Application of precipitation prediction model based on multi-model coupling in agricultural irrigation

Bo Wang*, Xingchen Lan, Yue Peng, Zitong Qiu, Ran Peng

Sichuan Agricultural University, Sichuan, China

ABSTRACT

Agricultural irrigation faces escalating challenges in water management due to the ongoing impacts of climate change. In this study, our goal is to enhance the precision of precipitation prediction to offer more dependable support for agricultural irrigation decision-making. We adopted a multi-model fusion algorithm based on long and short-term memory networks and integrated various data sources, including meteorological station observations and satellite remote sensing gridded data, to construct a comprehensive precipitation prediction model.

The objective of this research is to develop an efficient and accurate precipitation prediction model that can provide scientific decision support for agricultural irrigation. Through rigorous comparison of different models, we identified the optimal combination to improve the model's robustness and accuracy. Our experimental results reveal that multi-model fusion exhibits higher accuracy and stability in precipitation prediction compared to a single model.

Our study further validates the substantial advantages of multi-model fusion in enhancing prediction accuracy and emphasizes the critical role of integrating data from multiple sources for optimal model performance. By furnishing more reliable predictive information for agricultural irrigation decision-making, this study introduces new methodologies and ideas for enhancing agricultural water use efficiency and addressing the challenges posed by climate change.

In terms of innovations, this study leverages a multi-model fusion algorithm grounded in long and short-term memory networks, and integrates multi-source data to offer a comprehensive and reliable solution for precipitation prediction. This approach provides valuable insights for future similar studies and contributes to the advancement of agricultural water management practices.

Keywords: Precipitation Prediction, LSTM, SVM, CEEMDAN, Agricultural Irrigation

1 INTRODUCTION

In recent years, precipitation forecasting has been a challenging issue in the field of meteorology due to its complexity and uncertainty caused by multiple factors. Traditional forecasting methods rely on statistics and dynamics, but the development of artificial intelligence and big data has brought new opportunities for precipitation forecasting. Models based on Support Vector Machine (SVM) and Random Forest algorithms have shown strong capabilities, providing significant support for agriculture, water resource management, and disaster prevention and control. Recently, neural networks based on deep learning technologies, especially Long Short-Term Memory (LSTM) networks, have performed exceptionally well in precipitation forecasting [1]. This paper explores LSTM-based precipitation forecasting methods, including basic LSTM models, SVM-LSTM, LSTM with multi-head attention mechanisms, and LSTM after Complex Empirical Mode Decomposition with Adaptive Noise (CEEMDAN). We also investigate the application of Convolutional LSTM methods, combining Convolutional Neural Networks (CNN) with LSTM to comprehensively model spatiotemporal information, potentially enhancing forecasting performance further. In conclusion, deep learning methods provide new insights for precipitation forecasting, particularly in the realm of daily precipitation forecasting, which warrants further research and exploration.

2 SOURCES AND CLEANSING OF DATA

2.1 Sources and characteristics of data

For the station data, we used the precipitation observation data collected at the Ya'an rainfall station (No. 60612000) and combined it with the latitude and longitude coordinates of the point. This information was organized through a script combined with the nc data of the coordinate point to obtain a multifactor dataset, which was then used for training and testing on multiple coupled models.

Table 1: Meteorological data of Ya'an, 2013/7/30-2022/12/31

Date	Rainfall	Temperature	Humidity
2013/07/30	7.80	22.46	0.12
2013/07/31	2.63	21.01	0.13
......	......	......	......
2022/12/27	0.00	25.22	0.39
2022/12/28	0.00	23.04	0.42

2.2 Data cleansing

When cleaning time series data, common methods include handling missing values (such as filling or deleting missing data), dealing with outliers (such as correcting or removing outlier data points), removing duplicate data, smoothing data (such as filtering or averaging), adjusting timestamp formats (such as unifying time units or time zones), handling abnormal timestamps (such as adjusting or marking abnormal timestamps), data type conversion (such as converting text data to numerical data), and data normalization (such as scaling data to specific ranges or standardizing distributions). These methods aim to improve data quality and accuracy, providing a reliable data foundation for subsequent analysis and modeling.

