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