ISSN 2616-5775 Vol. 5, Issue 9: 1-6, DOI: 10.25236/AJCIS.2022.050901

Study on Urban Rainfall Trend Based on Neural Network and Grey Correlation Analysis Model

Yi Wu[1], Lei Huang[2], Anqi Chen[1], Cai Chen[1]

[1]School of Automotive and Traffic Engineering, Hubei University of Arts and Sciences, Hubei, China
[2]College of Science, Liaoning Technical UniversitFux, Liaoning, China

Abstract: *This paper is based on a quantitative analysis of the 2021 flood event in Zhengzhou City. The precipitation data of more cities in China are collected and compiled for many years, and the precipitation trends of the cities they collect are analyzed. It also collects weather data from more cities, uses various methods to forecast and analyze cities that may experience extreme rainfall in the future, and compares and analyzes the forecast results.*

Keywords: *Quadratic exponential smoothing method; LSTM neural network model; Elm Algorithm*

1. Introduction

In the context of global warming, many parts of China have been affected by extreme weather recently, causing serious economic losses and casualties. Strong rainfall occurs frequently. For example, the once-in-a-thousand-year rainstorm in Zhengzhou on July 20, 2021 and other disasters deserve our deep thoughts. In this paper, the quadratic exponential smoothing method and LSTM model are used to predict the future rainfall trend and rainfall in typical cities such as "Beijing and Guangzhou" in order to predict natural disasters and reduce losses. This paper uses the clustering method to obtain a rainfall information tree for each site, which can see the distribution of rainfall differences at the site. Based on this, the weather statistics of Beijing from 2005-2021 and Guangzhou from 2015-2021 are collected to analyze the annual precipitation. By testing, the time series is a stable white noise series. The ARMA order P and Q are determined to be 2, i.e. ARMA (2,2). This model is then used to predict the rainfall for the next ten years. The Guangzhou city data failed the time stability test, and the autocorrelation coefficient and partial autocorrelation coefficient points were within two standard deviations, so the quadratic exponential smoothing method was used. The trend of urban precipitation is alternately increasing and decreasing precipitation. Meanwhile, elm algorithm and LSTM neural network model algorithm were used to predict Beijing and Guangzhou as cities with possible extreme rainfall in the future. The accuracy of elm algorithm in testing Beijing weather data was 93.333% for the training set and about 37.5% for the test set, and 96.67% for the training set and 40.5% for the test set for Guangzhou weather. The LSTM model is used to predict the precipitation in Beijing for the next two years. The annual precipitation for the next three years (2022, 2023 and 2024) is 794.6, 1017 and 533.2, respectively, and the precipitation for Guangzhou in 2022 and 2023 is 3012 and 3324, respectively.

2. Symbol and Assumptions

2.1. Symbol Description

Symbol Description is shown in Table 1.

Table 1: Symbol Description

Symbol	Meaning
$X^{\hat{}}$	Merge all points in the first two clusters
D_{JM}	Distance from each point x to the center point
T	forecast period
$X^{\hat{}}_{T+t}$	the predicted value at time $t + T$.
X_i	original sequence data
a	Smoothing parameters
$S^{(i)}_t$	Smoothing value for the ith time
γ	Correlation degree
ρ	Resolution coefficient

ISSN 2616-5775 Vol. 5, Issue 9: 1-6, DOI: 10.25236/AJCIS.2022.050901

2.2. Fundamental assumptions

(1) It is assumed that the historical weather data collected in Beijing, Shanxi, Taiyuan and other cities are true and effective.

(2) It is assumed that the data after outlier processing is correct and effective, and can correctly reflect the correlation between various features.

3. Model construction and solving

3.1. Establish ARMA model

Conditions for establishing ARMA model:

In practical problems, most of the non-stationary series can become stationary time series after one or more differences, so the model can be established.

Set X_1 X_2 X_N as original sequence data, Then the primary smoothing value $S^{(1)}{}_{t.}$ and the secondary smoothing value $S^{(2)}{}_t$ are respectively[1].

$$\begin{cases} S_t^{(1)} = aX_t + (1-a)S_{t-1}^{(1)} \\ S_t^{(2)} = aS_t^1 + (1-a)S_{t-1}^{(2)} \end{cases} \tag{1}$$

Where a is the smoothing parameter.

$$\hat{X}_{T+t} = A_t + B_t \cdot T \tag{2}$$

$$\begin{Bmatrix} A_t = 2S_t^{(1)} - S_t^{(2)} \\ B_t = \dfrac{a}{1-a}\left(S_t^{(1)} - S_t^{(2)}\right) \ t = 1, 2, \cdots N \end{Bmatrix} \tag{3}$$

3.2. Solution of ARMA model

The cities selected in this question are Beijing and Guangzhou, both of which have experienced rainstorms, such as the severe rainstorm in Beijing on July 21, 2012, the severe rainstorm in North China on July 19, 2016 and the severe rainstorm in Guangzhou on May 3, 2018.