3 MODEL SELECTION AND METHODOLOGY

3.1 LSTM (Long Short-Term Memory) model

LSTMs address the limitations of traditional RNNs with long sequence data, which struggle to capture complex dependencies across multiple time steps due to gradient vanishing or explosion issues. LSTMs solve this with a gating mechanism, effectively handling long-term dependencies.

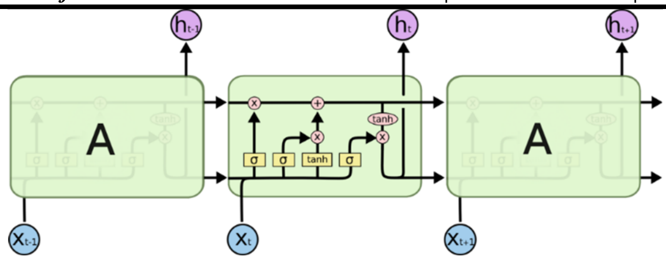

Fig.1: Structure of LSTM network

Input Gate: The input gate allows the model to selectively receive new information. This is done by using a sigmoid activation function to control the weight of the input information and a tanh activation function to determine new candidate values [2]. This mechanism allows the LSTM to selectively update its internal state and thus be more flexible in adapting to different input patterns.

Input Gate:

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

Forget Gate: Determines which past information should be forgotten by using a sigmoid activation function. This allows the LSTM to retain important information while discarding less relevant data, improving memory performance.

Output Gate: Controls the output of the current time step using sigmoid and tanh activation functions, enabling selective output of the internal state. This helps the LSTM adapt to different prediction tasks by focusing on essential information [3].

Oblivion Gate:

$$f_t = \sigma\left(W_f[x_t, h_{t-1}] + b_f\right) \tag{2}$$

Output Gate:

$$o_t = \sigma\left(W_f[x_t, h_{t-1}] + b_o\right) \tag{3}$$

LSTM's gating mechanisms excel at capturing long-term dependencies in time series data, making it highly effective in various applications. In precipitation forecasting, LSTM accurately captures spatio-temporal precipitation characteristics, improving forecast accuracy significantly. Future studies will explore leveraging LSTM properties and enhanced methods to better predict approaching precipitation and address flash flood challenges more effectively [4].

3.2 BiLSTM model

BiLSTM (Bidirectional Long Short-Term Memory) is a deep learning model used for time series prediction, and it offers several clear advantages over traditional LSTM (Long Short-Term Memory) models. Firstly, BiLSTM can simultaneously consider past and future context information, enabling a more comprehensive capture of patterns and trends in time series data, thereby enhancing prediction accuracy and stability. Secondly, BiLSTM is more flexible in modeling time series data, as it can adapt to sequences of different lengths and structures, showcasing stronger generality and adaptability [5]. Additionally, the introduction of bidirectional structure in BiLSTM makes it easier to avoid gradient vanishing or exploding issues compared to unidirectional LSTM models, facilitating the training of deeper models. as shown in the following figure.

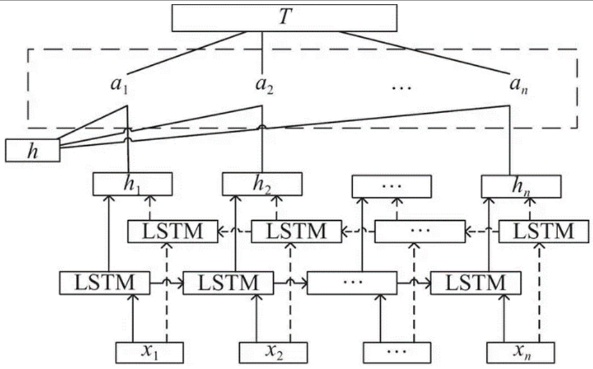

Fig.2: Structure of BiLSTM network

3.2 Muti Head Attention

Attention Mechanism (AM) in deep learning mimics human visual and cognitive systems, allowing neural networks to focus on relevant input data parts, enhancing model performance and generalization. It automatically learns and prioritizes crucial information in the input [5].

