



A Novel Coupled Model for Monthly Rainfall Prediction Based on ESMD-EWT-SVD-LSTM

Ziyu Li^{1,2,3} · Xianqi Zhang^{1,2,3}

Received: 1 January 2024 / Accepted: 1 March 2024 / Published online: 19 March 2024
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract

Precise predicting of rainfall is paramount for effective water resource management, ecological conservation, and the prevention of droughts and floods. Influenced by numerous variables, the process of rainfall is complex and the rainfall series exhibit high degrees of nonlinearity, making it challenging for traditional statistical prediction models to accurately capture the characteristics of rainfall series. Therefore, this paper proposes a new coupled model for predicting monthly rainfall based on Extreme-Point Symmetric Mode Decomposition (ESMD), Empirical Wavelet Transform (EWT), Singular Value Decomposition (SVD) and Long Short-Term Memory Neural Network (LSTM). By training and evaluating the ESMD-EWT-SVD-LSTM model on Kaifeng City's monthly rainfall data from 2009 to 2020 and comparing its predictions with those of the ESMD-SVD-LSTM, SVD-LSTM, LSTM models, the analysis reveals that: the quadratic decomposition of ESMD-EWT and SVD denoising can further reduce the complexity of rainfall data, obtain more predictable feature IMFs, and enhance the precision in LSTM predicting; in comparison with alternative models, the ESMD-EWT-SVD-LSTM coupled model shows the highest accuracy in predicting results, with MAE of 4.96, RMSE of 6.13, and SI of 0.12, indicating that the ESMD-EWT-SVD-LSTM model has strong nonlinear process learning ability and accuracy in regional monthly rainfall prediction. This study can offer dependable scientific grounding and technical assistance for regional rainfall predicting, water resources planning, and disaster mitigation.

Keywords Extreme-point symmetric mode decomposition · Empirical wavelet transform · Singular value decomposition · Long short-term memory neural network · Rainfall prediction

Highlights

- (1) ESMD can decompose rainfall series into multiple real feature IMFs.
- (2) Combining ESMD and EWT methods for data preprocessing can further simplify and extract information from rainfall series.
- (3) The IMFs after ESMD-EWT decomposition and SVD processing can effectively improve the prediction accuracy of LSTM.
- (4) The coupled model of ESMD-EWT-SVD-LSTM has a better prediction performance and strong applicability for nonlinear rainfall series.

Extended author information available on the last page of the article

1 Introduction

Rainfall plays a crucial role in the terrestrial water cycle, as global warming leads to increased climate instability, extreme rainfall and drought events are occurring frequently, introducing numerous uncertainties into the sustainable growth of the local economy and the well-being of its inhabitants. (Aydin and Sevgi 2022; Fatimi et al. 2023; Zhuang et al. 2023). Therefore, scientific and accurate rainfall prediction based on hydrological characteristics is important for effective hydrological planning and development of strategies to manage rainfall risks.

The hydrological processes, such as rainfall, are characterized by complexity and fluctuation, making their precise forecasting a key issue in contemporary research (Moraga et al. 2022; Li et al. 2012; Zhou et al. 2017). During the past few years, scholars have extensively researched rainfall predicting and attained remarkable outcomes (Wu et al. 2021; Yang et al. 2022). Currently, the common practice for rainfall prediction is analyzing available rainfall data and constructing corresponding prediction models (Chen et al. 2022). The GP, SVR, ANN and other classical time series modeling methods are now effectively used in hydrological forecasting (Chadalawada et al. 2017; Kajewska Szkudlarek 2020; Devi et al. 2016). However, owing to the randomness of rainfall and the influence of climate, the actual rainfall may exhibit significant fluctuations within a single year, rendering the traditional statistical methods inadequate for accurate quantitative prediction of rainfall (Olmo and Bettolli 2022). As computing power increases, deep learning neural networks have found widespread applications in various fields (Han et al. 2022; Afshari Nia et al. 2023; Chen and Wang 2022; Hewamalage et al. 2021). Byun et al. (2023) analysed cloud image data using the CNN algorithm for regional rainfall prediction. Zhang et al. (2021) predicted the daily runoff of a river using the RNN model. However, due to its algorithmic reasons, the RNN model was prone to gradient explosion. In response to this issue, the Long Short-Term Memory (LSTM) model was proposed (Gao et al. 2020). Ni et al. (2020) confirmed the effectiveness of LSTM in predicting time series by examining precipitation data. Chen et al. (2021) conducted an analysis of the rainfall and runoff processes in the Shishengchuan basin and discovered that the LSTM model had the capability to produce highly accurate rainfall predictions. Zhu et al. (2023) combined high-resolution radar rainfall data to propose an LSTM-based approach to radar rainfall nearing prediction.

Rainfall time series are always volatile and nonlinear, which seriously affects the performance of prediction models (Chen et al. 2022; Nourali 2023; Ehteram et al. 2023). Therefore, the choosing of techniques capable of handling non-stationary time series is the essential foundation for the subsequent enhancement of rainfall prediction accuracy. EMD algorithm is a traditional signal feature extraction method, but it is prone to mode mixing (Johnny et al. 2022; Zheng et al. 2021). As technology evolves and modernizes, methods such as ESMD, EWT, and SVD have been proposed and effectively used in the feature recognition process of prediction data. Gao et al. (2023) applied ESMD to decompose denoised wind speed time series adaptively and proposed FO-BSO-LSSVM for multistep wind speed prediction. Karijadi et al. (2023) utilized the EWT method for secondary processing of high-frequency IMFs and combined it with the LSTM method to achieve wind power prediction. Zhang and Chen (2022) extracted the features of wind energy data by CEEMDAN-SVD and performed wind speed prediction.

