RESEARCH ARTICLE
ATMOSPHERIC PHYSICS

Earth and Planetary Physics
7: 471–486, 2023
doi: 10.26464/epp2023049

Deep learning-based subseasonal to seasonal precipitation prediction in southwest China: Algorithm comparison and sensitivity to input features

GuoLu Gao[1,2], Yang Li[3], XueYun Zhou[1,2], XiaoMing Xiang[2,4*], JiaQi Li[1,5], and ShuCheng Yin[6]

[1]Ya'an Meteorological Observatory, Sichuan Meteorological Bureau, Ya'an Sichuan 625000, China;

[2]Plateau and Basin Rainstorm, Drought and Flood Key Laboratory of Sichuan Province, Chengdu 610000, China;

[3]College of Atmospheric Science, Chengdu University of Information Technology, Chengdu 610225, China;

[4]Meteorological Observation and Data Centre, Sichuan Meteorological Bureau, Chengdu 610072, China;

[5]Leshan Meteorological Observatory, Sichuan Meteorological Bureau, Leshan Sichuan 614000, China;

[6]Quanzhou Meteorological Observatory, Fujian Meteorological Bureau, Quanzhou Fujian 362000, China

Key Points:

- A subseasonal to seasonal precipitation prediction model was established based on deep learning.
- A sensitivity-based method was used to explain the physical basis of deep learning.
- The recurrent neural network model has better performance in different terrains.

Citation: Gao, G. L., Li, Y., Zhou, X. Y., Xiang, X. M., Li, J. Q., and Yin, S. C. (2023). Deep learning-based subseasonal to seasonal precipitation prediction in southwest China: Algorithm comparison and sensitivity to input features. *Earth Planet. Phys.*, *7*(4), 471–486. http://doi.org/10.26464/epp2023049

Abstract: The prediction of precipitation at subseasonal to seasonal (S2S) timescales remains an enormous challenge because of the gap between weather and climate predictions. This study compares three deep learning algorithms, namely, the long short-term memory recurrent (LSTM), gated recurrent unit (GRU), and recurrent neural network (RNN), and selects the optimal algorithm to establish an S2S precipitation prediction model. The models were evaluated in four subregions of the Sichuan Province: the Plateau, Valley, eastern Basin, and western Basin. The results showed that the RNN model had better performance than the LSTM and GRU models. This could be because the RNN model had an advantage over the LSTM model in the transformation of climate indices with positive and negative variations. In the validation of test datasets, the RNN model successfully predicted the precipitation trend in most years during the wet season (May–October). The RNN model had a lower prediction bias (within ±10%), higher sign accuracy of the precipitation trend (~88.95%), and greater accuracy of the maximum precipitation month (>0.85). For the prediction of different lead times, the RNN model was able to provide a stable trend prediction for summer precipitation, and the time correlation coefficient score was higher than that of the National Climate Center of China. Furthermore, this study proposed a method to measure the sensitivity of the RNN model to different input features, which may provide unprecedented insights into the nonlinear relationship and complicated feedback process among climate systems. The results of the sensitivity distribution are as follows. First, the Niño 4 and Niño 3.4 indices were equally important for the prediction of wet season precipitation. Second, the sensitivity of the snow cover on the Tibetan Plateau was higher than that in the Northern Hemisphere. Third, an opposite sensitivity appeared in two different patterns of the Indian Ocean and sea ice concentrations in the Arctic and the Barents Sea.

Keywords: recurrent neural network; long short-term memory recurrent; sensitivity analysis; artificial intelligence explainability; complex terrain; southwest China

1. Introduction

The prediction of precipitation at subseasonal to seasonal (S2S) timescales has an important influence on government decision making, the prevention of floods and droughts, water management, agricultural irrigation, health and disease, and abundant economic activities (White et al., 2017). The short-term initial field is as important as the long-term external forcing from the Earth's climate system for the S2S prediction. The prediction sources of S2S precipitation could be from atmospheric circulation (initial field conditions), natural variation of the underlying surface (ocean surface temperature, sea ice concentration, land snow cover, soil moisture), changes in vegetation, or anthropogenic

First author: G. L. Gao, gaoguolu123@outlook.com
Correspondence to: X. M. Xiang, micxiang@qq.com
Received 07 FEB 2023; Accepted 26 MAY 2023.
Accepted article online 28 JUN 2023.

activities, among others (Brunet et al., 2010; Vitart et al., 2017). It is still difficult for climate models to predict S2S precipitation (Li CC et al., 2018; De Andrade et al., 2019), especially the prediction of precipitation at the subseasonal timescale, called the "predictability desert" (Vitart et al., 2012).

Sichuan, located in southwest China, is one of the most densely populated areas in China and the world (Lu RY and Ye H, 2011). As the transition zone between the Tibetan Plateau and Yangtze River, it is known as a mountainous region with a complex terrain, geological disasters, and fragile ecosystems (Nie YB and Sun JQ, 2021). Therefore, it is important to improve the accuracy of S2S precipitation prediction in this region. Figure 1a shows the complex terrain in this area, including the Tibetan Plateau, Panxi Valley, and Sichuan Basin (hereafter the Plateau, Valley, and eastern and western Basin). The altitude varies drastically, from 500 m to more than 4,000 m. Because the terrain and atmospheric circulation systems have complex influences on rainfall in Sichuan, the precipitation prediction of the climate model in Sichuan is more difficult than in other parts of China (Ruan CQ et al., 2015; Guo Q et al., 2021).

With the development of observation and computer science, deep learning has been successfully applied in many fields (Wu Y

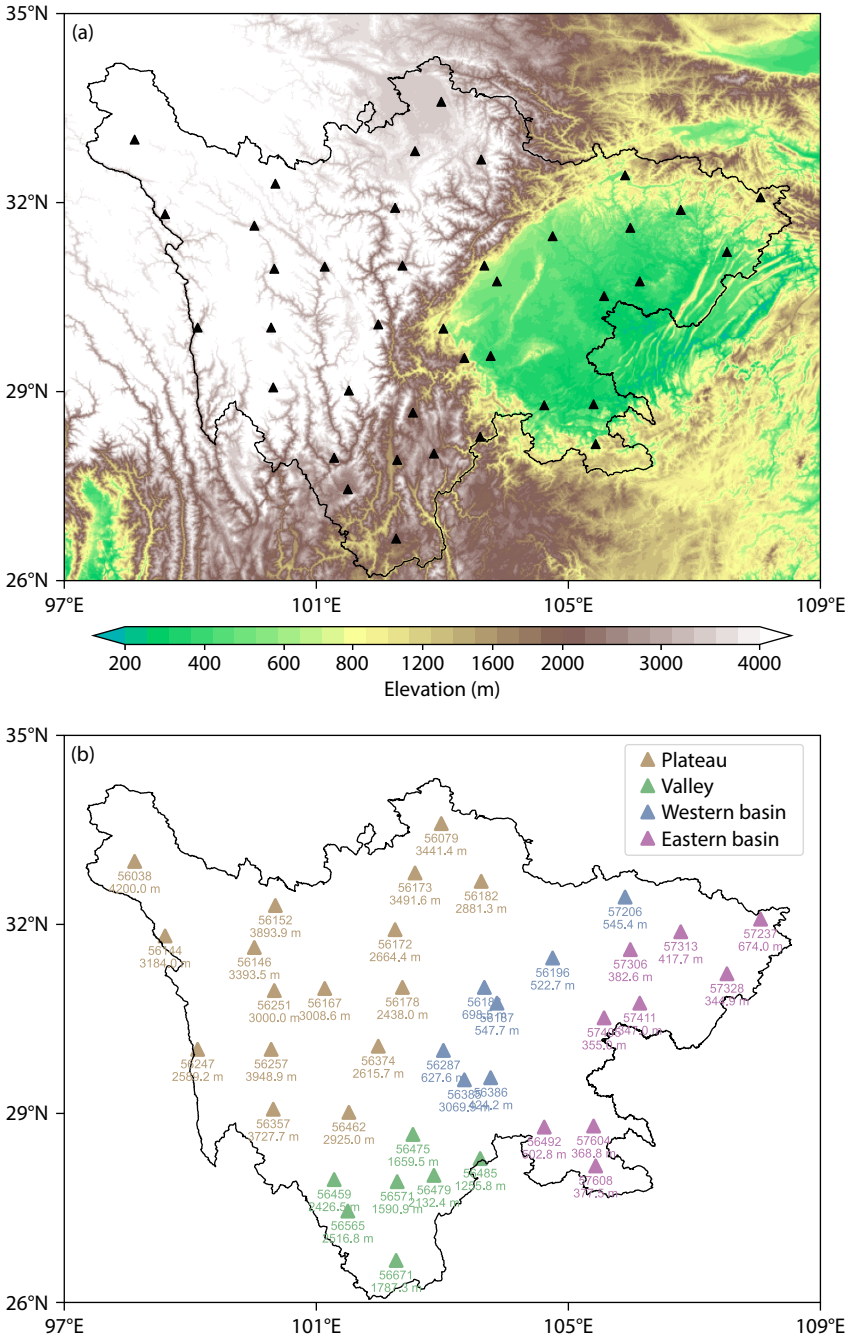

Figure 1. (a) The topography in Sichuan. (b) The distribution, altitude, and number of stations in the study area. Different colors represent different subregions. The elevation data are available at www.ngdc.noaa.gov/mgg/global/global.html.