Multi-head attention extends AM by enabling different weight distributions in various attention subspaces. It employs multiple attention heads, each learning a distinct weight matrix. These heads produce weighted outputs, combined linearly to obtain the final multi-head attention output.

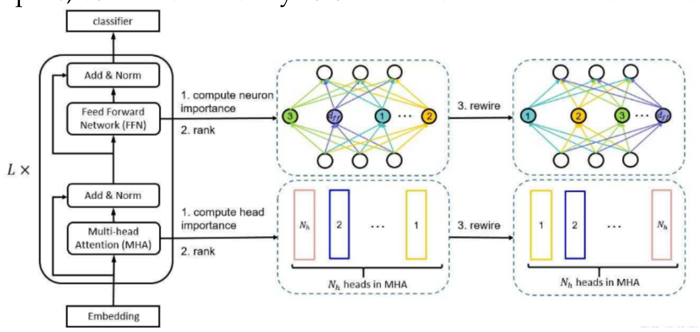

Fig.3: Multi-attention mechanism network structure

Introducing multi-head attention into LSTM models helps enhance the model's ability to capture long-term dependencies, increases flexibility in representing different features, alleviates attention bottlenecks when processing large amounts of information, and improves the model's interpretability and explainability by observing the parts each head focuses on, aiding in a deeper understanding of the model's workings and decision-making processes.

3.3 SVM

SVMs are mainly used in linear separability problems where the input data and the learning objective are given: each sample of the input data contains multiple features and thus forms a feature space, while the learning objective is a binary variable representing the negative class and the positive class [6].

If there exists a decision boundary in the feature space where the input data are located, the Hyperplane Separate the learning objectives by positive and negative classes and make the point-to-plane distance of any sample greater than or equal to one:

$$a_{ij} = \begin{cases} a_{ij} + \left(a_{ij} - a_{max}\right) \times f(g) \\ a_{ij} + \left(a_{min} - a_{ij}\right) \times f(g) \end{cases} \tag{4}$$

$$P = 1 - \frac{6\sum d^2{}_i}{n(n^2 - 1)} \tag{5}$$

Then the classification problem is said to be linearly differentiable and the parameters are the normal vector and the intercept of the hyperplane, respectively [7].

The decision boundary that satisfies this condition actually constructs 2 parallel hyperplanes as interval boundaries to discriminate the classification of the samples:

$$w^T X_i + b \geq +1 \quad => \quad y_i = +1 \tag{6}$$

$$w^T X_i + b \leq -1 \quad <= \quad y_i = -1 \tag{7}$$

In the SVM-MutiHeadAttention-LSTM model, SVM is mainly responsible for binary classification of precipitation sequences, defining daily precipitation greater than 0.1 mm as a rainy day and vice versa as a sunny day. Such a binary classification task helps to transform complex precipitation patterns into simpler judgement problems.

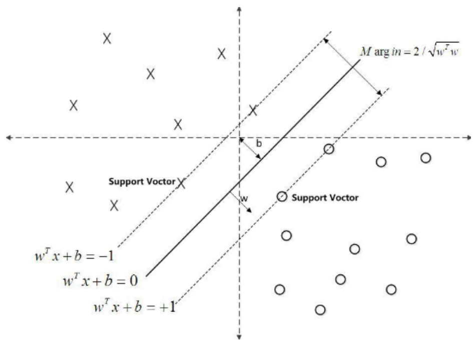

Fig.4: SVM binary classification result map

3.4 Ceemdan

CEEMDAN is introduced into the time series prediction model to handle the data's nonlinearity and non-stationarity more effectively. It is an improved empirical mode decomposition method that decomposes nonlinear and non-stationary time series, simplifying complex data into manageable components. This enhances the model's ability to model time series data and improve prediction accuracy. By incorporating CEEMDAN, the model captures dynamic time series characteristics better, enhancing overall performance and robustness [8]. The principle of modal decomposition is shown below:

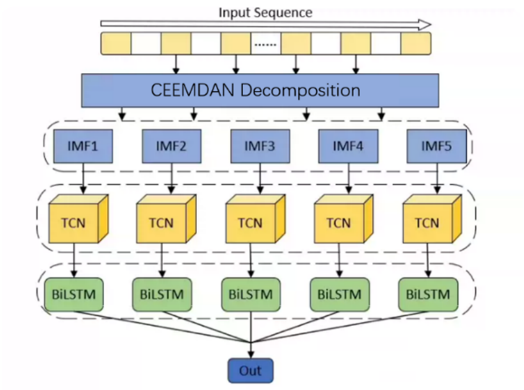

Fig.5: Schematic diagram of collective empirical modal decomposition

Let Ei(.) be the $i\,th$ eigenmode component obtained after EMD decomposition, the $j\,th$ eigenmode component obtained by CEEMDAN decomposition is $\overline{c\iota(t)}$, ϑ^j is the Gaussian white noise signal satisfying the standard normal distribution, j = 1, 2, 3, … N is the number of times of adding white noise, ε is the standard table of white noise, and y(t) is the signal to be decomposed. The steps of CEEMDAN decomposition are as follows:

Gaussian white noise is added to the signal to be decomposed y(t) to obtain the new signal $y(t)+(-1)^q\varepsilon\vartheta^j(t)$, where q=1,2. EMD decomposition of the new signal is performed to obtain the first order eigenmode component C1.

$$E\left(y(t)+(-1)^q c v^j(t)\right)=C_1^j(t)+r^j \tag{8}$$

The 1st eigenmode component of the CEEMDAN decomposition is obtained by averaging over the resulting N modal components:

$$\overline{C_1(t)}=\frac{1}{N}\Sigma_{j=1}^{N}C_1^j(t) \tag{9}$$

Calculate the residuals after removing the first modal component:

$$r_1(t)=y(t)-\overline{C_1(t)} \tag{10}$$

Add positive and negative paired Gaussian white noise in $r_1(t)$ to get a new signal, and use the new signal as a carrier for EMD decomposition to get the first-order modal component D_1^j, which can get the 2nd eigenmode component of CEEMDAN decomposition:

$$\overline{C_2(t)}=\frac{1}{N}\sum_{j=1}^{N}D_1^j(t) \tag{11}$$

Calculate the residuals after removing the second modal component:

$$r_2(t)=r_1(t)-\overline{C_2(t)} \tag{12}$$

Repeat the above steps until the residual signal obtained is a monotonic function and cannot be further decomposed, the algorithm ends. At this point the number of eigenmode components obtained is K. The original signal $y(t)$ is decomposed as:

$$y(t)=\Sigma_{k=1}^{K}\overline{C_k(t)}+r_k(t) \tag{13}$$

3.5 Conv LSTM

Conv LSTM, introduced in the research paper "Convolutional LSTM Network: a Machine Learning Approach for Precipitation Nowcasting," stands out as a powerful tool for handling gridded precipitation data and forecasting future precipitation events. Its notable effectiveness lies in its ability to model spatio-temporal relationships accurately, a task that traditional time series models often struggle with when dealing with gridded precipitation data [9]. Unlike these conventional models, ConvLSTM integrates Convolutional Neural Networks (CNN) and Long Short-Term Memory Networks (LSTM), enabling it to comprehensively capture and process spatio-temporal information. This comprehensive approach significantly improves the accuracy and timeliness of precipitation predictions.