The efficiency of ESMD, EWT, SVD, and LSTM methods in data processing and prediction has been verified in practical engineering, however, these methods are less applied in the field of hydrology and the practical effects of the coupled algorithms have not been

fully explored. Therefore, for the practical needs of rainfall prediction, a coupled model based on EMD-EWT-SVD-LSTM is proposed. Firstly, for the rainfall observation sequence of Kaifeng City, Henan Province, ESMD-EWT is applied to perform quadratic decomposition to simplify the rainfall sequence features and obtain the characteristic IMFs; Secondly, the high-frequency IMFs obtained after quadratic decomposition are subjected to SVD noise reduction to obtain the main features; Thirdly, the final predicted rainfall sequence is obtained by predicting and integrating by the LSTM model for all the IMFs; Finally, the results of ESMD-EWT-SVD-LSTM prediction are juxtaposed with alternative models to validate applicability and accuracy of the model. This study explores a new regional rainfall prediction method, which can improve the accuracy effectively and offer a reliable scientific basis for regional water resources planning and disaster mitigation.

2 Method and Methodology

2.1 Extreme-Point Symmetric Mode Decomposition

Wang and Li (2013) proposed ESMD based on the EMD algorithm, which is applicable across various engineering fields involving data processing, such as information science, atmospheric science, and economics.

The ESMD algorithm works as follows:

Step 1: Find all maxima and minima in the data X and denote them as $E_i (i = 1, 2, 3, \dots, n)$ in order.

Step 2: Connect neighboring extreme points with the help of a straight line and mark the midpoint of each section of the line as $F_i (i = 1, 2, 3, \dots, n)$.

Step 3: Complement the midpoints F_0, F_n of the left and right boundaries by linear interpolation.

Step 4: Construct the interpolation lines $L_1, \dots, L_p (p \geq 1)$ for the obtained midpoints $n + 1$ and calculate the average value of the interpolation lines according to $L^* = (L_1 + \dots + L_p) / p$.

Step 5: Repeat the above steps for $X - L^*$ until $|L^*| \leq \varepsilon$, where ε is a predetermined tolerance error, usually taken as $\varepsilon = 0.001\delta_0$, δ_0 is as follows:

$$\delta_0^2 = \frac{1}{N} \sum_{i=1}^N \left(x_i - \frac{1}{N} \sum_{i=1}^N x_i \right)^2 \quad (1)$$

where: N is length of data.

In addition to the allowable error ε , the maximum number of decompositions K is also an adjustable parameter. When the decomposition process makes $|L^*| \leq \varepsilon$ or reaches the predetermined maximum number of decompositions K , the first empirical mode M_1 is decomposed.

Step 6: Repeat the above steps for $X - M_1$, obtain M_2, M_3, \dots in order, and finally the remaining extreme points are composed as residual component R .

Step 7: Calculate the variance ratio δ / δ_0 to obtain the optimal number of decompositions K , where δ is the relative standard deviation of $X - R$ and δ_0 is the standard deviation of the original series X .

Step 8: In the integer interval $[Kmin, Kmax]$, find out the maximum number of filtering times K corresponding to the minimum variance ratio δ / δ_0 as an optimal number of

decompositions, and accordingly repeat the first to sixth steps to output the signal decomposition results.

2.2 Empirical Wavelet Transform

Gilles integrated wavelet transform theory with the EMD method and introduced the empirical wavelet transform (EWT) method, which can effectively avoid the phenomenon of mode mixing (Gilles 2013; Peng et al. 2022; Liu et al. 2022).

The Eq.(2) is established to distinguish the optimal decomposition order N .

$$\begin{cases} E_N^l = \sqrt{\frac{\sum_{i=1}^n x_i^2(i)}{n}}, l = 1, 2, \dots, N \\ E_N = \sum_{l=1}^N E_N^l \\ \theta_{N,N-1} = \frac{|E_N - E_{N-1}|}{E_{N-1}} \end{cases} \tag{2}$$

where: E_N^l is the energy of each IMF, E_N is the sum of the energy of the N IMFs obtained by decomposition, and $\theta_{N,N-1}$ represents the energy difference of the IMF under different N values.

When $\theta_{N,N-1}$ suddenly becomes larger, it shows that the EWT decomposition process appears as the phenomenon of over-decomposition, resulting in false components. At this time, the corresponding decomposition time $N - 1$ is the optimal decomposition order of EWT.

2.3 Singular Value Decomposition

SVD, as a traditional and effective method for reducing noise, has been extensively utilized in the domain of extracting feature information from complex signals (Zhang et al. 2022; Zhang et al. 2018).

Define the optimal singular value decomposition order as r , this paper selects the order at which the singular entropy increment stabilizes as the optimal singular value decomposition order (Zhang et al. 2018).

2.4 Long Short-Term Memory Networks

In view of the fact that RNN is difficult to deal with long sequence data (Elsaraiti and Merabet 2021), Hochreiter and Schmidhuber (1997) proposed the LSTM model. Assuming σ stands for the sigmoid activation function; \tanh represents the hyperbolic tangent function. The formulae for σ and \tanh are as follows (Bakhshi Ostadkalayeh et al. 2023; Johnny et al. 2022):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{4}$$

The signal undergoes a series of transformations from its input to its output as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (5)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (6)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (7)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (8)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (9)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (10)$$

where: W_f , W_i , W_o , W_c indicates the respective weight vectors of the forgotten gate, the input gate, the export gate, and the memory cell; ‘ \cdot ’ denotes the dot product operation between two vectors; x_t denotes the input to the network at time t ; b_f , b_i , b_o , b_c represent the bias vectors of the forgotten gate, the input gate, the export gate, and the memory cell, respectively; h_{t-1} denotes the hidden state at the moment $t-1$; c_{t-1} indicates the cellular state of $t-1$; x_t denotes the input at the moment of t ; f_t denotes the forgetting gate; i_t denotes the input gate; o_t denotes output gate.

3 ESMD-EWT-SVD-LSTM Coupled Prediction Model

3.1 Model Steps

Taking into account the non-stationary of rainfall series and the benefits of different methods, this research suggests a coupled ESMD-EWT-SVD-LSTM model for predicting rainfall. The flowchart is illustrated in Fig. 1, and the crucial procedures are detailed underneath.

1) ESMD decomposition. The ESMD method is employed to break down the initial series of rainfall, leading to a finite number of IMFs consisting of components M and a residual component R .

