

Article

Not peer-reviewed version

Application Raindrop Size Distribution Base on LSTM in Precipitation Forecasting in Guizhou

Wang Fu-Zeng , [An Xue-Jiao](#) ^{*} , Wang Qiu-Song , Gu Xiao-Ping , [Li Zi-Xin](#)

Posted Date: 30 August 2023

doi: 10.20944/preprints202308.2068.v1

Keywords: Raindrop Size Distribution; Microphysical Characteristics; LSMT Neural Network; Precipitation Forecasting



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Application Raindrop Size Distribution Base on LSTM in Precipitation Forecasting in Guizhou

Wang Fu-Zeng ¹, An Xue-Jiao ^{1,*}, Wang Qiu-Song ¹ and Gu Xiao-Ping ², Li Zi-Xin ¹

¹ College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China

² Guizhou Ecological Meteorology and Satellite Remote Sensing Center, Guiyang, Guizhou 550002.

* Correspondence: 1229072941@qq.com

Abstract: Raindrop size distribution (RSD) is an index for reflecting precipitation characteristics. Analyzing the differences of RSD plays an important role in understanding the precipitation microphysical processes and improving quantitative precipitation prediction of radar. In this paper, the RSD data of Dafang (57708) station, Majang (57828) station, and Luodian (57916) station (with an altitude of 1722.7m, 985.0m, and 441.5m, respectively) in Guizhou are analyzed according to the precipitation microphysical characteristics. First, Particles with particle size less than 1mm contributes the highest value to the density of particle number, and decrease with the descending altitude. Second, the GAMMA distribution fit shows a better fit compare to M-P distribution fit, and GAMMA distribution fit increases with the ascending altitude. Third, the mass-weighted average diameter is not sensitive enough to the change of precipitation intensity with correlation coefficients of 54.84%, but there is a obvious relationship between the average volume diameter and the precipitation intensity with correlation coefficients of 69.15%. Then, the precipitation prediction model is established using LSTM neural network after fusing the representative microphysical characteristics of RSD with radar and rain gauge data. The precipitation prediction model is applied to the Dafang (57708) site to predict precipitation for time range of 0-180 minutes, and it is found that for the prediction of convective cloud and stratiform cloud precipitation, the 60-minute prediction results are the most consistent with the actual precipitation, where the correlation coefficients are 92.87% and 92.57%, respectively. Conclusively, the results demonstrate that combining with the RSD base data could improve the reliability of precipitation forecasting.

Keywords: raindrop size distribution; microphysical characteristics; LSMT neural network; precipitation forecasting

0. Introduction

Precipitation is closely related to human life, and impact on crop production, the transportation and the natural environment. Therefore, quantitative precipitation estimation (QPE) has always been the focus of meteorological and hydrological fields (WU Jiang-sheng et al. [1]). The uneven and unstable spatiotemporal distribution of precipitation are the important causes of natural disasters such as floods. The large-scale macroscopic climate conditions and small-scale microscopic physical processes are both important factors determining the formation of precipitation (HUANG Xingyou et al. [2] and LIN Hou-bo et al. [3]). From 20:00 on August 9, 2019, to 20:00 on August 13, 2019, a heavy precipitation hit Shandong Province, Shen Gaohang et al. [4] used RSD data and double-polarized Doppler radar data to analyze the minute precipitation of Zhangqiu Station during this time, It has played a certain role in forecasting. Therefore, as an important indicator of reflecting the changing process and internal mechanism of precipitation, RSD plays a significance role in evaluating the effect of weather modification and improving the accuracy of radar quantitative prediction.

The study of RSD has a long history. Since the 1960s, research has been carried out at home and abroad and a series of important results have been achieved. In recent years, many meteorologists have studied the application of RSD and Doppler weather radar products. The results show that the highest frequency of occurrence of raindrops is about 0.5 mm diameter (CAI Zhao et al. [5]). Zeng Guangyu et al. [6] established the CSU-LPA algorithm based on the data of S-band dual polarization

radar, rain gauge and laser RSD measurements. Rivelli Zea Lina et al. [7] analyzed the RSD in watersheds in America use the covariability between the physical parameters of vertical raindrop size distributions, suggests that precipitation observed in Córdoba may confuse existing methods using droplet size distributions to determine rain types. François Mercier et al. [8] proposed a new framework for retrieving vertical raindrop size distribution and wind vertical profiles during light rain events to better characterize the microphysical processes of rainfall. According to previous studies, the improvement of precipitation prediction accuracy mainly depends on improving the quality of measurement data or seeking combinations of physical variables which have a strongly correlation with the predicted results.

At present, the precipitation forecasting methods mainly include the weather radar detection echo extrapolation method, the conceptual model forecasting method, the numerical model forecasting method, etc. Zhu Ping et al. [9] used the ground-based measurements of volume scanning of Doppler weather radar to extrapolate the echo data of its intensity data. Liu Guozhong et al. [10] through the mathematical statistics and climate analysis methods obtained the distribution characteristics of the significant increases and drop in 24 hours maximum and minimum temperature over Baise weather station, and obtained its interval division and categories. Cheng Conglan et al. [11] integrated radar extrapolation proximity forecasting and mesoscale numerical model forecasting technology, and they conducted the nowcasting experiments on lightning, heavy rain and precipitation, respectively. Through experimental analysis, it is concluded that these three methods have large errors in forecasting. To solve this problem, we propose a novel method based on the neural networks.

With the development of the machine learning (ML), more and more technologies in this field are applied to meteorology field. Numbers of scholars at home and abroad have achieved tons of outstanding works using machine learning to forecast weather. Since the neural network is a method with strong nonlinear adaptive information processing ability, it can extract the feature information from the input training samples and store the conversion relationship between the input and the output inside the network to achieve the goal of the recognition and prediction (Elshaboury N et al. [12]). For example, Suhail Sharadqah et al. [13] tested rainfall data using NAR models. Gunathilake Miyuru B et al. [14] used artificial neural network in combined with remote sensing information for precipitation estimation. Zhang Shuai et al. [15] compared prediction results by building long short-term memory network models, feed-forward neural network models, integrated moving average autoregressive models, and wavelet neural network models. All reflect the superiority of artificial neural networks in predicting precipitation.

At present, most meteorological precipitation forecasts are using single detection data source, and there is a strong dependence on the reliability of the single detection data. Therefore, this paper aims to establish a multi-source data fusion precipitation prediction model based on weather radar RSD and automatic rain gauge data, in order to reduce the impact of single data quality on precipitation calculation and improve the accuracy of precipitation prediction.

1. Data sources

Guizhou Province is located in the east of the Yunnan-Guizhou Plateau in southwest China. The terrain in Guizhou Province is high in the west and low in the east, which slopes from the middle to the east, south and north, with an average altitude of about 1100 meters. Its landforms are mainly of plateau mountains, hills, and basins, of which the mountain and hilly types account for 92.5% of the total area. It is reflecting the complex geomorphology of Guizhou Province and high-frequency precipitation characteristics. So this paper selects the circle research area centered at Guiyang radar with the radius of 150 km, where the radar detection data quality in this range is high and the RSD site is evenly distributed. The OTT-Parsivel disdrometer is used in this paper, which is an instrument for optical measurement of RSD based on modern laser technology, and the weather radar is a new generation of Doppler weather radar (CINRADCD). The topographic map and the distribution of RSD stations and radar stations in the study area are shown in Figure 1.