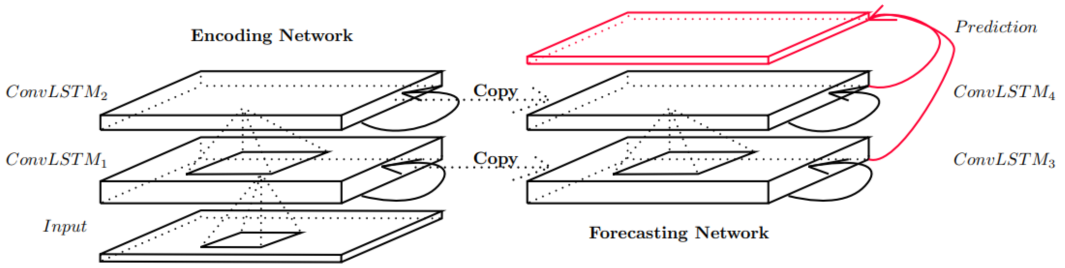

Fig.6: Schematic diagram of ConvLSTM

4 EXPERIMENTS AND ANALYSIS OF RESULTS

4.1 LSTM(BiLSTM) training and prediction results

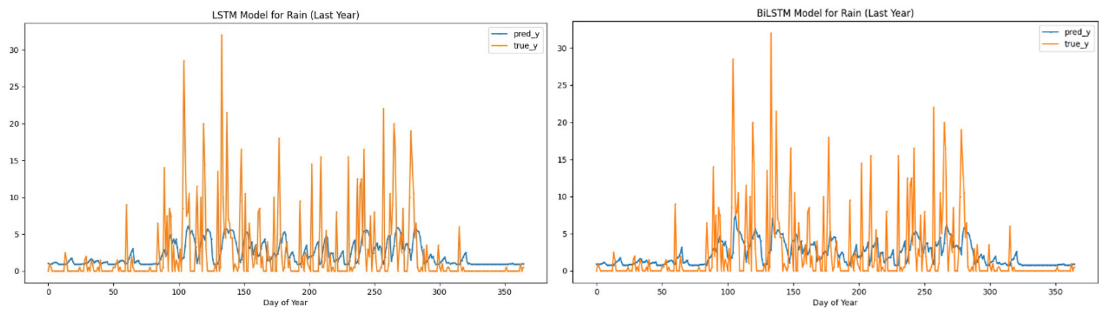

Fig.7: LSTM & BiLSTM prediction results

BiLSTM's enhanced predictive capability stems from its adeptness at synthesizing insights from both past and future data points, a feat that traditional LSTM models struggle to match. This superiority is attributed to BiLSTM's bidirectional information flow and advanced memory mechanisms, enabling it to capture nuanced sequence dependencies crucial for precipitation forecasting [10]. Furthermore, the incorporation of fusion techniques with multiple models alongside BiLSTM promises to elevate prediction accuracy to new heights, paving the way for more reliable and nuanced precipitation forecasts.

4.2 MutiheadAttention-LSTM(BiLSTM) training and prediction results

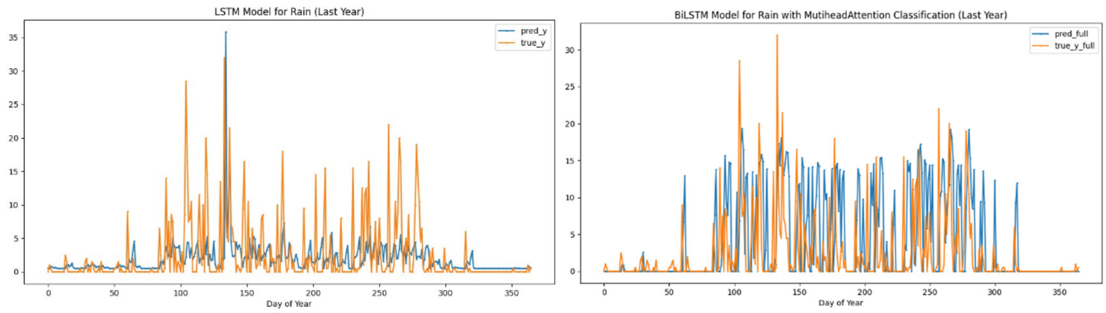

Fig.8: MutiheadAttention-LSTM prediction results

From the prediction results, it can be seen that the model with the introduction of multihead attention mechanism performs better than simple LSTM or BiLSTM models. Moreover, the prediction results from the multihead attention mechanism combined with BiLSTM are better than those combined with LSTM [11]. However, there is still room for improvement in prediction accuracy, especially in terms of classifying rainy and sunny day data. Therefore, the next step involves using SVM for binary classification to separate rainy and sunny day data, followed by separate time series predictions for each. This will help improve the model's prediction capabilities and accuracy [12].