et al., 2022), such as automatic driving, medical assistance, facial recognition, and natural language processing. For the prediction of S2S precipitation, deep learning is also becoming an alternative approach (Mayer and Barnes, 2021; Weyn et al., 2021). In the time-series prediction project (Pascanu et al., 2013), the recurrent neural network (RNN) and its variants (i.e., long short-term memory recurrent [LSTM] and gated recurrent unit [GRU]) had a unique advantage over the traditional artificial neural network (ANN) because the RNN fit the data with a period of time rather than with a single moment (Pascanu et al., 2013). In meteorological data, the climate states in the past are closely related to the climate states at certain times in the future. Therefore, the RNN and its variants are suitable for predicting meteorological data with the characteristics of time-series variation. The two variants of the RNN, the LSTM and GRU, introduce different activation functions and gating structures to process data with long time series (Dey and Salem, 2017; Sherstinsky, 2020).

Recent research has indicated that the RNN and its variants perform well in predicting precipitation. Misra et al. (2018) inputted the U.S. National Centers for Environmental Prediction reanalysis circulation data into the LSTM, multiple regression methods, and the ANN and found that the LSTM had the smallest bias in its daily precipitation predictions for two basins in Canada and India. Kumar et al. (2019) directly inputted the observed precipitation into the RNN and LSTM and found that the two methods could predict the monthly precipitation. Chhetri et al. (2020) inputted the precipitation and sunshine hours observed by the station into a mixed model with LSTM and GRU and found that this mixed model could achieve a lower bias in its monthly precipitation prediction. Shen HJ et al. (2020) inputted a large amount of climate model output data into the LSTM and found that the prediction score of the LSTM was higher than the climate model for summer precipitation in China in 2014–2015. The aforementioned studies used different input features to train the RNN, LSTM, or GRU model. However, research is limited on whether a performance difference exists among the three models when the input features are the same.

In addition to the improvements in deep learning algorithms, scientists are paying increasing attention to the "explainability" of phenomena by artificial intelligence (Schneider et al., 2022; Sun ZH et al., 2022). A few researchers are inclined to improve the prediction accuracy of the deep learning model by further extracting the physical signals of the input features. Raissi et al. (2019) observed that a neural network with physical information constraints had better performance than a data-driven neural network in less training time. Gao GL et al. (2021) found that after synoptic classification, atmospheric circulation data could improve the prediction accuracy of a rainstorm. Zhou L and Zhang RH (2022) found that after principal oscillation pattern analysis, sea surface temperature (SST) data could improve the prediction of El Niño–Southern Oscillation (ENSO) signals. Questions that remain to be addressed are how deep learning understands these input features and which input features play a more important role.

To gain a better understanding of these two issues, we compared the performance of three deep learning algorithms (i.e., RNN, LSTM, and GRU) in S2S precipitation prediction and analyzed the impact of two activation functions in the algorithms on the input features. The time scope of S2S in this study extended from 1 month to 1 year into the future. In the deep learning models, we similarly used certain climate signals as the input features, such as atmospheric teleconnections, SST anomalies, snow cover, and sea ice concentrations. In particular, the time-decomposed historical precipitation was also used as an input feature reflecting terrain and climatological characteristics to solve the precipitation prediction problem under different terrains. Furthermore, we proposed a sensitivity-based approach to calculate which climate factors were important in the RNN model.

This work is organized as follows. Section 2 describes the data and methods, which include how to build the model, decompose the precipitation sequence, and calculate the sensitivity. Section 3 shows the principal results and includes a discussion of the differences among the RNN, LSTM, and GRU for prediction evaluation in the RNN model and the sensitivity distribution of the RNN model to input features. The conclusions are presented in Section 4.

2. Data and Methods

2.1 Datasets

2.1.1 Station data and subregion division

The monthly observations of precipitation are from the Big Data Platform of the China Meteorological Administration (http://10.194.90.120:8088/cmadaas/). The elevation and topographic characteristics are primarily considered to divide Sichuan into different subregions. Three distinguishing subregions can be found: (1) the Tibetan Plateau, with an elevation of more than 3,000 m; (2) the Yunnan Plateau or Valley, with an elevation ranging from 1,000 to 2,000 m; and (3) the Sichuan Basin, with an elevation below 1,000 m. Obvious differences exist in the annual total, seasonal cycle, and interannual variation of precipitation in different terrains (shown in a later figure). Previous research has also reported an anti-correlation of the precipitation decadal trend in the eastern and western Basins (Chen YR et al., 2009; Xiao C et al., 2016); thus, the Sichuan Basin is further divided into the eastern Basin and the western Basin in this study. This classification is similar to that of Sun SL et al. (2016). Figure 1b shows the distribution, altitude, and number of stations in the study area.

2.1.2 Climate index datasets

Parts of the climate indexes connected with precipitation anomalies in Sichuan and southwest China were selected for input features, including the atmospheric teleconnection (Sun JQ and Wang HJ, 2012; He SP et al., 2017), the SST over the Pacific Ocean (Li CY et al., 2008; Yang S et al., 2018), the SST over the Indian Ocean (Saji et al., 1999; Jia XL and Li CY, 2005), the sea ice concentration over the Arctic and Arctic margins (Vihma, 2014; Wu BY et al., 2015), and the snow cover content over the Tibetan Plateau and Northern Hemisphere (Zhang RN et al., 2017b; You QL et al., 2020). A description of the datasets and the websites where they can be found is listed in Table 1.

The Madden–Julian Oscillation, an important prediction source (Ren HL et al., 2015), was not selected as one of the features

Table 1. Description of the climate factor indexes and websites where the data can be found.

	Index	Description[a]	Website
Land	SnowCover_NH	Extent of the snow cover over the Northern Hemisphere	www.ncei.noaa.gov
	SnowCover_TP	Extent of the snow cover over the Tibetan Plateau	
Ocean	SeaIce_Arctic	Sea ice concentration poleward of 60°N	www.cmdp.ncc-cma.net/cn/monitoring.htm
	SeaIce_BL	Sea ice concentration range over the Barents–Laptev Sea (72.5°N–83.5°N, 30.5°N–149.5°E)	
	TIOD	The tropical Indian Ocean SST dipole mode	www.psl.noaa.gov
	SIOD	The subtropical Indian Ocean SST dipole mode	
	Niño_4	The SST anomalies in the Niño 4 region	www.climatedataguide.ucar.edu
	Niño_1+2	The SST anomalies in the Niño 1 + 2 region	
	Niño_3.4	The SST anomalies in the Niño 3.4 region	
Atmosphere	AO	The Arctic Oscillation index	www.cpc.ncep.noaa.gov
	NAO	The North Atlantic Oscillation index	

[a]SST, sea surface temperature.

inputted into the RNN model. The timescale of phase changes and life periods of the Madden–Julian Oscillation are different from the indices above, which made it incompatible with the RNN model.

2.2 Methods

2.2.1 Time decomposition of precipitation

To predict precipitation across different terrains in this study, we used time-decomposed historical precipitation as the input features. The precipitation observed at a station actually contains considerable geographical and climatological information about that place, which includes solar radiation, water vapor and heat transported by the monsoon, altitude and latitude, the interannual variability of climate, and the influence of global warming. Guo Y et al. (2012) adopted a similar method to decompose the precipitation sequence, but at different timescales. The precipitation sequence (Pseq) was composed of the season (Ps), trend (Pt), and residuals of the previous two (Pr):

$$\mathrm{Pseq} = \mathrm{Ps} + \mathrm{Pt} + \mathrm{Pr}. \tag{1}$$

The Ps term was obtained by calculating the climatological average. The precipitation anomalies (Pt + Pr) were then equal to Pseq minus Ps. The Pt term was obtained by a locally weighted regression (LOESS). The expression $w_i\left(ax_i^2 + bx_i + c\right)$ approximated by the least squares method is the Pt term in Equation (2):

$$J(a,b,c) = \frac{1}{n}\sum_{i=0}^{n} w_i\left[ax_i^2 + bx_i + c - (\mathrm{Pseq} - \mathrm{Ps})\right]^2, \tag{2}$$

where x_i represents the time sequence, w_i represents the weight coefficient, and the letters a, b, and c represent the coefficients of a quadratic polynomial calculated by the Jacobi least squares method. The term w_i was calculated by the distance (d_i) from the current time point:

$$w_i = \left[1 - \left(\frac{d_i}{d_{\max}}\right)^3\right]^3. \tag{3}$$

The time width of each subinterval ($d_{\max}$) was set to 120 months, which meant that the trend of each time point considered the precipitation data for the adjacent 10 years. More calculation details for the LOESS are provided in Cleveland and Devlin (1988). Finally, the residual of the Ps and Pt terms was the Pr term. The Pr term contained some nonlinear perturbations with smaller time scales, which were still useful for predicting precipitation.