2) EWT quadratic decomposition. For the $M1$ with complex frequency among the IMFs of ESMD decomposition, quadratic decomposition by EWT algorithm is applied, and the main information of the IMFs is further extracted. The quadratically decomposed EM of the $M1$ is obtained.

3) Noise reduction. For the high-frequency IMFs that remain after the ESMD decomposition, the SVD method is employed to minimize noise, while the low-frequency IMFs are directly incorporated into both training and prediction samples.

4) Prediction. All the IMFs after quadratic decomposition and SVD noise reduction are segregated into training sample and prediction sample, and the LSTM method is used in this paper to predict and reconstruct the rainfall series.

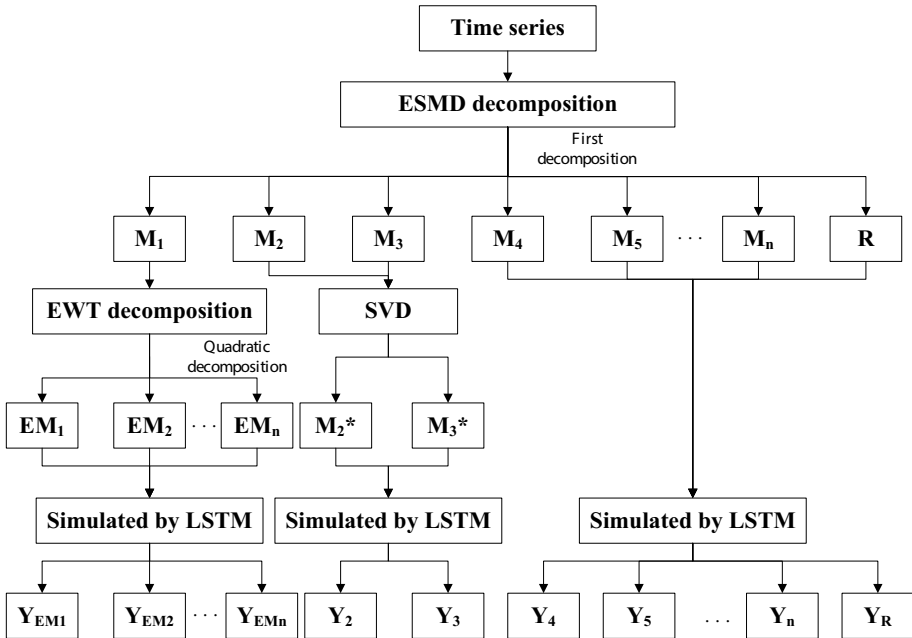


Fig. 1 The flowchart of ESMD-EWT-SVD-LSTM coupled prediction model

By combining quadratic decomposition with noise reduction in signal processing, it is possible to significantly simplify and extract key features from monthly rainfall data, thus enhancing the precision of rainfall predictions.

3.2 Model Verification

To further assess the predictive accuracy of the ESMD-EWT-SVD-LSTM model, we employ MAE, RMSE, and SI as evaluation metrics (Johny et al. 2022; Zhang et al. 2023). The specific equations are as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \tag{11}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}} \tag{12}$$

$$SI = \frac{1}{\mu} \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}} \tag{13}$$

where: \hat{y}_i is the predicted value at time i ; y_i is the actual observed value at time i ; μ is the average of the observations; N is data length.

4 Example Analysis

4.1 Area Profile

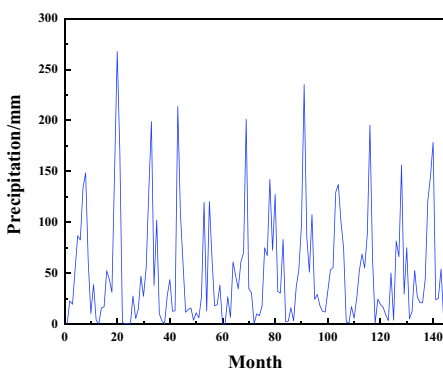
Kaifeng, located in the central region of China and the eastern part of Henan Province, is renowned as the ancient capital of eight dynasties, making it an essential city for protection and development in China. Rainfall serves as the primary water resource for Kaifeng, exerting a direct influence on both the local ecological environment and the daily lives of its residents. Therefore, predicting monthly rainfall is crucial for water resources planning and sustainable development in Kaifeng.

This research utilizes the recorded monthly rainfall data from five counties and four districts in Kaifeng spanning 2009 to 2020 as foundational data (Data source: the Water Resources Bulletin of Henan Province, <https://slt.henan.gov.cn/bmzl/szygl/szygb/>).

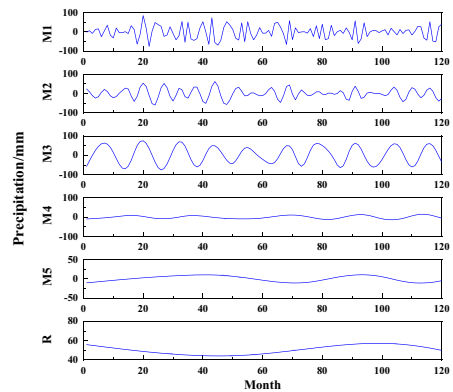
4.2 Data Preprocessing

Before the prediction, it is essential to carry out the division of training samples and prediction samples. Here, 120 monthly rainfall series from 2009 to 2018 in Kaifeng City are used as training samples, and 24 monthly rainfall series from 2019 to 2020 are used as prediction samples to compare and validate the prediction results (Data source: the Water Resources Bulletin of Henan Province, <https://slt.henan.gov.cn/bmzl/szygl/szygb/>). According to the ESMD decomposition steps, the rainfall series is shown in Fig. 2.

