-3 to Sc-6, with small relative prediction errors.

Although the prediction performances of Sc-1 and Sc-2 were unsatisfactory, the precipitation prediction is determined by several subcomponents, with different proportions of subcomponents, as shown in the precipitation decomposition results of the five

decomposition methods and Section 4.2. In general, the relative prediction error of the five forecasting models was within the acceptable range.

Comparisons were made with traditional prediction models, including the radial basis function (RBF) neural network, the long short-term memory (LSTM) neural network, and the ENN. The TVF-EMD, REMD, ESMD, and CEEMD are based on the EMD. We also assessed the prediction performance of the EMD-ENN and EEMD-ENN models. The comparison of the prediction performance of the ten models is shown in Table 1.

Table 1. The comparison of the prediction performance of the ten models.

Prediction	Model									
Indices	CENN	EEENN	EENN	ENN	ESENN	LSTM	RBF	RENN	TEENN	WTENN
MAE(mm)	306.26	519.9	1076	1012	217.12	546.5	893.9	197.42	86.58	246.5
MAPE(%)	14.01	22.61	43.44	41.96	9.93	21.69	36.31	8.35	3.68	10.78
R^2	0.7	0.7	0.06	0.6	0.64	0.16	0.06	0.7	0.93	0.82
RMSE(mm)	411.89	581.6	1276	1120	273.88	631.4	967.1	250.85	96.48	276.42

Note. CENN=CEEMD-ENN; EEENN=EEMD-ENN; EENN=EMD-ENN; ENN=Elman neural network; ESENN=ESMD-ENN; LSTM= long short-term memory neural network; RBF= radial basis function neural network; RENN=REMD-ENN; TEENN=TVF-EMD-ENN; WTENN=WT-ENN.

The TEENN had the best prediction performance, with the optimal RMSE, R^2, MAPE, and MAE values of 96.48 mm, 0.93, 3.68%, and 86.58 mm, respectively. The prediction performances of the CENN, RENN, ESENN, WTENN, EEENN, and TEENN models were better than that of the EENN model. Among the seven decomposition-reconstruction prediction models, the prediction performance of the EENN model was the worst, with RMSE, R^2, MAPE, and MAE values of 1276 mm, 0.06, 43.44%, and 1076 mm, respectively. In addition, we found that the prediction performance of the decomposition-reconstruction model was better than that of the neural network models without decomposition. The prediction performances of the LSTM, ENN, and RBF models were the three lowest among the ten prediction models.

In general, the TEENN model was optimal, and we recommend TVF-EMD decomposition for precipitation prediction when the data are significantly influenced by climate factors and anthropogenic influences.

4.4. Improving the Prediction Performance of the Model

We found that the TVF-EMD provided the best prediction performance, and the other four methods performed poorly for predicting the Sc-1(figure 20). Thus, we used the TVF-EMD model and performed a secondary decomposition (figure 20) of Sc-1 for the other four methods to determine if the prediction performance of the REMD-ENN, CEEMD-ENN, WT-ENN, and ESMD-ENN models could be improved.
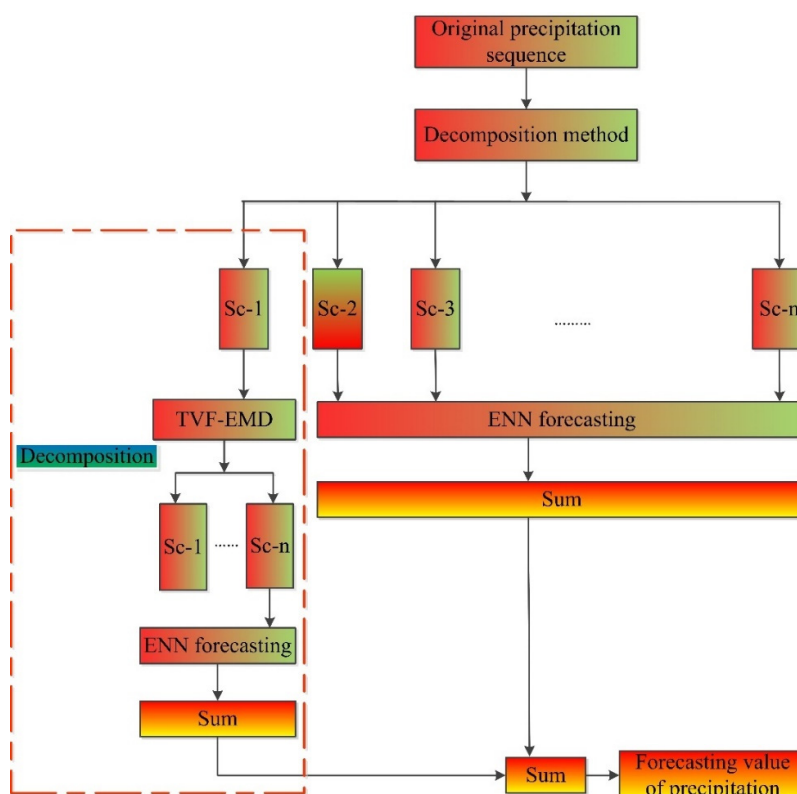
Figure 20. The flowchart for the secondary decomposition of Sc-1.