2.2.2 Model configurations and hyperparameters

Figure 2 shows how our model predicted the S2S precipitation at a station. The prediction range of S2S in this study extended from 1 month to 1 year into the future. First, the climate signals and time-decomposed historical precipitation were normalized by the *Z*-score method to input into the RNN, LSTM, or GRU model (Figure 2a, green box). The *Z*-score was calculated in terms of Equation (4). The input features of 12 time steps were then calculated through 4 × 14 hidden layers in chronological order and passed into a fully connected layer to obtain precipitation in the dry and wet seasons (Figure 2b, blue box). Finally, the precipitation of the two seasons was merged to obtain the output precipitation for the next 12 months (Figure 2c, purple box). PyTorch provided the code support for this model (www.pytorch.org).

The structure of the RNN is similar to that of a standard multilayer perceptron, with the distinction that the hidden units connect the current input with the past input, thus enabling it to discover the correlations among input features with a time delay (Pascanu et al., 2013). Developed from the RNN, the LSTM adds some special units called memory blocks in the hidden layers, including the input gate, output gate, and forget gate. These gates control the information flow in a timeline via different activation functions and matrix operations (Sak et al., 2014). Similar to the LSTM units, the GRU adopts gated units to modulate the information flow, but with different locations and with different activation functions of the input gate and reset gate (Cho et al., 2014). To find the optimal loss value, we set a dynamic learning rate with an initial value of 0.01. This learning rate decreased continuously as the loss value

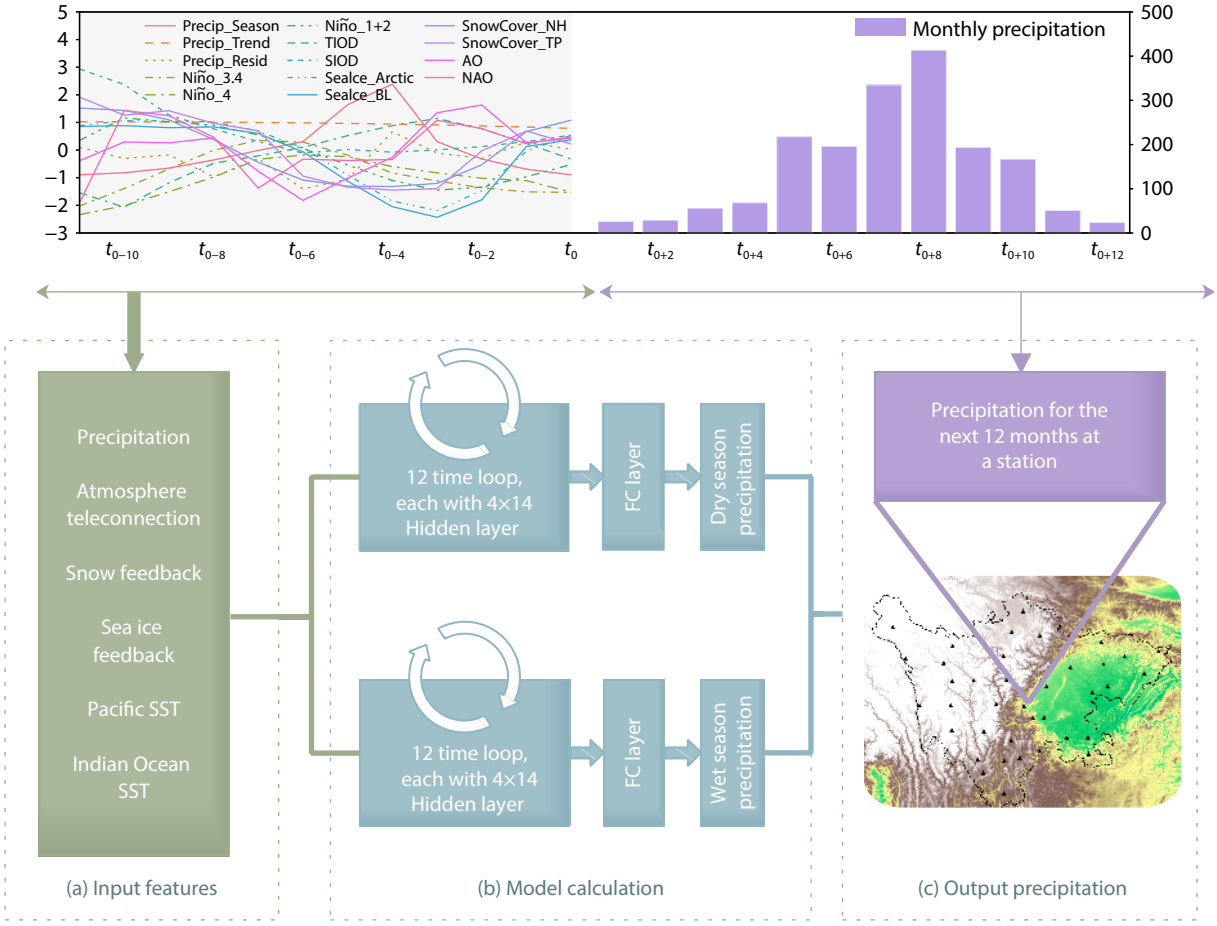

Figure 2. Conceptual view and calculation process of the recurrent neural network (RNN) model. The RNN, long short-term memory recurrent (LSTM), and gated recurrent unit (GRU) models have the same input (a) and output (c) formats. The only difference is in the calculation process of neurons in the hidden layer (b). The left and right *y*-axes in the top panel express the normalized climate indices and predicted precipitation, respectively. TIOD, tropical Indian Ocean sea surface temperature (SST) dipole mode; SIOD, subtropical Indian Ocean SST dipole mode; BL, Barents–Laptev Sea; NH, Northern Hemisphere; TP, Tibetan Plateau; AO, Arctic Oscillation index; NAO, North Atlantic Oscillation index; FC layer, fully connected layer.

converged. When the learning rate was less than 0.0001, the training process stopped. The Adam algorithm was used as an optimizer of loss value, which added first- and second-order momentum to control the direction of loss gradient descent on the basis of the stochastic gradient descent (Kingma and Ba, 2014). To avoid overfitting, an L2 regularization penalty term was also added to the loss function.

In the hidden layer of Figure 2b, two kinds of loss functions, the mean square error (MSE) and mean absolute error (MAE), were adapted to fit the weight and bias parameters of the model of the wet season (May–October) and dry season (November–April of the next year). The MSE loss function was used to fit precipitation in the wet season because we preferred to predict precipitation anomalies far from average. In contrast, precipitation in the dry season was scarce and stable; thus, the MAE loss function was adapted to fit precipitation in the dry season.

It is worth noting that the input features, outputs, and hyperparameters of the RNN, LSTM, and GRU models were completely consistent. The difference between the three models was only in the calculation method of the neurons in the hidden layer:

$$Z = \frac{X - E[X]}{\sigma(X)}, \tag{4}$$

$$\mathrm{MSE\_loss}(\hat{y}, y) = \frac{1}{n}\sum_{i=0}^{n}(\hat{y}_i - y_i)^2, \tag{5}$$

$$\mathrm{MAE\_loss}(\hat{y}, y) = \frac{1}{n}\sum_{i=0}^{n}|\hat{y}_i - y_i|. \tag{6}$$

2.2.3 Sensitivity of the RNN model to input features

The method of sensitivity analysis was first proposed by Garson (1991) to understand the ANN model. This method can help us understand the importance of the input features to the model by analyzing the weight and bias parameters of the hidden layer (Dimopoulos et al., 1995; Olden and Jackson, 2002; Koh and Liang P, 2017).

On the basis of ideas from the research cited above, we generalized the method of sensitivity analysis to the RNN model. In the hidden layer of Figure 2b, there are 4 hidden layers in each time step, each layer has 14 neurons, and these 14 neurons correspond to 14 input features. The calculation method of each neuron in the RNN

model is shown in Equation (7):

$$h_t = \tanh\left(w_i x + b_i + w_h h_{t-1} + b_h\right). \quad (7)$$

In Equation (7), x represents the input features at the current time step, h_{t-1} indicates the output from the previous time step. Terms x and h_{t-1} are multiplied by their respective weight parameter (w) and bias parameter (b) to obtain the output h_t of the current time step. Term h_t reflects how the RNN model understands these input features. For example, if the RNN model considers the Niño 3.4 index important, then the model will assign a larger weight and parameter to produce an h_t with a larger absolute value.