From Fig. 2(a), it can be observed that due to the influence of the temperate monsoon climate, the peak and trough values of Kaifeng’s rainfall exhibit obvious periodicity and uncertainty, the annual variation in rainfall is significant with complex characteristics. From Fig. 2(b), we can clearly observe that the ESMD effectively decomposes the original



(a) Monthly rainfall data for Kaifeng City from 2009 to 2020



(b) The ESMD results of Kaifeng from 2009 to 2018

Fig. 2 Rainfall series graph

monthly rainfall series into M components representing information at different scales and a R component representing the trend. From the $M1$ to R , the frequency decrease gradually and the wavelength become longer. The highest frequency $M1$ still contains rich frequency information, which is not favorable for the next prediction. Consequently, the EWT method is utilized for the quadratic decomposition of the $M1$ to further simplify the component features. The EWT decomposition process is illustrated in Fig. 3.

From Fig. 3(a), it can be seen that when $N > 4$, the IMF energy difference $\theta_{N, N-1}$ starts to increase, therefore, $N=4$ is chosen as the optimal decomposition order for the EWT algorithm. From Fig. 3(b), it's evident that the EWT decomposition effectively decomposes the $M1$ into feature EMs which include distinct frequency information from the frequency domain perspective and ensures that each feature EM aligns with the frequency information present in $M1$. Therefore, through the EWT decomposition, several EM components with simpler frequency characteristics are obtained, which can provide basic data for the subsequent prediction analyses.

The monthly rainfall time series is characterized by its high volatility, non-linearity, and non-stationary, the obtained IMFs also exhibit significant volatility, which is not conducive to prediction and analysis. To address this, we utilize SVD denoising on the IMFs to refine the precision of the prediction model. The process is demonstrated using $M2$ as an example and the SVD noise reduction results is shown in Fig. 4.

From Fig. 4(a), the optimal decomposition order of the $M2$ is determined to be $r=10$ which corresponds to retaining 90% of the singular value. In Fig. 4(b), the blue curve illustrates the data before SVD processing, while the red curve illustrates the data after SVD processing. Comparing the red and blue curves, it can be found that the fluctuation peaks of the $M2$ are reduced in some months after SVD processing and the overall fluctuation curves are more homogeneous. By applying SVD noise reduction, the original signal characteristics are represented more comprehensively, and some detailed fluctuations are more accurately represented.

4.3 Rainfall Prediction

This study utilizes the monthly rainfall data of Kaifeng City from 2009 to 2018 as a training sample for the LSTM model and predicts the monthly rainfall for the period 2019 to

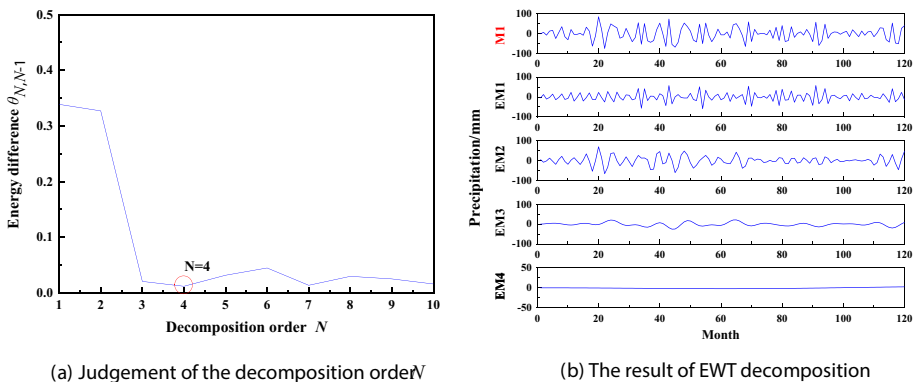


Fig. 3 The EWT decomposition process

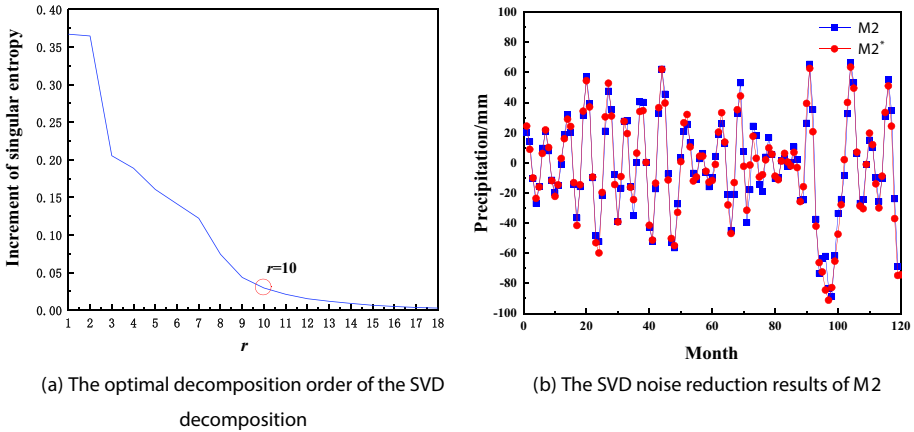


Fig. 4 The SVD noise reduction results

2020 (Data source: the Water Resources Bulletin of Henan Province, <https://slt.henan.gov.cn/bmzl/szygl/szygb/>). The predictions of IMFs are generated using LSTM model, the corresponding results are presented in Fig. 5.

From Fig. 5, although there are still a small number of points in each IMF that do not match completely, the proportion of these points is very low and does not affect the final prediction results, which indicates that the model exhibits a high degree of predictive accuracy.

The relative error is calculated between the predicted and actual values, as detailed in Table 1.

According to Table 1, the M1 and M2 have the largest prediction errors with the average relative positive error ranging from 40 to 50% and the average relative negative error ranging from -50% to -30%. This suggests that the non-smoothing of the M1 and M2 is high, the sequence's fluctuations in complex components can notably reduce prediction accuracy and cause a significant relative error. Figure 6 shows the prediction of rainfall data.

From Fig. 6, the predicted rainfall series can correspond to the actual observation series with small relative errors, and the prediction results are accurate without lag. This indicates that the coupled rainfall prediction model based on ESMD-EWT-SVD-LSTM has a good ability to process data and prediction, the prediction results are accurate and can be effectively used in practical engineering.