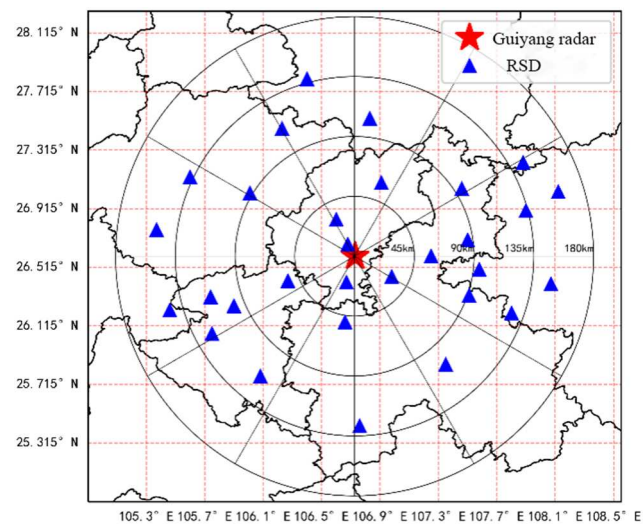


Figure 1. Study area radar and RSD site distribution.

The OTT-Parsivel disdrometer measures particle sizes from 0.2 mm to 5 mm for liquid precipitation types and 0.2 mm to 25 mm solid precipitation types, and can measure precipitation particles at speeds from 0.2 to 20 m/s. It classifies the measured particles in two dimensions of diameter (D) and velocity (V), with 32 intervals set in each, resulting in a total of $32 \times 32 = 1024$ particle types. In addition, the disdrometer stores the retrieved precipitation intensity, retrieved radar reflectance sensor status, and original particle species data as binary data files every 5 minutes.

CINRAD/CD type new generation Doppler weather radar has the following specifications: wavelength range 1.5-3.75m, frequency range 4000-8000MHz, beam width of about 1°, reflectivity factor distance resolution of 250m, maximum detection distance of 250km. The radar common body sweep mode is VCP21, that is, a body scan can scan 9 elevation angles in 6 minutes, the elevation angles are 0.5°, 1.5°, 2.4°, 3.4°, 4.3°, 6.0°, 9.9°, 14.6°, and 19.5°.

2. Analysis of raindrop spectrum characteristics of precipitation

2.0. Sketch out

Raindrop size distribution plays a crucial role in the study of microphysical precipitation processes because it has a profound impact on precipitation formation, the evolution of precipitation processes, and the quantitative estimation of precipitation by weather radar. In this paper, the data of RSD in Guizhou Province from March to September in 2020 are used to analyze the difference of RSD under different geographical situation, and to provide basis of data selection for the subsequent precipitation prediction with the use of the neural networks. Table 1 shows the basic information of the selected RSD stations. All three stations characterize by the subtropical monsoon humid climate, and the mountainous topography and landforms. Three stations with different altitudes are selected to study the influence of altitude on precipitation particle spectrum distribution through statistical analysis of data samples.

Table 1. Basic information of RSD site.

Site number	Site name	Geographic coordinates	altitude
57708	Dafang	105.60°E, 27.13°N	1722.7m
57828	Majang	107.58°E, 26.50°N	985.0m
57916	Luodian	106.76°E, 25.43°N	441.5m

2.1. Spectral parameter calculation

1. Particle number density

Particle density $N(D)$ represents the total number of raindrop particles per unit volume, with the unit $m^{-3}mm^{-1}$, and is calculated as follows:

$$N(D) = \sum_{i=1}^{32} \sum_{j=1}^{32} \frac{n_{ij}}{A \times \Delta T \times V_j} \quad (1)$$

where n_{ij} represents the number of particles in i -th diameter bin and the number of particles in the j -th velocity bin; A is the sampling base area of RSD, equalling to $5400mm^2$; ΔT is the sampling time 60s; V_j is the velocity value of the sampled particle in m/s.

2. Radar reflectivity factor

The radar reflectance factor, in mm^6/m^3 , is calculated as follows:

$$Z = \sum_{i=1}^{32} N(D_i) D_i^6 \quad (2)$$

3. Precipitation intensity

The precipitation intensity is the precipitation per unit time, with the unit mm/h. The calculation formula is as follows:

$$I = \frac{6\pi}{10^4} \sum_{i=1}^{32} D_i^3 V(D_i) N(D_i) \quad (3)$$

where D_i represents the diameter of the sampled particle, $N(D_i)$ is the number of particles at the current particle diameter and particle velocity (Pruppacher H R et al. [16]).

4. Average diameter

The average diameter is the sum of the diameters of all raindrops divided by the total number of raindrops, which is calculated as follows:

$$D_1 = \frac{\sum_{i=1}^{32} N(D_i) D_i}{\sum_{i=1}^{32} N(D_i)} \quad (4)$$

5. Mass-weighted average diameter

The mass-weighted average diameter is the average diameter of the diameter weighted mass of all particles in a unit volume relative to the total mass of particles, in mm, and the calculation formula is as follows:

$$D_m = \frac{\sum_{i=1}^{32} N(D_i) D_i^4}{\sum_{i=1}^{32} N(D_i) D_i^3} \quad (5)$$

6. Average volume diameter

The average volume diameter represents the diameter of an equivalent raindrop whose volume is equal to the average raindrop volume, in mm, and is calculated as follows:

$$D_v = \left[\frac{\sum_{i=1}^{32} N(D_i) D_i^3}{\sum_{i=1}^{32} N(D_i)} \right]^{\frac{1}{3}} \quad (6)$$

7. M-P distribution and GAMMA distribution

M-P distribution, which can be used to describe the spectral distribution of raindrops using an exponential function, is calculated as follows:

$$N(D) = N_0 \times \exp(-\lambda D) \quad (7)$$

where the particle number density parameter N_0 is in $\text{mm}^{-1}\text{m}^{-3}$, and the particle scale parameter λ is in mm^{-1} (Marshall J S. [17]).

In order to describe the RSD more accurately, Ulbrich proposed a method based on the M-P distribution to treat the RSD as a GAMMA distribution, and the calculation formula is as follows:

$$N(D) = N_0 \times D^\mu \times \exp(-\lambda D) \quad (8)$$

where the form factor μ is a dimensionless parameter. The function curve is curved upwards when $\mu > 0$. The function curve is curved downwards when $\mu < 0$. (Ulbrich C W. [18]) The formula becomes an M-P distribution when $\mu = 0$.

