Improving annual precipitation prediction using extreme learning machine with wavelet packet decomposition

Hua Wang
Changsha University of Science and Technology
Wenchuan Wang (wangwen1621@163.com)
North China University of Water Resources and Electric Power https://orcid.org/0000-0003-1367-5886
Yu-jin Du
North China University of Water Resources and Electric Power
Dong-mei Xu
North China University of Water Resources and Electric Power
Yi-duo Zhang
North China University of Water Resources and Electric Power

Research Article

Keywords: Precipitation prediction, Extreme learning machine, Artificial neural network, Wavelet packet decomposition, Hybrid intelligent computing

Posted Date: March 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-317538/v1

Improving annual precipitation prediction using extreme learning machine with wavelet packet decomposition

Hua Wang
School of Hydraulic Engineering, Changsha University of Science and Technology, Changsha, 410114, People's Republic of China
E-mail: 66181460@qq.com

Wen-chuan Wang
College of Water Resources, Henan Key Laboratory of Water Resources Conservation and Intensive Utilization in the Yellow River Basin, North China University of Water Resources and Electric Power, Zhengzhou, 450046, People's Republic of China
Corresponding author, E-mail: wangwen1621@163.com; wangwenchuan@ncwu.edu.cn

Yu-jin Du
College of Water conservancy, Henan Key Laboratory of Water Resources Conservation and Intensive Utilization in the Yellow River Basin, North China University of Water Resources and Electric Power, Zhengzhou, 450046, People's Republic of China
E-mail: 2366036835@qq.com

Dong-mei Xu
College of Water Resources, Henan Key Laboratory of Water Resources Conservation and Intensive Utilization in the Yellow River Basin, North China University of Water Resources and Electric Power, Zhengzhou, 450046, People's Republic of China
E-mail: xudongmei@ncwu.edu.cn

Yi-duo Zhang
College of Water Resources, Henan Key Laboratory of Water Resources Conservation and Intensive Utilization in the Yellow River Basin, North China University of Water Resources and Electric Power, Zhengzhou, 450046, People's Republic of China
E-mail: 877025930@qq.com

Abstract: Accurate precipitation prediction can help plan for different water resources management demands and provide an extension of lead-time for the tactical and strategic planning of courses of action as well as activity. In this paper, a novel merged precipitation prediction framework (ELM-WPD) is proposed on the Extreme learning machine (ELM) with wavelet packet decomposition (WPD). The model can be

described as the following: (a) we use the WPD to decompose the original precipitation data into several sub-layers; (b) ELM model is employed to complete the forecasting calculation for the decomposed series; (c) the results are integrated to complete the final prediction. Four quantitative indexes (RMSE, MAE, R and NSEC) are employed for the comparison criteria. The results are compared with Back-propagation neural network (BPNN), autoregressive integrated moving average model (ARIMA), ELM, BPNN-WPD model, ARIMA-WPD, indicating that the ELM-WPD model has better performance than other models used in this paper. Hence, the proposed method can provide higher accuracy and reliability for annual precipitation forecasting and can be extended to similar situations for precipitation forecasting.

Key words: Precipitation prediction; Extreme learning machine; Artificial neural network; Wavelet packet decomposition; Hybrid intelligent computing

1 Introduction

Precipitation prediction modeling is the key in many practical applications, i.e., agriculture (Wei et al. 2005), streamflow forecasting (Kisi and Cimen 2011; Wang et al. 2013), flood prediction (Wang et al. 2019), water resources management (Hartmann et al. 2016), urban flooding predictions (Nguyen and Bae 2020), facilities maintenance and control (Benedetto 2002) and so on (Ortiz-García et al. 2014). Accurate precipitation prediction can help plan for different environmental water demands and provide an extension of lead-time for the tactical and strategic planning of courses of action as well as activity. In hydrology, the forecasts of precipitation are commonly obtained from numerical weather prediction models or by applying various techniques,

such as time series regression, artificial neural network (ANN), support vector machines (SVMs), etc. But, numerical weather prediction models are still unable to provide quantitative precipitation forecasts at enough spacial and time resolution compatible with the time-space variability of precipitation processes (Kuligowski and Barros 1998). Soft computing approaches have several advantages: they are easy to operate and can carry out large-scale data operation, and can adjust the model structure and parameters according to the characteristics of the watershed, and their performance is indeed very competitive compared to numerical models (Ortiz-García et al. 2014).

The Box-Jenkins models (Box et al. 1994), which can be considered as the most conventional and comprehensive technology for time series prediction, contains AR (auto-regressive), ARMA (autoregressive moving average), and ARIMA (auto-regressive integrated moving average), etc. These methods have been extensively utilized in non-stationary data analysis and prediction in recent decades (Mohammadi et al. 2006; Pektaş and Kerem Cigizoglu 2013; Valipour et al. 2013; Wang et al. 2009; Wang et al. 2015). Many studies in which time series analysis methods have been successfully applied in precipitation (rainfall) modeling can be found in recent decades. Rainfall forecasting is approached assuming that hourly rainfall follows an autoregressive moving average (ARMA) process (Burlando et al. 1993). The performance of the SSA (singular spectrum analysis) and MA are investigated as data-preprocessing technology, and combined with forecasting model to improve model accuracy for precipitation prediction (Wu and Chau 2013). A hybrid the maximal overlap discrete wavelet transform (MODWT) and the ARMA model is presented for

84 daily rainfall prediction, and the combined model is an effective way to improve
85 forecasting accuracy (Zhu et al. 2014). The Box-Jenkins method is applied to forecast
86 short term monthly rainfall, and model is fit to both generate the reliable future forecasts
87 and describe the past precipitation data (Papalaskaris et al. 2016). Seasonal
88 Autoregressive Integrated Moving Average (SARIMA) model is adopted to predict
89 monthly precipitation at thirty station in Bangladesh with twelve months lead-time
90 (Mahmud et al. 2017). The Box–Jenkins forecasting method with ARIMA model is
91 used to predict the changes in precipitation for projected years (Al Balasmeh et al. 2019).
92 In recent years, Artificial Intelligent (AI) such as Artificial Neural Networks (ANN)
93 have been a powerful technology to solve the forecasting problems. Hung et al. (2009)
94 developed ANN technique to improve rainfall forecast performance for Bangkok,
95 Thailand, with lead times of 1 to 6h. Nastos et al. (2014) developed ANNs to forecast
96 the maximum daily precipitation for the next coming year in Athens, Greece. Abbot and
97 Marohasy (2014) used ANNs to forecast Queensland monthly rainfall, and ANNs
98 forecasts are superior to Australian official forecasts. Abbot and Marohasy (2017)
99 skillful achieved seasonal and monthly rainfall forecasts by mining the historical
100 climate data using ANNs. A comparison of several advanced AI methods, namely
101 ANFIS (Adaptive Network based Fuzzy Inference System) optimized with PSO
102 (Particle Swarm Optimization), SVM and ANN, for the forecasts of daily rainfall were
103 introduced by Pham et al. (2020).
104 In the last few years, many works have been done to enhance the prediction ability
105 of soft computing approaches by data preprocessing technology such as wavelet

106 analysis. The WT (Wavelet Transform) is a multi-resolution signal identify and analysis
107 tool, which can provide a time-frequency representation of signal (Balasubramanian et
108 al. 2018; Feng et al. 2015). Time series modelling techniques with wavelet (WT) has
109 seen to made enormous interests in hydrologic data analysis and prediction. Wavelet
110 regression (WR) technique is proposed for short-term runoff forecasting, and results
111 illustrate that the performance of WR is better than those of ARMA and ANN models
112 (Kisi 2010). The accuracy of wavelet and SVM hybrid model is investigated in monthly
113 streamflow prediction, and the experiment results indicated that the hybrid model can
114 improve the prediction accuracy (Kisi and Cimen 2011). A new wavelet-SVM hybrid
115 model is proposed for daily rainfall forecasting, and the results indicated that the model
116 can dramatically improve the forecasting accuracy of single SVM (Kisi and Cimen
117 2012). A wavelet predictor–corrector model is developed by (Zhou et al. 2008) for the
118 simulation and prediction of monthly discharge time series. The SSA-ARIMA model is
119 developed to forecast mid- to long-term streamflow and results shows that the
120 conjunction model can provide the best performance (Zhang et al. 2011). Karthikeyan
121 and Nagesh Kumar (2013) ascertained the predictability of four non-stationary runoff
122 sites using EMD (Empirical Mode Decomposition) and wavelet based ARMA. The
123 wavelet neural network was studied to predict the monthly rainfall series (Venkata
124 Ramana et al. 2013). Three different methods (generalized regression neural network;
125 radial basis function; feed forward back propagation) based WT were utilized for daily
126 rainfall forecasting (Partal et al. 2015). The wavelet analysis-support vector machine
127 coupled model (WA-SVM) is evaluated for 1, 3 and 6 months ahead rainfall forecasting,