5 Discussion

In this study, the ESMD-EWT-SVD-LSTM coupled model has achieved good application in the prediction of rainfall. To further validate the feasibility and superiority of the ESMD-EWT-SVD-LSTM coupled model, ESMD-SVD-LSTM, SVD-LSTM, and LSTM are applied to compare the prediction ability respectively. The indicators such as MAE, RMSE, SI, average relative error are introduced to quantitatively evaluate the prediction results and the comparison are shown in Table 2.

Table 2 demonstrates that the ESMD-EWT-SVD-LSTM model yields superior predictive outcomes and boasts a higher accuracy, with the MAE of 4.96, the RMSE of 6.13, the

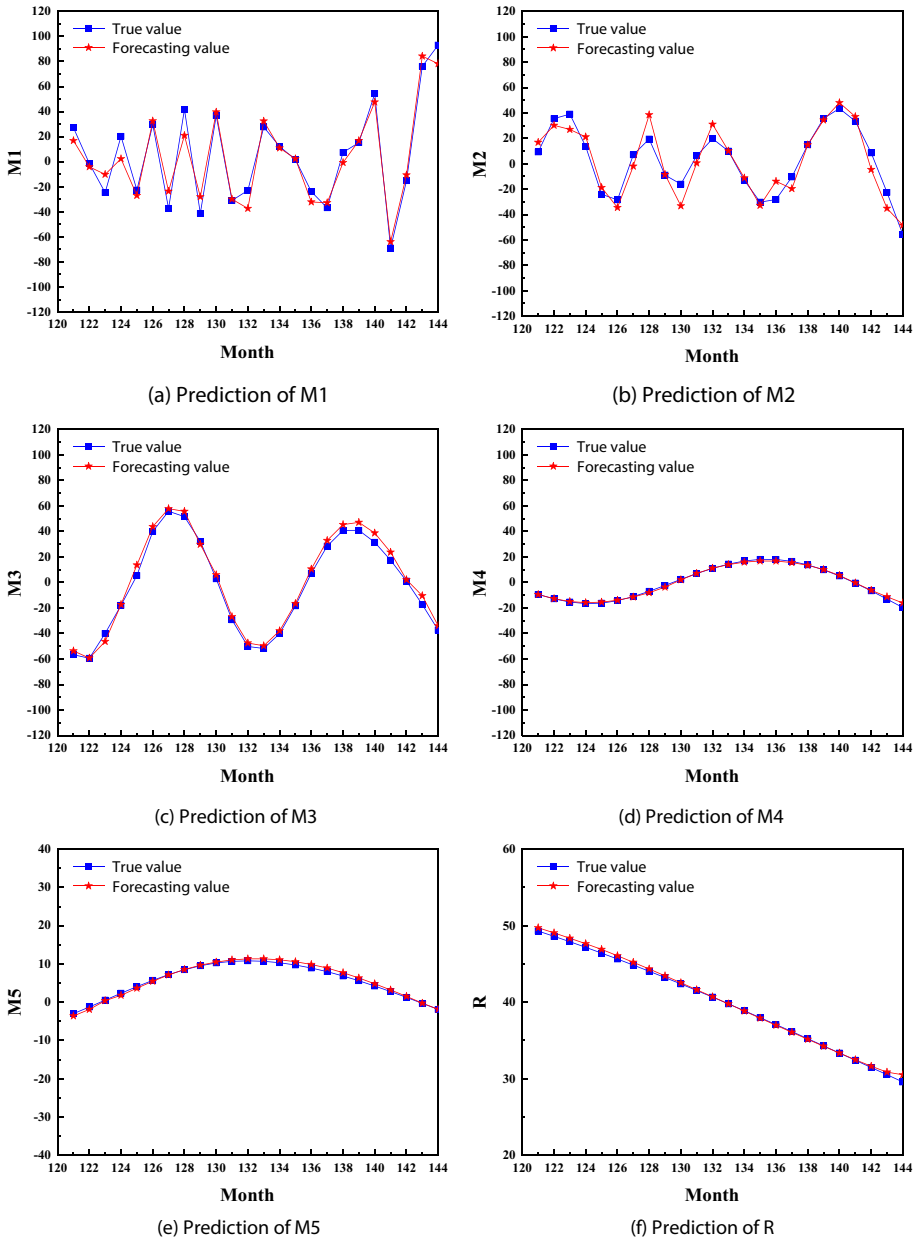
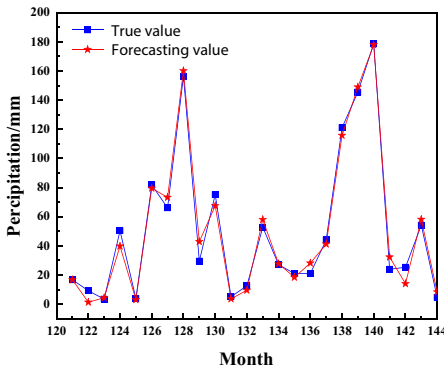


Fig. 5 The prediction results of LSTM

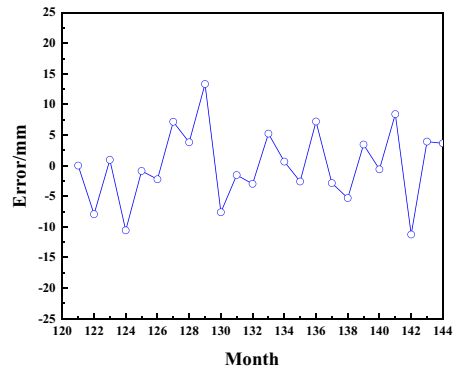
SI of 0.12, and the average relative error less than 1%. Taking the MAE as an example for analysis, comparing the prediction results of ESMD-EWT-SVD-LSTM and ESMD-SVD-LSTM, it can be found that after the quadratic decomposition of ESMD-EWT, the MAE is reduced by 56.53%, which indicates that reducing the complexity of the original rainfall sequence through multiple decompositions enhances the accuracy of LSTM. Besides,