2.2. Raindrop size distribution analysis

1. Particles are distributed and proportion

In Figure 2, the x-axis represents the particle size, with the unit mm. The y-axis represents the final falling velocity of the particle, in m/s. The color scale represents the particle number density in units. And it is arranged from top to bottom in descending order of site altitude, and the results shown are the average data of RSD inversion from March to August in 2020. The following conclusions are drawn from the figure: first, the particle distributions of three stations are similar, showing that the particles are concentrated in the smaller particle size range with the peak value being 0.312mm. Second, the final falling velocity of raindrops is almost the same and its mainly distributed in the range of 1-5 m/s; Third, the particle number density spectrum width increases with the decrease of altitude.

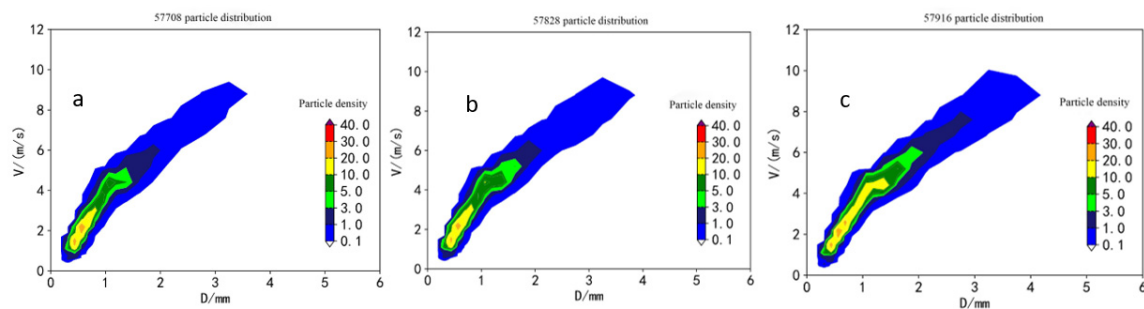


Figure 2. Particle distribution (a. Dafang (57708) Station b. Majiang (57828) Station c. Luodian (57916) Station.

2. Proportion of particle number density to precipitation intensity

Here, the sample of RSD is divided into three ranges ($D < 1\text{mm}$, $1 < D < 2\text{mm}$, $D > 2\text{mm}$) according to the particle size, then Statistical analyses are performed separately. As shown in Figure 3, small raindrops are an important component of the size spectrum in the three stations. With the decrease of altitude, the ratio proportion of small particle over total particles decreases, while the number of large size particles increases. The key reason may be that the collision and growth process of raindrops during the falling process consumes those small-sized particles and increases the concentration of those large-sized particles. Moreover, in terms of precipitation contribution rate, in Dafang (57708) station and Majiang (57828) station, the precipitation intensity are mainly provided by small size particles. While particles larger than 1mm particle size can provide a higher precipitation intensity with need to only a lower particle number density, which is caused by the sensitivity of precipitation intensity to particle diameter.

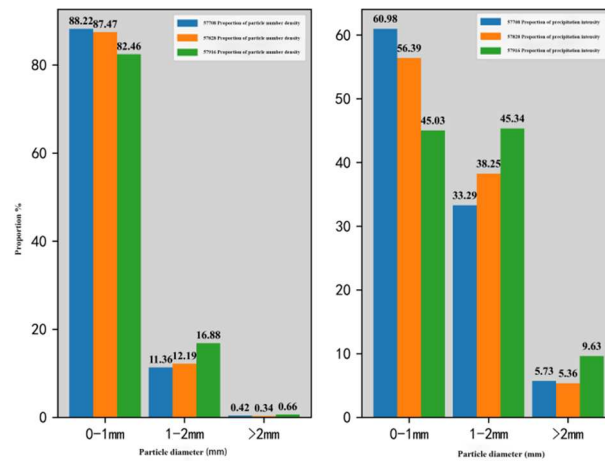


Figure 3. Particle density and intensity proportion.

3. Z-I relationship.

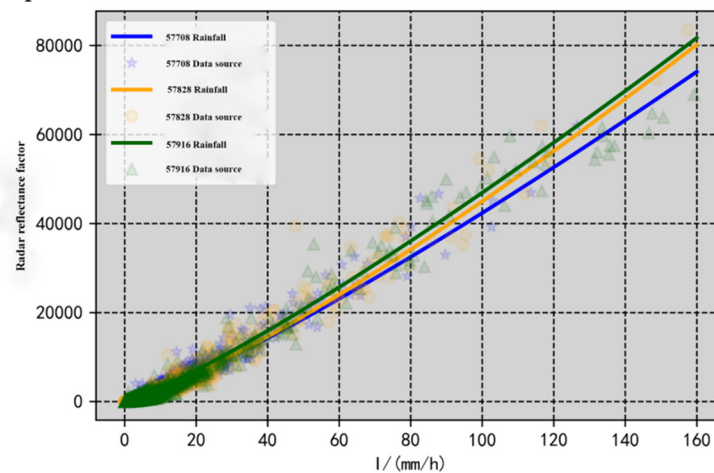


Figure 4. Z-I diagram.

Table 4. are the precipitation samples from March to September in 2020 at the three stations, and the curve is the fitting result of source data using the least square method. The site fitting curve of Dafang (57708) is $Z = 176.56 \times I^{1.19}$, the fitting curve of Majiang (57828) station is $Z = 155.94 \times I^{1.23}$, and the fitting curve of Luodian (57916) station is $Z = 204.78 \times I^{1.18}$. Compared with the classical radar quantitative precipitation estimation curve $Z = 300 \times I^{1.40}$, the coefficient values are quite different. If the classical Z-I curve is used for precipitation estimation, the precipitation will be underestimated to a certain extent under the same radar reflectance factor.

4. Raindrop size distribution fitting

Figure 5 are the fitting curves of the M-P and GAMMA distribution by using the least squares method, in which the x-axis represents the particle diameter and the y-axis represents the particle number density. Furthermore, the data sample is the precipitation base data of the three stations from March to September in 2020, and the distribution of precipitation particles in the rainfall process of the three places is analyzed by the average particle density spectrum. It can be seen from Figure 5 that the precipitation particle spectra of the three places are unimodal, with a peak of 0.312mm. In addition, the raindrop spectra of Dafang (57708) station with a higher altitude have a higher particle concentration of small particle size ($D < 1\text{mm}$), while Luodian (57916) station with a lower altitude occupies a higher particle concentration in a different size bin ($1 < D < 3\text{mm}$).