Statistical map of precipitation in Beijing from 2005 to 2021 is shown in Figure 1. As the time series is stable (through ADFtest test), ARMA (2,2) model is determined through AIC order determination to predict the precipitation trend in Beijing after ten years, as shown Figure 2.

Determine whether there is autocorrelation is shown in Figure 3.

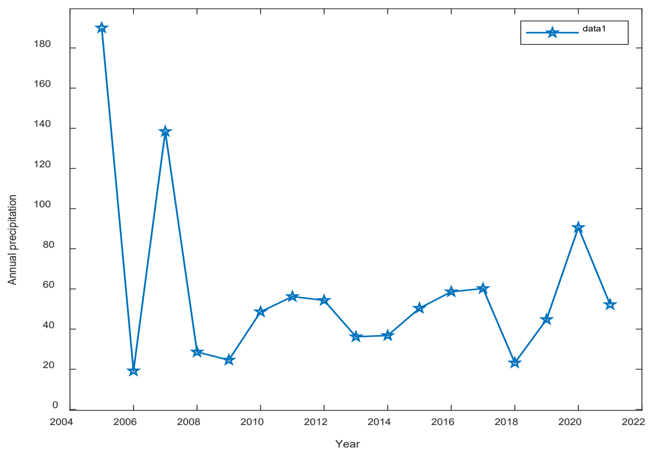

Figure 1: Statistical map of precipitation in Beijing from 2005 to 2021

ISSN 2616-5775 Vol. 5, Issue 9: 1-6, DOI: 10.25236/AJCIS.2022.050901

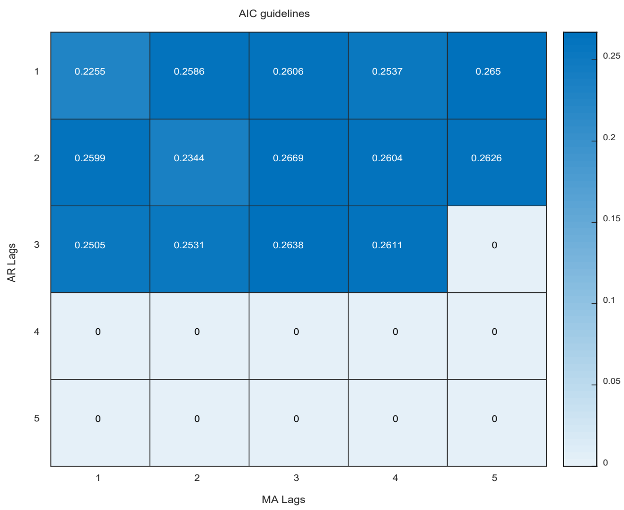

Figure 2: AIC order diagram

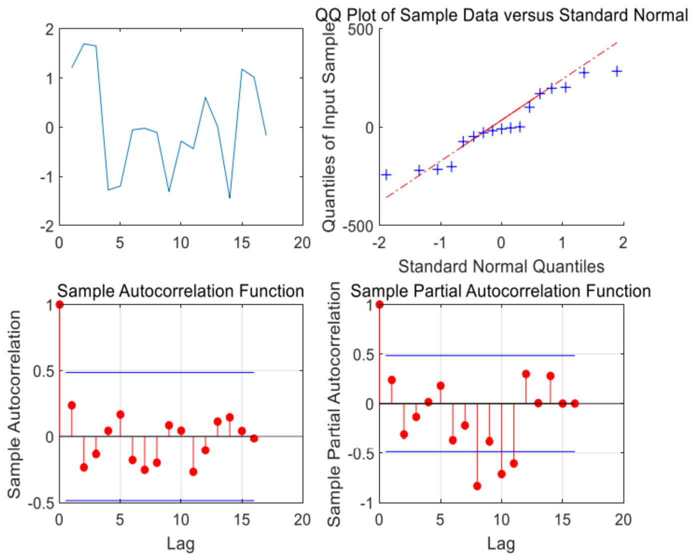

Figure 1: Determine whether there is autocorrelation

It can be roughly judged from the graph that the res sequence is a white noise sequence and there is no autocorrelation.

