Rainfall time series prediction based on the DWT-SVR-Prophet hybrid model

Dongsheng Li  
Capital Normal University

Jinfeng Ma  
(jfma@rcees.ac.cn)  
Research Center for Eco-Environmental Sciences

Kaifeng Rao  
Research Center for Eco-Environmental Sciences

Xiaoyan Wang  
Capital Normal University

Ruonan Li  
Research Center for Eco-Environmental Sciences

Yanzheng Yang  
Research Center for Eco-Environmental Sciences

Hua Zheng  
Research Center for Eco-Environmental Sciences

Research Article

Keywords: rainfall time series prediction, wavelet decomposition, machine learning, DWT-SVR-Prophet coupled model

Posted Date: February 8th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2549227/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

The discrete wavelet method can be used to decompose rainfall time series into subseries of different frequencies. It would be worthwhile to investigate whether combining forecasting results from different frequency subseries could improve the accuracy of rainfall prediction. A novel DWT-SVR-Prophet (DSP) hybrid model for rainfall prediction is proposed in this paper. First, the rainfall time series is decomposed into high-frequency and low-frequency subseries using discrete wavelet transform (DWT). The SVR and Prophet models are then used to predict high-frequency and low-frequency subsequences, respectively. Finally, the predicted rainfall is determined by summing the predicted values of each subsequence. A case study in China is conducted from January 1, 2014, to June 30, 2016. The results show that the DSP model provides excellent prediction, with RMSE, MAE, $R^2$ values of 6.17, 3.3, and 0.75, respectively. The DSP model yields higher prediction accuracy than the three baseline models considered, with the prediction accuracy ranking as follows: DSP > SSP > Prophet > SVR. In addition, the DSP model is quite stable, and can achieve good results when applied to rainfall data from various climate types, with RMSEs ranging from 1.24 to 7.31, MAEs ranging from 0.52 to 6.14 and $R^2$ values ranging from 0.62 to 0.75. The proposed model may provide a novel approach for rainfall forecasting and is readily adaptable to other time series predictions.

1. Introduction

Rainfall, as an essential process in the hydrological cycle, is one of the most studied components of hydrological and climate science, as it directly or indirectly affects our society (Hammad et al. 2021; Tran-Anh et al. 2018). Accurate rainfall prediction is vital in daily life, risk assessment, natural disaster prevention, and water resource planning and management (Ni et al. 2020). The dynamic complexity and nonstationarity of measured hydrological data create significant challenges regarding rainfall prediction (Adaryani et al. 2022). Models used for hydrometeorological time series prediction can be divided into physical process-driven models and data-driven models (He et al. 2015). The former require complex equations to be solved with large amounts of data, and the latter learn long-term patterns of physical phenomena directly from data and can be quickly developed and easily implemented (Zhang et al. 2015).

Statistical and machine learning methods are often used to develop data-driven models. Traditional statistical models, such as the autoregressive moving average (ARIMA) and multiple linear regression (MLR) models, only yield satisfactory results when predicting linear or near-linear time series, and fail to capture nonlinear and nonstationary factors in hydrometeorological time series (Zhang et al. 2016). Machine learning techniques have a powerful learning capability ideal for modeling linear and nonlinear relationships in data without necessarily understanding the physical mechanisms associated with the data (Nourani et al. 2014). Therefore, various machine learning models have been applied for hydrometeorological time series prediction, such as the artificial neural network (ANN), support vector regression (SVR), Prophet and random forest (RF) models. These models provide satisfactory predictions of nonlinear hydrological and meteorological processes (Aksoy et al. 2018; Chen et al. 2018; Karevan et al. 2020).
Although these single methods have achieved breakthroughs in prediction performance, their shortcomings have also been exposed. For example, the SVR model has limitations in predicting highly nonstationary hydrological time series at different scales; notably, prediction performance is influenced by the size of the data set and the parameters and kernel functions used (Liu et al. 2014). The strong nonlinear mapping capability of neural networks has led to their increased application in climate data prediction. However, as the amount of data increases, the structure of neural networks becomes more complex, significantly increasing the processing speed and leading to convergence to local minima, resulting in low prediction accuracy (Aditya-Satrio et al. 2021).

In recent years, time series decomposition algorithms have been applied for feature extraction and prediction involving hydrometeorological time series. The combination of time series decomposition methods and machine learning has facilitated the development of hybrid models and improved prediction accuracy (Vivas et al. 2022; Zhang et al. 2022). Apaydin et al. (2021) used singular spectrum analysis to process monthly flow data and combined this approach with neural networks for prediction. The results showed that this method yields higher prediction accuracy than other single neural network models and can achieve more accurate river flow predictions. Ravansalar et al. (2017) constructed a wavelet linear genetic programming (WLGP) approach to predict monthly flows at two stations; they compared the WLGP model with an LGP model, a neural network model, a hybrid wavelet neural network model (WANN) and an MLR model. The results showed that the WLGP model could significantly improve the accuracy of flow predictions and other hydrological predictions. However, the decomposed time series components usually exhibit different characteristics, and it is difficult for a single prediction method to accurately predict all components. According to the characteristics of time series components, some scholars have used different models for prediction, to further improve prediction performance (Wang et al. 2021; Xiang et al. 2018).