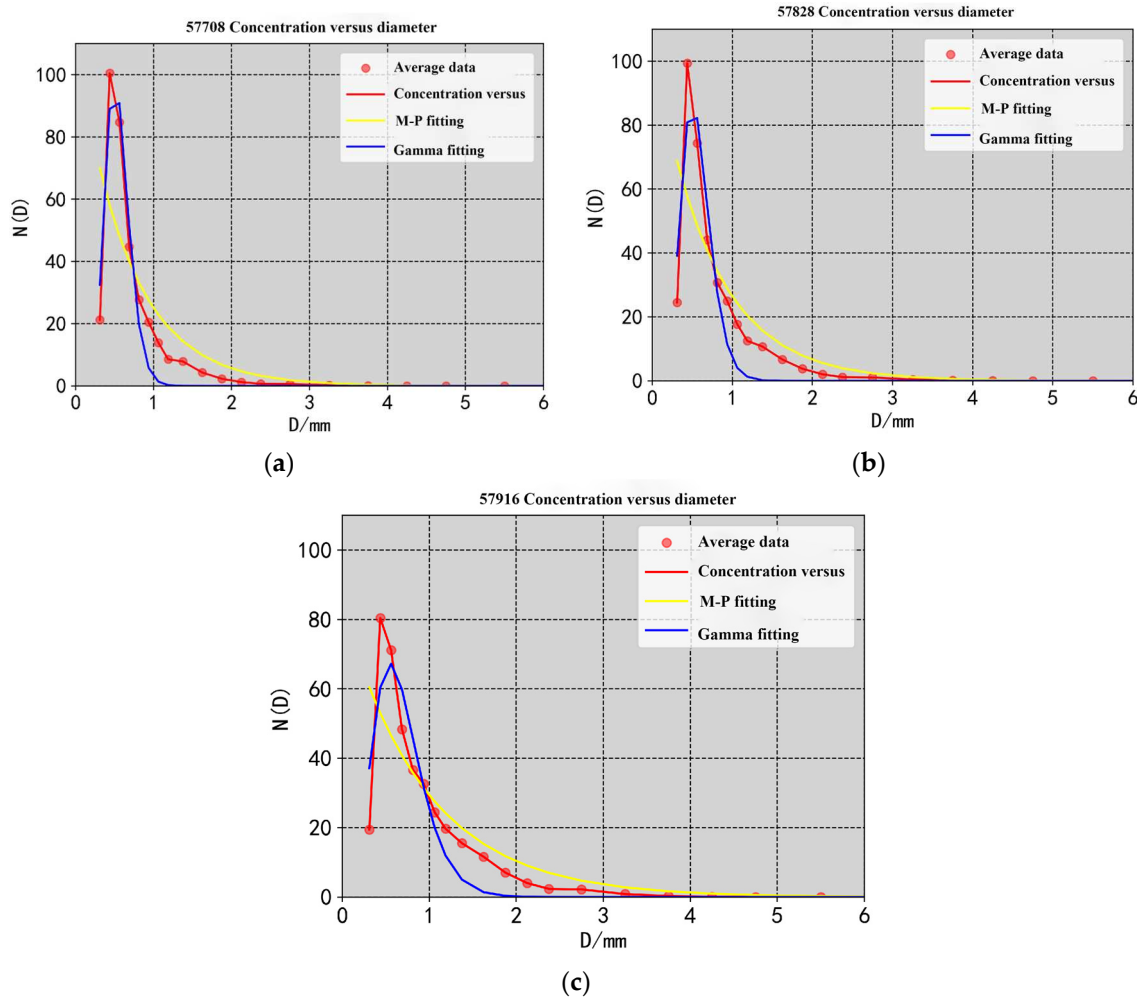


Figure 5. Average RSD, M-P distribution and Gamma distribution fitting(a. Dafang (57708) Station b. Majiang (57828) Station c. Luodian (57916) Station).

5. Particle diameter fitting

In Figure 6, the X-axis represents the precipitation intensity (PI), the Y-axis represents the precipitation particle diameter, and the simulated value of PI ranges from 0 to 150mm/h. The following conclusions are drawn from the figure: first, the fitting curve of the average diameter of rainfall in the three stations in 2020 has a small variation with the increase in rain intensity. And only has significant changes when the PI is low, and basically maintains around 1mm without much fluctuation in the area with large PI. Second, the average diameter is the lowest of the three diameters in the range of the simulated value of PI; the mass-weighted average diameter and average volume diameter fit curve amplitude are larger, where the average of the latter is relatively larger than that of the former.

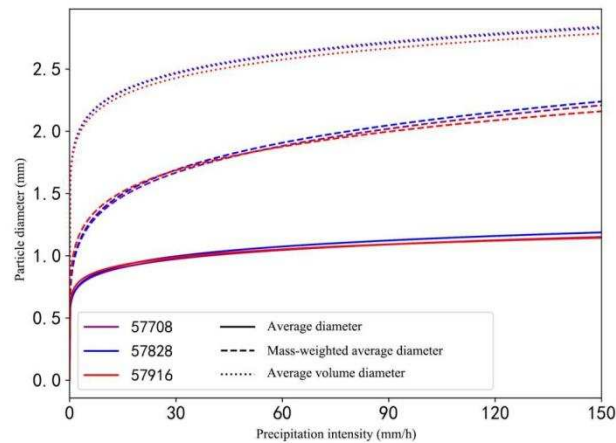


Figure 6. Fitting relation of particle diameter(57708, 57828, 57916).

Figure 7. shows a precipitation process in June at Dafang (57708) station. During the early and late precipitation periods where the precipitation intensity remains low, the two curves coincide with each other. But when the precipitation intensity is high, the performance of the average volume diameter is significantly stronger than that of the mass-weighted average diameter. According to the calculation, the correlation coefficients between the mass-weighted average diameter, and the average volume diameter and the precipitation intensity are 54.84% and 69.15%, respectively, indicating that the average volume diameter has a stronger correlation with the precipitation intensity.

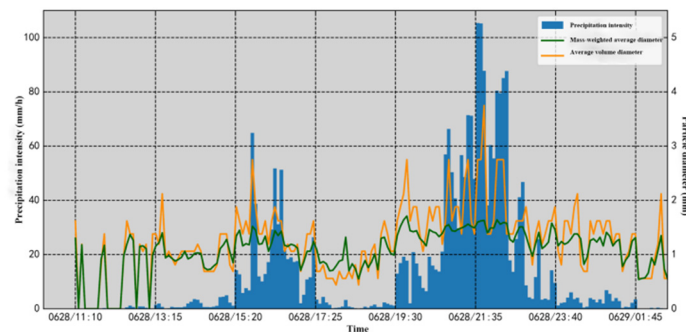


Figure 7. Mass-weighted average diameter and average volume diameter change with precipitation intensity.

Based on the comparative analysis of the particle distribution and microphysical characteristic parameters of the rainfall process mentioned above, it can be concluded that in the precipitation process of the three stations, first, the particle landing velocity distribution is very similar. Second, the particle size less than 1mm contributes the most to the particle density and decreases with the descending altitude. Third, the GAMMA distribution fit effect is better in the two distribution fittings, and average volume diameter than mass-weighted average diameter agree better with variation in precipitation intensity.

3. Neural networks predict precipitation

3.1. Predicting neural network configuration

1. Prediction network structure

The prediction model adopts two independent LSTM neural networks, in which the predicted values of precipitation and benchmark are given by prediction and benchmark network. The actual prediction value is obtained by subtracting the output value of the prediction network and the benchmark network, which the basic structure of the network is shown in Figure 8.

