Journal of Information and Computing Science
J. Info. Comput. Sci., Vol. 19, No. 1, pp. 41-52

DOI: 10.4208/JICS-2024-003
June 2024

Application of Improved LSTM-WBLS Model in Daily Precipitation Forecast

Ying Han[1], Jian Guan[1], Yunzhong Cao[1] and Jia Luo[2,*]

[1] *School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China*
[2] *Hubei Public Meteorological Service Center, Wuhan 430074, China*

Abstract. The popular Long Short-Term Memory (LSTM) based precipitation prediction models suffer from overfitting and time lag. Broad Learning System (BLS), which does not require multiple iterations, helps to solve the above disadvantages of LSTM. Weighted Broad Learning System (WBLS) reduces the impact of noise and outliers on precipitation prediction accuracy by introducing a weighted penalty factor constraint to assign sample weights in the BLS. Thus, a LSTM-WBLS daily precipitation prediction model is proposed in this paper. The daily precipitation at Badong station in Hubei province is selected for empirical study. And the influence of air pressure, temperature, humidity, wind speed and sunshine on precipitation is considered. The experimental results demonstrate that the LSTM-BLS model has significantly improved the prediction accuracy in the evaluation indexes of RMSE, MAE and R2 compared with existing prediction models. The prediction accuracy of the new model outperforms existing models at different time steps, proving its stability. In particular, the direct calculation of weights by WBLS does not make any reduction in operational efficiency of LSTM-WBLS.

AMS subject classifications: 62M20, 92B20

Key words: Precipitation forecast, Long Short-Term Memory (LSTM) network, Broad Learning System (BLS), Weighted Broad Learning System (WBLS), Multi-factor predication.

1 Introduction

Short-term heavy rainfall can cause heavy rain and flood, and then cause secondary

Translated from *Journal of Nanjing University of Information Science & Technology*, 2023, 15(2): 180-186.
*Corresponding author. *Email addresses:* **hanyingcs@163.com** (Y. Han), **jeeaaan@qq.com** (J. Luo).

disasters such as flash floods and mudslides, which seriously threaten people's life and property safety. Therefore, mastering the regularity of precipitation and accurately predicting daily precipitation are of great guiding significance for the research and control of flood disasters [1].

There are two kinds of precipitation prediction methods: process-based method and data-driven method. The advantage of process-based precipitation prediction method is that it can explain the physical process of precipitation clearly, but the complexity of the physical process increases the difficulty of modeling, and a series of hypotheses are needed to solve the model. The data-driven method is empirical, which does not need to analyze the physical process of precipitation, but only predicts the precipitation based on historical data, and the model is simple and easy to operate.

Statistical methods and machine learning are the most common data-driven precipitation prediction methods. In terms of statistical methods, the most popular forecasting method in recent years is based on the Auto Regressive Integrated Moving Average (ARIMA) model [2-3]. The results show that when the precipitation time series is linear or close to linear, the statistical model can produce satisfactory prediction results, but when the time series is non-linear, the prediction results are often unsatisfactory. In view of this, machine learning methods suitable for complex nonlinear process modeling are widely used in precipitation prediction. Hartigan et al. [4] used Random Forest (RF) and Support Vector Regression (SVR) to predict precipitation and temperature in Sydney basin. Xiang et al. [5] predicted the data of 34 meteorological observation stations in Chongqing by using the dual-system cooperative influence model of decision tree and FR. Peng et al. [6] built a mixed model for daily precipitation prediction based on extreme learning machine and gene expression. Gou et al. [7] combined the advantages of genetic algorithm and BP neural network to study the prediction method of daily precipitation level in Tianjin. Rostam et al. [8] used a variety of optimization algorithms to optimize the multi-layer perceptron algorithm in order to explore any meaningful relationship between the large-scale climate index and precipitation in the Iranian capital.