The Figure 4 shows the forecast of precipitation in Beijing in the next ten years. It can be seen that there will be no severe rainstorm in Beijing in the next ten years.

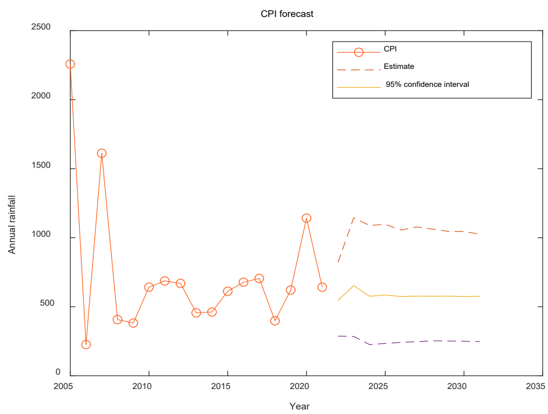

Figure 4: Forecast of precipitation in Beijing in the next ten years

ISSN 2616-5775 Vol. 5, Issue 9: 1-6, DOI: 10.25236/AJCIS.2022.050901

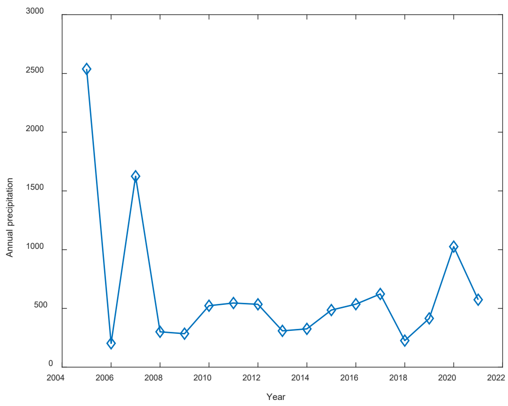

Figure 5: Forecast of precipitation in Guangzhou. the next ten years

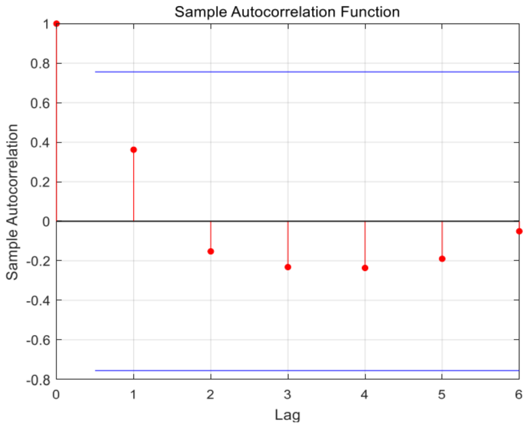

Figure 6: Adftest test for autocorrelation in coefficient

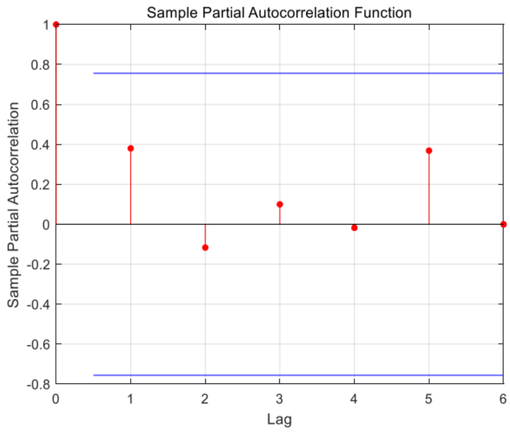

Figure 7: ADFtest test for partial autocorrelation coefficient

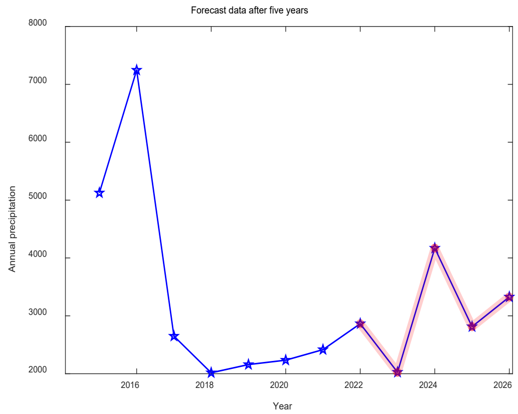

Figure 8: Forecast of rainfall in Guangzhou in the next five years

Forecast of precipitation in Guangzhou in the next ten years, adftest test for autocorrelation coefficient and ADFtest test for partial autocorrelation coefficient are shown in Figure 5-7[2].