and results indicate that the WA-SVM model provides high accuracy and reliability (Feng et al. 2015). Wavelet transform and Frank copula function are combined in a mutual information-based input variable selection for non-linear rainfall forecasting models (Abdourahamane et al. 2019). The performance enhancement model is developed for rainfall forecasting over the Langat River Basin through the integration of WT and convolutional neural network (Chong et al. 2020). The main deficiency of the WT is that the number of frequency bands is restricted, and WPD is a further extension of WT technique (Ben messaoud et al. 2016). The WPD can provide a more complete wavelet packet tree, and which extracts the features of original signal more comprehensively (Balasubramanian et al. 2018). The WPD has been proved to exhibit good performance for time series forecasting, such as wind speed forecasting (Liu et al. 2018; Yu et al. 2018), River Stage Forecasting (Seo et al. 2016). However, few studies investigated the performance of combined WPD and time series models for prediction in hydrology field, which needs further exploration.

Motivated by the principle of "decomposition and ensemble" (Guo et al. 2012; Tan et al. 2018), the raw annual rainfall data can be decomposed into different components. Each component can be predicted with the purpose of fine results and easy forecasting task, and the forecasting results of all components are aggregated to generate the final prediction (Wang et al. 2015; Wang et al. 2013). In this paper, a novel merged precipitation prediction framework is presented based on the Extreme learning machine (ELM) with WPD. The framework can be described as the following: (a) we use the WPD to decompose the original precipitation data into several sub-layers; (b) ELM

model is employed to complete the forecasting calculation for the decomposed series; (c) Integrate the results of (b) are integrated to complete the final prediction. To ascertain the performance of the proposed framework, six models are employed for benchmark comparison as: ARIMA, ARIM-WPD, BPNN, BPNN-WPD, ELM, and ELM-WPD. Validation of the models was done using four criteria, which are RMSE (root mean square errors), R (coefficient of correlation), NSEC (Nash-Sutcliffe efficiency coefficient), and MAE (mean absolute error).

2 Methods

2.1 Wavelet Packet Decomposition WPD

WPD is identical to WD except that the former extends the abilities of the later (Alickovic et al. 2018). The three-layer binary trees of WPD are illustrated in Fig. 1. The WPD splits the signal into approximation coefficients and detail coefficients by mother wavelet function. The decomposition levels and mother wavelet function have a deep influence on the performance of the WPD. The WPD contains continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The CWT is as follows:

$$CWT_{\mathrm{x}}(a,b)=\left\langle x(t),\psi_{a,b}(t)\right\rangle=\int x(t)\psi^{*}((t-b)/a)/\sqrt{a}dt \tag{1}$$

where $x(t)$ is the input signal, a the scale parameter, b the translation parameter, $*$ the complex conjugate, $\psi(t)$ the mother wavelet function. The a and b in DWT are:

$$\begin{cases} a=2^{i} \\ b=j2^{i} \end{cases} \tag{2}$$

where i and j denotes the scale parameter and translation parameter, respectively.

Insert Fig. 1.

2.2 Extreme Learning Machine (ELM)

ELM proposed by Guang-Bin et al. (2004) is a single hidden layer feed-forward network (SLFN) with the characteristic that without adjusting the internal parameters, which is a modified version of traditional ANN (Yaseen et al. 2019). ELM selects the hidden threshold randomly in training the network, and the calculation of output weight does not need complicated iterative, which greatly heighten the training speed.

For the vector (x_i, y_i), $x_i = [x_{i1}, x_{i2}, \cdots, x_{im}]^T \in R^m$, $y_i = [y_{i1}, y_{i2}, \cdots, y_{in}]^T \in R^n$, ELM can be modeled by:

$$\sum_{i=1}^{L} \beta_i G(\omega_i \cdot x_j + \alpha_i) = z_j, j = 1,\mathrm{L}\ ,N \tag{3}$$

The variables in Eq. (3) denote: β_i (the weight connecting the nodes of hidden layer and output layer), $G(x)$ (activation function), $\omega_i = [\omega_{i1}, \omega_{i2}, \cdots \omega_{im}]$ (the weight vector connecting the nodes of input layer and output layer), x_j (input vectors), α_i (threshold of the hidden node), z_j (output value), L (hidden nodes number).

The goal of ELM is to minimize the error between the original and predicted values, $\sum_{j=1}^{N} \| \ z_j - y_j \| = 0$, there exist β_i, ω_i and α_i. So,

$$\sum_{i=1}^{L} \beta_i G(\omega_i \cdot x_j + \alpha_i) = \mathrm{y}_j, j = 1,\mathrm{L}\ ,N \tag{4}$$

Eq. (4) can be expressed into the following equations:

$$H\beta = Y \tag{5}$$

where

$$H(\omega_i,\alpha_i,x_j)=\begin{bmatrix} g(\omega_1\cdot x_1+\alpha_1)\,\mathrm{L} & g(\omega_L\cdot x_1+\alpha_L) \\ \mathrm{M} \quad \mathrm{L} \quad \mathrm{M} & \\ g(\omega_1\cdot x_N+\alpha_1)\,\mathrm{L} & g(\omega_L\cdot x_N+\alpha_L) \end{bmatrix}_{N\times L} \tag{6}$$

$$\beta=\begin{bmatrix} \beta_1^T \\ \mathrm{M} \\ \beta_L^T \end{bmatrix}_{L\times 1} \text{ and, } Y=\begin{bmatrix} \mathrm{y}_1^T \\ \mathrm{M} \\ y_N^T \end{bmatrix}_{N\times 1} \tag{7}$$

H stands for the output matrix of hidden, β the output weight, Y the target output. Thus the minimum norm least-squares solution $\hat{\beta}$ of Eq. (5) is:

$$\hat{\beta}=H^{+}Y \tag{8}$$

where H^{+} represents the MPGI (Moore-Penrose generalized inverse) of H and Y is the target.

2.3 Back-propagation Neural Network (BPNN)

BPNN proposed by McCelland and Rumelhart (1986) is a commonly multilayered ANN. BPNN includes input nodes, hidden layers and output layer. The characteristic of BPNN is the forward transmission of the signal and the reverse transmission of the error. BPNN minimize the global error according to gradient descent approach (Wang et al. 2017).The connection weights are modified based on the errors between the input data and output data in the reverse transmission. Forward and afterword propagation are repeated until the errors reach the expected precision. In this study, Levenberg-Marquardt (LM) method, sigmoid function and perelin formula are adopted as the training function, transfer function and output function, respectively. The mathematical formula of BPNN model can be expressed as follows:

$$\mathrm{x}_j^m = f\left(\sum_{i=1}^{m} x_i^{m-1} w^m + \beta^m\right) \tag{8}$$

where x_i^{m-1} represents the input data of node I in layer m-1, w^m the weight of x_i^{m-1}, β^m the bias of layer m, x_j^m the output value of node I in layer m; $f(x)$ denotes the transfer function of layer m, which is written as:

$$f(x) = \tan sig(x) = (e^x - e^{-x}) / (e^x + e^{-x}) \tag{9}$$

The purelin formula is as the following:

$$f(x) = x \tag{10}$$

2.4 ARIMA

ARIMA, proposed by Box (2013), is normally utilized to time series analysis. The mathematical forecasting equation of ARIMA is linear in which the predictors include autoregressive (AR) terms and moving average (MA) terms. The representation of ARIMA is ARIMA (p, d, q), and the representation of SARIMA (seasonal ARIMA) is ARIMA (p, d, q)$\times$(P, D, Q)$_s$, where (p, d, q) represents the non-seasonal order and (P, D, Q)$_s$ denotes the seasonal order. ARIMA model can be expressed as:

$$y_{t+1} = \mu + \phi_1 y_t + \phi_2 y_{t-1} + \mathrm{L} + \phi_p y_{t-p+1} + \varepsilon_{t+1} - \theta_1 \varepsilon_t - \theta_2 \varepsilon_{t-1} - \mathrm{L} - \theta_q \varepsilon_{t-q+1} \tag{11}$$

where y_t is the time series, ϕ_i the AR coefficient, θ_i the MA coefficient, μ the model parameter, p the order of AR component, q the order of MA component, d the order of differentiation.