To remove the influence of x and h_{t-1} changes on h_t, we first assumed that there was no feed-forward signal in the RNN model and that h_{t-1} was a matrix of all zeros. There was then no magnitude difference between the input features; x was a unit matrix of all ones. On the basis of the assumptions above, h_t was determined by only the weight and bias parameters. Therefore, h_t could be considered the sensitivity of the RNN model to the input features. The results of the sensitivity analysis are shown in Section 3.5.

2.2.4 Model evaluation

In this study, we divided the data into three subsets to train and test the model. The historical data from 1982 to 2020 were used as the training and testing datasets. Continuously updated data from 2021−2022 were used as revalidation of the model. The historical data included 445 samples, which were randomly divided into training and testing subsets at a proportion of 80% and 20%, respectively. Computer-generated random numbers were used to select and randomly train and test the subsets without intervention of the subject.

The following methods were used to evaluate the RNN model:

(1) *Bias and BiasP*: Bias represents the difference between observation and prediction, whereas BiasP represents the percentage of difference. The subscript of BiasP_{05} represents the 5% quantile in the BiasP distribution for all samples.

(2) *Sign accuracy*: When the trends of prediction and observation were the same, it was correct. The sign accuracy is equal to the correct numbers on the trend sign divided by the total number of predictions.

(3) *Accuracy of the maximum precipitation month* (*AMPM*): If the maximum precipitation month (MPM) of observation and prediction were the same, the prediction was correct. The AMPM is equal to the correct numbers divided by the total number of predictions.

(4) *Temporal correlation coefficient* (*TCC*): The TCC represents a correlation of the time sequences of observation and prediction. Its calculation formula is listed in Equation (8):

$$\mathrm{TCC}_i = \frac{\sum_{j=1}^{n}\left(x_{i,j}-\bar{x}_i\right)\left(y_{i,j}-\bar{y}_i\right)}{\sqrt{\sum_{j=1}^{n}\left(x_{i,j}-\bar{x}_i\right)^2}\times\sqrt{\sum_{j=1}^{n}\left(y_{i,j}-\bar{y}_i\right)^2}}. \quad (8)$$

Here, x and y are the sequences of observation and prediction, respectively. The subscripts i and j refer to the different stations and months, respectively.

3. Results and Discussion

3.1 Time Decomposition of the Precipitation Sequence

According to the method described in Section 2.2.1, the historical precipitation sequences were divided into the terms Ps, Pt, and Pr. Figure 3a shows the Ps terms, which are closely related to the topography and generally decrease with altitude. This is because the huge topography limits low-level water vapor transport on the Plateau. The maximum precipitation occurs in the western Basin as a consequence of the trumpet and windward slope topography, which make this area a hot spot of precipitation (Gao GL et al., 2021). In spite of the altitude difference, which is more than 1,000 m (Figure 1b), the eastern Basin and Valley have a similar annual precipitation. Plentiful rivers, including the Jinsha River, Dadu River, and Yalong River, can support water vapor sources for precipitation in an elevated topography.

Except for the annual precipitation, the MPM is different across the four subregions. The MPM occurs in June–July on the Plateau, in July in the Valley and eastern Basin, and in July–August in the western Basin. The MPM is affected by the subseasonal progress of the Tibetan Plateau summer monsoon, East Asian summer monsoon, and Western Pacific subtropical high (Zhou CY et al., 2011; Zhang Q et al., 2017a).

No obvious periodicity was observed for the terms Pt and Pr, as shown in Figure 3b, c, which means it was difficult to predict future precipitation through the autoregressive model. Figure 3b shows a decreasing trend in the Basin in the late 1990s, which is consistent with previous research (Zhai PM et al., 2005; Ma ZF et al., 2013). Note that a sharply increasing trend appears in the 2010s, except for the Valley. This result suggests that some complex regional differences exist in the precipitation trend against the background of global warming.

Figure 3 shows that the four subregions in Sichuan have different climatological precipitation patterns, which are related to the latitude, altitude, and topography. The term Pr in Equation (1) could be regarded as a comprehensive index of the geographic position. The terms Pt and Pr bring the interannual and decadal signals into the input features. The last figure shows that the three features in the RNN model have a greater sensitivity than the Niño 3.4 index.

3.2 Fitting Ability of the RNN, LSTM, and GRU Models

The predictions of the three deep learning algorithms (the RNN, LSTM, and GRU) were compared in this study. Table 2 shows that the superiority of the RNN increased with the depth of the hidden layers. When the three models had only one hidden layer, the loss values were relatively high and very close. With the increase in the depth of the hidden layers, the loss values dropped sharply in the RNN and GRU models, in contrast with only a slight drop in the LSTM model. The RNN was obviously superior to the LSTM and slightly superior to the GRU. However, it should be noted that deeper hidden layers did not necessarily mean better performance. The loss values of the three models increased when the depth of the models was greater than five hidden layers. In recent research, Zhou KQ et al. (2021) reported a similar result: graphed convolutional networks tended to undergo a performance drop when the model became deep. This phenomenon similarly

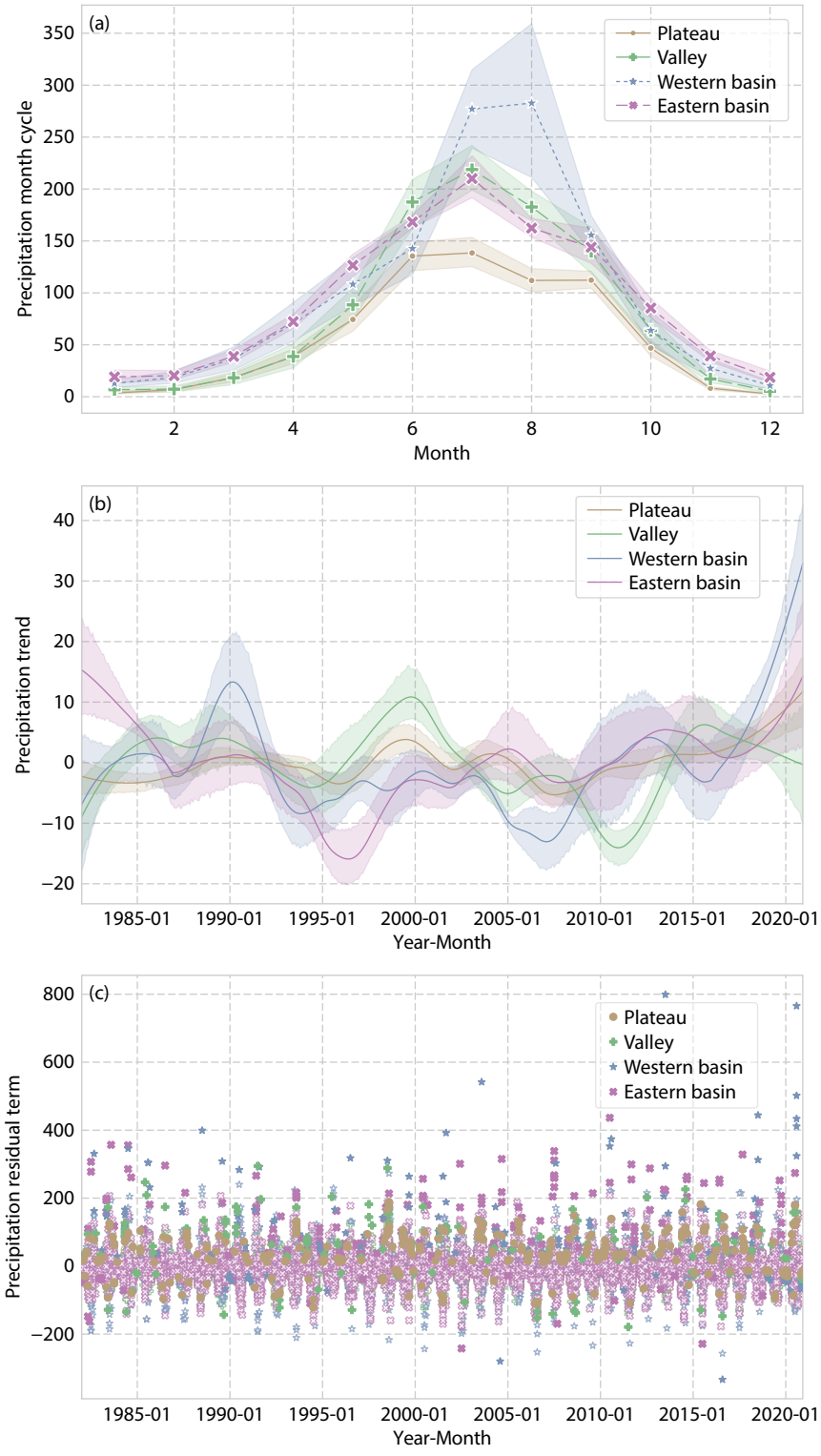

Figure 3. The (a) seasonal (Ps), (b) trend (Pt), and (c) residual (Pr) terms of the precipitation sequence. The precipitation data cover the period from January 1982 to December 2020. In panels (a) and (b), the curves represent the average, and the shading represents the 95% confidence level. In panel (c), the filled polygons represent anomalies that twice exceeded the standard deviation.

appeared in the RNN, LSTM, and GRU models, which should be considered by researchers when building a deep learning model with better performance. Therefore, the four hidden layers were appropriate for training the RNN model in this study.