Since Sc-1 has a certain influence on the prediction performance of the REMD-ENN CEEMD-ENN, WT-ENN, and ESMD-ENN models, only the prediction performance of each model before and after Sc-1 re-decomposition is compared.

Figure 21. The improvement in the relative error of the Sc-1 and annual precipitation prediction after TVF-EMD decomposition.

Note. CENN=CEEMD-ENN; CTEENN=CEEMD-TVF-EMD-ENN; EENN=ESMD-ENN; ETENN=ESMD-TVF-EMD-ENN; RENN=REMD-ENN; RTEENN=REMD-TVF-EMD-ENN; WTENN=WT-ENN; WTEENN=WT-TVF-EMD-ENN.

As shown in Figure 21, after the secondary decomposition of the Sc-1 of the TVF-EMD, the prediction performances of Sc-1 and precipitation improved to varying degrees,

and the absolute value of the relative prediction error decreased to varying degrees. The most significant improvement in the prediction performance of Sc-1 was observed for the CEEMD-ENN model in 2017 and 2018; the absolute values of the relative error decreased from 220% to 35.56% and 175.2% to 12.66%, respectively. For the ESMD-ENN model, the prediction performance of Sc-1 improved most significantly in 2016 and 2017, with decreases in the absolute value of the relative error from 259.4% to 40.7% and 205.9% to 5.32%, respectively. The prediction performance of the REMD-ENN model improved most significantly in 2018, and the absolute value of the relative prediction error dropped from 74.66% to 0.69%. The prediction performance of the Sc-1 of the WT-ENN model improved most significantly in 2016 and 2017. The absolute value of the relative prediction error decreased from 305.1% to 3.26% and from 489.5% to 69.42, respectively.

From the perspective of improving the performance of annual precipitation prediction, the prediction performance of the CEEMD-ENN model improved most significantly in 2016 and 2018. The absolute value of the relative prediction error decreased from 16.42% to 2.18% and 40.53% to 2.44%, respectively. The prediction performance of the ESMD-ENN model improved most significantly in 2017 and 2018. The absolute value of the relative prediction error decreased from 5.57% to 2.18% and 26.9% to 10.19%, respectively. The prediction performance of the REMD-ENN model improved most significantly in 2016 and 2018. The absolute value of the relative prediction error decreased from 14.42% to 6.35% and 17.93% to 3.39%, respectively. The prediction performance of the WT-ENN model improved most significantly in 2015 and 2017. The absolute value of the relative prediction error decreased from 7.67% to 3.23% and 22.01% to 15.53%, respectively.

After Sc-1 was decomposed again by TVF-EMD, the improvements in the performance indices of the Sc-1 for the precipitation prediction of the four models (REMD-ENN, CEEMD-ENN, WT-ENN, ESND-ENN) are shown in Figure 22.

	CENN	CTEENN	ESENN	ESTEENN	RENN	RTEENN	WTENN	WTTEENN
MAE (mm)	306.3	72.04	217.1	186.3	197.4	108.3	246.5	171.7
MAPE(%)	14.01	3.19	9.93	6.33	8.35	4.35	10.78	7.84
R^2	0.7	0.95	0.64	0.9	0.7	0.92	0.82	0.89
RMSE (mm)	411.9	8.68	273.9	200.2	205.8	136.4	276.4	201.1

Figure 22. The secondary decomposition of Sc-1 improved the annual precipitation prediction of the four models.

Note. CENN=CEEMD-ENN; CTEENN=CEEMD-TVF-EMD-ENN; EENN=ESMD-ENN; ETENN=ESMD-TVF-EMD-ENN; RENN=REMD-ENN; RTEENN=REMD-TVF-EMD-ENN; WTENN=WT-ENN; WTEENN=WT-TVF-EMD-ENN.