4.3 SVM-MutiheadAttention-LSTM(BiLSTM) training and prediction results

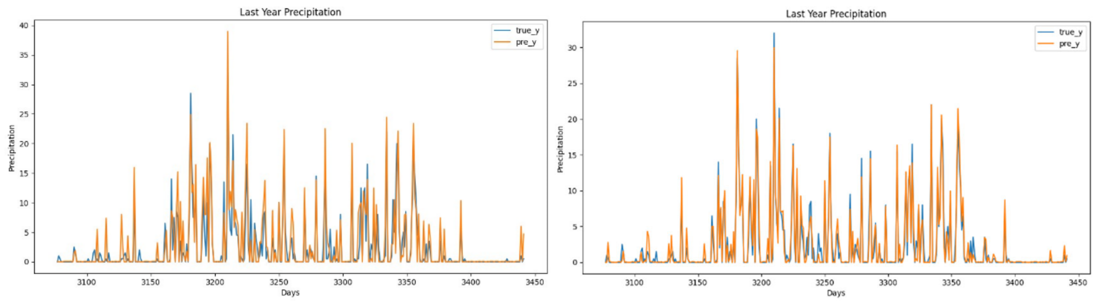

Fig.9: SVM-MutiheadAttention-LSTM prediction results

Combining the SVM binary classification model with a multi-attention mechanism-based temporal prediction model has shown good performance, enhancing the prediction capabilities for both sunny and rainy weather, reducing data interference, and improving time series feature extraction. In the future, signal decomposition can be explored to further enhance the predictive performance of the model. By integrating SVM with MutiHeadAttention-LSTM, the model is able to capture the dynamics in time-series data while classifying sunny and rainy weather to better understand and predict precipitation events [13]. This combination helps improve the robustness of the model, enabling it to adapt more flexibly to different meteorological conditions and time series variations, providing more reliable support and reducing the risk of flash floods. In future experiments, we will further validate the performance of this model in precipitation prediction, aiming to provide more innovative and practical solutions.

4.4 SVM- MutiheadAttention-CEEMDAN training and prediction results

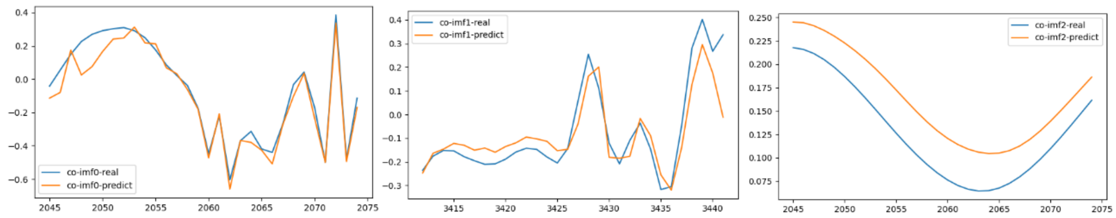

Fig.10: Prediction results of the first signal obtained after modal decomposition of precipitation data

Integrating CEEMDAN into BiLSTM greatly enhances the model's capability to address the complex time-frequency characteristics, noise, and nonlinear structures inherent in data. The superior performance observed in the SVM-MutiheadAttention-CEEMDAN-BiLSTM model compared to the SVM-MutiheadAttention-CEEMDAN-LSTM model highlights this enhancement. CEEMDAN's decomposition into intrinsic modal functions (IMFs) facilitates noise reduction and crucial feature extraction, complementing BiLSTM's bidirectional learning approach. This combined framework significantly improves the model's understanding of spatio-temporal relationships and predictive accuracy, particularly in challenging environments with diverse sources of noise.

4.5 ConvLSTM training and prediction results

ConvLSTM stands out due to its advanced capabilities in processing and predicting precipitation forecasts, a step above traditional LSTM models. The integration of convolutional operations empowers ConvLSTM to extract nuanced temporal and spatial features concurrently, resulting in a thorough understanding of intricate patterns and trends inherent in time series data. Additionally, ConvLSTM excels in capturing complex spatio-temporal relationships, especially crucial when dealing with gridded precipitation data. This refined modeling approach not only enhances prediction accuracy but also ensures robust and stable performance, marking a significant leap forward in the realm of precipitation forecasting technology [错误!未找到引用源。].