Table 1 The relative error of the IMFs

	Relative error				Average relative positive error	Average relative negative error
	Maximum relative error		Minimum relative error			
	Positive	Negative	Positive	Negative		
M1	190.88%	-3.24%	7.95%	-108.70%	40.32%	-35.91%
M2	105.96%	-0.34%	1.51%	-153.48%	49.78%	-44.01%
M3	151.05%	-0.20%	3.61%	-40.03%	41.97%	-8.92%
M4	55.91%	-0.87%	1.04%	-28.01%	11.12%	-8.04%
M5	49.75%	-0.74%	-1.27%	-28.84%	11.93%	-11.67%
R	3.19%	-0.01%	0.00%	-0.24%	0.79%	-0.16%



(a) The prediction of rainfall data



(b) The prediction error

Fig. 6 The prediction of rainfall data

Table 2 Comparison table of indicators of each model

	MAE	RMSE	SI	Average relative error
ESMD-EWT-SVD-LSTM	4.96	6.13	0.12	-0.34
ESMD-SVD-LSTM	11.41	13.96	0.27	20.43
SVD-LSTM	22.20	27.15	0.53	155.71
LSTM	16.84	20.12	0.39	65.73

comparing the SVD-LSTM and LSTM prediction results, it can be found that although the SVD can further extract the signal features, the application of the SVD in the case of complex original rainfall sequence will lead to the interference or loss of the data features, which results in an increase of the MAE by 31.83% of the prediction results. The comparison of the model prediction results is provided in Fig. 7.

As can be seen from Fig. 7, the trend and period of the ESMD-EWT-SVD-LSTM prediction results can effectively coincide with the actual observation data, and the prediction

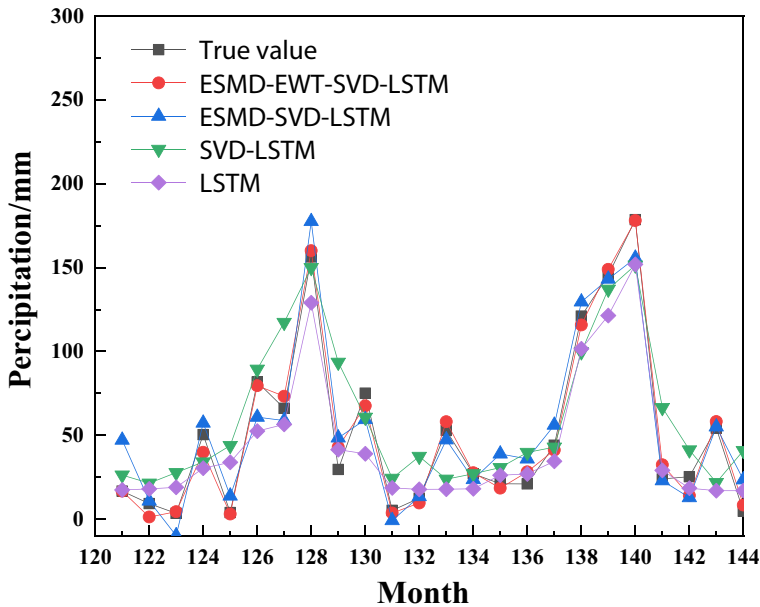


Fig. 7 Comparison of model prediction results

accuracy exceeds that of the other three models, which has a great potential in rainfall prediction.

Comparing the current research in the field of rainfall prediction (Byun et al. 2023; Chen et al. 2022; Johnny et al. 2022), all the technique proposed by these researchers can be effectively applied to actual projects. However, in the context of the randomness, volatility, and non-linearity inherent in rainfall data, the central challenge for further investigation in the field of rainfall prediction becomes how to effectively diminish the complexity of rainfall data and enhance the precision of predictive outcomes. On the basis of previous studies and fully considering the particularity of rainfall data, we put forward ESMD-EWT-SVD-LSTM coupled model. Different from the traditional direct prediction or prediction after a single decomposition, the coupled model addresses the characteristics of rainfall data by performing a quadratic decomposition with ESMD-EWT and denoising with SVD, continuously reducing data complexity and enhancing predictability, ultimately achieving accurate predicting results. In this study, we apply the ESMD-EWT-SVD-LSTM coupled model to predict the monthly rainfall data in Kaifeng City and assess the validity of the model by comparing it with alternative models. The results show that the ESMD-EWT-SVD-LSTM coupled prediction model has better theoretical advantages and accuracy in monthly rainfall prediction, and can be effectively applied to practical projects. The results show that the ESMD-EWT-SVD-LSTM coupled prediction model has better theoretical advantages and accuracy in monthly rainfall prediction, and can be effectively applied to practical projects.

Kaifeng, a city in China, exemplifies a warm temperate continental monsoon climate with complex and seasonal variations in monthly rainfall. The ESMD-EWT-SVD-LSTM coupled model has been effectively applied to the prediction of monthly rainfall in Kaifeng, indicating that this model has a strong ability to analyze and predict complex rainfall series and can be applied to rainfall predicting in regions with similar climates in China and

around the world. It is particularly worth mentioning that, based on the theoretical advantages of the ESMD-EWT-SVD-LSTM coupled model, it is feasible to perform rainfall prediction tasks well in complex rainfall areas different from Kaifeng, but the specific practical application effects require further research and exploration.

6 Conclusion

In this study, we present the ESMD-EWT-SVD-LSTM coupled prediction model to address the challenge of monthly rainfall prediction. By applying this approach to the analysis of monthly rainfall in Kaifeng City from 2009 to 2020, the results demonstrate that the predicted rainfall series closely aligns with the actual observation series, exhibiting small relative errors and without any lag. Comparing the predictive outcomes of the ESMD-EWT-SVD-LSTM model against those from the ESMD-SVD-LSTM, SVD-LSTM, and LSTM models, it is found that the ESMD-EWT-SVD-LSTM has the best prediction results, with MAE of 4.96, RMSE of 6.13, SI of 0.12, and average relative error less than 1%. This suggests that the coupled ESMD-EWT-SVD-LSTM monthly rainfall prediction model possesses strong predictive capabilities and can be effectively applied to regional water resources scheduling.