However, traditional machine learning methods cannot capture the long-term memory of the input sequence [9], thus affecting the prediction accuracy. Long Short-Term Memory (LSTM) networks overcome these shortcomings. Wang Ziyue et al. [10] used the sentence state LSTM model to identify the speaker's intention. Wang Peng et al. [11] predicted the ultra-short-term probability of wind power based on small wavelength short-term memory network. Luo Jia et al. [12] combined LSTM and BLS to analyze public emotional tendency in sudden meteorological disaster events. In terms of precipitation forecast, Nguyen et al. [13] improved radar-based rainfall forecast by using LSTM. Shen Haojun et al. [14] used LSTM to study the summer precipitation in China. Ni et al. [15] put forward two kinds of improved LSTM models (WD-LSTM and CNN-LSTM), and discussed their application in runoff and rainfall prediction respectively. Kang et al. [16] selected a multi-input variable LSTM model to forecast the daily precipitation in Jingdezhen, Jiangxi Province.

Although the precipitation prediction model based on LSTM has shown strong

advantages, the existing models have not solved the problem of time lag in the prediction. This is mainly due to the need for cyclic weight adjustment in LSTM training. It is noted that the new Broad Learning System (BLS) has the advantages of direct weight calculation, simple and fast operation, and can be used to improve LSTM. However, noise and outliers have adverse effects on the model, so Weighted Broad Learning System (WBLS) is proposed by applying the weighted penalty factor to the BLS. By automatically assigning appropriate weights to each sample, highly reliable samples are given higher weights, while suspicious outliers get lower weights. Therefore, the influence of abnormal samples on modeling is reduced. Combining the advantages of the two algorithms, the LSTM-WBLS daily precipitation prediction model is proposed in this paper.

In order to verify the new model effectively, this paper selected Badong Station in Hubei province to conduct an empirical study of daily precipitation prediction. In terms of prediction accuracy, compared with existing precipitation prediction models, this model has the best performance in three evaluation indexes: root mean square error (RMSE), mean absolute error (MAE) and determination coefficient (R^2). In terms of stability, by analyzing the influence of time step size of 1, 3 and 5 d on the prediction accuracy of each model, it is proved that although the prediction accuracy of all models will decrease with the increase of time step size, the model in this paper still performs best in three evaluation indexes of RMSE, MAE and R2 under different time step size. In terms of operational efficiency, the LSTM-WBLS model added with WBLS has no decrease in operational efficiency compared with the LSTM model because of its convenient and fast computation.

2 Model principle and structure

In this paper, the basic structure and principle of LSTM and WBLS are presented first, and then the LSTM-WBLS prediction model based on multiple factors is presented.

2.1 LSTM principle and structure

The structure of LSTM is shown in Figure 1 [17].

In Figure 1, $\boldsymbol{x}_t$ is the input vector, $\boldsymbol{i}_t$ is the input state in time step t, $\boldsymbol{f}_t$ is the forgotten state in time step t, $\boldsymbol{o}_t$ is the output state in time step t, $\boldsymbol{h}_t$ and $\boldsymbol{C}_t$ are respectively the hidden state and cell state in time step t, and h_{t-1} and $\boldsymbol{C}_{t-1}$ are respectively the hidden state and cell state in time step $t-1$. Nonlinearity is added in the form of tanh and sigmoid activation functions σ.

The LSTM principle is as follows:

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_f \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_f\right), \tag{1}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_i \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_i), \tag{2}$$

$$\widetilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_c \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_c), \tag{3}$$

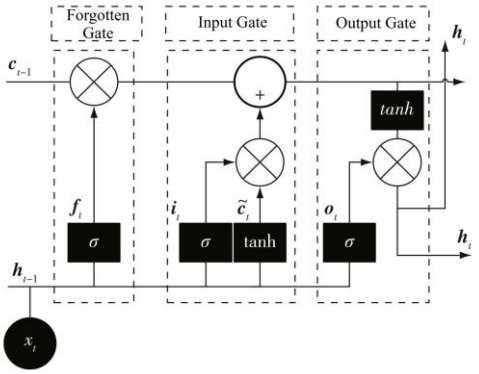

Figure 1: Structure principle of LSTM

$$\boldsymbol{C}_t = \boldsymbol{f}_t * \boldsymbol{C}_{t-1} + \boldsymbol{i}_t * \widetilde{\boldsymbol{C}}_t, \tag{4}$$