2.5 Framework of the proposed hybrid model

The methodological architecture of merged precipitation prediction framework (ELM-WPD) is presented in Fig. 2. The content of the model is expressed as:

Step 1: Data pre-processing. (a) the WPD is adopted to decompose the original precipitation series into several sub-series; (b) the data both in original series and sub-

series are partitioned into training and testing sets; (c) all data are tested by the ADF (Augmented Dickey-fuller Test); (d) normalize all data between [0, 1] by:

$$x_i^{'} = \frac{x_i - \min\limits_{1\le i\le n}\{x_i\}}{\max\limits_{1\le i\le n}\{x_i\} - \min\limits_{1\le i\le n}\{x_i\}} \tag{12}$$

where x_i is the original data, x_i' the normalized data series, n the number of data.

Step 2: Model building. (a) the PACF (partial autocorrelation function) and precipitation theory are employed to select the number of input variables; (b) set the values of model parameter, such as the number of hidden layers in ANN and the optimized parameters for ARIMA; (c) the ELM, ARIMA and BPNN are used as forecasting tools to model and predict each decomposed sub-sequence separately; (d) the results are integrated to complete the final prediction.

2.6 Evaluation indicators

The results of the models are evaluated with respect to four numerical indicators. These indexes include root mean square errors (RMSE), mean absolute error (MAE). Nash-Sutcliffe efficiency coefficient (NSEC), and coefficient of correlation (R). Their equations are provided below.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_e(i) - y_o(i))^2} \tag{13}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_e(i) - y_o(i)\right| \tag{14}$$

$$NSEC = 1 - \frac{\sum_{i=1}^{n}(y_e(i) - y_o(i))^2}{\sum_{i=1}^{n}(y_o(i) - \overline{y_o})^2} \tag{15}$$

$$R = \frac{\sum_{i=1}^{n}(y_o(i) - \overline{y_o})(y_e(i) - \overline{y_e})}{\sqrt{\sum_{i=1}^{n}(y_o(i) - \overline{y_o})^2 \sum_{i=1}^{n}(y_e(i) - \overline{y_e})^2}} \quad (16)$$

where $y_e(i)$, $y_o(i)$, $\overline{y_e}$ and $\overline{y_o}$ are the estimated value, observed value, mean estimated value, and mean observed value of precipitation, respectively.

Insert Fig. 2.

3 Case study

3.1 Study area and used data

The annual rainfall data of Jinsha weather station on Chishui river in the northwest of Guizhou Province were used. Fig. 3 displays the precipitation area of this paper. Jinsha county, with about 2524 km^2 area, spans 105º47′-105º44′ E and 27º07′-27º46′ N, covering a total of 2524 km^2. There are four distinct seasons, abundant rainfall, uneven distribution of rainfall, and frequent occurrences of disasters such as mudslides and landslides, which endanger the lives and property of the people. Therefore, it is necessary to carry out high-precision precipitation simulation and forecast around Jinsha County, to provide theoretical guidance for regional water resources management and flash flood early warning.

The annual rainfall data from 1958 to 2016 are studied. Fig.4 shows rainfall data for Jinsha station, data from 1958 to 2011 are adopted for training and the final 5 y are selected for testing. The statistical parameters of the original data are list in Table 1, and the original data shows relatively obvious skewness, indicating that the difficulty of

modeling is very high.

Insert Fig. 3.

Insert Fig. 4.

3.2 Decomposition results

The WPD is adopted to split the original annual precipitation data into eight sub-series. The frequency of sub-series is different, and each sub-series plays different role in the original dataset. The decomposed results for Jinsha station using WPD at level 3 are shown in Fig. 5.

Insert Fig. 5.

3.3 The selection of input variable

The selection of input variables is a prior step for the final predicted results. In this paper, two methods are utilized to select the input combinations: (a) trial and error method; (b) PACF statistical approach. PACF values for all series of Jinsha station are shown in Fig. 6, where PACF denotes the original rainfall series and the others are the sub-series. We conduct twelve ANN models with different input combinations. Table 1 list the input variables for different series with respect to the information in Fig. 7 and trial-and-error method， where $q_{(t)}$ represents the estimated value of rainfall and $q_{(t-p)}$ is the rainfall at time *t-p*.

Insert Fig. 6.

Insert Table 1.

294 3.4 Model development

295 To verify the ELM-WPD model, these six models, that is, ELM, BPNN, ARIMA,
296 ELM-WPD, BPNN-WPD, ARIMA-WPD, are employed for comparison. Detailed
297 information relating to these models are presented in this part.

298 (1) ELM and BPNN models

299 For conventional ELM and BPNN models, the observed rainfall values were
300 adopted as the target value. The nodes' number of input and output layers are equal to
301 the number of input variables and one, respectively. The best hidden nodes' number is
302 determined by trial-and-error method. In this paper, the best hidden nodes' number in
303 ELM and BPNN are set to twenty and eight, respectively. The LM algorithm is adopted
304 to train the BPNN model, and the training epoch is set to 500. Selecting sigmoidal
305 function as the activation function of ELM, and MPGI method is adopted to determine
306 the hidden output weights after randomly setting its hidden threshold and weight vector
307 between the hidden layer and input layer.

308 (2) ARIMA

309 First, ADF (augmented Dickey-Fuller) unit root test is employed to judge the
310 stationarity of input series. If the dataset is nonstationary, difference and MA can be
311 used to smooth it. The results are shown in Table 2. Value h=1 means that the test rejects
312 the null hypothesis of unit root. The significance level of pValue is determined as 0.05,
313 when the t-statistic value is less than the critical value and pValue <0.05, the test results
314 can be considered as feasible and it will reject the null hypothesis. It can be seen from
315 Table 2 that the sample dataset is stationary without single root effect. Subsequently,

316 the optimal ARIMA (p, d, q) structure is chiefly determined according to the minimum
317 BIC (Bayes information criteria) value. The obtained best ARIMA models are shown
318 in Table 3.

319 Insert Table 2.

320 Insert Table 3.

321 (3) ANN-WPD and ARIMA-WPD

322 For the ELM-WPD, BPNN-WPD and ARIMA-WPD models, the original rainfall
323 is split into eight sub-series using WPD, and several special hybrid models are
324 reconstructed for each subsequence. For the WPD model, the selection of the
325 appropriate wavelet basis function is very significant. Symlet wavelet is an improved
326 approximate symmetric wavelet function based on Daubechies wavelet, which can
327 avoid signal distortion during decomposition and reconstruction. Thus, a three-scale
328 and four order Symlet wavelet is considered as the wavelet basis function in this paper.

329 3.5 Comparative analysis

330 The forecast of the series was implemented using the above models with the input
331 data of original annual precipitation and extracted subseries. The assessment results
332 obtained by different models in training and test phases are shown in Tables 4 and 5.
333 Fig 8 exhibits rainfall forecasting results using six models.

334 Table 4 indicates the following: (a) when comparing the single models with the
335 hybrid model, the hybrid models evidently attain better performance during both
336 training and testing periods; (b) the ELM outperforms other single methods (c) the
337 ARIMA-WPD demonstrates the optimal performance in terms of all measures in the

training period, and ELM-WPD attains the best forecasting accuracy in terms of all measures during the testing phase; (d) ARIMA provides the worst results during both training and testing periods; (e) there is no significant difference in the results between BPNN and ELM and that between BPNN-WPD and ELM-WPD.

Table 5 lists the improvements obtained by several models. Compared with BPNN-WPD and ELM-WPD, ARIMA-WPD can exhibit the best improvements in terms of all measures during both training and testing periods. For example, in the training period, the ARIMA-WPD method improves the ARIMA method with 85.20 and 85.08% reduction in MAE and RMSE, and improvements of the predicted results in terms of NSEC and R are 607.65 and 138.81%. In the testing phase, the ARIMA-WPD method improves the ARIMA method with 73.67 and 74.8% reduction in MAE and RMSE, and the improvements of the predicted results in terms of NSEC and R are 268.63 and 413.25%. Meanwhile, we can see that the improvements of the BPNN-WPD is superior to ELM-WPD, whilst there is no dramatic difference between them.