Although the proposed ESMD-EWT-SVD-LSTM model has demonstrated efficacy in monthly rainfall prediction, the current model solely focuses on the rainfall time series during the prediction process and overlooks the physical mechanisms that contribute to changes in rainfall. How to explore the connection between predicting rainfall and other pertinent factors will be a crucial focus for future studies.

Acknowledgements We are particularly grateful to the anonymous reviewers and editors for their comments.

Author Contributions Conceptualization: Z. X. Methodology and analysis: L. Z; Z. X. Writing—original draft preparation: L. Z; Z. X. Writing—review and editing: L. Z; Z. X. Supervision: Z. X.

Funding This work was supported by Program for Innovative Research Team (in Science and Technology) in University of Henan Province (24IRTSTHN012) and National Natural Science Foundation of China (51779093).

Data Availability The data that support the findings of this study are available on request.

Declarations

Ethical Approval All research activities conducted for this article have obtained ethical approval from the relevant institutional review board or ethics committee. This ensures that the research complies with ethical standards and safeguards the rights and well-being of the participants.

Consent to Participate Participants involved in this study have provided informed consent to voluntarily participate. Clear and comprehensive information regarding the study's purpose, procedures, and potential risks has been communicated, allowing participants to make informed decisions about their involvement.

Consent to Publish Authors have obtained explicit consent from study participants, when applicable, for the publication of any identifiable information. This ensures respect for individuals' privacy and confidentiality, aligning with ethical publishing practices.

Competing Interests Authors disclose any competing interests that could influence the interpretation or presentation of the research. This includes financial interests, relationships, or affiliations that may be perceived as influencing the work.

References

- Afshari Nia M, Panahi F, Ehteram M (2023) Convolutional neural network-ANN-E (Tanh): a new deep learning model for predicting rainfall. *Water Resour Manag* 37(4):1785–1810. <https://doi.org/10.1007/s11269-023-03454-8>
- Aydin MC, Sevgi B (2022) Flood risk analysis using GIS-based analytical hierarchy process: a case study of Bitlis Province. *Appl Water Sci* 12(6):122. <https://doi.org/10.1007/s13201-022-01655-x>
- Bakhshi Ostadkalayeh F, Moradi S, Asadi A, Moghaddam Nia A, Taheri S (2023) Performance improvement of LSTM-based deep learning model for streamflow forecasting using Kalman filtering. *Water Resour Manag* 37(8):3111–3127. <https://doi.org/10.1007/s11269-023-03492-2>
- Byun J, Jun C, Kim J, Cha J, Narimani R (2023) Deep learning-based rainfall prediction using cloud image analysis. *IEEE Trans Geosci Remote Sens* 61:1–11. <https://doi.org/10.1109/TGRS.2023.3263872>
- Chadalawada J, Havlicek V, Babovic V (2017) A genetic programming approach to system identification of rainfall-runoff models. *Water Resour Manag* 31:3975–3992. <https://doi.org/10.1007/s11269-017-1719-1>
- Chen C, Zhang Q, Kashani MH, Jun C, Bateni SM et al (2022) Forecast of rainfall distribution based on fixed sliding window long short-term memory. *Engineering Applications of Computational Fluid Mechanics* 16(1):248–261. <https://doi.org/10.1080/19942060.2021.2009374>
- Chen G, Wang WC (2022) Short-term precipitation prediction for contiguous United States using deep learning. *Geophys Res Lett* 49(8):e2022GL097904. <https://doi.org/10.1029/2022GL097904>
- Chen YC, Gao JJ, Bin ZH, Qian JZ, Pei RL, Zhu H (2021) Application study of IFAS and LSTM models on runoff simulation and flood prediction in the Tokachi River basin. *J Hydroinf* 23(5):1098–1111. <https://doi.org/10.2166/hydro.2021.035>
- Devi SR, Arulmozhivarman P, Venkatesh C, Agarwal P (2016) Performance comparison of artificial neural network models for daily rainfall prediction. *Int J Autom Comput* 13:417–427. <https://doi.org/10.1007/s11633-016-0986-2>
- Ehteram M, Ahmed AN, Sheikh Khozani Z, El-Shafie A (2023) Convolutional neural network-support vector machine model-Gaussian process regression: a new machine model for predicting monthly and daily rainfall. *Water Resour Manag* 37(9):3631–3655. <https://doi.org/10.1007/s11269-023-03519-8>
- Elsaraiti M, Merabet A (2021) Application of long-short-term-memory recurrent neural networks to forecast wind speed. *Appl Sci* 11(5):2387. <https://doi.org/10.3390/app11052387>
- Fatimi AS, Anwar E, Shaikh T (2023) The precedent set by unprecedented rainfall: lessons to be learnt from disastrous flooding in Pakistan. *Disaster Med Public Health Prep* 17:e411. <https://doi.org/10.1017/dmp.2023.79>
- Gao B, Huang X, Shi J, Tai Y, Zhang J (2020) Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks. *Renew Energy* 162:1665–1683. <https://doi.org/10.1016/j.renene.2020.09.141>
- Gao Y, Wang B, Chen F, Zhang W, Zhou D, Wu F, Chen D (2023) Multi-step wind speed prediction based on LSSVM combined with ESMD and fractional-order beetle swarm optimization. *Energy Rep* 9:6114–6134. <https://doi.org/10.1016/j.egy.2023.05.034>
- Gilles J (2013) Empirical wavelet transform. *IEEE Trans Signal Process* 61(16):3999–4010. <https://doi.org/10.1109/TSP.2013.2265222>
- Han L, Liang H, Chen H, Zhang W, Ge Y (2022) Convective precipitation nowcasting using U-net model. *IEEE Trans Geosci Remote Sens* 60:1–8. <https://doi.org/10.1109/TGRS.2021.3100847>
- Hewamalage H, Bergmeir C, Bandara K (2021) Recurrent neural networks for time series forecasting current status and future directions. *Int J Forecast* 37(1):388–427. <https://doi.org/10.1016/j.ijforecast.2020.06.008>
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Johny K, Pai ML, Adarsh S (2022) A multivariate EMD-LSTM model aided with time dependent intrinsic cross-correlation for monthly rainfall prediction. *Appl Soft Comput* 123:108941. <https://doi.org/10.1016/j.asoc.2022.108941>
- Kajewska Szkudlarek J (2020) Clustering approach to urban rainfall time series prediction with support vector regression model. *Urban Water J* 17(3):235–246. <https://doi.org/10.1080/1573062x.2020.1760319>
- Karijadi I, Chou SY, Dewabharata A (2023) Wind power forecasting based on hybrid CEEMDAN-EWT deep learning method. *Renew Energy* 218:119357. <https://doi.org/10.1016/j.renene.2023.119357>
- Li CZ, Zhang L, Wang H, Zhang YQ, Yu FL, Yan DH (2012) The transferability of hydrological models under nonstationary climatic conditions. *Hydrol Earth Syst Sci* 16(4):1239–1254. <https://doi.org/10.5194/hess-16-1239-2012>