Firstly, the source data are fed into the LSTM layer for training. Then the Dropout layer randomly discards half of the connected neurons to increase the network generalization ability. Finally, output the training results through the BP layer and the single neuron Dense layer without activation function.

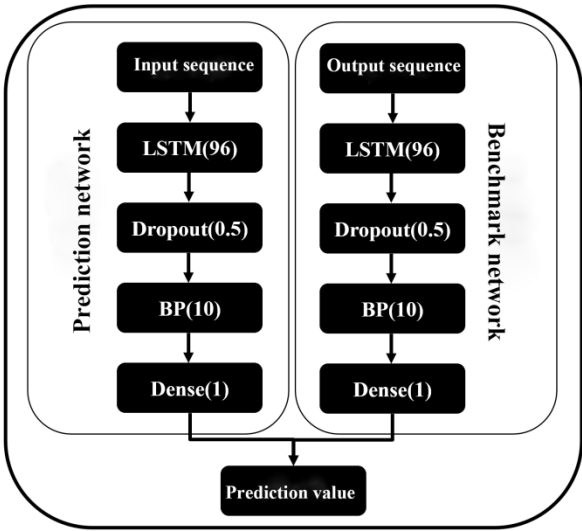


Figure 8. The basic structure of prediction neural network.

2. Prediction dataset construction

Neural network training requires feature values and prediction values, where the former is the input that the neural network needs to learn and the latter is the output that calculated by the neural network after a series of changes in the weight coefficients. In this experiment, the selection of features first takes into account the physical quantities with a strong correlation with precipitation intensity in the inversion parameters of the RSD, which can be obtained from equation (3). The particle diameter, particle fall velocity and particle number density parameters are selected here as the input values due to heir usage in the precipitation intensity inversion of the RSD. In terms of radar reflectance intensity, considering the distance between the RSD site and the radar station, the data in the 1-3 layer body sweep are selected as the feature input. Since the data of the rain gauge are usually regarded as the true values from the precipitation measuring, they are selected as the true value here. The parameters used in our research are shown in Table 2, and the prediction values are measured by rain gauges.

Table 2. Estimation of neural network inputs.

Feature value	estimated values
Radar reflectance intensity (layer 1-3)	A rain gauge measures precipitation
Particle number density retrieved by RSD	
Average particle velocity retrieved by RSD	
Average volume diameter retrieved by RSD	

If the prediction values is ahead time of the feature values, the neural network is for precipitation prediction. After the data of the tipping bucket rain gauge, RSD and weather radar are matched in time and space, their projections on the ground coincide. The start time and end time of the corresponding samples are the same, of which the RSD and weather radar data are used as feature values, with the tipping bucket rain gauge data as the predictive value. The prediction dataset is constructed as shown in Figure 9.

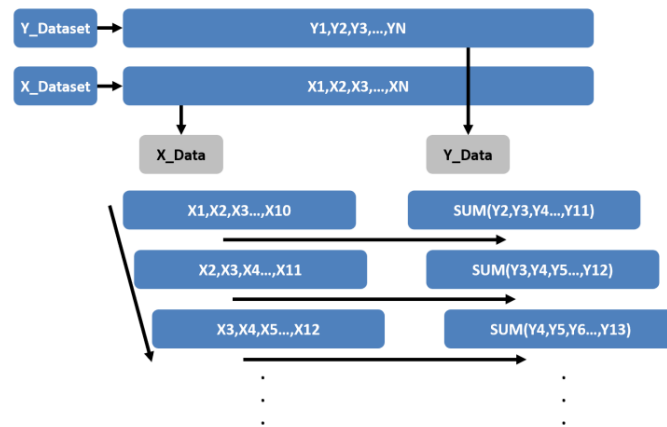


Figure 9. Predictive dataset construction.

In Figure 9 $X_Dataset$ represents the sequences set of the feature values, and $Y_Dataset$ represents the sequences set of the prediction values, in which $X1, X2, X3... XN$ corresponds to $Y1, Y2, Y3... YN$. And the input dataset of the neural network is constructed based on two datasets above. Since precipitation is a continuous process in time and the intensity is also continuously changed, When constructing the dataset need to consider the input feature values to form a continuous time series. However, the prediction value considering the convergence speed and training accuracy of the neural network uses the accumulation of the time series as the input.

In Figure 9, X_Data and Y_Data are the input data sets of the neural network. Firstly, in the first row of the first column of data, X_Data takes $X1$ to $X10$ as the features. In the second column of the first row, Y_Data takes the sum of $Y2$ to $Y11$ as the prediction value. The starting time of prediction value accumulation intervals is ahead of the feature value for one radar body sweep time. Therefore, this set of data is the input of the first 10 radar sweep time to predict the rainfall after 1 radar sweep time. Similarly, for the purpose of prediction after N radar body sweep times, X_Data remains unchanged and Y_Data only need to move the starting point for N times. Secondly, taking the sample count into account, it can be seen in the vertical direction of the block diagram that the starting point of each group is shifted back one sample time in turn, so that the time resolution of the constructed sample will be a single radar body sweep time. The dataset constructed by sliding at this time resolution can not only maximize the use of samples, but also improve the ability of neural networks to calculate the start and stop of precipitation.

Although the two neural network structures are the same, there are differences in the input sequence. Where the X_Data of the input are the same, but the the input Y_Data are different, as shown in Figure 10.

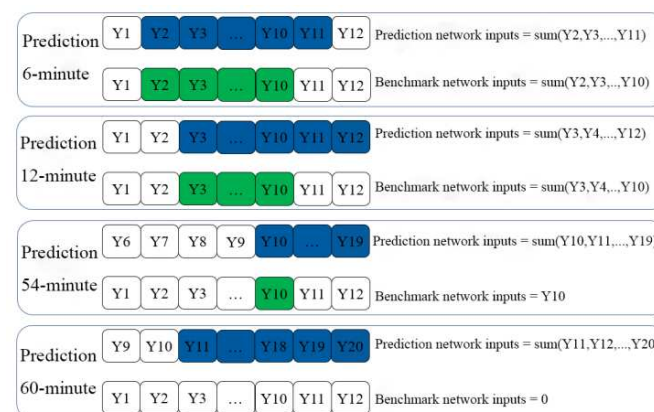


Figure 10. Prediction network and benchmark network inputs.

In Figure 10, the prediction input of the prediction network keeps the accumulation of the 10-time body scan data unchanged, and slides for one body scan time sequentially along the time axis according to the prediction time. The prediction value of the benchmark network remains unchanged at the last one Y10, and the starting time is moved back along the time axis by 1 time according to the prediction time. The method of subtracting the training results from the prediction network and the benchmark network has the following advantages over the method of direct prediction: (1) The use of dual network training can reduce the systematic error in the training process. (2) Compared with direct prediction, the dual network is more accurate in predicting within 60 minutes.

3. Parameter configuration

Neural network training is a very complex process. With the increase of network depth, neurons, feature values and other factors, activation function, network hyperparameters to the neural network convergence speed and training accuracy is crucial. And some of the parameters in the LSTM network as shown in Table 3.