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_o \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_o), \tag{5}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t * \tanh(\boldsymbol{C}_t), \tag{6}$$

where $\boldsymbol{W}_f, \boldsymbol{W}_i, \boldsymbol{W}_c, \boldsymbol{W}_o$ represent the corresponding weight vectors of the forgetting gate, input gate, memory unit and output gate respectively; $\boldsymbol{b}_f, \boldsymbol{b}_i, \boldsymbol{b}_c, \boldsymbol{b}_o$ represent the deviation variables of the forgetting gate, input gate, memory unit and output gate respectively; $*$ is the Hadamard product of the matrix.

2.2 WBLS principle and structure

Chen et al. [18] proposed BLS in early 2019 and WBLS in 2020 to reduce the impact of abnormal samples on modeling [19].

Suppose $\boldsymbol{X}$ contains N samples, each with M dimensions, and $\boldsymbol{Y}$ is the output matrix belonging to $\mathbf{R}^{N\times Q}$, where Q is the dimension of the output. The structure of WBLS is shown in Figure 2[19].

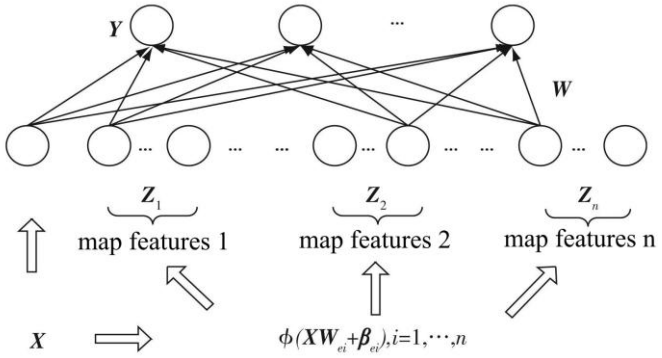

Figure 2: Structure principle of WBLS

The n-TH feature is mapped by equation (7) to generate p nodes:

$$\boldsymbol{Z}_i = \phi(\boldsymbol{X}\boldsymbol{W}_{ei} + \boldsymbol{\beta}_{ei}), i = 1,2,\cdots,n, \tag{7}$$

where $\boldsymbol{W}_{ei}$ and $\boldsymbol{\beta}_{ei}$ are randomly generated weights and biases. Finally, the WBLS model can be expressed as

$$\boldsymbol{Y} = [\boldsymbol{X}|\boldsymbol{Z}_1,\cdots,\boldsymbol{Z}_m]\boldsymbol{W} = \boldsymbol{H}\boldsymbol{W}. \tag{8}$$

2.3 LSTM-WBLS daily precipitation prediction model

Through the above analysis, combining the advantages of deep learning and weighted width learning, this paper proposes the LSTM-WBLS daily precipitation prediction model. The overall architecture of the model is shown in Figure 3.

The detailed operation steps of LSTM-WBLS multi-factor daily precipitation prediction model are as follows:

1) Input of precipitation series data

There are some problems in the collection process of daily precipitation data, such as human error or machine failure, which leads to abnormal values in the collected data. In order to reduce the impact of deleting outliers on the model prediction results, the operation of filling 0 values for outliers is carried out in this paper. The daily precipitation data of multiple factors has the problem of inconsistent index data scale, and the value range of each dimension is too different, which will not easily converge in the training process and affect the learning process of the algorithm. Therefore, data need to be normalized. In this paper, maximum and minimum normalization is selected as shown in equation (9):

$$x_{\text{scaled}} = \frac{x - x_{\min}}{x_{\max}}, \tag{9}$$

$x_{\min}$ indicates the minimum value and $x_{\max}$ indicates the maximum value of the current data.