- Liu Q, Yang J, Zhang K (2022) An improved empirical wavelet transform and sensitive components selecting method for bearing fault. *Measurement* 187:110348. <https://doi.org/10.1016/j.measurement.2021.110348>
- Moraga JS, Peleg N, Molnar P, Fatichi S, Burlando P (2022) Uncertainty in high-resolution hydrological projections: partitioning the influence of climate models and natural climate variability. *Hydrol Process* 36(10):e14695. <https://doi.org/10.1002/hyp.14695>
- Ni L, Wang D, Singh VP, Wu J, Wang Y, Tao Y, Zhang J (2020) Streamflow and rainfall forecasting by two long short-term memory-based models. *J Hydrol* 583:124296. <https://doi.org/10.1016/j.jhydrol.2019.124296>
- Nourali M (2023) Improved treatment of model prediction uncertainty: estimating rainfall using discrete wavelet transform and principal component analysis. *Water Resour Manag* 37(11):4211–4231. <https://doi.org/10.1007/s11269-023-03549-2>
- Olmo ME, Bettolli ML (2022) Statistical downscaling of daily precipitation over southeastern South America: assessing the performance in extreme events. *Int J Climatol* 42(2):1283–1302. <https://doi.org/10.1002/joc.7303>
- Peng L, Wang L, Xia D, Gao Q (2022) Effective energy consumption forecasting using empirical wavelet transform and long short-term memory. *Energy* 238:121756. <https://doi.org/10.1016/j.energy.2021.121756>
- Wang JL, Li ZJ (2013) Extreme-point symmetric mode decomposition method for data analysis. *Adv Adapt Data Anal* 5(03):1350015. <https://doi.org/10.1142/s1793536913500155>
- Wu S, Hu X, Zheng W, He C, Zhang G, Zhang H, Wang X (2021) Effects of reservoir water level fluctuations and rainfall on a landslide by two-way ANOVA and K-means clustering. *Bull Eng Geol Environ* 80(7):5405–5421. <https://doi.org/10.1007/s10064-021-02273-8>
- Yang J, Xiang Y, Sun J, Xu X (2022) Multi-model ensemble prediction of summer precipitation in China based on machine learning algorithms. *Atmosphere* 13(9):1424. <https://doi.org/10.3390/atmos13091424>
- Zhang J, Chen X, Khan A, Zhang YK, Kuang X, Liang X, Taccari ML, Nuttall J (2021) Daily runoff forecasting by deep recursive neural network. *J Hydrol* 596:126067. <https://doi.org/10.1016/j.jhydrol.2021.126067>
- Zhang J, Hou G, Ma B, Hua W (2018) Operating characteristic information extraction of flood discharge structure based on complete ensemble empirical mode decomposition with adaptive noise and permutation entropy. *J Vib Control* 24(22):5291–5301. <https://doi.org/10.1177/1077546317750979>
- Zhang J, Li Z, Huang J, Cheng M, Li H (2022) Study on vibration-transmission-path identification method for hydropower houses based on CEEMDAN-SVD-TE. *Appl Sci* 12(15):7455. <https://doi.org/10.3390/app12157455>
- Zhang X, Chen H, Wen Y, Shi J, Xiao Y (2023) A new water level prediction model based on ESMD–VMD–WSD–ESN. *Stoch Env Res Risk A* 37:3221–3241. <https://doi.org/10.1007/s00477-023-02446-9>
- Zhang Y, Chen Y (2022) Application of hybrid model based on CEEMDAN, SVD, PSO to wind energy prediction. *Environ Sci Pollut Res* 29:22661–22674. <https://doi.org/10.1007/s11356-021-16997-3>
- Zheng J, Su M, Ying W, Tong J, Pan Z (2021) Improved uniform phase empirical mode decomposition and its application in machinery fault diagnosis. *Measurement* 179:109425. <https://doi.org/10.1016/j.measurement.2021.109425>
- Zhou Z, Ouyang Y, Li Y, Qiu Z, Moran M (2017) Estimating impact of rainfall change on hydrological processes in Jianfengling rainforest watershed China using BASINS-HSPF-CAT modeling system. *Ecol Eng* 105:87–94. <https://doi.org/10.1016/j.ecoleng.2017.04.051>
- Zhu K, Yang Q, Zhang S, Jiang S, Wang T, Liu J, Ye Y (2023) Long lead-time radar rainfall nowcasting method incorporating atmospheric conditions using long short-term memory networks. *Frontiers in Environmental Science* 10:1054235. <https://doi.org/10.3389/fenvs.2022.1054235>
- Zhuang X, Fan Y, Li Y, Wu C (2023) Evaluation climate change impacts on water resources over the upper reach of the Yellow River Basin. *Water Resour Manag* 37(6-7):2875–2889. <https://doi.org/10.1007/s11269-023-03501-4>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Ziyu Li^{1,2,3} · Xianqi Zhang^{1,2,3}

✉ Xianqi Zhang
zxqi@163.com

¹ North China University of Water Resources and Electric Power, Zhengzhou 450046, China

² Collaborative Innovation Center for Efficient Utilization of Water Resources, Zhengzhou 450046, China

³ Technology Research Center of Water Conservancy and Marine Traffic Engineering, Zhengzhou 450046, China