In addition, the prediction accuracy is different in terms of different phrase. The performance of ARIMA method during the testing period is far inferior to that during the training period. The ARIMA-WPD method is in the highest level in the training period, but the performance in the testing period is the worst among several hybrid models. This kind of model usually has poor generalization ability. However, it is considered that the model has strong generalization capability for a reliable performance when it shows modest during the training period and performs well in the testing period, e.g., BPNN, BPNN-WPD, ELM, ELM-WPD.

The performances of all forecasting models developed in this paper during the training and testing phases is shown in Fig. 7. One can clearly see that the performances of the hybrid models are better than those single methods as their trend line are very close to the original data line. Meanwhile, the algorithm prior to improvement is difficult to capture the drastic changes in rainfall.

The following should be considered before analyzing the above. The forecasting performance of the testing phase plays a greater role than that of the training phase. Because the training phase is utilized to train the model, and its performance is measured by the data related to the modeling. But the testing dataset does not participate in modeling, so its performance can truly reflect the model application efficiency. Based on the above considerations, we can draw the following conclusions from the above analysis. Firstly, the ELM-WPD, proposed in this paper, can attain the best performance in terms of all statistical measures. Secondly, WPD is suitable for decomposing the annual rainfall series as it can overcome the complicated abrupt change of precipitation. Thirdly, there are obvious difference between the ARIMA, BPNN, and ELM models, which means the important role of training tool in modeling. Finally, the annual rainfall series decomposed by WPD as input in modeling can substantially enhance forecasting accuracy.

Insert Table 4.

Insert Table 5.

Insert Fig. 7.

4 Conclusions

Improving accuracy of long-term annual precipitation forecasting is an important yet challenging work in the management of water resources. We proposed a novel merged precipitation prediction framework (ELM-WPD). First, WPD is adopted to decompose the original annual precipitation series into several sub-series. Second, ANN models are utilized to predict each sub-series. Finally, the results are integrated to complete the final prediction. The annual precipitation series collected from Jinsha weather station of Guizhou Province in China is taken to develop the empirical study. The results obtained in this paper reveal that the presented framework can effectively enhance forecasting accuracy with respect to four numerical indicators. Thus, the proposed method based on data decomposition can be particularly relevant for precipitation prediction, and provides more accurate and reliable results, hence may be a promising alternative for long-term precipitation prediction.

Authors' Contributions

Hua Wang: Methodology, Software, Writing- original draft. **Wen-chuan Wang**: Conceptualization, Writing - original draft. **Yu-jin Du**: Program implementation, Writing - original draft. **Dong-mei Xu**: Formal analysis. **Yi-duo Zhang:** Data curation, Program implementation.

Funding

The authors are grateful to acknowledge the funding support the Project of key science and technology of the Henan province (202102310259; 202102310588), and

Henan province university scientific and technological innovation team (No:18IRTSTHN009).

Availability of data and materials

All authors made sure that all data and materials support our published claims and comply with field standards.

Ethics declarations

Ethics Approval: All authors kept the 'Ethical Responsibilities of Authors'.

Consent to Participate: All authors gave explicit consent to participate in this work.

Consent to Publish: All authors gave explicit consent to publish this manuscript.

Conflict of interest: The authors declare that they have no conflict of interest.

References

Abbot J, Marohasy J (2014) Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks Atmospheric Research 138:166-178 doi:https://doi.org/10.1016/j.atmosres.2013.11.002

Abbot J, Marohasy J (2017) Skilful rainfall forecasts from artificial neural networks with long duration series and single-month optimization Atmospheric Research 197:289-299 doi:https://doi.org/10.1016/j.atmosres.2017.07.015

Abdourahamane ZS, Acar R, Serkan S (2019) Wavelet-copula-based mutual information for rainfall forecasting applications Hydrological Processes 33:1127-1142 doi:10.1002/hyp.13391

Al Balasmeh O, Babbar R, Karmaker T (2019) Trend analysis and ARIMA modeling for forecasting precipitation pattern in Wadi Shueib catchment area in Jordan Arabian Journal of Geosciences 12 doi:10.1007/s12517-018-4205-z

Alickovic E, Kevric J, Subasi A (2018) Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction Biomedical Signal Processing and Control 39:94-102 doi:https://doi.org/10.1016/j.bspc.2017.07.022

Balasubramanian G, Kanagasabai A, Mohan J, Seshadri NPG (2018) Music induced emotion using wavelet packet decomposition—An EEG study Biomedical Signal Processing and Control 42:115-128 doi:https://doi.org/10.1016/j.bspc.2018.01.015

Ben messaoud Ma, Bouzid A, Ellouze N (2016) Speech enhancement based on wavelet packet of an improved principal component analysis Computer Speech & Language 35:58-72 doi:https://doi.org/10.1016/j.csl.2015.06.001

Benedetto A (2002) A decision support system for the safety of airport runways: the case of heavy

436 rainstorms Transportation Research Part A: Policy and Practice 36:665-682
437 doi:https://doi.org/10.1016/S0965-8564(01)00029-5
438 Box G (2013) Box and Jenkins: Time Series Analysis, Forecasting and Control. In: Mills TC (ed) A
439 Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century.
440 Palgrave Macmillan UK, London, pp 161-215. doi:10.1057/9781137291264_6
441 Box GEP, Jenkins GM, Reinsel GC (1994) Time Series Analysis: Forecasting and Control. 3rd ed. .
442 NJ: Prentice Hall, Englewood Cliffs
443 Burlando P, Rosso R, Cadavid LG, Salas JD (1993) Forecasting of short-term rainfall using ARMA
444 models Journal of Hydrology 144:193-211 doi:http://dx.doi.org/10.1016/0022-1694(93)90172-6
445 Chong KL, Lai SH, Yao Y, Ahmed AN, Jaafar WZW, El-Shafie A (2020) Performance Enhancement
446 Model for Rainfall Forecasting Utilizing Integrated Wavelet-Convolutional Neural Network Water
447 Resources Management 34:2371-2387 doi:10.1007/s11269-020-02554-z
448 Feng Q, Wen XH, Li JG (2015) Wavelet Analysis-Support Vector Machine Coupled Models for
449 Monthly Rainfall Forecasting in Arid Regions Water Resources Management 29:1049-1065
450 doi:10.1007/s11269-014-0860-3
451 Guang-Bin H, Qin-Yu Z, Chee-Kheong S Extreme learning machine: a new learning scheme of
452 feedforward neural networks. In: 2004 IEEE International Joint Conference on Neural Networks (IEEE
453 Cat. No.04CH37541), 25-29 July 2004 2004. pp 985-990 vol.982. doi:10.1109/IJCNN.2004.1380068
454 Guo Z, Zhao W, Lu H, Wang J (2012) Multi-step forecasting for wind speed using a modified EMD-
455 based artificial neural network model Renewable Energy 37:241-249
456 Hartmann H, Snow JA, Su B, Jiang T (2016) Seasonal predictions of precipitation in the Aksu-
457 Tarim River basin for improved water resources management Global and Planetary Change 147:86-96
458 doi:https://doi.org/10.1016/j.gloplacha.2016.10.018
459 Hung NQ, Babel MS, Weesakul S, Tripathi NK (2009) An artificial neural network model for
460 rainfall forecasting in Bangkok, Thailand Hydrol Earth Syst Sci 13:1413-1425 doi:10.5194/hess-13-
461 1413-2009
462 Karthikeyan L, Nagesh Kumar D (2013) Predictability of nonstationary time series using wavelet
463 and EMD based ARMA models Journal of Hydrology 502:103-119
464 doi:http://dx.doi.org/10.1016/j.jhydrol.2013.08.030
465 Kisi O (2010) Wavelet regression model for short-term streamflow forecasting Journal of
466 Hydrology 389:344-353 doi:https://doi.org/10.1016/j.jhydrol.2010.06.013
467 Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly
468 streamflow forecasting Journal of Hydrology 399:132-140
469 doi:https://doi.org/10.1016/j.jhydrol.2010.12.041
470 Kisi O, Cimen M (2012) Precipitation forecasting by using wavelet-support vector machine
471 conjunction model Engineering Applications of Artificial Intelligence 25:783-792
472 doi:https://doi.org/10.1016/j.engappai.2011.11.003
473 Kuligowski RJ, Barros AP (1998) Experiments in Short-Term Precipitation Forecasting Using
474 Artificial Neural Networks Monthly Weather Review 126:470-482
475 Liu H, Mi X, Li Y (2018) Comparison of two new intelligent wind speed forecasting approaches
476 based on Wavelet Packet Decomposition, Complete Ensemble Empirical Mode Decomposition with
477 Adaptive Noise and Artificial Neural Networks Energy Conversion and Management 155:188-200
478 doi:https://doi.org/10.1016/j.enconman.2017.10.085
479 Mahmud I, Bari SH, Rahman MTU (2017) Monthly rainfall forecast of Bangladesh using

autoregressive integrated moving average method Environmental Engineering Research 22:162-168 doi:10.4491/eer.2016.075