Among the many time series decomposition methods, seasonal decomposition and wavelet decomposition have been widely used (Adib et al. 2021; He et al. 2022; Khan et al. 2020). In the seasonal decomposition method, a time series is divided into trend, periodic and residual terms, which represent different typical characteristics, and the predictability of the trend and periodic terms’ is generally high (Zhu et al. 2022). The residual term represents irregular fluctuations in time series with certain volatility and randomness, and thus, a machine learning model with excellent learning ability must be applied. This may result in poor prediction performance in cases with detailed features. In contrast, the wavelet transform approach, specific frequencies are filtered at each scale, and the original data are decomposed into single-frequency components, thus smoothing the sequence and making it more conducive to model fitting and the prediction of each component (Chong et al. 2020). The SVR model is based on the structural risk minimization criterion, which can capture the nonlinear features in data, is characterized by good generalization ability, and can improve prediction accuracy and speed (Shamshirband et al. 2015). This approach is suitable for high-frequency component forecasting. The Prophet model includes simple and intuitive parameters and yields a good prediction effect; notably, it can provide accurate predictions in cases with periodic data, many outliers and large trend changes (Huang et al. 2022). Thus, this model has advantages in the prediction of low-frequency components. Theoretically, the combination of DWT
with SVR and the Prophet model can improve the accuracy of time series prediction, and the performance of this method deserves further exploration. No study has yet applied such a combined model for rainfall time series prediction.

In this paper, a hybrid DWT-SVR-Prophet (DSP) model is proposed for rainfall time series prediction. In this model, we first decompose the original data into high-frequency components and low-frequency components using DWT, then, SVR is applied for high-frequency component prediction, and the Prophet model is used for low-frequency component prediction. Finally, the prediction results of each subseries are combined to obtain the final prediction results. The effectiveness of the DSP method is verified by using daily rainfall time series as a case study. Compared with three baseline methods, the superiority of the DSP approach is verified. Moreover, the DSP model is used to predict rainfall from stations in different climate types to further verify its universality. Next, the seasonal decomposition method is compared with DWT to demonstrate the advantages of the proposed approach for data preprocessing. Overall, the DSP method decomposes complex rainfall time series into subseries with a single fluctuation frequency, and different models are established for different components. This method improves the accuracy of rainfall time series prediction.

The rest of this paper is organized as follows. Section 2 presents the prediction method used in the DSP model, as well as the parameter optimization and model evaluation schemes. Detailed experimental results are given in Section 3. Section 4 discusses the applicability of the DSP model, the advantages of the DSP model in rainfall prediction, and improvements in the prediction accuracy compared to other models. Finally, a summary is presented in Section 5.

2. Methodology

2.1 Hybrid model based on DWT-SVR-Prophet

We propose a DSP model for rainfall time series prediction, and the framework shown in Fig. 1.

(i) Data preparation and preprocessing: We construct a dataset with daily rainfall data from the National Meteorological Center. The validity and superiority of the model introduced in this paper are verified using this dataset. We preprocess the measured rainfall data to ensure the fitting effect of the applied machine learning model. This process is described in detail in Section 2.2.

(ii) DWT processing: The rainfall time series are decomposed using DWT to obtain high-frequency subsequences with high randomness and volatility and low-frequency subsequences with high periodicity (see Section 2.2.1). This approach allows us to choose forecasting models based on the characteristics of each subseries.

( ) Hyperparameter optimization: The hyperparameters of the three methods in the coupled model are optimized to obtain the optimal prediction effect. Notably, the parameters of the DWT method are determined by referring to previous literature, and a grid search method is used to set the
hyperparameters of the SVR and Prophet models. The specific process of parameter selection is described in detail in Section 2.5.

(iv) **Rainfall prediction**: The optimized SVR model and Prophet model are used to predict high-frequency subseries and low-frequency subseries, respectively. The prediction results for each subseries are summed to obtain the final prediction results.

### 2.2 Data preprocessing

The data used in this paper were obtained from the National Weather Science Data Center (http://data.cma.cn). To construct the dataset, we selected daily rainfall data (from 1 January 2014 to 30 June 2016) from five stations in different climate zones. The rainfall statistics are shown in Table 1. Notably, station 59855, station 58345 and station 57348 are associated with high average annual rainfall and a high rainfall frequency. Station 54823 is characterized by moderate annual rainfall and the lowest rainfall frequency. Moreover, station 56018 displays the smallest annual average rainfall but a high rainfall frequency.