The normalized data is divided into training set, verification set and test set, and the training set is taken as the input of the LSTM model.

2) LSTM-based training

LSTM has five layers. The first four layers are LSTM network layer, which aims to fully extract the timing features of the input sequence. The fifth layer is a fully connected layer, which aims to convert the features extracted from the first-time step to the last time step of the LSTM network layer into fixed feature vectors. To prevent LSTM from overfitting during training, Dropout processing is performed after each layer of LSTM.

3) Forecasts based on WBLS

First, the output of the fully connected layer is taken as the first n group mapping feature set $\boldsymbol{Z}^n = [\boldsymbol{Z}_1, \boldsymbol{Z}_2, \cdots, \boldsymbol{Z}_n]$. Each group contains p nodes, then the number of hidden layer nodes of the mapping feature is np. Equation (8) can be obtained by combining with input $\boldsymbol{X}$.

Secondly, the WBLS connection weight can be calculated through the weighted ridge regression algorithm, and the solution problem is shown in equation (10):

$$\min_{W} f(\boldsymbol{W}) = \min_{W} \|\theta \boldsymbol{H}\boldsymbol{W} - \theta \boldsymbol{Y}\|_F^2, \tag{10}$$

where $\|\cdot\|_F$ refers to the F norm, θ is the weighted penalty factor of the sample weight, and $\boldsymbol{H}$ is the hidden layer, represented by equation (11).

$$\boldsymbol{H} = [\boldsymbol{X}|\boldsymbol{E}^m] = (\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_L) \in \mathbf{R}^{K \times L}, \tag{11}$$

where: $\boldsymbol{h}_j \in \mathbf{R}^K, j = 1, \cdots, L$ is the j-th node of hidden layer $\boldsymbol{H}$, $L = np + mq$ is the number of hidden layer nodes, K is the number of input time series.

Equation (10) is a least squares problem, which is a convex optimization estimate of **W**, aiming to find the output weight **W** that minimizes the training error. It is solved, and the results are shown in equation (12):

$$\boldsymbol{W} = \boldsymbol{H}^{+}\boldsymbol{Y}, \tag{12}$$

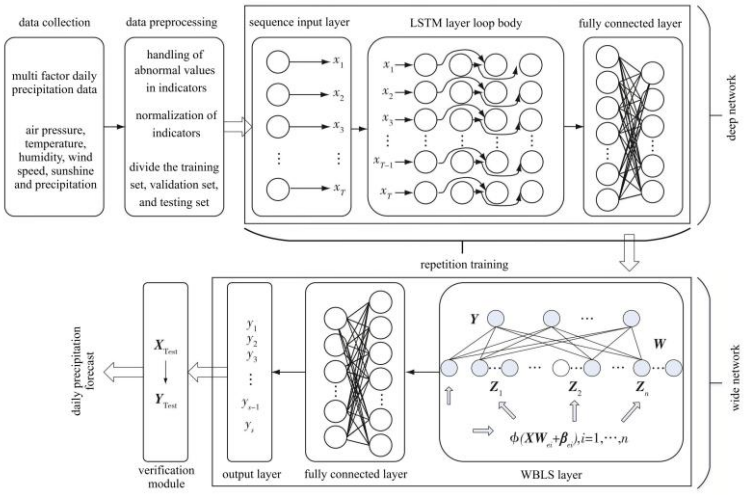

Figure 3: Overall framework of the LSTM-WBLS model

where $\boldsymbol{H}^{+}$ is the pseudo inverse of $\boldsymbol{H}$. But in general, the generalization error of the above solution may be large, especially for some pathological problems. In order to improve the generalization ability of the network, the F norm regular term is introduced into the original formula to prevent the network from overfitting, and the formula (13) is obtained:

$$\min_{W} f(\boldsymbol{W}) = \min_{W} \|\theta\boldsymbol{H}\boldsymbol{W} - \theta\boldsymbol{Y}\|_F^2 + C\|\boldsymbol{W}\|_F^2, \tag{13}$$