To analyze the reason for the differences noted above, we ensured the consistency of the hyperparameters in the three models. Hence, the differences in prediction performance were determined only by the calculation method of the neural cells. The LSTM and GRU added the gated structure on the basis of the RNN neural cells. These gated neurons control the information

Table 2. The optimal loss values of the gated recurrent unit (GRU), long short-term memory recurrent (LSTM), and recurrent neural network (RNN) models with different hidden layers.

Hidden layers		1	2	3	4	5	6
Dry season	GRU	0.3143	0.2898	0.2913	0.2958	0.3036	0.5598
	LSTM	0.3129	0.2937	0.2956	0.7649	0.7649	0.7649
	RNN	0.3230	0.2887	0.2805	0.2844	0.3133	0.3234
Wet season	GRU	0.1791	0.0607	0.0517	0.0862	0.1663	0.6626
	LSTM	0.2378	0.2226	0.2139	1.0348	1.0353	1.0351
	RNN	0.2179	0.0886	0.0556	0.0488	0.0479	0.0663

transport from the previous time step, and the input features transform from the current time step, thereby effectively reducing the gradient disappearance problem of the RNN (Pascanu et al., 2013). Information control is mainly achieved through sigmoid and hyperbolic tangent (tanh) activation functions. The RNN model uses only the tanh function (Equation (7)). The LSTM and GRU both use the sigmoid function to control the information from the previous time step. The difference between the LSTM and GRU is that the GRU still uses the tanh function to transform the current input features, which is the same as with the RNN (Dey and Salem, 2017; Sherstinsky, 2020). Therefore, the difference between the sigmoid and tanh functions could have resulted in better performance of the RNN than the LSTM and GRU.

Figure 4 shows how the Niño 3.4 index was transformed with the sigmoid and tanh functions. The Niño 3.4 index, whose positive and negative values respectively represent El Niño and La Niña, is a crucial climate signal. However, the negative values disappear in the sigmoid transform (solid purple line) but remain in the tanh transform (dashed green line). The sigmoid function cannot learn the ENSO cycle properly, which is why the LSTM could not converge the loss values. In the GRU model, the input features were correctly transformed by the tanh function, but the information from the previous time step was limited by the sigmoid function. Therefore, the GRU was superior to the LSTM but inferior to the RNN. This result also suggests that the physical foundation of the input features plays an important role in the prediction of deep learning.

3.3 Validation in the Test Datasets

According to the results in Section 3.2, the RNN model had the optimal loss value. Therefore, the following sections show only the detailed predictions of the RNN model. In the validation of the test datasets, the precipitation bias, sign accuracy of the precipitation trend, and AMPM were evaluated in this study.

Figure 5 shows the accumulated bias between prediction and observation at the seasonal time scale. In the dry season, 5% to 95% quantiles in the boxplot are all within a range of ±80 mm. The exact statistical data are from −38.57 to 34.03 mm in the Plateau, from −66.50 to 40.76 mm in the Valley, from −79.83 to 56.74 mm in the western Basin, and from −70.26 to 66.97 mm in the eastern Basin. In the wet season, the 5% to 95% quantiles in the boxplot are almost within a range of ±100 mm. The exact statistical data are from −62.30 to 37.22 mm in the Plateau, from −89.21 to 58.58 mm in the Valley, from −123.40 to 94.77 mm in the western Basin,

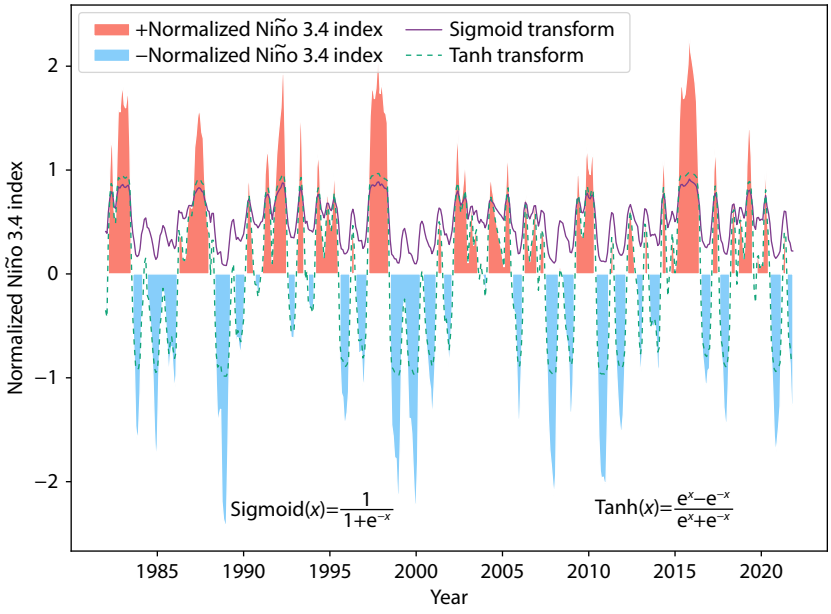

Figure 4. A case of how the normalized Niño 3.4 index was transformed with the sigmoid and tanh activation functions. The results of the sigmoid and tanh transformations are shown by the solid purple line and the dashed green line, respectively.

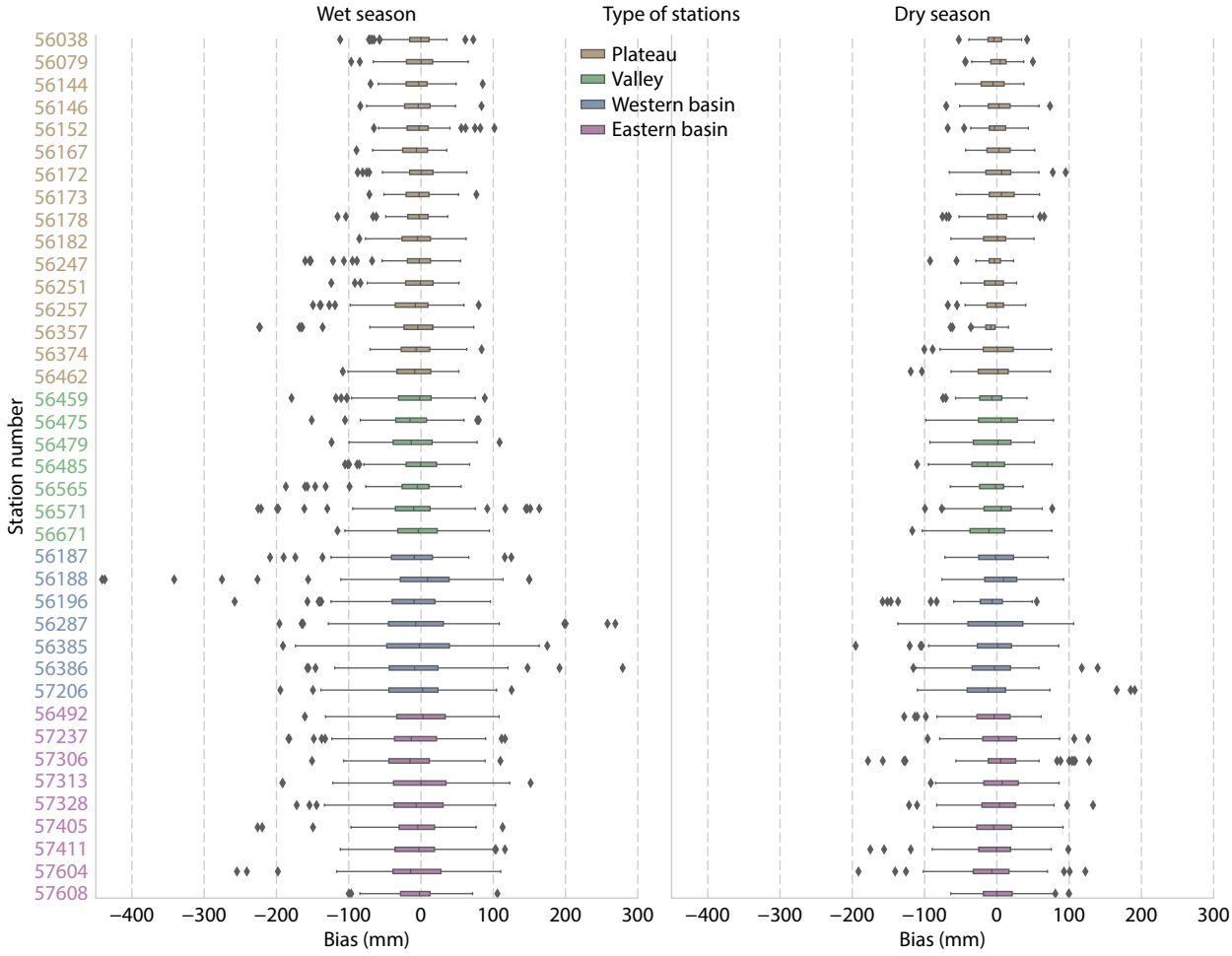

Figure 5. Boxplot showing the bias distribution in the wet (left) and dry (right) seasons. The station number is the label of vertical coordinates, and the different colors represent the different subregions.

and from −94.74 to 69.24 mm in the eastern Basin. Although the absolute bias in the wet season is larger than that in dry season, the relative bias is more accurate in the wet season. The ascending order of the BiasP interval ($BiasP_{05}$–$BiasP_{95}$) is as follows: the Plateau (from −8.83% to 6.73%) < the Valley (from −8.88% to 7.37%) < the eastern Basin (from −9.53% to 8.16%) < the western Basin (from −10.62% to 10.67%).