McCelland J, Rumelhart D (1986) Parallel Distributed Processing. MIT Press, Cambridge, MA, USA

Mohammadi K, Eslami HR, Kahawita R (2006) Parameter estimation of an ARMA model for river flow forecasting using goal programming Journal of Hydrology 331:293-299 doi:http://dx.doi.org/10.1016/j.jhydrol.2006.05.017

Nastos PT, Paliatsos AG, Koukouletsos KV, Larissi IK, Moustris KP (2014) Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece Atmospheric Research 144:141-150 doi:https://doi.org/10.1016/j.atmosres.2013.11.013

Nguyen DH, Bae D-H (2020) Correcting mean areal precipitation forecasts to improve urban flooding predictions by using long short-term memory network Journal of Hydrology 584:124710 doi:https://doi.org/10.1016/j.jhydrol.2020.124710

Ortiz-García EG, Salcedo-Sanz S, Casanova-Mateo C (2014) Accurate precipitation prediction with support vector classifiers: A study including novel predictive variables and observational data Atmospheric Research 139:128-136 doi:https://doi.org/10.1016/j.atmosres.2014.01.012

Papalaskaris T, Panagiotidis T, Pantrakis A (2016) Stochastic Monthly Rainfall Time Series Analysis, Modeling and Forecasting in Kavala City, Greece, North-Eastern Mediterranean Basin Procedia Engineering 162:254-263 doi:https://doi.org/10.1016/j.proeng.2016.11.054

Partal T, Cigizoglu HK, Kahya E (2015) Daily precipitation predictions using three different wavelet neural network algorithms by meteorological data Stochastic Environmental Research and Risk Assessment 29:1317-1329 doi:10.1007/s00477-015-1061-1

Pektaş AO, Kerem Cigizoglu H (2013) ANN hybrid model versus ARIMA and ARIMAX models of runoff coefficient Journal of Hydrology 500:21-36 doi:http://dx.doi.org/10.1016/j.jhydrol.2013.07.020

Pham BT, Le LM, Le T-T, Bui K-TT, Le VM, Ly H-B, Prakash I (2020) Development of advanced artificial intelligence models for daily rainfall prediction Atmospheric Research 237:104845 doi:https://doi.org/10.1016/j.atmosres.2020.104845

Seo Y, Kim S, Kisi O, Singh VP, Parasuraman K (2016) River Stage Forecasting Using Wavelet Packet Decomposition and Machine Learning Models Water Resources Management 30:4011-4035 doi:10.1007/s11269-016-1409-4

Tan Q-F, Lei X-H, Wang X, Wang H, Wen X, Ji Y, Kang A-Q (2018) An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach Journal of Hydrology 567:767-780 doi:https://doi.org/10.1016/j.jhydrol.2018.01.015

Valipour M, Banihabib ME, Behbahani SMR (2013) Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir Journal of Hydrology 476:433-441 doi:http://dx.doi.org/10.1016/j.jhydrol.2012.11.017

Venkata Ramana R, Krishna B, Kumar SR, Pandey NG (2013) Monthly Rainfall Prediction Using Wavelet Neural Network Analysis Water Resources Management 27:3697-3711 doi:10.1007/s11269-013-0374-4

Wang JJ et al. (2017) Application of BP Neural Network Algorithm in Traditional Hydrological Model for Flood Forecasting Water 9 doi:10.3390/w9010048

Wang W-C, Chau K-W, Cheng C-T, Qiu L (2009) A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series Journal of Hydrology 374:294-306

524 doi:https://doi.org/10.1016/j.jhydrol.2009.06.019
525 Wang W-c, Chau K-w, Xu D-m, Chen X-Y (2015) Improving Forecasting Accuracy of Annual
526 Runoff Time Series Using ARIMA Based on EEMD Decomposition Water Resources Management
527 29:2655-2675 doi:10.1007/s11269-015-0962-6
528 Wang W-c, Xu D-m, Chau K-w, Chen S (2013) Improved annual rainfall-runoff forecasting using
529 PSO-SVM model based on EEMD Journal of Hydroinformatics 15:1377-1390
530 doi:10.2166/hydro.2013.134
531 Wang X, Kinsland G, Poudel D, Fenech A (2019) Urban flood prediction under heavy precipitation
532 Journal of Hydrology 577:123984 doi:https://doi.org/10.1016/j.jhydrol.2019.123984
533 Wei H, Li J-L, Liang T-G (2005) Study on the estimation of precipitation resources for rainwater
534 harvesting agriculture in semi-arid land of China Agricultural Water Management 71:33-45
535 doi:https://doi.org/10.1016/j.agwat.2004.07.002
536 Wu CL, Chau KW (2013) Prediction of rainfall time series using modular soft computingmethods
537 Engineering Applications of Artificial Intelligence 26:997-1007
538 doi:https://doi.org/10.1016/j.engappai.2012.05.023
539 Yaseen ZM, Sulaiman SO, Deo RC, Chau KW (2019) An enhanced extreme learning machine
540 model for river flow forecasting: State-of-the-art, practical applications in water resource engineering
541 area and future research direction Journal of Hydrology 569:387-408 doi:10.1016/j.jhydrol.2018.11.069
542 Yu C, Li Y, Xiang H, Zhang M (2018) Data mining-assisted short-term wind speed forecasting by
543 wavelet packet decomposition and Elman neural network Journal of Wind Engineering and Industrial
544 Aerodynamics 175:136-143 doi:https://doi.org/10.1016/j.jweia.2018.01.020
545 Zhang Q, Wang B-D, He B, Peng Y, Ren M-L (2011) Singular Spectrum Analysis and ARIMA
546 Hybrid Model for Annual Runoff Forecasting Water Resources Management 25:2683-2703
547 doi:10.1007/s11269-011-9833-y
548 Zhou H-c, Peng Y, Liang G-h (2008) The Research of Monthly Discharge Predictor-corrector Model
549 Based on Wavelet Decomposition Water Resources Management 22:217-227 doi:10.1007/s11269-006-
550 9152-x
551 Zhu L, Wang Y, Fan Q (2014) MODWT-ARMA model for time series prediction Applied
552 Mathematical Modelling 38:1859-1865 doi:https://doi.org/10.1016/j.apm.2013.10.002
553

554 **Figure captions**

555 Fig. 1. Sketch map of the WPD method.

556 Fig. 2. The framework of the proposed hybrid model.

557 Fig. 3. Study area and location of the Jinsha weather station.

558 Fig. 4. Annual rainfall time series in Jinsha station.

559 Fig. 5. Decomposed results for annual rainfall in Jinsha station.

560 Fig. 6. PACF values for all series.

561 Fig. 7. Observed and forecasted rainfall during training and testing period by six methods.

562

563

Table captions

Table 1. The number of Input variables for different series.

Table 2. ADF test results.

Table 3. ARIMA models based on BIC.

Table 4. Performance measures of models.

Table 5. Comparison results of model prediction performance.