<table>
<thead>
<tr>
<th>Station (ID)</th>
<th>Rainfall (mm)</th>
<th>Rainfall frequency (%)</th>
<th>Climatic type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Per year</td>
<td>Per month</td>
<td>Daily maximum</td>
</tr>
<tr>
<td>Hainan (59855)</td>
<td>1972.2</td>
<td>182.6</td>
<td>253.1</td>
</tr>
<tr>
<td>Jiangsu (58345)</td>
<td>1362.7</td>
<td>126.2</td>
<td>154.8</td>
</tr>
<tr>
<td>Chongqi (57348)</td>
<td>881.9</td>
<td>81.7</td>
<td>113.6</td>
</tr>
<tr>
<td>Shandon (54823)</td>
<td>642.5</td>
<td>59.5</td>
<td>127.1</td>
</tr>
<tr>
<td>Qinghai (56018)</td>
<td>466.3</td>
<td>43.2</td>
<td>31.3</td>
</tr>
</tbody>
</table>

One of the prerequisites for a data-driven model is ensuring that the data quality meets the relevant modeling requirements (Luo et al. 2018). Therefore, before performing data decomposition, the measured rainfall data need to be preprocessed, such as filling missing values and normalizing the data. There were no missing values in the dataset used in this paper. To avoid the influence of extreme rainfall values and improve the accuracy of the machine learning algorithm, the maximum-minimum scaling method (Ponnoprat 2021) was adopted to normalize daily rainfall data to between 0 and 1. The specific formula is as follows:

\[ P_{\text{norm}} = \frac{P_i - P_{\text{min}}}{P_{\text{max}} - P_{\text{min}}} \]
where $P_{\text{norm}}, P, P_{\text{min}}$ and $P_{\text{max}}$ are the normalized, measured, minimum and maximum values of rainfall, respectively.

2.3 Methods used in the DSP model

2.3.1 Discrete wavelet transform

Wavelet transform is a time-frequency analysis method with multiresolution characteristics. The characteristics of an original sequence in different frequency bands is obtained by changing the corresponding scale (Ma et al. 2022). The main wavelet transform methods include continuous wavelet transform (CWT) and discrete wavelet transform (DWT). DWT is based on simple processing steps, avoids the redundancy problem of CWT and is more suitable for processing time series data (Liu et al. 2014). Therefore, DWT is chosen for rainfall time series decomposition, and the corresponding expression is:

$$
W(p, q) = 2^{-\frac{p}{2}} \sum_{t=0}^{T} \psi \left( \frac{t - q \cdot 2^p}{2^p} \right) \cdot x(t)
$$

where $t$ is the time parameter, $T$ is the signal length, $p$ is the scale parameter, and $q$ is the offset parameter.

The specific decomposition process is shown in Fig. 2. The rainfall time series is decomposed into $m$ subseries of different frequencies ($D_1, D_2, ..., D_m$, and $A_m$), which are used as the inputs of the coupled model. The frequency of the subsequences decreases from $D_1$ to $A_m$ in the above order.

2.3.2 Support vector regression model

SVR is an extended approach based on the support vector machine (SVM) concept. The benefits of this approach include structural risk minimization and the ability to use small sample sizes. It is an application of SVM in the field of regression (Essam et al. 2022). Its regression process is as follows.

The basic form of linear regression is:

$$
f(x) = \omega^T \cdot x + b
$$

For a given sample set $\{(x_i, y_i), i = 1, 2, ..., N\}$, $x_i \in R^n$ is the input quantity and $y_i \in R$ is the output quantity. Consider a mapping form, such that $\phi(x)$ is the eigenvector of $x$ after mapping to higher dimensions, yielding the following linear regression function:
where \( \omega \) is the coefficient, and \( b \in \mathbb{R} \) is the deviation. \( \omega \) after \( b \) is learned, the model is established.

According to the structural risk minimization criterion, the problem is transformed into an objective function \( R \) minimization problem, which can be expressed as:

\[
\min_{(\omega, e)} R(\omega, e) = \frac{1}{2} \omega^T \cdot \omega + C \sum_{i=1}^{n} e_i^2 \quad \text{s.t.} \quad y_i = \omega^T \varphi(x_i) + b + e_i \tag{5}
\]

where the equation after \( \text{s.t.} \) denotes the constraint that the objective function \( R \) should satisfy when minimized; \( e_i \in \mathbb{R} \) is the error variable, to be determined based on model training; and \( C \) is the penalty coefficient, which is greater than 0.

To solve the above minimization problem for the objective function \( R \), the Lagrangian function \( L \) is constructed.

\[
L(\omega, b, e_i, \alpha_i, \tilde{\alpha}_i) = R(\omega, e) - \sum_{i=1}^{n} \alpha_i \left[ \omega^T \varphi(x_i) + b + e_i - y_i \right] + \sum_{i=1}^{n} \tilde{\alpha}_i \left[ \omega^T \varphi(x_i) + b + e_i - y_i \right]
\]

The above equation is substituted into Eq. (3), and the following expression is obtained:

\[
\omega = \sum_{i=1}^{n} (\tilde{\alpha}_i - \alpha_i) \cdot x_i
\]

where \( \tilde{\alpha}_i, \alpha_i \) are the Lagrange multiplier operators, and can make the sample of \( \tilde{\alpha}_i - \alpha_i \neq 0 \) that is the support vector of SVR.