The statistics above illustrate that the RNN model can predict precipitation in most years, but for a few years (i.e., the samples in the <5% or >95% quantiles), the BiasP is more than ±10%. The prediction bias from summer (June–August) contributes the majority of the accumulated bias. Therefore, prediction at a subseasonal time scale in June–August was further evaluated. Figure 6 shows that most of the dots are distributed along locations adjacent to the diagonal line. The slope of the regression equation is all greater than 0.82, which means that prediction is very close to observation in summer. But it should be noted that the RNN model underestimates precipitation greater than 600 mm and overestimates precipitation below 200 mm. These samples represent precipitation anomalies far from average.

Figure 7 shows the sign accuracy of the precipitation trend in different months. In the wet season, the accuracy averages 83.20% –92.25%, which is higher than that of 58.60%–64.32% in the dry season. It is difficult for the RNN model to predict the precipitation trend in the dry season because an error value of approximately 10 mm may change the mathematical sign of the precipitation trend in the dry season, when the monthly precipitation is usually less than 50 mm. Precipitation in the Tibetan Plateau is rare; thus, the sign accuracy is also lower than that in the Sichuan Basin (shown by the colored dots in Figure 7).

As illustrated in Section 3.1, the average MPM has a regional difference. In fact, the MPM also has an interannual change that is effected by climatic variability (Zhang Q et al., 2017a). The MPM in the Plateau usually occurs in July (frequency of 0.34), followed by June (0.29), August (0.18), and September (0.15). The MPM in the Valley usually occurs in July (0.45), followed by August (0.25) and June (0.24). The frequency distribution of the MPM in the eastern Basin is similar to that in the Valley. The maximum frequency is in July (0.39), followed by June (0.20) and August (0.19). The MPM in the western Basin is mainly concentrated in July (0.44) and August (0.41).

The RNN model successfully predicted the MPMs in most years of the test datasets. Figure 8 shows the joint distribution between prediction and observation. The accumulated frequency located at the diagonal line indicates the AMPM (as detailed in Section 2.2.4). The AMPM was 0.86, 0.85, 0.90, and 0.88 in the Plateau,

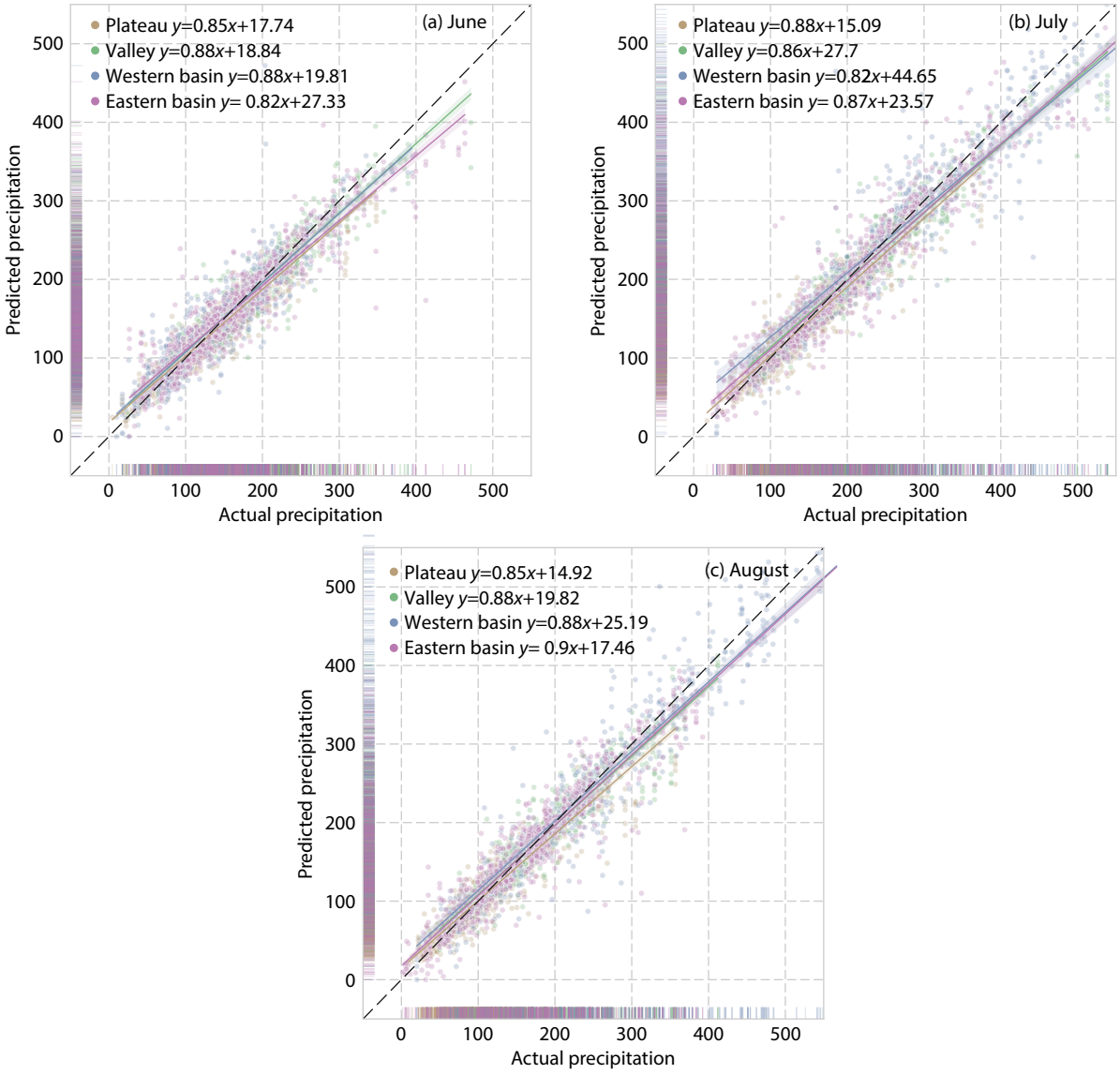

Figure 6. The joint distribution of prediction and observation in June (a), July (b) and August (c). If the prediction is equal to the observation, the dot is placed on the dashed black diagonal line. The colors of the dots represent the different subregions. The colored solid lines represent the linear regression fit between prediction and observation. The regression equations are listed in the upper left corner of the axis. The shadows behind the lines illustrate the 95% confidence interval. The marginal segment lines of the axis show the individual distributions of prediction and observation.

Valley, western Basin, and eastern Basin. If the prediction was completely correct, the AMPM was 1.

3.4 Updated Validation in 2021–2022

To further evaluate the prediction performance with complex terrain in this study, we expanded the original datasets (1982–2020). The updated observation precipitation (January 2021–August 2022) was used again to validate the RNN model. Figure 9 shows the continually updated prediction in 2021–2022. The colored lines represent 12 sliding prediction results, for which the initial forecast time was from January 2021 to December 2021.

In Figure 9, the RNN model shows stable prediction performance at different lead times. Although some biases between prediction and observation still existed, the RNN model was successfully able to predict the precipitation trend and MPM at four stations that belong to the four subregions. Here, we used the TCC score (as detailed in Section 2.2.4) to demonstrate the temporal correlation of curves between prediction and observation. The TCC scores almost exceeded 0.7 at the Ganzi, Huili, Ya'an, and Dachuan stations, which is higher than that of the multimodel ensemble prediction system from the NCC of China (www.cmdp.ncc-cma.net/pred/vescom.php).

It is important to note that the prediction by the RNN model was not a simple reproduction of the historical average. For example, the MPM in Ganzi is June in the historical average, but it was July in 2021 and June in 2022. The MPM in Ya'an is August in the historical average, but it was July in 2021 and August in 2022. These changes were accurately predicted by the RNN model.