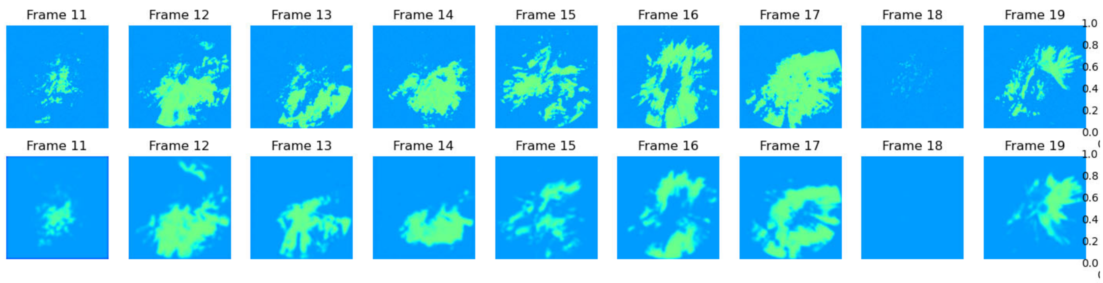

Fig.11: Precipitation forecast frame

5 MODEL EVALUATION AND PERFORMANCE COMPARISON

5.1 Model assessment methodology

The performance of an LSTM model can be evaluated using a variety of metrics and methods. Commonly used evaluation metrics include Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and correlation coefficient. Additionally, comparing the predicted values with the actual values through scatter plots or line graphs can provide visual insights into the model's predictive performance. Furthermore, methods such as cross-validation, time series cross-validation, or hold-out validation can be used to assess the model's generalization and robustness across different datasets [错误!未找到引用源。].

5.1.1 Mean Squared Error (MSE)

Definition:MSE is the average of the squares of the differences between predicted and actual values.

Advantages:It is sensitive to outliers and is a widely used indicator for evaluating regression models.

Calculations:

$$MSE = \frac{1}{n}\Sigma_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{14}$$

5.1.2 Root Mean Squared Error (RMSE)

Definition:RMSE is the square root of MSE and is used to normalise the magnitude of the error

Calculations:

$$RMSE = \sqrt{MSE} \tag{15}$$

5.1.3 Mean Absolute Error (MAE)

Definition:MAE is the average of the absolute values of the differences between the predicted and actual values.

Calculation method.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{16}$$

5.1.4 Coefficient of Determination.R²)

Definition:R2 indicates the extent to which the model explains the variance of the target variable. The value ranges from 0 to 1, with closer to 1 indicating a better model.

Calculations:

$$R^2 = 1 - \frac{\Sigma_{i-1}^{n}(y_i - \hat{y}_i)^2}{\Sigma_{i-1}^{n}(y_i - \bar{y}_i)^2} \tag{17}$$

5.1.5 Relative bias (BIAS)

Definition: Relative deviation is the average of the differences between predicted and actual values, and represents the average deviation of the predicted value relative to the actual value.

Calculations:

$$BIAS = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \tag{18}$$

5.2 Performance comparison

Table 2: performance evaluation indicators

Model Name	MSE	BIAS	RMSE	MAE	R^2
LSTM	12.6898	0.3458	3.5623	0.05683	0.1486
MUTIHEADATTENTION-LSTM	12.3633	0.3064	3.5161	0.04895	0.2089
SVM-MUTIHEADATTENTION-LSTM	10.9677	0.2485	3.3117	0.03965	0.2765
SVM-MUTIHEADATTENTION-CEEMDAN-LSTM	8.1253	0.1904	2.8505	0.03564	0.4955
Bilstm	12.2847	0.2745	3.5050	0.05054	0.2064
MUTIHEADATTENTION-Bilstm	12.2819	0.2286	3.5046	0.04256	0.3589
SVM-MUTIHEADATTENTION-Bilstm	9.1314	0.1867	3.0218	0.02356	0.5456
SVM-MUTIHEADATTENTION-CEEMDAN-Bilstm	7.5145	0.1564	2.7413	0.01958	0.6574

By coupling multiple models with different advantages based on the data and model features of a single model, overall prediction capability can be enhanced and prediction errors reduced. It is believed that accurate precipitation prediction through coupled models can provide data support and pre-planning for agricultural irrigation, ensuring the normal operation of agricultural production and the healthy growth of agricultural products. The specific relationship is shown in the table below:

Tabel 3: Relationship between water use for agricultural irrigation and precipitation

Conversion table of precipitation h to irrigation v per hectare a1, or a2 per acre					
Calculation formula	$v = ha_1$	$v = ha_2$		$v = ha_1$	$v = ha_2$
	$a_1 = 10000m^2$	$a_2 = 666.7m^2$		$a_1 = 10000m^2$	$a_2 = 666.7m^2$
Precipitation: h	Irrigation per hectare	Irrigation per acre	Precipitation: h	Irrigation per hectare	Irrigation per acre
mm	m^3	m^3	mm	m^3	m^3
1	10	0.67	25	250	16.67
2	20	1.33	30	300	20.00
3	30	2.00	35	350	23.33

6 CONCLUSION

We have integrated BiLSTM, SVM, CEEMDAN, and Multihead Attention technologies to develop a sophisticated precipitation prediction model. Through comparative analysis of various coupling methods, we have identified the BiLSTM+SVM+CEEMDAN+Multihead Attention approach as highly advantageous for precipitation forecasting.

This coupling method effectively leverages BiLSTM's bidirectional learning and time-series processing capabilities to capture spatiotemporal relationships and long-term dependencies, thus enhancing model accuracy. Additionally, incorporating SVM and Multihead Attention mechanisms enhances model flexibility and robustness. The introduction of CEEMDAN's time-frequency decomposition technology further improves the model's adaptability and generalization to time-series data.

We strongly recommend adopting the BiLSTM+SVM+CEEMDAN+Multihead Attention approach for precipitation forecasting, anticipating significant advancements in meteorological research and applications. Furthermore, integrating ConvLSTM technology has substantially improved the efficiency of handling radar echo data, offering additional avenues for enhancing precipitation prediction accuracy.

REFERENCES

[1] Ding, D., Sun, N., Li, A., Li, Z., & Zhang, Y. (2024). Research on vehicle battery data cleaning method based on OOA-VMD-ATGRU-GAN. Physica Scripta.

[2] Linghu, T. I. A. N., & Bingxia, Y. U. A. N. (2024). Prediction of ion battery remaining life of energy storage system based on data preprocessing and computer VMD-LSTM-GPR. Energy Storage Science and Technology, 13(1), 336.

[3] Ning, X., Shi, Z., & Xu, J. (2023). Health state estimation of lithium-ion battery based on health factor and PSO-LSTM. Journal of Power Supplies, 1-13.

[4] Li, S. (2023). Research on metro passenger flow prediction with improved LSTM model based on attention mechanism. Industrial Control Computer, 36(11), 124-125, 128.

[5] Jin, H., Cao, P., Jin, W., et al. (2023). False alarm identification of oil chromatography online monitoring device based on LSTM-MACNN. Power Information and Communication Technology, 21(11), 55-62.

[6] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019, December). The performance of LSTM and BiLSTM in forecasting time series. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 3285-3292). IEEE.

[7] Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics & Proteomics, 15(1), 41-51.

[8] Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. Physica A: Statistical Mechanics and Its Applications, 519, 127-139.

[9] Li, J., Shi, Y., Zhang, T., Li, Z., Wang, C., & Liu, J. (2024). Radar precipitation nowcasting based on ConvLSTM model in a small watershed in north China. Natural Hazards, 120(1), 63-85.

[10] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). A comparative analysis of forecasting financial time series using ARIMA, LSTM, and BiLSTM. arXiv preprint arXiv:1911.09512.

[11] Tang, J., Yang, L., Zhao, J., Qiu, Y., Deng, Y., & Shen, S. (2021, August). Research on RFID indoor positioning algorithm based on attention. In 2021 IEEE International Conference on Electronic Technology, Communication and Information (ICETCI) (pp. 140-143). IEEE.

[12] Kim, S., Hong, S., Joh, M., & Song, S. K. (2017). Deeprain: ConvLSTM network for precipitation prediction using multichannel radar data. arXiv preprint arXiv:1711.02316.

[13] Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. Geoscientific Model Development Discussions, 2022, 1-10.