Formula (13) is a ridge regression problem. C represents a further constraint on the sum of squares of weights, which can be finally solved by equation (14) to obtain the weight $\boldsymbol{W}$:

$$\boldsymbol{W} = (C\boldsymbol{I} + \boldsymbol{H}^{T}\theta^{2}\boldsymbol{H})^{-1}\boldsymbol{H}^{T}\theta^{2}\boldsymbol{Y}. \tag{14}$$

4) Output forecast results

The weight $\boldsymbol{W}$ obtained by calculation in equation (14) is combined with the hidden layer $\boldsymbol{H}$ to get the final prediction result.

3 Case Analysis

3.1 Study area and data set description

Badong County, which belongs to the Tujia and Miao Autonomous Prefecture of Enshi, Hubei Province, is located in the southwest of Hubei Province. It has a subtropical monsoon climate, warm and rainy, hot and humid and foggy, with four distinct seasons. The average temperature of the hottest month is generally higher than 22 ℃, and the temperature of the coldest month is between 0 ℃ and 15 ℃. Most of the annual precipitation in Padang ranges from 800 mm to 1 600 mm. The weather in Padang is

aperiorical and the seasonal variation of precipitation is very significant, so it is difficult to predict its daily precipitation.

Data for this article were obtained from the website of the National Meteorological Center. The data range is the observed precipitation observation data of meteorological observation station in Badong Region from 2000 to 2020. A total of 7 671 days of data were set as the training set, verification set and test set with the ratio of 7∶2∶1, and the test set was the precipitation data of recent years.

3.2 Parameter setting and evaluation index

Daily precipitation is mapped to $S \times \tau \times D$ tensor data as input to the model. Where S is the samples, τ is the time steps, and D is the features. The model in this paper is the input of six dimensions of air pressure, temperature, humidity, wind speed, sunshine and precipitation, and the output of one dimension of precipitation. So D is 6.

Dropout is used to exit some neurons to prevent overfitting, and the P-value of random discard ratio is determined. Then through the fully connected layer, its output is taken as the mapping feature of WBLS layer, and together with the input X, the hidden layer H is formed. Finally, the output weight W. N_1 is the number of each mapping feature node, N_2 is the number of mapping feature, and C is L_2 regularization parameter. In this paper, the validation set is used to test the hyperparameters of this model, and the value is the optimal value selected after several experiments. The parameters used in this article are shown in Table 1.

Table 1: Main parameters of LSTM-WBLS model

layer	parameter	value
Input	Timesteps	1,3,6
LSTM	Layers	4
	Number of neurons	{256,128,64,32}
	Loss	mse
	Optimizer	Adam
	Batch_size	64
	epochs	80
Dropout	P	0.2
Dense	Layers	1
	Number of neurons	20
	Activation	relu
WBLS	N_1	20
	N_2	30
	C	le-5

RMSE, MAE and R^2 were selected to evaluate the accuracy of the algorithm. RMSE is very sensitive to the error of predicted value and can reflect the accuracy of prediction. MAE can avoid the problem of error canceling each other, and can accurately reflect the actual prediction error. R^2 is often used to determine the degree of fit of regression equations, with a value between 0 and 1, with a larger value indicating better predictive performance of the model.

$$E_{\mathrm{RMSE}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\tilde{y}(i) - y(i)\right)^2}, \tag{15}$$

$$E_{\mathrm{MAE}} = \sum_{i=1}^{n}\left|\left(\tilde{y}(i) - y(i)\right)\right|, \tag{16}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\tilde{y}(i) - y(i)\right)^2}{\sum_{i=1}^{n}\left(\bar{y}(i) - y(i)\right)^2}, \tag{17}$$

where y_i represents true monthly precipitation, $\tilde{y}$ represents predicted monthly precipitation, and $\bar{y}$ represents average monthly precipitation.