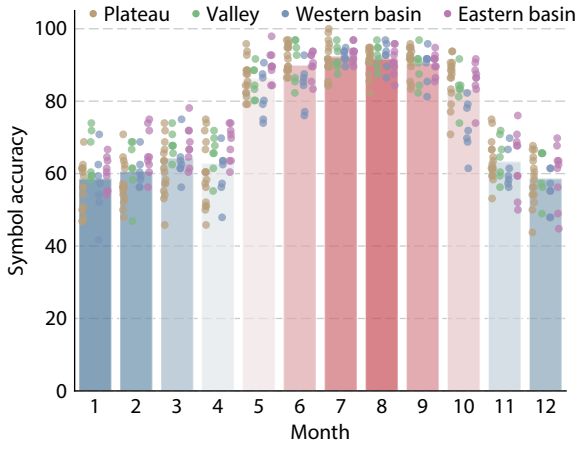

Figure 7. Histogram representing the average sign accuracy at different stations. The dots represent the accuracy at different stations. The colors of the dots represent the different subregions. The warm and cold colors in the histogram indicate the wet and dry seasons, respectively.

Compared with prediction of the Beijing Climate Center Climate System Model (BCC-CSM) from the NCC, the RNN model showed a more stable prediction for the monthly precipitation. For example, for precipitation in August 2021, the RNN model predicted a trend of more precipitation from January to July 2021. The BCC-CSM, with a huge adjustment in different lead times, predicted a trend of more precipitation in the Sichuan Basin in March–June and less precipitation in January, February, and July. For precipitation in August 2022, the BCC-CSM likewise had the opposite prediction at different lead times (more details are available via www.cmdp.ncc-cma.net/cn/monitoring.htm).

At the seasonal time scale, the NCC predicted less precipitation in the summer of 2021 in the eastern Basin, but the observation was more than average precipitation (Zhao JH et al., 2022). The NCC and Institute of Atmospheric Physics of China predicted more precipitation in the summer of 2022 in southwestern China (Peng JB et al., 2022), but the observation was less than average precipitation. The RNN model also had an overestimation in the western Basin and Plateau in 2022, but the prediction in 2021 was closer to the observation.

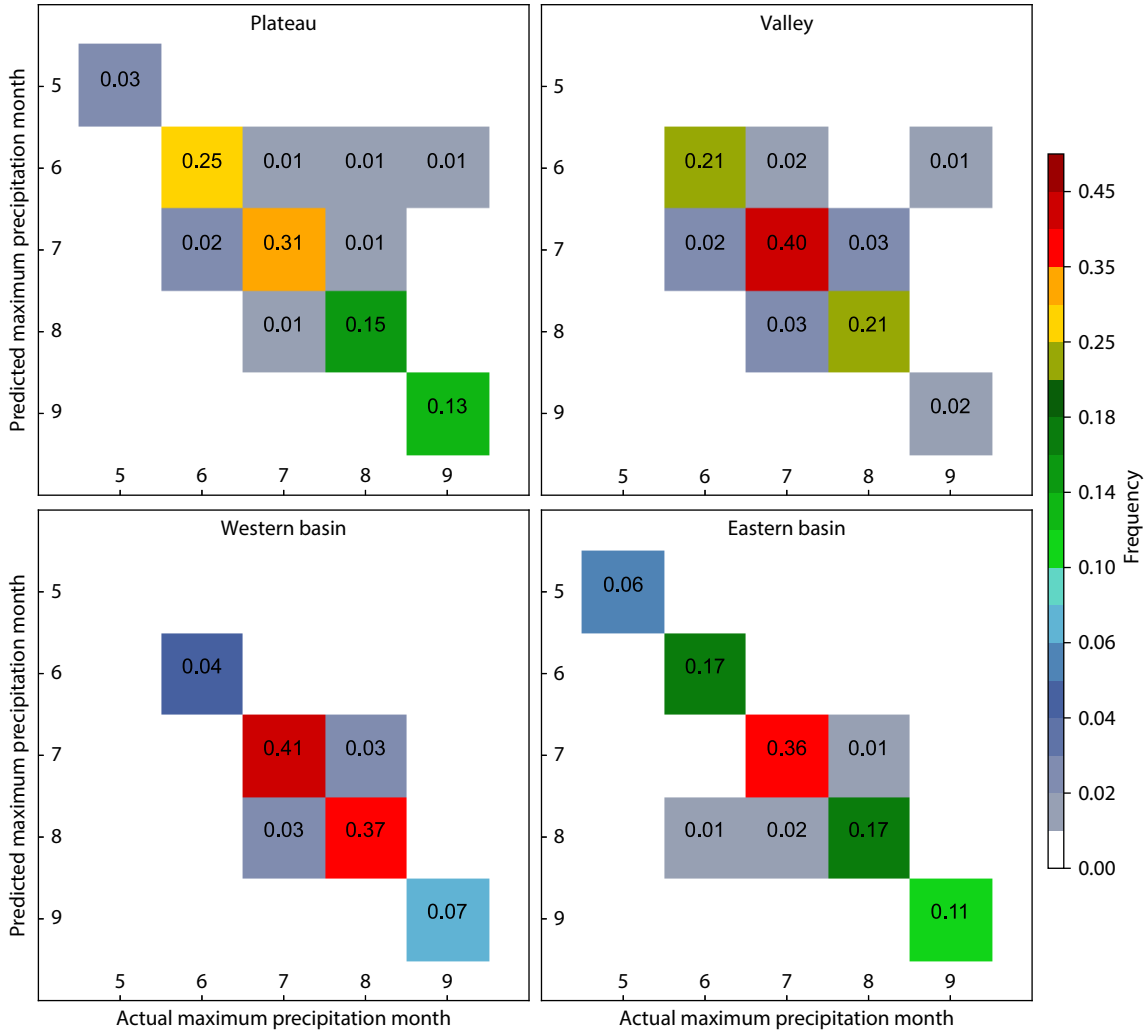

Figure 8. Heatmap representing the frequency distribution between prediction and observation in the four subregions. The values of the diagonal line represent the correct frequency when prediction and observation are the same. A frequency below 0.01 is represented by the white color.