Finally, by satisfying the Karush-Kuhn-Tucker (TTK) condition in the SVR dual problem and substituting the resulting formula into Eq. (3), the SVR solution can be obtained as follows:

\[
f(x) = \sum_{i=1}^{n} (\tilde{\alpha}_i - \alpha_i) \cdot x_i^T x + b
\]
where bias term $b = y_i - \sum_{i=1}^{n} (-\alpha_i) \cdot x_i^T x$. The SVR solution when considering the form of feature mapping is:

$$f(x) = \sum_{i=1}^{n} (\tilde{\alpha}_i - \alpha_i) \cdot k(x_i, x) + b$$

where $k(x_i, x)$ is the kernel function. The commonly used kernel functions are the linear kernel function and the radial basis kernel function.

### 2.3.3 Prophet model

The Prophet model is an open-source decomposable time series prediction model that was developed by Facebook (Taylor et al. 2018). It is based on time series decomposition and machine learning fitting, which are used to predict the future trends of time series. The model consists of four main components:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t)$$

where $t$ is the current time; $y(t)$ is the current value; $g(t)$ is the trend term, which represents a nonperiodic variation of the time series; $s(t)$ is the period term, which reflects the cyclical or seasonal variation in the time series; $h(t)$ is the holiday-event term, which can be interpreted as an additional influence term; and $\varepsilon(t)$ is the error term, which obeys a normal distribution.

The trend term $g(t)$ can be based on either a logistic regression function or a segmented linear function. The segmented linear modelling equation is as follows:

$$g(t) = \left(k + a(t)^T \delta\right) t + \left(m + a(t)^T \gamma\right)$$

where $k$ denotes the growth rate of the model; $\delta$ is the change in $k$; $m$ is the offset; $t$ is the timestamp; and $a(t)$ is the indicator function. Additionally, $a(t)^T$ is the transpose vector of $a(t)$, and $\gamma$ is the offset of the smoothing process, which serves to make the function segment continuous. The logistic regression model is:

$$g(t) = \frac{c}{1 + \exp(-k(t - m))}$$
where \( c \) denotes the bearing capacity of the model; and the definitions of \( k \) and \( m \) are the same as those above.

The periodic term \( s(t) \), which models the periodicity of the time series using a Fourier series, is as follows:

\[
s(t) = \sum_{n=1}^{N} \left[ a_n \cos \left( \frac{2\pi nt}{P} \right) + b_n \sin \left( \frac{2\pi nt}{P} \right) \right]
\]

The parameters \( P \) and \( N \) can be adjusted according to the length of the selected period. For example, \( P = 365.25 \) and \( N = 10 \) when the period is annual, and \( P = 7 \) and \( N = 3 \) when the period is weekly.

The holiday-event model \( h(t) \) is:

\[
h(t) = \sum_{i=1}^{L} k_i \cdot 1_{\{t \in D_i\}}
\]

where \( Z(t) = \left( 1_{\{t \in D_1\}}, \ldots, 1_{\{t \in D_L\}} \right) \); \( k = (k_1, \ldots, k_L)^T \); \( L \) denotes the number of holiday events; \( D_i \) represents the time range of the holiday-event; \( Z(t) \) is set to 1 when time \( t \) is within the range of a holiday-event and equals 0 otherwise; and \( k_i \) indicates the influence of different holiday-events on the time series prediction; \( k \) obeys normal distribution.

### 2.4 Hyperparameter optimization

The prediction effect of the coupled model mainly depends on the selection of the model parameters, and the main parameters of each model are shown in Table 2. To obtain the optimal prediction effect, the three model parameters are optimized, and the detailed methods are as follows.

#### 2.4.1 DWT

The main parameters of DWT include the wavelet function and decomposition level. Among the standard wavelet functions, Daubechies (db) family wavelets are commonly used in hydrometeorology (Quilty et al. 2021; Wu et al. 2021). Nalley et al. (2012) used db5-db10 wavelets for DWT-based analyses of rainfall time series. Altunkaynak et al. (2015) performed a 3-level decomposition of rainfall time series and used the result as the input to a prediction model. Referring to the existing studies, we select db7 and 3-level decomposition for DWT data processing.

#### 2.4.2 SVR
We use a grid search method for parameter search optimization in the SVR model. Then, a 10-fold cross-validation technique is applied to the training set to prevent overfitting.

(i) The radial basis function (rbf) is chosen as the kernel function of the SVR model. First, the ranges of values and search steps are set for the main parameters $C$ and $\gamma$, and all parameter combinations within the given ranges are obtained $(\gamma_x, C_y), (x = 1, 2, \ldots, M; y = 1, 2, \ldots, N)$.

(ii) All parameter combinations are applied to rainfall predictions, and the best parameter combination is selected based on effect evaluation $(\gamma_j, C_k)$.

(iii) To ensure the stability of the search result, the adjacent interval of the optimal parameter combination is selected as the new search range $\gamma \in (\gamma_{j-1}, \gamma_{j+1}), C \in (C_{k-1}, C_{k+1})$. Then, the search step size is reduced by factor of 2 (or another multiple), and the optimal parameter combination is again obtained. If the result is unstable, the process is continued until a stable result, i.e., the optimal combination of parameters, is obtained.