Figures

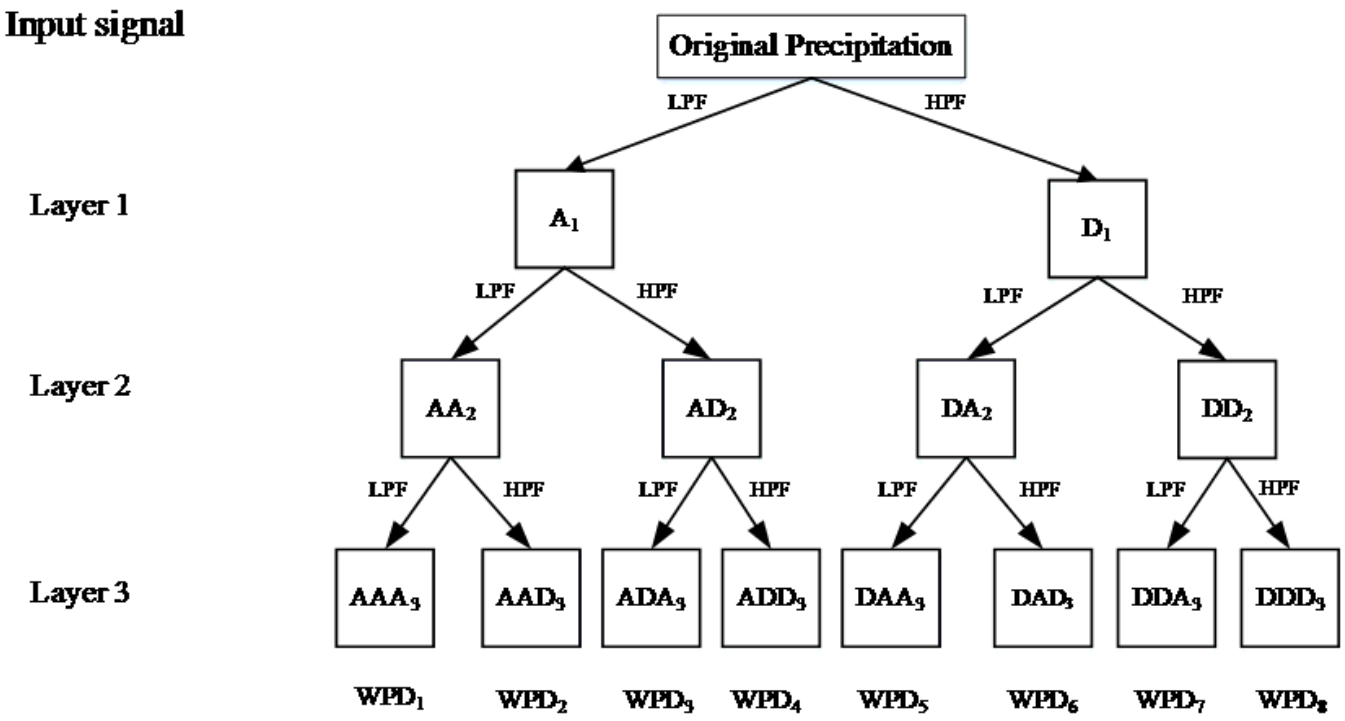

Figure 1

Sketch map of the WPD method.

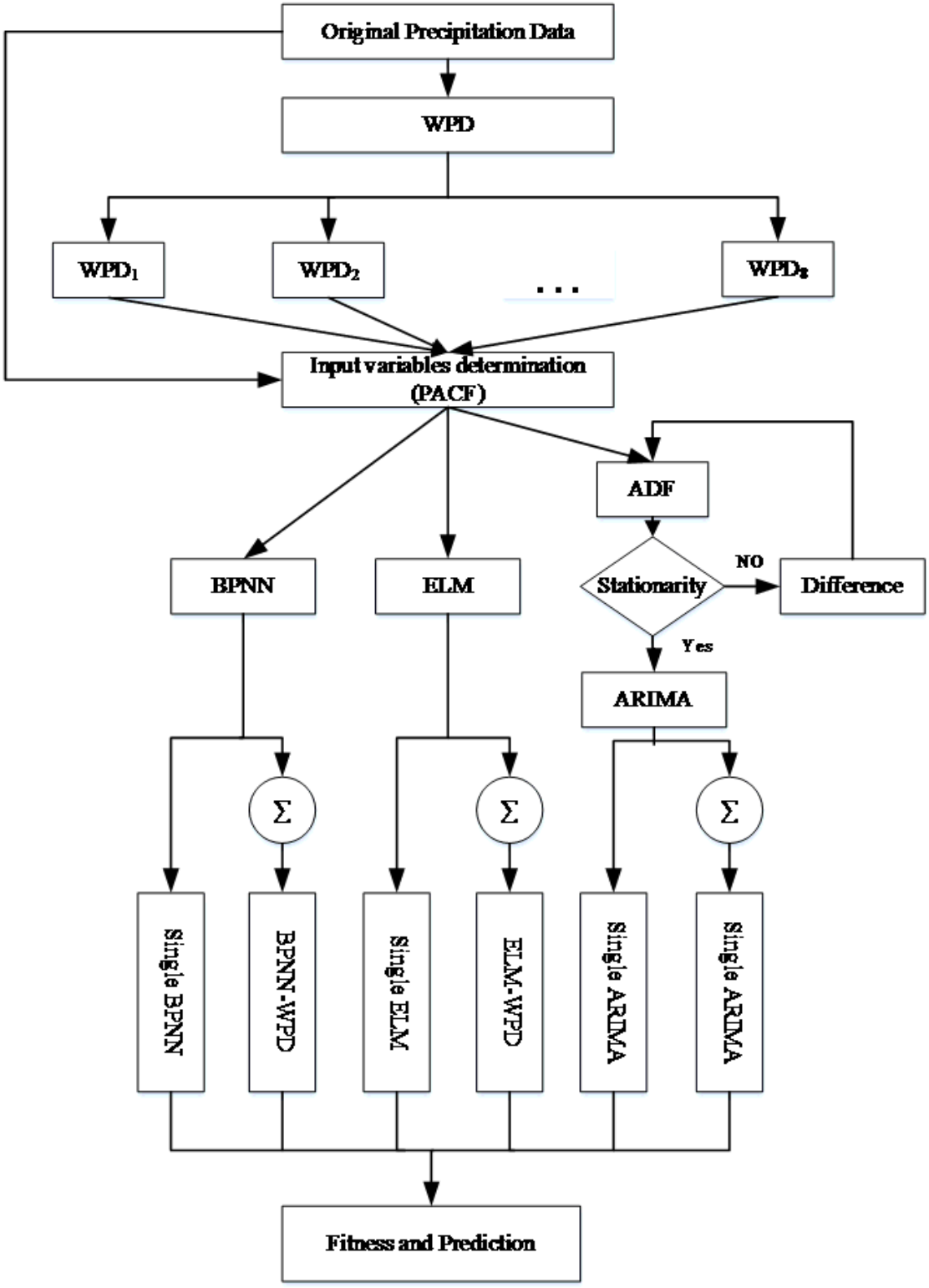

Figure 2

The framework of the proposed hybrid model.