After the secondary decomposition of the Sc-1 by the TVF-EMD, the prediction performance of annual precipitation of the four methods improved. The RMSE of CENN decreased from 411.9 mm to 8.68 mm, which was the most significant improvement. The R² of ESENN improved most obviously, from 0.64 to 0.9. The MAPE of CENN decreased

from 14.01% to 3.19%, which was the most significant improvement. The MAE of CENN improved most obviously, from 306.3 mm to 72.04 mm.

In general, the secondary decomposition of the Sc-1 by the TVF-EMD significantly improved the prediction performance of the models. This proposed secondary decomposition prediction model shows applicability for other time-series predictions in hydrology.

5. Conclusions

The annual precipitation in Guangzhou is substantially influenced by human activities and climatic factors, and the time series shows significant non-stationarity and volatility. The following conclusions can be drawn based on the results of the precipitation prediction.

The subcomponents obtained by the TVF-EMD decomposition were relatively stable and were less affected by the sunspots, NAO index, and Nino 3.4 index than those of the other models. The VCR of the subcomponents obtained from the TVF-EMD was relatively balanced, and the VCR of the Sc-1 was significantly smaller than that of the other four decomposition methods. The PCC between the precipitation and the Sc-1 of TVF-EMD was smaller than that of the other four decomposition methods. The TVF-EMD-ENN prediction model is superior to the other four prediction models.

The TVF-EMD-ENN model had the best and most stable prediction performance among the five decomposition-reconstruction prediction models. The prediction performance of the decomposition-reconstruction prediction model was better than that of traditional machine learning models. The prediction performance of the Sc-1 and precipitation improved significantly after the secondary decomposition of the Sc-1 by the TVF-EMD, and the MAE, MAPE, R^2, and RMSE showed various degrees of improvement. We first recommend the TVF-EMD for precipitation prediction to deal with non-stationary data.

Under the interference of climate change and human activities, the precipitation series presents a certain non-stationarity. Therefore, the development of hybrid prediction models by introducing signal decomposition methods may lead to more accurate and stable prediction results, and may also contribute to research related to hydrological time series prediction, to address a wide range of issues related to effective water resource management.

Acronyms

TVF-EMD	Time-varying filter-based empirical mode decomposition
REMD	Robust empirical mode decomposition
CEEMD	Complementary ensemble empirical mode decomposition
ESMD	Extreme-point symmetric mode decomposition
WT	Wavelet transform
ENN	Elman neural network
VCR	Variance contribution rate
PCC	Pearson correlation coefficient
WTC	Wavelet transform coherence
NAO	North Atlantic Oscillation
Sc-1	The first subcomponent
RMSE	Root mean square error
MAE	Mean absolute error
MAPE	Mean absolute percentage error
R^2	Coefficient of determination
TVF-EMD-ENN	Time-varying filter-based empirical mode decomposition and Elman neural network
REMD-ENN	Robust empirical mode decomposition and Elman neural network
CEEMD-ENN	Complementary ensemble empirical mode decomposition and Elman neural network
ESMD-ENN	Extreme-point symmetric mode decomposition and Elman neural network
WT-ENN	Wavelet transform and Elman neural network
Tn	The Nth subcomponent obtained from TVF-EMD
TEn	The absolute prediction error of the Nth subcomponent obtained from TVF-EMD-ENN
Rn	The nth subcomponent obtained from REMD
REn	The absolute prediction error of the Nth subcomponent obtained from REMD-ENN
Cn	The nth subcomponent obtained from CEEMD
CEn	The absolute prediction error of the Nth subcomponent obtained from CEEMD-ENN
Xn	The nth subcomponent obtained from the WT
XEn	The absolute prediction error of the Nth subcomponent obtained from WT-ENN
EEn	The absolute prediction error of the Nth subcomponent obtained by ESMD-ENN
En	The nth subcomponent obtained from ESMD
COI	Cone of influence
CENN	CEEMD-ENN
CTEENN	CEEMD-TVF-EMD-ENN
ESENN	ESMD-ENN
ESTEENN	ESMD-TVF-EMD-ENN
RENN	REMD-ENN
RTEENN	REMD-TVF-EMD-ENN
WTENN	WT-ENN
WTTEENN	WT-TVF-EMD-ENN
LSTM	Long short-term memory neural network
RBF	Radial basis function neural network

References

1.　Swain, S.; Patel, P.; Nandi, S. A multiple linear regression model for precipitation forecasting over Cuttack district, Odisha, India. In Procceding of the 2nd International Conference for Convergence in Technology (I2CT), Mumbai, India, Apr 1, **2017**; pp. 355–357, doi:10.1109/I2CT.2017.8226150.