In this paper, $C \in [2^{-10}, 2^{10}], \gamma \in [2^{-10}, 2^{10}]$, and the step size is 2.

2.4.3 Prophet model

A grid search method is used to search for the optimal parameters of the Prophet model, and the basic process is the same as that used for the SVR model. The following initial range settings are used for the parameters of the Prophet model.

(i) Both linear and logistic "Growth" parameters, and additive and multiplicative "Seasonality mode" parameters are considered.

(ii) The monthly period term is summed with the "Add_seasonality" function in the Prophet model, with "period" = 30.5. Then, the initial range of "Year_seasonality" and "Seasonality_prior_scale" is set to [1,100] with a step size of 5.

(iii) The "Changepoint_prior_scale" parameter has a range of [0.01,20] and the corresponding step size is 0.5.
### Table 2
Hyperparameters of the DWT, SVR and Prophet models

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>Parameters description</th>
<th>Default value</th>
<th>Optimization method</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWT</td>
<td>Wavelet name</td>
<td>Wavelet basis function</td>
<td>—</td>
<td>From</td>
</tr>
<tr>
<td></td>
<td>Level</td>
<td>Wavelet decomposition level</td>
<td>—</td>
<td>previous research</td>
</tr>
<tr>
<td>SVR</td>
<td>Kernel</td>
<td>Kernel function</td>
<td>rbf</td>
<td>Grid search</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>Penalty coefficient</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>γ</td>
<td>Kernel function coefficient</td>
<td>auto</td>
<td></td>
</tr>
<tr>
<td>Prophet</td>
<td>Growth</td>
<td>Function in the trend model</td>
<td>linear</td>
<td>Grid search</td>
</tr>
<tr>
<td></td>
<td>Changepoint_prior_scale</td>
<td>Trend flexibility</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year_seasonality</td>
<td>Year flexibility</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Seasonality_prior_scale</td>
<td>Seasonality flexibility</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Seasonality mode</td>
<td>Model learning style</td>
<td>additive</td>
<td></td>
</tr>
</tbody>
</table>

#### 2.5 Evaluation metrics

The mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination ($R^2$) are chosen as evaluation indicators, and the formulae are shown below:

\[
MAE = \frac{1}{N} \sum_{k=1}^{N} |y'_k - y_k|
\]  
\[
RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (y'_k - y_k)^2}
\]  
\[
R^2 = 1 - \frac{\sum_{k=1}^{N} (y'_k - \bar{y})^2}{\sum_{k=1}^{N} (y'_k - \bar{y})^2}
\]

where $y'_k$ is the ith predicted value and $y_k$ is the ith true value. Additionally, $\bar{y}$ is the average of the true values, and $N$ is the number of true values.
2.6 Open-Source Libraries

Our proposed model is implemented in a Windows environment using the Python language. The DWT, SVR and Prophet models are implemented by using the packages "pywt", "sklearn.svm" and "fbprophet", respectively. The data processing and visualization steps are based on "numPy", "panda" and "matplotlib", respectively.

3. Results

We divided the dataset into a test set and a validation set at a ratio of 8:2. Subsequently, a case study was conducted using the DSP model. Three baseline models were introduced for comparison. The single-station DSP model predictions, a comparison of the prediction model, and the multistation DSP model predictions are detailed in Sections 3.1, 3.2, and 3.3, respectively.

3.1 The DSP model provides accurate predictions of rainfall

To verify prediction accuracy of the DSP model proposed in this paper, we used data from station 57348 for rainfall prediction experiments. This station is located in Chongqing, China, with a subtropical, humid climate and abundant rainfall. The average annual rainfall at the station totals 881.9 mm over an average of 131 days. The distribution characteristics of the rainfall time series and the DWT decomposition results are shown in Fig. 3. The final prediction results were obtained using the DSP model.

The SVR model was used to predict three high-frequency components (D1, D2, and D3), and the subsequence with the highest frequency D1 was used as an example for parameter selection (Table 3) and prediction (results in Fig. 4). Figure 4 shows that the SVR model displays an excellent fitting ability for the high-frequency subsequence and yields satisfactory prediction results. The corresponding RMSE = 4.3709, MAE = 2.5513, and R² = 0.6757.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter values</th>
</tr>
</thead>
<tbody>
<tr>
<td>kernel function</td>
<td>rbf</td>
</tr>
<tr>
<td>C (penalty variable)</td>
<td>1024</td>
</tr>
<tr>
<td>γ(kernel function parameter)</td>
<td>0.03125</td>
</tr>
</tbody>
</table>

The Prophet model was applied to predict the low-frequency component A1, and the prediction results are shown in Fig. 5. The resulting RMSE, MAE and R² were 4.2192, 2.7367 and 0.3075, respectively. The model can simulate the basic trend of A1, and the prediction results are within the acceptable range. The main parameters of the Prophet model are shown in Table 4.
Table 4
Parameters setting of the Prophet model for the A1 sequence

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>Linear</td>
</tr>
<tr>
<td>Changepoint_prior_scale</td>
<td>1</td>
</tr>
<tr>
<td>Year_seasonality</td>
<td>9</td>
</tr>
<tr>
<td>Seasonality_prior_scale</td>
<td>60</td>
</tr>
<tr>
<td>Seasonality mode</td>
<td>Additive</td>
</tr>
</tbody>
</table>

The prediction results of all components are summed to obtain the final prediction time series. Figure 6 shows the scatter plots of the real and predicted rainfall values based on the validation set, and Fig. 7 compares the true and predicted values. The RMSE, MAE and $R^2$ of the DSP model are 6.1704, 3.2901 and 0.7518, respectively, indicating that among the studied models, the DSP model provides the best prediction of the basic trend of daily rainfall at station 57348. The predicted rainfall is generally similar to the actual rainfall, with satisfactory prediction accuracy.