Forecast of rainfall in Guangzhou in the next five years is shown in Figure 8. The increase and decrease of precipitation occur alternately.

3.3. Elm algorithm for rainfall prediction

(1) After importing the weather data of Beijing, the first typical city, into Matlab, run the program.

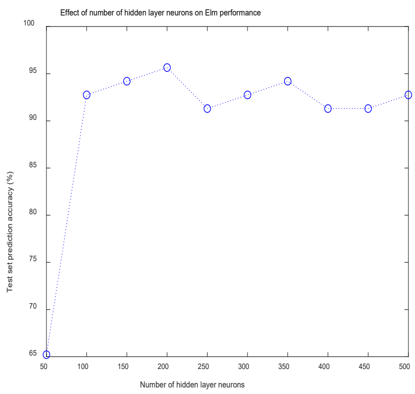

Figure 9: Effect of the number of hidden neurons on Elm performance

ISSN 2616-5775 Vol. 5, Issue 9: 1-6, DOI: 10.25236/AJCIS.2022.050901

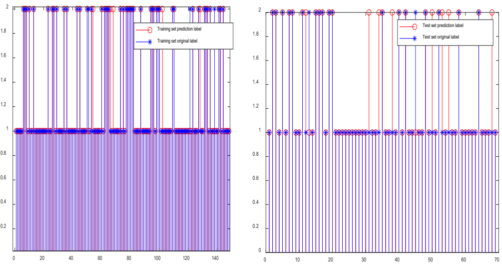

Figure 10: comparison between simulation training set and test set

Effect of the number of hidden neurons on Elm performance and comparison between simulation training set and test set are shown in Figure 9-10.

The model is trained with the data of the training set, and then the error on the test set is used as the generalization error of the final model in dealing with the real scene. Through the test set, the trained model calculates the error on the test set, which can be considered as the approximation of the generalization error, and the final effect of verifying the model can be obtained [3]. The correct rate of training set is 93.333%, and the correct rate of test set is approximately 92.5%.

(2) After importing the weather data of the second typical city Guangzhou and Guangzhou into Matlab, run the program.

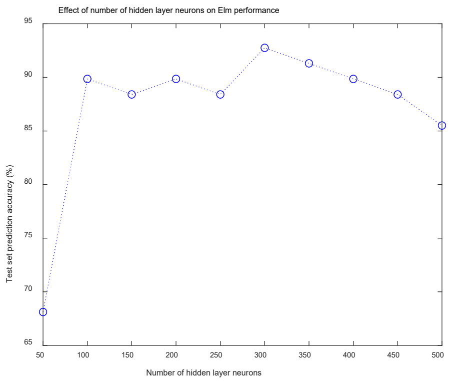

Figure 11: Effect of the number of hidden neurons Elm performance

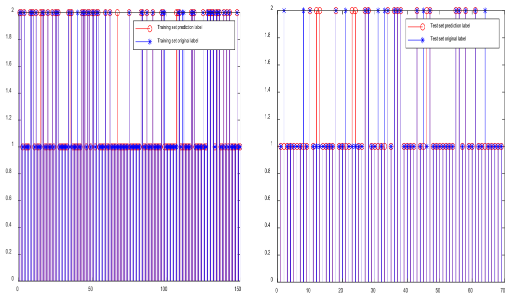

Figure 12: Comparison between simulation on training set and test set

Effect of the number of hidden neurons on Elm performance and comparison between simulation training set and test set are shown in Figure 11-12.

The correct rate of training set is 96.67%, and that of test set is 40.5%. By using a large number of data sets to train the model, the error of the model on the data set iteratively trains the model, and a reasonable model fitting the data set is obtained; Apply the trained and adjusted model to this problem. The smaller the error of the prediction result of the model on the real data, the better, and the best prediction result is obtained.