2.　Rahman, M.A.; Yunsheng, L.; Sultana, N. Analysis and prediction of rainfall trends over Bangladesh using Mann–Kendall, Spearman's rho tests and ARIMA model. Meteorol. *Atmos. Phys*. **2017**, *129*, 409–424, doi:10.1007/s00703-016-0479-4.

3.　Amiri, S.S.; Mottahedi, M.; Asadi, S. Using multiple regression analysis to develop energy consumption indicators for commercial buildings in the US. *Energy Build*. **2015**, *109*, 209–216.

4.　Maria, Cleofé, Valverde, Ramírez, and, & Haroldo, et al. Artificial neural network technique for rainfall forecasting applied to the so Paulo region. *J. Hydrol*. **2005**, *301*(1-4), 146-162, doi:10.1016/j.jhydrol.2004.06.028.

5.　Venkata Ramana, R., Krishna, B., Kumar, S.R. et al. Monthly Rainfall Prediction Using Wavelet Neural Network Analysis. *Water Resources Management*. **2013**, *27*, 3697–3711. doi:10.1007/s11269-013-0374-4.

6.　Mislan; Haviluddin; Hardwinarto, S.; Sumaryono; Aipassa, M. Rainfall monthly prediction based on artificial neural network: A case study in Tenggarong station, east Kalimantan-Indonesia. Procedia Comput. Sci. 2015. *59*, 142–151, doi:10.1016/j.procs.2015.07.528.

7.　Wu, C.L.; Chau, K.W.; Fan, C. Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. *J. Hydrol*. 2010, *389*, 146–167, doi:10.1016/j.jhydrol.2010.05.040.

8.　Kim, T.; Heo, J.H.; Jeong, C.S. Multireservoir system optimization in the Han River basin using multi-objective genetic algorithms. *Hydrol. Process*. **2006**, *20*, 2057–2075.

9.　Maryam, S.; Jan, A.; Ahmad, F.F.; Yagob, D.; Kazimierz, A. A wavelet-sarima-ann hybrid model for precipitation forecasting. *J. Water Land Devt*. **2016**, *28*, 27–36, doi:10.1515/jwld-2016-0003.

10.　Chau, K.W.; Wu, C.L. A hybrid model coupled with singular spectrum analysis for daily rainfall prediction. *J. Hydroinform*. **2010**, *12*, 458–473, doi:10.2166/hydro.2010.032.

11.　Tan, Q.-F.; Lei, X.-H.; Wang, X.; Wang, H.; Wen, X.; Ji, Y.; Kang, A.-Q. An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach. *J. Hydrol*. **2018**, *567*, 767–780, doi:10.1016/j.jhydrol.2018.01.015.

12.　Chen, R.; Jia, H.; Xie, X.; Wen, G. Sparsity-Promoting Adaptive Coding with Robust Empirical Mode Decomposition for Image Restoration. In Pacific Rim Conference on Multimedia; Springer: Cham, Switzerland, **2017**, doi:10.1007/978-3-319-77380-3_36.

13.　Mumtaz. A.; Ramendra. P.; Yong. X.; Zaher, M.Y. Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. *J. Hydrol*. **2020**, *584*, 124674, doi:10.1016/j.jhydrol.2020.124647.

14.　Ji, Y.; Dong, H.T.; Xing, Z.X.; Sun, M.X.; Fu, Q.; Liu, D. Application of the decomposition prediction reconstruction framework to middle and long-term runoff forecasting. *Water Supply*. 2020, doi:10.2166/ws.2020.337.

15.　Arab, A.M.; Amerian, Y.; Mesgari, M.S. Spatial and temporal monthly precipitation forecasting using wavelet transform and neural networks, Qara-Qum catchment. Iran. *Arab. J. Geosci*. **2016**, *9*, 1–18, doi:10.1007/s12517-016-2446-2.

16.　Partal, T.; Cigizoglu, H.K. Prediction of daily precipitation using wavelet—neural networks. Hydrol. Sci. J. 2009, *54*, 234–246, doi:10.1623/hysj.54.2.234.