3.2 The prediction accuracy of the DSP model is higher than that of the baseline models

To validate the superiority of the DSP model, four experiments were conducted involving daily rainfall prediction at station 57348 using three baseline models and the DSP model. The three baseline models were the SVR, Prophet, and coupled SVR-Prophet models based on the seasonal decomposition method (SSP), proposed by Guo et al. (2021). The predicted and actual values of the four models based on the validation set are shown in Fig. 8. Notably, both the DSP model and the SSP model exhibit good performance, and their prediction results are generally in agreement with the actual values. The main difference is that the rainfall peak predicted by the SSP model is much lower than the observed value, and the corresponding prediction error is large. The DSP model displays better performance in fitting the detailed features of the rainfall series. The prediction accuracy of the four models was compared (Table 5), and the DSP model obtained the best evaluation results, followed by the SSP model. Compared with those of the SSP model, the RMSE and MAE of the DSP model were reduced by 46.3% and 55.9%, respectively, and the $R^2$ was improved by 67.4%. The results above verify the superiority of the DSP model.
Table 5
Comparison of the prediction accuracy of the DSP, SSP, SVR and Prophet models for rainfall at station 57348

<table>
<thead>
<tr>
<th>Metric</th>
<th>DSP</th>
<th>SSP</th>
<th>SVR</th>
<th>Prophet</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>3.2901</td>
<td>4.3931</td>
<td>9.5772</td>
<td>4.9510</td>
</tr>
<tr>
<td>R²</td>
<td>0.7518</td>
<td>0.4492</td>
<td>-0.3061</td>
<td>0.0348</td>
</tr>
</tbody>
</table>

To further investigate the reasons for the difference in prediction effectiveness between the DSP model and the SSP model, comparative experiments are designed. The main difference between the two models is that DWT and Seasonal_decompose are utilized for time series decomposition. According to the principles of the decomposition methods, most of the detailed features of rainfall time series are associated with the high-frequency components in DWT and the residual terms in the Seasonal_decompose approach. Therefore, the rainfall data from station 57348 are used as an example, and the SVR model is applied to predict the D1 subsequence with the highest frequency in DWT and the residual term of Seasonal_decompose at the same time. Then, we compare the distribution characteristics and prediction results for components in detail, as shown in Fig. 9. The prediction results of D1 fit the trend of the actual values, and good prediction performance is observed; in contrast, the prediction results based on the residual terms are poor, and variations in the components are not well predicted. D1 yields the smallest prediction error, with RMSE, MAE, and $R^2$ values of 4.3486, 2.5343, and 0.6754, respectively. These findings indicate why the DSP model is superior to the SSP model in terms of peak rainfall prediction.

3.3 The DSP model displays outstanding stability

To evaluate the generalization ability of the DSP model, we conducted prediction experiments with data from stations in different climate zones. Additionally, we compared the prediction results of the DSP model with those of three baseline models. Figure 10 compares the predicted and true at each station. The DSP model fits the general trend of data at different stations, and the prediction results are in good agreement with the actual values. The station-scale prediction accuracy metrics for the four models are given in Table 5. The ranges of RMSE and MAE of the DSP model are 1.24364 to 7.3116 and 0.5197 to 6.1431, respectively, and most $R^2$ values range from 0.6217 to 0.7518. We select $R^2$ as the evaluation index to further evaluate the performance of each model in different cases of rainfall prediction (see Fig. 11). We found that most of the $R^2$ values of the DSP model fluctuated slightly but remained high at different stations. The $R^2$ values of the SSP model were not stable and fluctuated considerably. Additionally, the $R^2$ values of the Prophet and SVR models were consistently low. Compared with other baseline models, the DSP model displays stronger generalization ability and yields stable prediction results.
Table 5
Comparison of the prediction accuracy evaluation indexes for the different models at 5 stations in different climate regions