3.4. Prediction of rainfall by LSTM neural network model

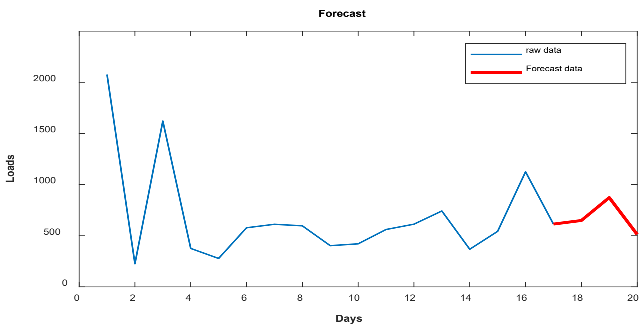

Figure 13: Prediction of precipitation in Beijing by LSTM algorithm

ISSN 2616-5775 Vol. 5, Issue 9: 1-6, DOI: 10.25236/AJCIS.2022.050901

In the LSTM model, the conventional neuron, that is, a unit that applies S-type activation to its linear combination of inputs, is replaced by a storage unit. Each storage unit is associated with an input gate, an output gate and an internal state that is transmitted to itself without interference across time steps[4-5].

Prediction of precipitation in Beijing by LSTM algorithm is shown in Figure 13. The red line represents the predicted value. The annual precipitation in the next three years (2022, 2023 and 2024) is 794.61017533.2 respectively.

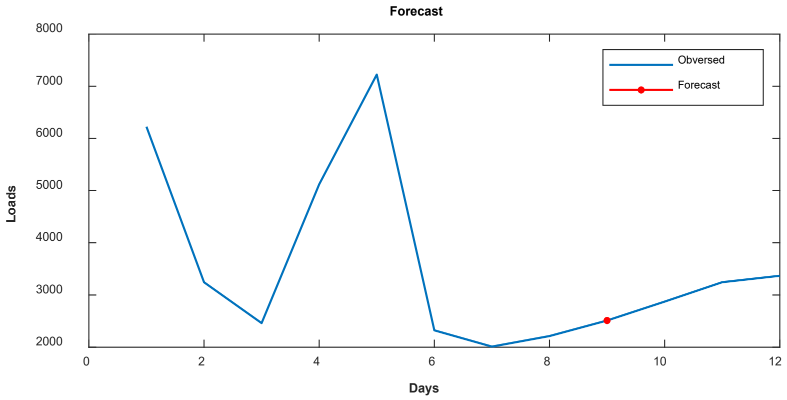

Figure 14: Forecast of precipitation in Guangzhou by LSTM algorithm

Forecast of precipitation in Guangzhou by LSTM algorithm is shown in Figure 14. The solver is ADMA, 500 rounds of training, the initial threshold is 1, and the initial learning rate is 0.005. In the 125th round, multiply the factor by 0.2 to reduce the learning rate.

It is predicted that the precipitation of Guangzhou in 2022 and 2023 will be 3012 and 3324 respectively.

The prediction result of LSTM is better than that of elm algorithm and time series algorithm.

4. Conclusion

This paper is based on a quantitative analysis of the flood events in Zhengzhou in 2021. Collect and organize the precipitation data of more cities in China over the years, and analyze the precipitation trends of the cities that you have collected. And collect more urban weather data, use a variety of methods to predict and analyze cities that may experience extreme rainfall in the future, and compare and analyze the prediction results.

The model applied in this paper has the following advantages.

Advantages of quadratic smoothing algorithm: It has the advantages of simple calculation, less sample size, strong adaptability and stable results. It can be used not only for short-term prediction, but also for medium and long-term prediction.

Elm algorithm advantages: Elm has the advantages of fast learning speed, high generalization accuracy, and will not fall into local minimum. It can adopt a variety of excitation functions (it can meet infinite differentiability).

References

[1] Zhu Wanli. Application of correlation analysis principle in Library and information analysis [J]. Jiangsu Science and technology information, 2019, 36 (01): 9-12.
[2] Qi Cuige, Tian Jianshe. Henan Water Conservancy and South-to-North Water Transfer, 2007(07): 16+20.
[3] Dong Wei. Correlation analysis between tree rings and precipitation in arid and semi-arid areas of Liaoning [J]. Heilongjiang water conservancy science and technology, 2020,48 (10): 7-9+64.
[4] Chang Xia, Liu Sai 'e. Computer and digital engineering, 2019, 47(08): 2002-2005.
[5] Zhong Shuixin. Research progress on influence mechanism and prediction methods of topography on precipitation [J]. Plateau elephant, 2020, 39 (05): 1122-1132.