17.　Qin, Y.H.; Li, B.F.; Sun, X.; Chen, Y., N.; Shi, X. Nonlinear response of runoff to atmospheric freezing level height variation based on hybrid prediction models. Hydrol. Sci. J. 2019, *64*, 1556–1572, doi:10.1080/02626667.2019.1662023.

18.　Li, H.; Li, Z.; Mo, W. A time varying filter approach for empirical mode decomposition. *Signal Process*. **2017**, *138*, 146–158, doi:10.1016/j.sigpro.2017.03.019.

19.　Zhang, X.; Liu, Z.; Miao, Q.; Wang, L. An optimized time varying filtering based empirical mode decomposition method with grey wolf optimizer for machinery fault diagnosis. *J. Sound Vib*. **2018**, *418*, 55–78, doi:10.1016/j.jsv.2017.12.028.

20.　Xu, Y.; Cai, Z.; Cai, X. An enhanced multipoint optimal minimum entropy deconvolution approach for bearing fault detection of spur gearbox. *J. Mech. Sci. Technol*. **2019**, *33*, 2573–2586, doi:10.1007/s12206-019-0505-9.

21.　Chen, P.C.; Wang, Y.H.; You, J.Y.; Wei, C.C. Comparison of methods for non-stationary hydrologic frequency analysis: Case study using annual maximum daily precipitation in Taiwan. *J. Hydrol*. **2017**, *545*, 197–211, doi:10.1016/j.jhydrol.2016.12.001.

22.　Yeh, J.R.; Shieh, J.S.; Huang, N.E. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Adv. Adapt. Data Anal*. **2010**, *2*, 135–156, doi:10.1142/S1793536910000422.

23.　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. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci*. **1998**, *454*, 903–995, doi:10.1098/rspa.1998.0193.

24.　Nejad, H.F.; Nourani, V. Elevation of wavelet denoising performance via an ANN-based streamflow forecasting model. *Int. J. Comput. Sci. Manag. Res*. **2012**, *1*, 764–77

25.　Adamowski, J.; Sun, K. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *J. Hydrol*. **2010**, *390*, 85–91, doi:10.1016/j.jhydrol.2010.06.033.

26.　Nourani, V.; Alami, M.T.; Aminfar, M.H. A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Eng. Appl. Artif. Intell*. **2009**, *22*, 466–472, doi:10.1016/j.engappai.2008.09.003.

27.　Benaouda, D.; Murtagh, F.; Starck, J.-L.; Renaud, O. Wavelet-based nonlinear multiScale decomposition model for electricity load forecasting. *Neurocomputing*. **2006**, *70*, 139–154.

28. Wang, J.L.; Li, Z.J. Extreme-point symmetric mode decomposition method for data analysis. *Adv. Adapt. Data Anal.* **2013**, *5*, 1137–1137, doi:10.1142/S1793536913500155.

29. Qin, Y.; Li, B.; Chen, Z.; Chen, Y.; Lian, L. Spatio-temporal variations of nonlinear trends of precipitation over an arid region of northwest China according to the extreme-point symmetric mode decomposition method. *Int. J. Climatol.* **2018**, *38*, 2239–2249, doi:10.1002/joc.5330.

30. Elman, J. Finding structure in time. *Cogn. Sci.* **1990**, *14*, 179–211, doi:10.1207/s15516709cog1402_1.

31. Song, C.; Zhang, X.; Hu, D.; Tuo, W. Research on a coupling model for groundwater depth forecasting. *Desalin. Water Treat.* **2019**, *142*, 125–135, doi:10.5004/dwt.2019.23458.

32. Pan, Y.; Li, Y.; Ma, P.; Liang, D. New approach of friction model and identification for hydraulic system based on MAPSO-NMDS optimization Elman neural network. *Adv. Mech. Eng.* **2017**, *9*, 1–14, doi:10.1177/1687814017744473.

33. Ardalani-Farsa, M.; Zolfaghari, S. Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks. *Neurocomputing.* **2010**, *73*, 2540–2553, doi:10.1016/j.neucom.2010.06.004.

34. Su, L.; Miao, C.; Duan, Q.; Lei, X.; Li, H. Multiple-wavelet coherence of world's large rivers with meteorological factors and ocean signals. *J. Geophys. Res. Atmos.* **2019**, *124*, 4932–4954, doi:10.1029/2018JD029842.

35. Grinsted, A.; Moore, J.C.; Jevrejeva, S. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process. Geophys.* **2004**, *11*, 561–566.