Table 3. LSTM network parameters.

LSTM network parameters	
Activation function	Relu
Dropout coefficient	0.5
Loss function	Mean Absolute Error
	Initial learning rate 0.1
Dynamic learning rate	Patience 50
	factor 0.1
Optimizer	Adam
Training batch size	32
Number of iterations	2000
The test set picks the scale	20%

3.2. Prediction result evaluation

1. Prediction of convective cloud precipitation process

A convective cloud precipitation event is selected on July 18, and the precipitation forecasts of 6 minutes, 18 minutes, 30 minutes, 60 minutes, 90 minutes and 120 minutes are carried out at the Dafang (57708) station, with the precipitation process began at 21:05 on July 18, 2020 lasting about 7 hours from that time on. In this case the precipitation is 37.3 mm, the precipitation intensity inverted by RSD peaked with the value being 26.78 mm/h.

Table 4 is the evaluation index of neural network precipitation prediction, and it can be seen from the table that the real-time correlation coefficient and the optimal average relative error of the prediction 60-minute network in the prediction time point reaches 0.9287 and 0.3897, respectively. The real-time prediction with time resolution more than 60 minutes is the same. Besides, the 60-minute prediction index is better than the 90-minute and 120-minute forecast as a whole. In the 60-minute prediction the good performance of the real-time correlation coefficient and the mean relative error indicates that it has a better performance in predicting the precipitation trend and the precipitation intensity consistency compared with other times.

Table 4. Prediction and evaluation index of neural network of convective cloud precipitation process.

Neural networks prediction time	Real-time correlation number	MRE	MAE	RMSE
6minutes	0.4763	0.6222	0.2660	0.5807
18minutes	0.8617	0.5129	0.4177	0.8843
30minutes	0.8874	0.5994	0.5456	1.0824

60minutes	0.9287	0.3897	0.8057	1.6419
90minutes	0.8203	0.4522	1.1601	2.5705
120minutes	0.6061	0.4077	1.6566	3.6103

Figure 11 shows that the deviation at the beginning of the two precipitation always tends to be negative, indicating that the prediction value is underestimated compared to the rain gauge detection value. That is, the rain gauge detects precipitation earlier than the neural network. Due to the principle and characteristics of the detection instrument, the neural network input feature parameter can capture the precipitation information of the priority rain gauge, but sometimes the rain doesn't exist near the measurement area at all. That's because the neural network input can only determine the occurrence of precipitation when there is enough precipitation information. If the actual raining time is short, the neural network can ignore the precipitation process easily.

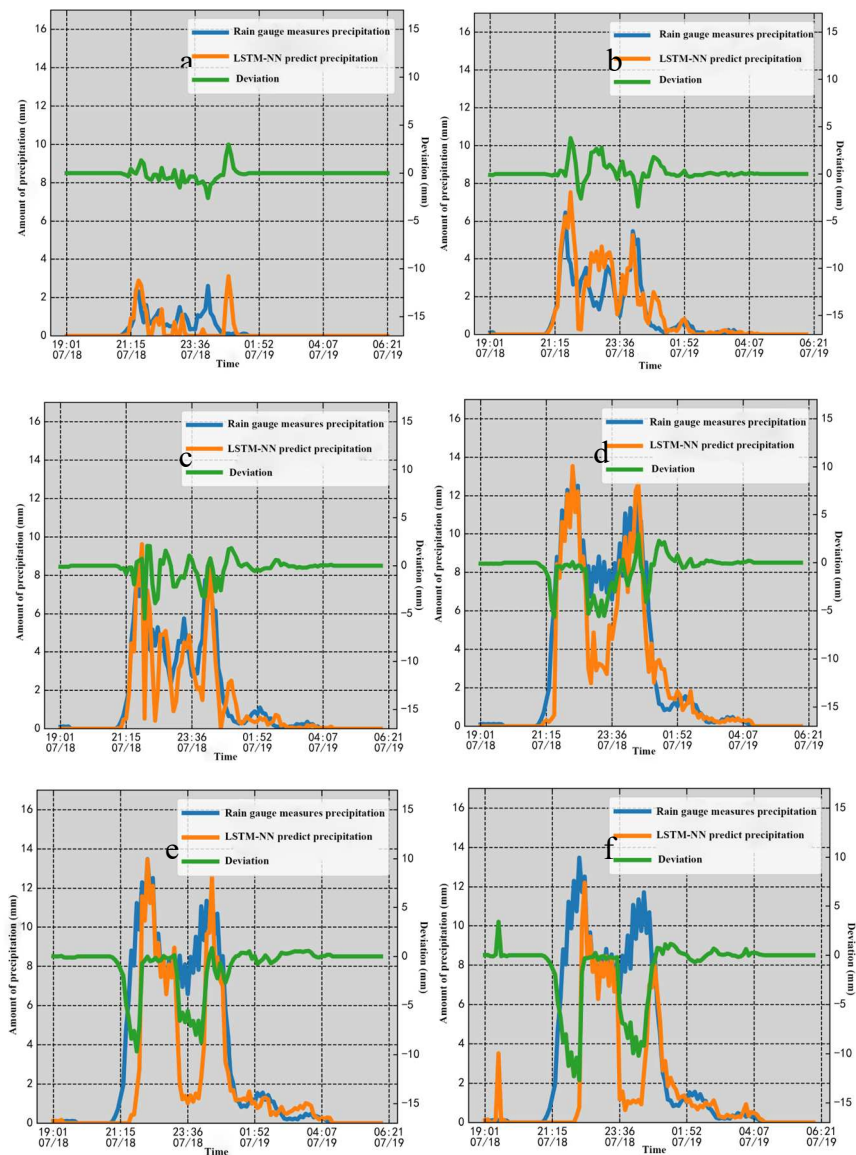


Figure 11. Real-time prediction of convective cloud precipitation process by neural network (a.6minutes b.18 minutes c.30 minutes d.60 minutes e.90 minutes f.120 minutes).

There is a significant difference between the rain gauge and the neural network prediction of rainfall during the precipitation process. The real-time correlation coefficient reflects the gap between the neural network's real-time rainfall prediction and the rain gauge detection result, which can be reflected in the time mismatch and numerical gap. Since the prediction is using the data of the first 1

hour to predict the future time. When the precipitation intensity changes drastically and the input feature value of the neural network has not yet received the change, but the ground rain gauge detection data will be reflected anyway, which causes a time mismatch, and it becomes more obvious as the predicted time increases. In order to find the cause of the numerical difference, the relationship between the features of the network model and the prediction values is found after controlling the variables: (1) Radar reflectance is mainly used to determine the start and end time of precipitation, and auxiliary revision of predicted precipitation intensity. (2) The average velocity mainly plays an auxiliary role in revising the predicted precipitation intensity and has little influence on the results. (3) The predicted precipitation intensity is mainly affected by the combined effect of average volume diameter and particle number density. The average volume diameter has a high weight on the amplitude of predicted precipitation intensity, and particle number density has a great influence on the fluctuation of predicted precipitation, so single value cannot predict precipitation well. Although the rain gauge and the RSD are geographically similar and coincident, it can be seen that the prediction values have a strong correlation with the current precipitation structure. However, the surface measurement data of the two instruments does not coincide with each other, the time series of the data cannot be completely consistent, which is also an important reason for the difference of real-time numerical prediction.