3.3 Comparison and analysis with existing models

The existing model and the model in this paper were compared and analyzed, taking the prediction length 1 d as an example, and the comparison results were shown in Table 2. Compared with the existing SVM [4], EEMD-ARIMA [3], LSTM [13], CNN-LSTM [15] and LSTM-BLS models, the RMSE value of the proposed model is reduced by 50.20%, 47.58%, 37.00%, 34.80% and 17.54%, respectively. The MAE value decreased by 55.29%, 53.19%, 49.20%, 48.00% and 22.72%, respectively. R^2 values increased by 0.209, 0.189, 0.078, 0.058 and 0.015, respectively. Obviously, the model presented in this paper is optimal in the three indexes, which proves the effectiveness and accuracy of the model presented in this paper.

Table 2: Comparison of evaluation indicators of each model

MODEL	RMSE/mm	MAE/mm	R^2
SVM	4. 191	2. 427	0. 711
EEMD-ARIMA	3. 981	2. 318	0. 731
LSTM	3. 313	2. 136	0. 842
CNN-LSTM	3. 201	2. 087	0. 861
LSTM-BLS	2. 531	1. 404	0. 905
LSTM-WBLS	2. 087	1. 085	0. 920

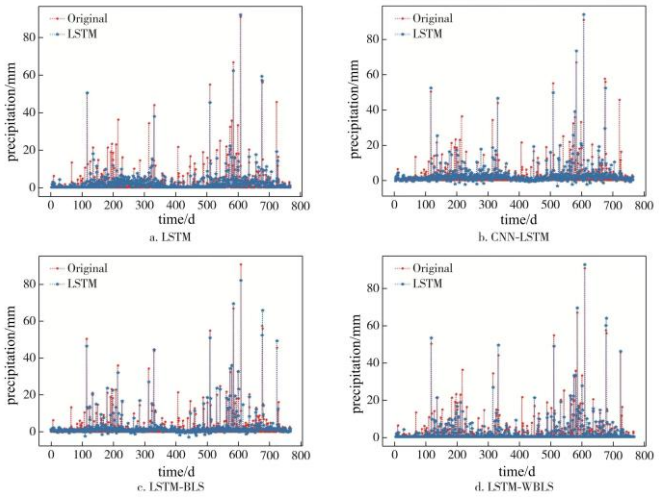

Figure 4: Visualized comparison of prediction results by LSTM series related models

In order to further verify the validity of the model in this paper, the prediction of LSTM series related models is visualized. The precipitation series of the test set was fitted with the predicted values of each model, and the comparison and visualization were shown in Figure 4. For the convenience of mapping, the first day corresponds to the true and predicted precipitation value on November 26, 2018, and has lasted for 767 days until December 31, 2020.

As can be seen from Figure 4, the prediction results of the model in this paper are significantly better than all existing models at the date of precipitation abrupt change. It should be noted that the existing LSTM-based models (Figure. 4a, 4b) have an inevitable lag in prediction, so they cannot be accurately predicted. Figure 4c basically solves the problem of lag due to the addition of BLS, but the adverse effects of noise and outliers on prediction still exist. In this paper, a weighted penalty factor is added to the model based on Figure 4c, and the prediction results are optimal (Figure 4d).

3.4 Comparison with single factor model

In order to further verify the effectiveness of the model presented in this paper, a comparison was made with the model with single-factor precipitation input, and the results were shown in Table 3. It can be seen that the prediction of multi-factor input is much higher than that of single-factor input model. The reason is that there are too many zeros in the data for the single-input model to accurately predict. Some of the data sets are shown in Table 4. Considering the influence of various meteorological factors, the model in this paper can accurately predict precipitation.