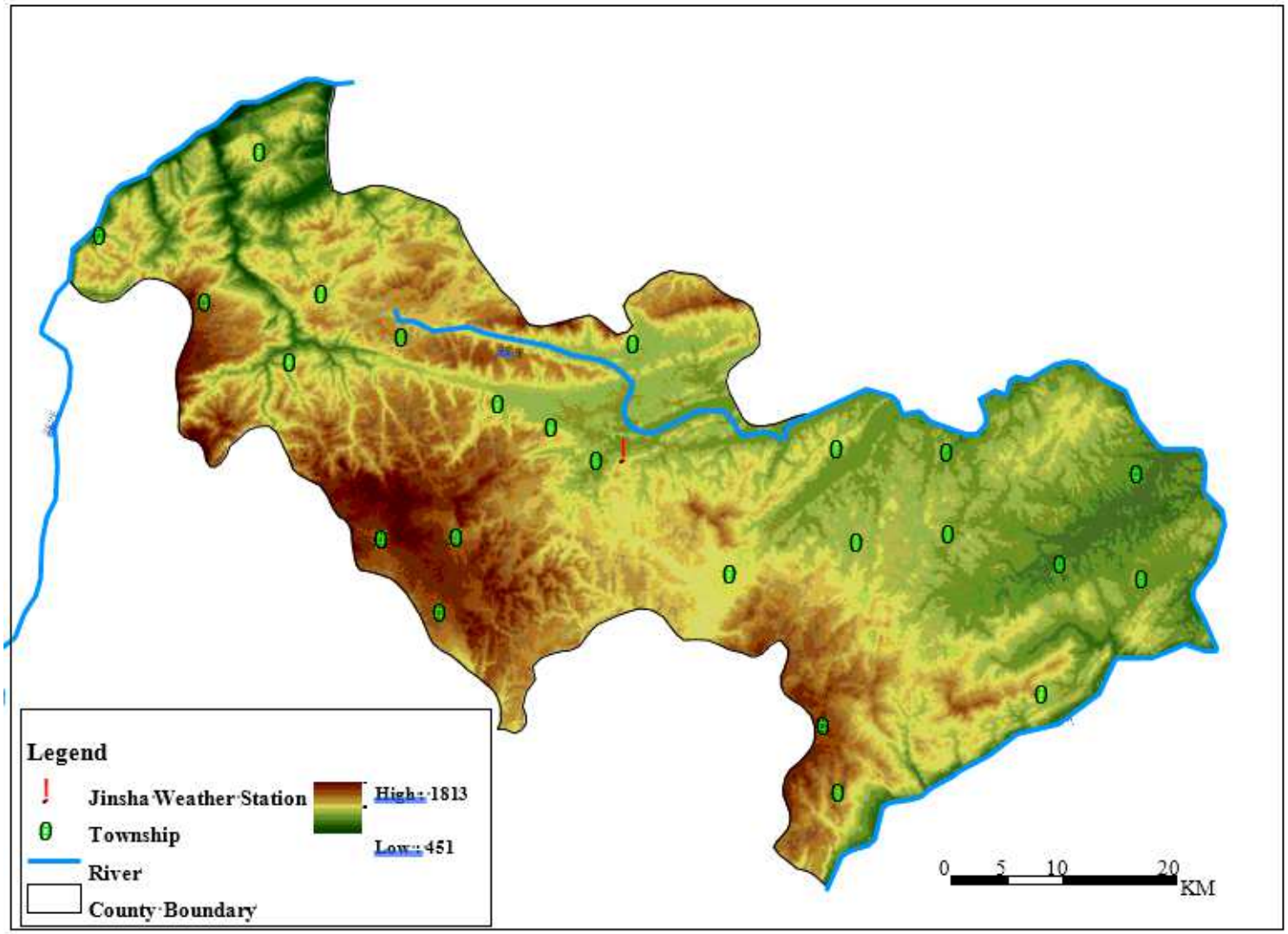

Figure 3

Study area and location of the Jinsha weather station. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area o bbnhjr of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

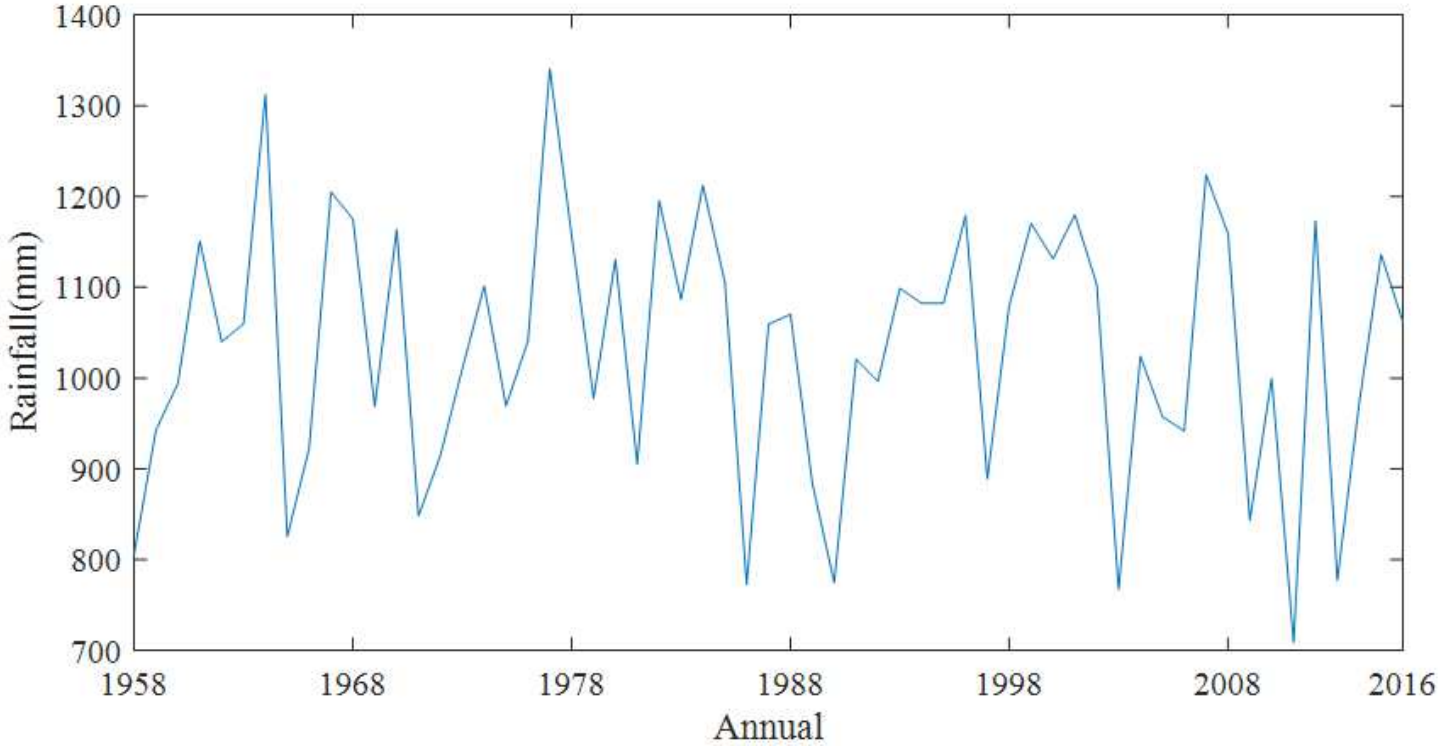

Figure 4

Annual rainfall time series in Jinsha station.

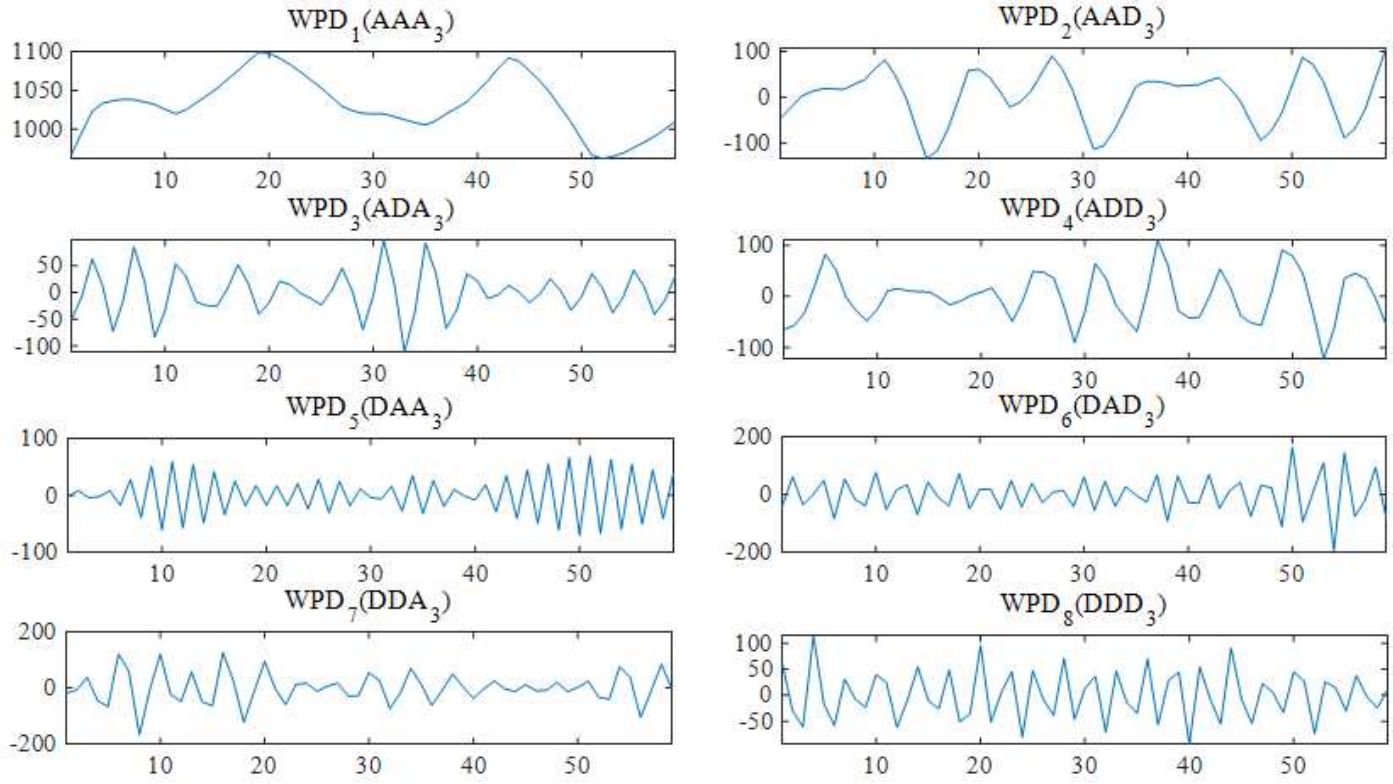

Figure 5

Decomposed results for annual rainfall in Jinsha station.