<table>
<thead>
<tr>
<th>Station</th>
<th>Metric</th>
<th>DSP</th>
<th>SSP</th>
<th>SVR</th>
<th>Prophet</th>
</tr>
</thead>
<tbody>
<tr>
<td>59855</td>
<td>RMSE</td>
<td>7.3116</td>
<td>13.5961</td>
<td>19.9489</td>
<td>13.0716</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>3.3035</td>
<td>7.4914</td>
<td>15.2908</td>
<td>4.4126</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.6632</td>
<td>-0.1647</td>
<td>-1.5074</td>
<td>-0.0766</td>
</tr>
<tr>
<td>56018</td>
<td>RMSE</td>
<td>1.2364</td>
<td>2.0519</td>
<td>2.9356</td>
<td>1.9336</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.5197</td>
<td>0.9194</td>
<td>2.7219</td>
<td>0.9101</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.6330</td>
<td>-0.0107</td>
<td>-1.0688</td>
<td>0.1024</td>
</tr>
<tr>
<td>58345</td>
<td>RMSE</td>
<td>8.8391</td>
<td>13.2151</td>
<td>17.0824</td>
<td>13.7878</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>6.1431</td>
<td>8.0558</td>
<td>12.7209</td>
<td>7.1224</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.6217</td>
<td>0.1543</td>
<td>-0.4131</td>
<td>0.0794</td>
</tr>
<tr>
<td>54823</td>
<td>RMSE</td>
<td>4.8553</td>
<td>4.9367</td>
<td>9.7700</td>
<td>5.7756</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>2.3643</td>
<td>2.8757</td>
<td>8.3983</td>
<td>2.4138</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.3214</td>
<td>0.2985</td>
<td>-1.7476</td>
<td>0.0398</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>3.2901</td>
<td>4.3931</td>
<td>9.5772</td>
<td>4.9510</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.7518</td>
<td>0.4492</td>
<td>-0.3061</td>
<td>0.0348</td>
</tr>
</tbody>
</table>

4. Discussion

4.1 The DSP model achieves accurate forecasts of rainfall time series

Since rainfall time series contain both periodic and nonstationary complex variations, hybrid methods are beneficial for overcoming the limitations of rainfall prediction (Fahad et al. 2023). Therefore, we propose a coupled DSP model for rainfall time series prediction, and satisfactory results were obtained (Fig. 6 and Fig. 7). Figure 8 shows that the DSP model exhibits the best prediction performance. The model accurately predicted the basic trends of the rainfall series and peak rainfall, and the predicted values generally matched the actual values. The Prophet model effectively fit the overall trend of the rainfall series, and the prediction curve was relatively flat, but the results of peak rainfall prediction were poor.
The SVR model can capture the variation characteristics of rainfall series, but it is influenced by the peak rainfall, and the predicted values in some periods with little rainfall are much higher than the true values.

The Prophet model can effectively fit the trend and period variations of time series, but the lack of consideration of residual autocorrelation leads to its poor fitting ability for complex models (Guo et al. 2021). The SVR model displays good generalization ability and is suitable for nonlinear prediction; however, it has limitations for general data. The DSP model successfully decomposes the periodic and nonstationary features in the rainfall time series using DWT and preserves them in different subseries. Moreover, the advantages of SVR and the Prophet model are combined to fit and predict subseries containing various features. Overall, the DSP model effectively fits the periodic and nonstationary factors in rainfall time series and achieves accurate predictions of rainfall time series.

4.2 The DSP model effectively captures the detailed features of rainfall time series

Achieving reliable predictions of rainfall time series is challenging, especially in the context of peak rainfall prediction (Pihrt et al. 2022). Figure 8 illustrates the good prediction accuracy of the DSP model and the SSP model, as well as the superior peak rainfall prediction capability of the DSP model. Based on the application of different data decomposition methods, Fig. 9 compares the distribution characteristics and prediction effects of the components of the two models. The components obtained in DWT are generally associated with a single frequency and yield better prediction results than those in other cases. Notably, DWT is used to decompose time series while filtering some frequencies, making each component smoother. The excellent scale decomposition ability of DWT reduces the difficulty of fitting each component used in machine learning. After the extraction of trend and period terms, the residual terms in the Seasonal_decompose method are characterized by strong randomness and instability, thus limiting the performance of subsequent prediction models. Therefore, the DSP model based on DWT for data decomposition can best capture the detailed features of rainfall sequences and achieve accurate predictions of peak rainfall.

4.3 Generalization of DSP models

Because the distribution characteristics of rainfall time series are influenced by the climate type, variations in the rainfall amplitude and frequency are impactful. Therefore, the rainfall time series from different climate types have different variation characteristics, affecting the results of prediction models. Multistation experiments involving the DSP model demonstrated its stable prediction ability (Fig. 9). Table 5 and Fig. 10 show that the DSP model yields high prediction accuracy at most stations. Only at station 54823 is there a significant decrease in accuracy, with an $R^2$ of 0.3214. Although the prediction accuracy at this station is low, the trend of the prediction results is similar to the actual trend, and the predictions provide reference significance to some extent. At different stations, the DSP model produced reliable predictions, and the prediction accuracy was higher than that of the compared models. The above results fully verify that the DSP model provides outstanding prediction accuracy and stability in applications involving rainfall time series prediction. For other fields (e.g., the energy field, transportation
field, and others), the complex time series of interest have the same characteristics as rainfall time series and contain both periodic and nonlinear variations. Theoretically, the DSP model can potentially achieve good performance in these fields and is worthy of further evaluation and application.