Table 3: Comparison with single factor model

mm

MODEL	RMSE	MAE
LSTM (single-factor)	7. 610	4. 008
LSTM-WBLS (single-factor)	6. 610	3. 890
LSTM-WBLS	2. 087	1. 085

Table 4: Part of the dataset in January 2001

pressure/ kPa	temperature/ ℃	precipitation/ mm	relative humidity/ %	wind speed/ (m/s)	sunshine/ h
97. 3	9. 2	0. 0	82	0. 3	1. 2
97. 4	8. 2	0. 0	84	1. 8	1. 5
98. 3	7. 9	0. 9	69	3. 0	0. 0
98. 1	4. 7	0. 1	83	1. 3	0. 0
97. 7	5. 9	0. 0	68	2. 0	1. 0
97. 8	7. 4	0. 0	75	1. 5	0. 6
97. 7	7. 5	1. 1	76	3. 8	0. 1
97. 8	6. 6	3. 3	82	3. 0	0. 0
98. 1	6. 3	0. 0	77	2. 8	0. 0
97. 9	6. 6	0. 0	67	2. 0	0. 5

3.5 Stability Analysis

Without changing the parameters in the model, the prediction length was set to 3 days and 5 days respectively, and the daily precipitation was predicted, as shown in Table 5. Combined with the prediction results with prediction length of 1 d, it can be seen that the prediction accuracy of all prediction models decreases with the increase of prediction length. However, the prediction accuracy of LSTM-WBLS model is still better than other models under different prediction lengths. This result verifies the stability of the proposed model.

Table 5: Comparison of evaluation indicators for each model under different prediction lengths

MODEL	RMSE/mm		MAE/mm		R^2	
	3 d	5 d	3 d	5 d	3 d	5 d
SVM	4. 325	4. 457	2. 601	2. 725	0. 701	0. 682
EEMD-ARIMA	4. 056	4. 216	2. 407	2. 421	0. 721	0. 705
LSTM	3. 408	3. 421	2. 234	2. 351	0. 831	0. 822
CNN-LSTM	3. 307	3. 392	2. 179	2. 306	0. 855	0. 843
LSTM-BLS	2. 641	2. 728	1. 503	1. 581	0. 889	0. 873
LSTM-WBLS	2. 145	2. 216	1. 202	1. 272	0. 909	0. 886

3.6 Calculation efficiency analysis

Operation efficiency is also the main evaluation index of the algorithm. When both LSTM-WBLS and LSTM training are guaranteed to achieve optimal results, the operation efficiency pair is shown in Table 6. As can be seen from Table 6, the training time of LSTM-WBLS is only about 2 s longer than that of LSTM, and the efficiency is not significantly reduced. The reason is that WBLS does not require a large number of operations, directly calculate the weight characteristics, so that LSTM-WBLS compared with LSTM, the operation efficiency will not have much decline.

Table 6: Comparison of operation efficiency between LSTM-WBLS and LSTM

MODEL	training time/s		
	time step=1d	time step=3d	time step=5d
LSTM	655. 925	945. 031	1 194. 134
LSTM-WBLS	657. 231	947. 132	1 196. 891

4 Conclusions

In view of the shortcomings of existing daily precipitation prediction models, a new LSTM-WBLS daily precipitation prediction model is proposed in this paper. Through empirical research, the model in this paper solves the lag problem in LSTM prediction by means of the characteristics of WBLS, which does not require a lot of training and

calculates weights directly by pseudo-inverse, and the operation efficiency does not decrease. By automatically assigning appropriate weights to each sample, the samples with high reliability are given higher weights, while the suspicious outliers are given lower weights, which reduces the influence of abnormal samples and improves the prediction accuracy and stability. This paper discusses the possibility of combining the advantages of deep learning and width learning in precipitation prediction, which provides a new idea for precipitation prediction. In this model, only historical meteorological data and specific daily precipitation data are considered, and geographical and geomorphic features will be added in the future to further improve the forecast accuracy of daily precipitation.

Acknowledgments

This work is supported by the Southern Marine Science and Engineering Laboratory of Guangdong Province (Zhuhai) Foundation (Grant No. SML2020SP007) and the National Natural Science Foundation of China (Grant No. 62076136).

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] H. P. Wang, B. Zhang, Z. H. Liu, et al., Chaos theory-based comparative study on monthly rainfall characteristics in Wuhan and Yichang during recent 60 years, *J. Nat. Disasters*, 2012, 21(6): 111-118.