Although the neural network on the real-time precipitation is different from the rain gauge, this error will be compensated before and after the time point. If the time series is stretched or predicted the total rainfall process, the predicted result of the neural network will be closer to the actual precipitation.

2.Prediction of Stratiform cloud precipitation process

A continuous stratiform cloud precipitation weather on July 19 is also selected, and the precipitation forecasts of 6 minutes, 18 minutes, 30 minutes, 60 minutes, 90 minutes and 120 minutes are carried out at the Dafang (57708) station. The precipitation began at 17:36 on July 19, 2020 lasting about 7 hours. The precipitation is 6.7 mm and the RSD inverted precipitation intensity comes with a peak of 4.98 mm/h.

Table 5. Prediction and evaluation index of neural network of stratiform cloud precipitation process.

Neural networks prediction time	Real-time correlation number	MRE	MAE	RMSE
6minutes	0.2743	0.7921	0.1347	0.2091
18minutes	0.5447	0.6970	0.2001	0.3379
30minutes	0.7783	0.6551	0.2481	0.3885
60minutes	0.9257	0.3312	0.2314	0.3906
90minutes	0.8856	0.3346	0.2919	0.5149
120minutes	0.8753	0.5496	0.3066	0.4596

Table 5 is the neural network precipitation evaluation prediction index. From the Table 5 it can be obtained that the characteristics of the stratiform cloud precipitation evaluation index are basically the same as the convective cloud precipitation. Besides, the real-time correlation coefficient and the average relative error are the best prediction results of 60 minutes, reaching 0.9357 and 0.3312 respectively. It can be concluded that the prediction 60 minutes neural networks have strong predictive ability for weak precipitation process. It can also be seen from Figure 12 that as the forecast time increases, the rain gauge detection data tends to smooth out, and the neural network prediction effect continues to improve. This is due to the fact that the rain gauge detection process in the weak precipitation process is greatly affected by the tipping bucket residual error, which itself has a certain deviation from the actual weather phenomenon, and the increase in the predicted time of the neural network can weaken this deviation to a certain extent. And we see that there is no lag error in stratiform precipitation as there is in convective precipitation. After checking the data set, it is found

that the weather radar and RSD detected from the precipitation information long before the rain gauge, because the echo was developed over the area. If the precipitation due to echo motion occurs during the convective cloud precipitation process, the weather radar will detect the echo of the earlier rain gauge at one or two individual sweep time, and the phenomenon of prediction lag will be easy to occur when the it is detected by time. Finally, in terms of the RSD, the relatively weak precipitation information can be captured by the higher detection accuracy, and the more complete precipitation process can be obtained for the precipitation with stable development.

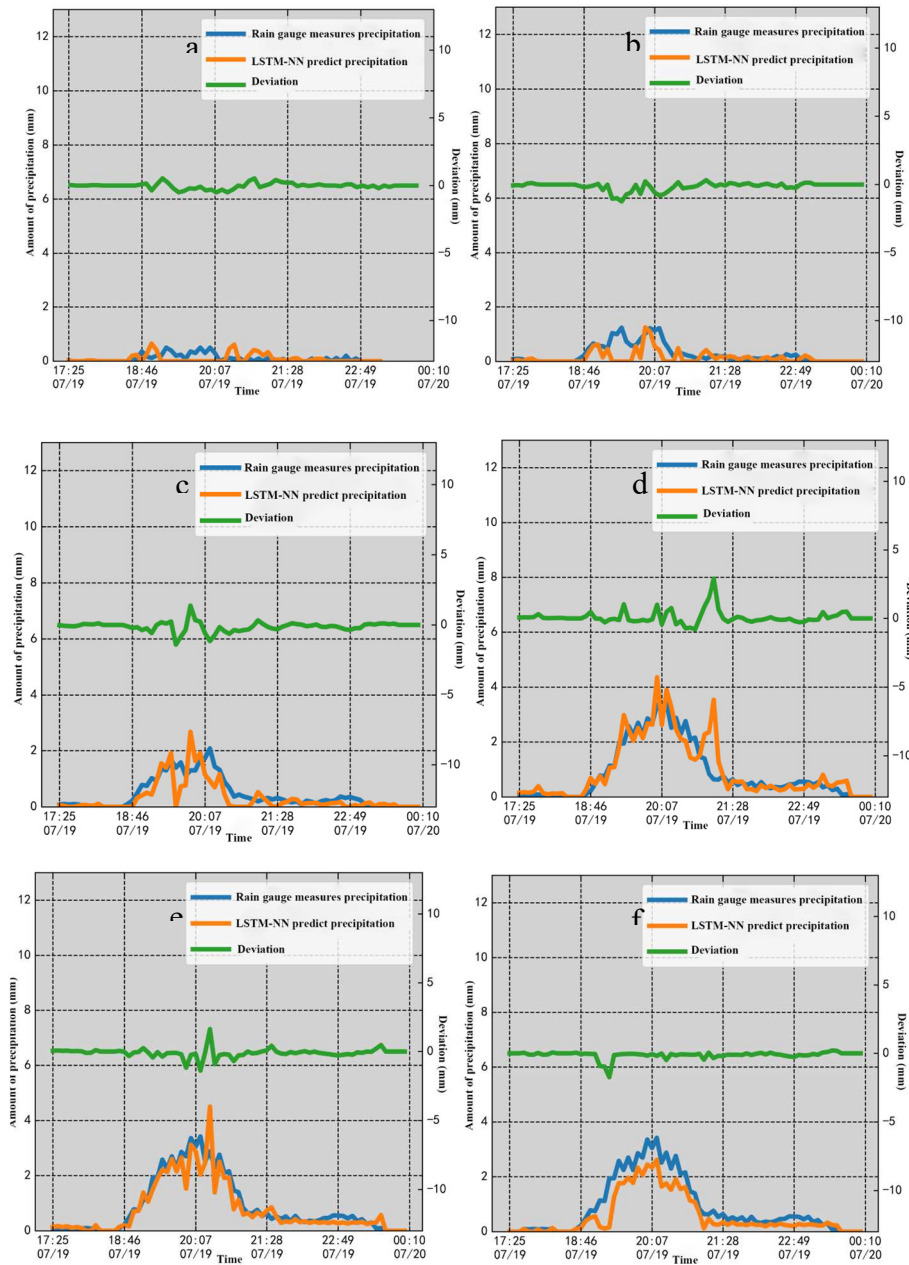


Figure 12. Real-time prediction of stratiform cloud precipitation process by neural network (a.6minutes b.18 minutes c.30 minutes d.60 minutes e.90 minutes f.120 minutes).