4.4 Improvement of the DSP model prediction accuracy

The feature extraction and decomposition effect of DWT influence the inclusion of single-variation features in subsequences, thus affecting the fitting results during machine learning and the prediction performance of the DSP model. The selection of the wavelet basis function and decomposition scale influences the quality of DWT feature extraction and decomposition. However, in this study, the wavelet basis function and decomposition level used in DWT are based on those reported in a previous study. Since we have yet to optimize the parameters of DWT, we cannot guarantee optimal decomposition. As shown in Fig. 11, the prediction accuracy of the DSP model decreased for station 54823. This station is located in an area with a temperate monsoon climate, and the rainfall frequency is 20.7%, much lower than that at other stations. Thus, the DWT parameters used in this paper may not be suitable for station 54823 based on its unique rainfall data characteristics, leading to the failure of the DWT to achieve effective feature extraction and decomposition and thus affecting the prediction accuracy of the DSP model. Moreover, the prediction accuracy of the A1 component in Fig. 5 is low, and there is a possibility that some high-frequency signals in the A1 component are not filtered, which affects the periodic fitting of the Prophet model. In subsequent work, the selection of DWT wavelet basis functions and decomposition levels for rainfall time series can be optimized with different features to adequately extract and decompose the different frequencies and characteristics of time series. The application of DWT decomposition after parameter optimization can further enhance machine learning and improve the accuracy and stability of rainfall time series prediction.

5. Conclusions

Accurate rainfall prediction is important to people's production and life. The ability of coupled models based on DWT to improve the accuracy of rainfall time series prediction is worthy of further study. In this paper, the DSP model is introduced for rainfall time series prediction. First, the DWT method is used to decompose rainfall time series into high-frequency and low-frequency components. An SVR model is then used to predict the high-frequency components, and the Prophet model is used to predict the low-frequency components. Finally, the prediction results for each component are summed to obtain the final rainfall predictions.

The results of case studies show that the DSP model can accurately predict rainfall time series with RMSE, MAE, and $R^2$ values of 6.17, 3.29, and 0.75, respectively. The DSP model significantly improves the prediction accuracy, and the results are compared with those of three baseline methods. Notably, the performance of the methods ranks as follows: DSP > SSP > Prophet > SVR. Compared with those of the SSP model, the RMSE and MAE of the DSP model are reduced by 46.3% and 55.9%, respectively, and $R^2$ is improved by 67.4%, verifying that the DSP model is an excellent prediction model. Moreover, the DSP model displays excellent prediction capability for peak rainfall events. The detailed feature prediction
results of DSP model display well, with RMSE = 4.35, MAE = 2.53, and $R^2 = 0.68$. The DSP model also exhibits good generalization capability. The model achieves reliable predictions of rainfall time series with different features. The calculated RMSEs ranged from 1.24 to 7.31, MAEs ranged from 0.52 to 6.14, and most $R^2$ values ranged from 0.62 to 0.75. The DSP model can be applied in other fields in which time series are also periodic and nonstationary (e.g., the transportation and energy fields). Thus, the prediction performance of DSP model deserves further study in additional applications.

**Declarations**

**Funding**

This work was supported by the National Key Research and Development Program of China (2022YFF1301803).

**Competing Interests**

The authors have no relevant financial or non-financial interests to disclose.

**Author Contributions**

Dongsheng Li mainly had the contribution of drafting, Dr. Jingfeng Ma and Hua Zheng had the contribution of designing the whole research framework and drafting, Kaifeng Rao and Dr. Xiaoyan Wang were responsible for related data processing, Dr. Ruonan Li and Yanzheng Yang contributed to the analysis. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Acknowledgments**

The authors are grateful to the funding support from the National Key Research and Development Program of China (2022YFF1301803).

**References**


https://doi.org/10.1016/j.jhydrol.2020.125380


**Figures**
Figure 1

Main flowchart of the hybrid model
Rainfall data → DWT → m subsequences

- High frequency
- Low frequency

Figure 2

Time series decomposition process

Figure 3

Rainfall time series and DWT decomposition results at station 57348
Figure 4

Scatter plot of the predicted and true values of the D1 sequence

- $R^2 = 0.6757$
- $MAE = 2.5513 \text{ mm}$
- $RMSE = 4.3709 \text{ mm}$
Figure 5

Scatter plot of the predicted and true values of the A1 sequence

R² = 0.3075
MAE = 2.7367 mm
RMSE = 4.2192 mm
Figure 6

Scatter plot of true and predicted rainfall values

- $R^2 = 0.7518$
- $MAE = 3.2901\ mm$
- $RMSE = 6.1704\ mm$
Figure 7

Comparison of true and predicted rainfall values
Figure 8

Comparison of true rainfall values and those predicted with the DSP, SSP, SVR and Prophet models at station 57348
Figure 9

Comparison of the variation characteristics and prediction results of different components (D1 and residual terms)
Figure 10

Comparison of the true values and those predicted with the DSP model: (a) Station 59855; (b) Station 56018; (c) Station 58345; (d) Station 54823
Figure 11

Comparison of $R^2$ values for 4 models based on predictions at different stations