[2] G. H. Zhang, Analysis and prediction of precipitation trend in Weinan city based on ARIMA model, *Value Eng.*, 2019, 38(34): 197-199.

[3] Y. Hu, and J. Wu, Analysis and prediction of precipitation spatial characteristics based on ARIMA model, *Jiangxi Sci.*, 2021, 39(1): 99-104.

[4] J. Hartigan, S. MacNamara, L. M. Leslie, et al., Attribution and prediction of precipitation and temperature trends within the Sydney catchment using machine learning, *Climate*, 2020, 8(10): 120.

[5] B. Xiang, C. F. Zeng, X. N. Dong, et al., The application of a decision tree and stochastic forest model in summer precipitation prediction in Chongqing, *Atmosphere*, 2020, 11(5): 508.

[6] Y. Z. Peng, H. S. Zhao, H. Zhang, et al., An extreme learning machine and gene expression programming-based hybrid model for daily precipitation prediction, *Int. J. Comput. Intell. Syst.*, 2019, 12(2): 1512-1525.

[7] Z. J. Gou, J. L. Ren, M. Xu, et al., Application of GA-BP algorithm based on Hadoop in precipitation forecast, *Comput. Syst. Appl.*, 2019, 28(9): 140-146.

[8] M. G. Rostam, S. J. Sadatinejad, and A. Malekian, Precipitation forecasting by large-scale climate indices and machine learning techniques, *J. Arid Land*, 2020, 12(5): 854-864.
[9] C. P. Shen, A trans-disciplinary review of deep learning research for water resources scientists, *Water Resour. Res.*, 2018, 54(11): 8558-8593.
[10] Z. Y. Wang, and X. Shao, Speaker intention recognition based on S-LSTM model and slot-gate, *J. Nanjing Univ. Inf. Sci. Technol. (Nat. Sci. Ed.)*, 2019, 11(6): 751-756.
[11] P. Wang, Y. H. Sun, S. W. Zhai, et al., Ultra-short-term probability prediction of wind power based on wavelet decomposition and long short-term memory network, *J. Nanjing Univ. Inf. Sci. Technol. (Nat. Sci. Ed.)*, 2019, 11(4): 460-466.
[12] J. Luo, L. H. Wang, S. S. Tu, et al., Analysis of public sentiment tendency in sudden meteorological disasters based on LSTM-BLS, *J. Nanjing Univ. Inf. Sci. Technol. (Nat. Sci. Ed.)*, 2021, 13(4): 477-483.
[13] D. H. Nguyen, J. B. Kim, and D. H. Bae, Improving radar-based rainfall forecasts by long short-term memory network in urban basins, *Water*, 2021, 13(6): 776.
[14] H. J. Shen, Y. Luo, Z. C. Zhao, et al., Prediction of summer precipitation in China based on LSTM network, *Clim. Chang. Res.*, 2020, 16(3): 263-275.
[15] L. L. Ni, D. Wang, V. P. Singh, et al., Streamflow and rainfall forecasting by two long short-term memory-based models, *J. Hydrol.*, 2020, 583: 124296.
[16] J. L. Kang, H. M. Wang, F. F. Yuan, et al., Prediction of precipitation based on recurrent neural networks in Jingdezhen, Jiangxi province, China, *Atmosphere*, 2020, 11(3): 246.
[17] S. Hochreiter, and J. Schmidhuber, Long short-term memory, *Neural Comput.*, 1997, 9(8): 1735-1780.
[18] C. L. P. Chen, Z. L. Liu, and S. Feng, Universal approximation capability of broad learning system and its structural variations, *IEEE Trans. Neural Netw. Learn. Syst.*, 2019, 30(4): 1191-1204.
[19] F. Chu, T. Liang, C. L. P. Chen, et al., Weighted broad learning system and its application in nonlinear industrial process modeling, *IEEE Trans. Neural Netw. Learn. Syst.*, 2020, 31(8): 3017-3031.