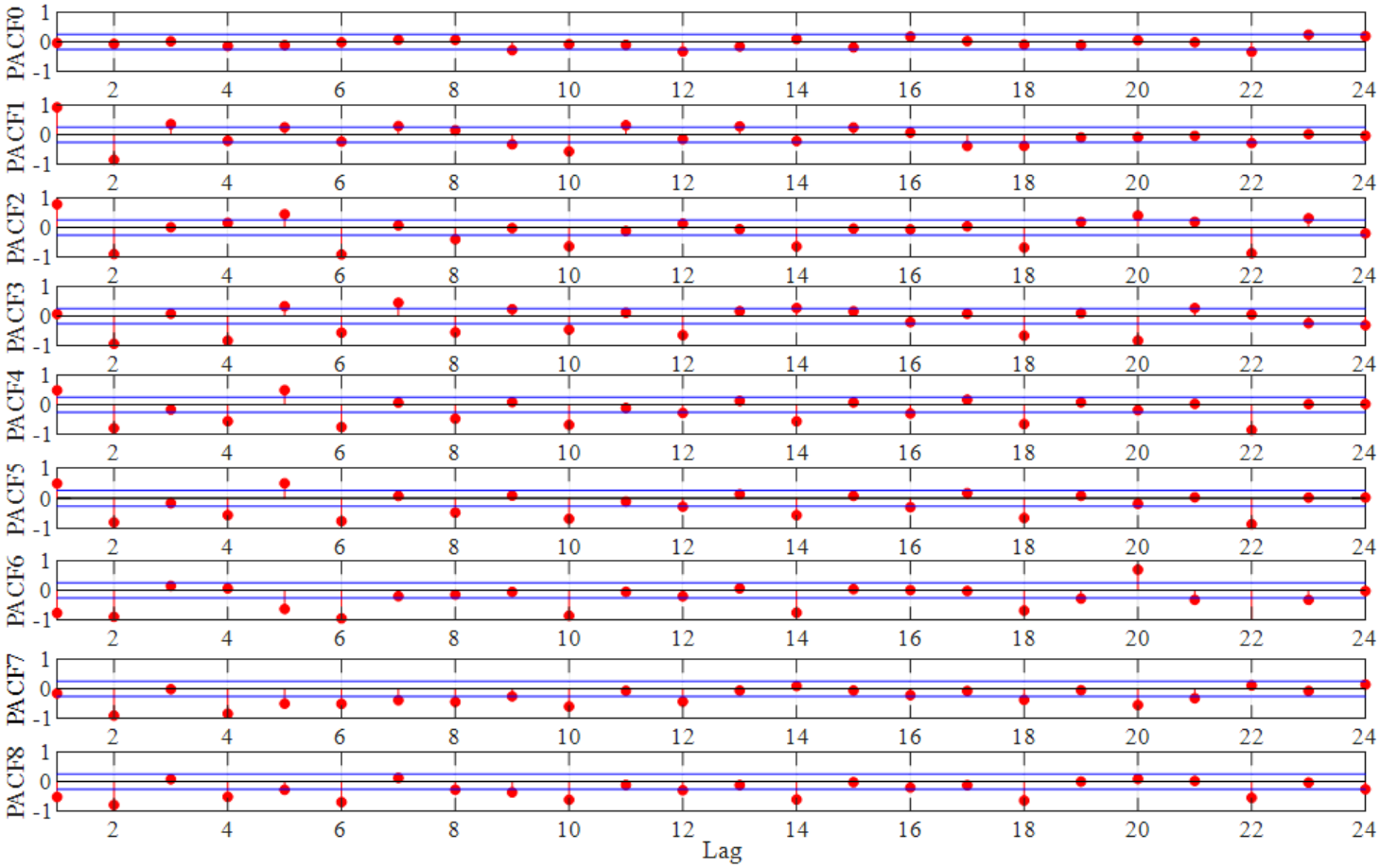

Figure 6

PACF values for all series.

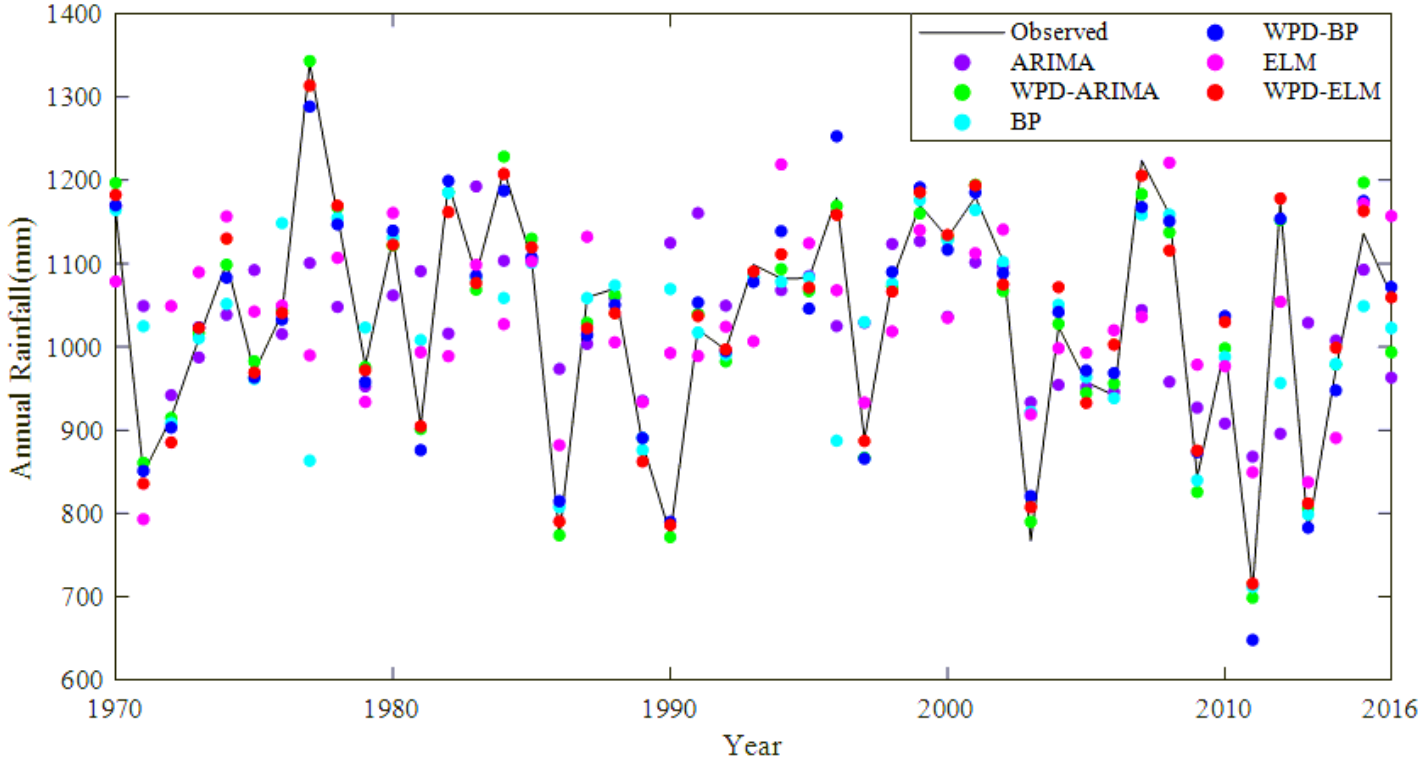

Figure 7

Observed and forecasted rainfall during training and testing period by six methods.