4. Conclusion

In this paper, the base data of the new generation Doppler weather radar (C-band) in Guizhou Province, the raw data of the OTT-Parsivel laser RSD and the precipitation data of the automatic meteorological observation station are used. Three RSD stations are selected for RSD analysis

according to the difference in altitude, then the three-source data are used to predict the precipitation through the LSTM neural network. The conclusions of the study are as follows:

(1) From the distribution characteristics of RSD in Dafang (57708) station, Majang (57828) station, and Luodian (57916) station (with an altitude of 1722.7m, 985.0m, and 441.5m, respectively) in Guizhou, we can conclude that the particle size of less than 1 mm contributes the highest to the density of particle numbers, reaching 88.22% and decrease with the descending altitude. In addition, in the fitting of RSD, the GAMMA distribution fit has more advantages than the M-P distribution fit in the small particle size region (D1mm), while the M-P distribution fit is more accurate in the large particle size region. In general, the GAMMA distribution fit effect is better, and increases with the ascending altitude. Finally, in terms of the correlation between particle size and precipitation intensity, the mass-weighted average diameter is not sensitive enough to the response to the change of precipitation intensity, but the correlation between the average volume diameter and the precipitation intensity is higher, with correlation coefficients being 54.84% and 69.15%, respectively.

(2) In terms of precipitation prediction, in the process of convective cloud precipitation the 60 minutes forecast index is better than the 90, 120 minutes forecast. Although the real-time precipitation neural network has a obvious difference from the rain gauge, this error will be compensated before and after the time point. If the time series is elongated or predict the total rainfall of a process, the prediction results of the neural network will be closer to the actual precipitation. Similarly, in the process of stratiform cloud precipitation, it is shown that the prediction 60-minute neural network has behaves better for the weak precipitation prediction. And with the increase of prediction time, the detection data of the rain gauge tends to be smooth and the predicted effect of the neural network continues to improve. In conclusion, in the prediction of convective cloud and stratiform cloud precipitation process the 60-minute prediction results have the highest consistency with the actual precipitation, and the correlation coefficients can reach 0.9287 and 0.9257, respectively.

Author Contributions: All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding: This work was supported by the Sichuan Provincial Natural Science Foundation (2022NSFSC0208) and the Open Fund Project of the Key Laboratory of Land Surface Processes and Climate Change in Cold and Arid Regions, Chinese Academy of Sciences (LPCC2020009).

Institutional Review Board Statement: Not applicable for studies not involving humans or animals.

Informed Consent Statement: Not applicable for studies not involving humans.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

Reference

1. WU Jiang-sheng, JIN Long, NONG Ji-fu, Forecast Research and Applying of BP Neural Network Based on Genetic Algorithms[J]. Mathematics In Practice and Theory ,2021,47(07):843-853.
2. HUANG Xingyou ,YIN Jianan,MA Lei,HUANG Zhaochu.Comprehensive Statistical Analysis of Rain Drop Size Distribution Parameters and Their Application to Weather Radar Measurement in Nanjing[J].Atmospheric Science,2019,43(03):691-704.
3. LIN Hou-bo, You Qing-long, JIAO Yang, MIN Jin-zhong. Spatial and Temporal Characteristics of the Precipitation over the Tibetan Plateau from 1961 to 2010 Based on High Resolution Grid-observation Dataset[J].Journal of Natural Resources,2015,30(02):271-281.
4. SHEN Gaohang, GAO Anchun, LI Jun.Application of Raindrop Spectrum and Dual Polarization Radar Data to a Heavy Rain Process[J].Meteorological ,2021,47(06):737-745.
5. CAI Zhao, LIU Jiufu, LIAO Aimin, LIAO Minhan, LIU Hongwei, WANG Huan, MA Tao, ZHUO Peng.Analysis of characteristics of raindrops based on distrometers[J].Journal of Advances in Science and Technology of Water Resources ,2021,41(05):93-98.
6. Zeng Guangyu , Guo Zeyong , Zhou Qinqiang , Zhang Honghao , Chen Xingdeng , Guo Jia.Application evaluation of dual polarization radar quantitative precipitation estimation algorithm based

- on laser raindrop spectrum[J].Meteorological Research and Application,2021,42(04):95-100.
7. Rivelli Zea Lina,Nesbitt Stephen W.,Ladino Alfonso,Hardin Joseph C.,Varble Adam. Raindrop Size Spectrum in Deep Convective Regions of the Americas[J]. Atmosphere,2021,12(8).
 8. François Mercier,Aymeric Chazottes,Laurent Barthès,Cécile Mallet. 4-D-VAR assimilation of disdrometer data and radar spectral reflectivities for raindrop size distribution and vertical wind retrievals[J]. Atmospheric Measurement Techniques,2016,9(7).
 9. Zhu Ping, Li Shengchen, Xiao Jianshe, Xu Liang, Jin Shiqiang.Study on Extrapolation Technique of Weather Radar Echo and Its Application to Nowcasting[J]. Meteorological ,2008,(07):3-9+129-132.
 10. Liu Guozhong, TANG Yuyong, Ban Ronggui, LU Xiaodan, Deng Ruyi, Huang Fei, Yang Yuhong.Brief Discuss on the conceptual model forecasting tool of the Maximum and minimum temperatures Sudden change[J].Meteorological Research and Application,2009,30(04):22-24+108.
 11. Cheng Conglan , Chen Mingxuan , Wang Jianjie, Cao Feng.Yang Hanxian.Short-term quantitative precipitation forecast experiments based on blending of nowcasting with numerical weather prediction[J].Acta Meteorologica Sinica ,2013,71(03):397-415.
 12. Elshaboury N,Elshourbagy M,Al-Sakkaf A,Abdelkader E M. Rainfall forecasting in arid regions using an ensemble of artificial neural networks[J]. Journal of Physics: Conference Series,2021,1900(1).
 13. Suhail Sharadqah,Ayman M Mansour,Mohammad A Obeidat,Ramiro Marbello,Soraya Mercedes Perez. Nonlinear Rainfall Yearly Prediction based on Autoregressive Artificial Neural Networks Model in Central Jordan using Data Records:1938-2018[J]. International Journal of Advanced Computer Science and Applications (IJACSA),2021,12(2).
 14. Gunathilake Miyuru B.,Senerath Thamashi,Rathnayake Upaka. Artificial neural network based PERSIANN data sets in evaluation of hydrologic utility of precipitation estimations in a tropical watershed of Sri Lanka[J]. AIMS GEOSCIENCES,2021,7(3).
 15. ZHANG Shuai, WEI Zheng-ying,ZHANG Yu-bin.Forecasting Rainfall with Recurrent Neural Network[J]. Water-saving irrigation,2017(05):63-66+71.
 16. Pruppacher H R, Klett J D, Wang P K. Microphysics of clouds and precipitation[J]. 1998.
 17. Marshall J S. The distribution of raindrops with size[J]. J. meteor., 1948, 5: 165-166.
 18. Ulbrich C W. Natural variations in the analytical form of the raindrop size distribution[J]. Journal of climate and applied meteorology, 1983: 1764-1775.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.