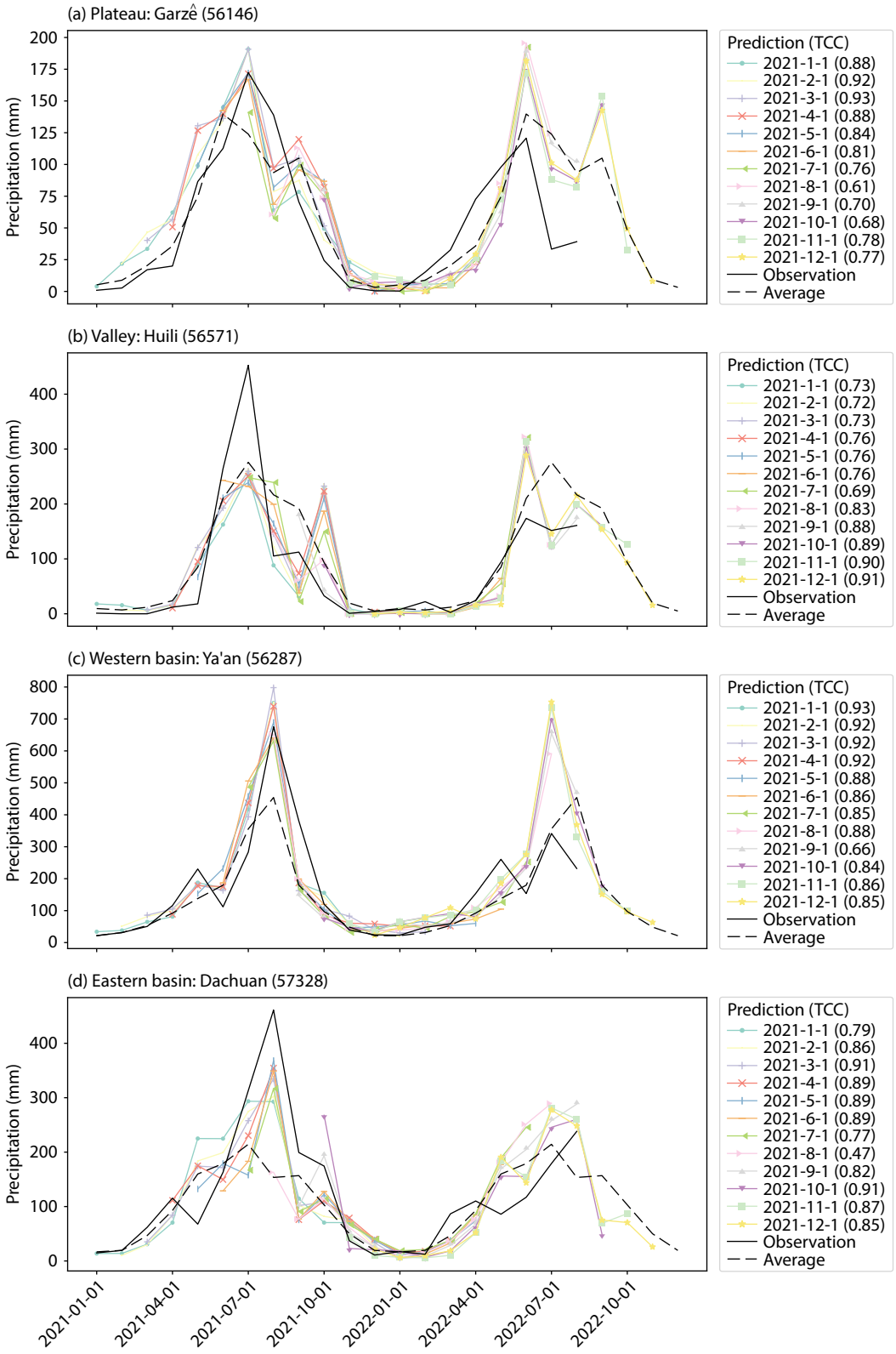

Figure 9. Precipitation curves at different stations by prediction and observation. The (a) Plateau, (b) Valley, (c) western Basin, and (d) eastern Basin are represented by the Garzê, Huili, Ya'an, and Dachuan stations. The station numbers are listed in the panel labels. The black solid and dashed lines indicate precipitation based on updated observations and the historical average, respectively. The colored lines represent the prediction of the RNN model at different lead times. The temporal correlation coefficient (TCC) scores and corresponding lead times are listed in the legend.

3.5 Sensitivity Distribution of the RNN Model to Input Features

The results above show that the RNN model was successfully able to predict the precipitation trend and MPM in most years in the test dataset and the updated datasets of 2021–2022. This result suggests that the RNN model learned some knowledge of climate variability from the input features. According to the method shown in Section 2.2.3, we calculated which input feature was more important or sensitive in the RNN model. Figure 10 shows the sensitivity distribution of the RNN model to input features at different stations. Note that in Figure 10, the absolute value of sensitivity represents importance, although a negative value does not equal unimportance.

Overall, the sensitivity distribution in the dry season was relatively consistent at different stations. The pattern of sensitivity distribution in the wet season was not clear. For example, in the dry season, the RNN model had a dominant positive sensitivity pattern for the Niño 3.4 index and a dominant negative sensitivity pattern for the Arctic Oscillation (AO) index, but in the wet season, the dominant sensitivity pattern disappeared at the different stations. Therefore, this result demonstrates that the influence of climate factors differs between the wet and dry seasons.

The RNN model remained highly sensitive to the decomposed precipitation signal, whether in the dry or wet season. It is important for precipitation prediction that the decomposed precipitation signal provides beneficial information as input features.

For the abnormal SST signal, the RNN model showed a greater sensitivity for the Niño 3.4 index in the dry season, but it obviously decreased in the wet season. In contrast, the sensitivity of the Niño 4 index increased in the wet season. This result was demonstrated by Guo Y et al. (2012), who used a method called the "optimal window" to seek a critical area of SST influencing the precipitation in southwest China. They also showed that the western Pacific warm pool makes a major contribution to the optimal window. For most stations, the RNN model considered that the Niño 3.4 and Niño 4 indexes were equally important for precipitation in the wet season. That result may have a relationship with the long-term trend of the "cold tongue" mode of the tropic Pacific SST and the morphological variation of the ENSO (Li Y et al., 2017a, 2019b; Wang R and Ren HL, 2020). In the Indian Ocean, the tropical Indian Ocean SST dipole (TIOD) pattern (Saji et al., 1999) and the subtropical Indian Ocean SST dipole (SIOD) pattern (Jia XL and Li CY, 2005) show opposite effects on precipitation in the wet season.

Similarly, the sea ice concentration in the Arctic and Barents Sea also shows an opposite effect on precipitation in the wet season. This may be caused by differences in the circulation patterns triggered by sea ice in different regions (Wu BY et al., 2015). On the

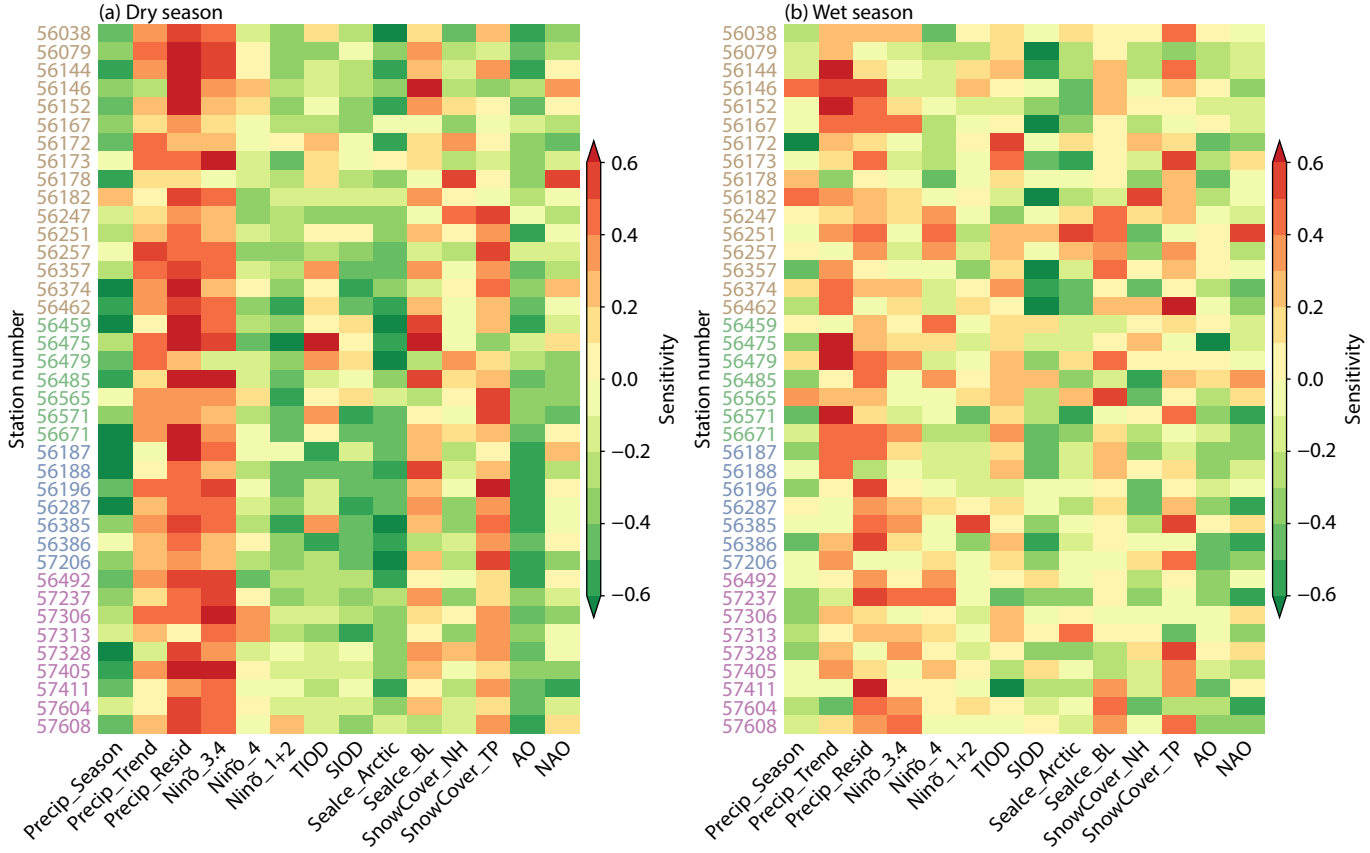

Figure 10. Sensitivity distribution of the recurrent neural network model to input features in the dry season (a) and wet season (b). The labels of the vertical coordinates are the station numbers, the same as in Figure 5. The larger the absolute value of the sensitivity, the more important the feature is. TIOD, tropical Indian Ocean sea surface temperature (SST) dipole mode; SIOD, subtropical Indian Ocean SST dipole mode; BL, Barents–Laptev Sea; NH, Northern Hemisphere; TP, Tibetan Plateau; AO, Arctic Oscillation index; NAO, North Atlantic Oscillation index.

other hand, the sea ice loss in the Arctic usually leads to a negative AO and North Atlantic Oscillation (NAO) phase (Vihma, 2014). Figure 10 shows a same-sign-oriented sensitivity pattern in the Arctic sea ice and AO index in the dry season. The sensitivity of the NAO index is relatively low. Wang P et al. (2020) similarly indicated that the importance of climate factors varies in different subregions in southwest China, but the importance of the AO index is always higher than that of the NAO index.

In both the dry and wet seasons, the sensitivity of snow cover on the Tibetan Plateau was higher than that on the whole Northern Hemisphere. This may have been caused by the snow melt on the Tibetan Plateau, which has a significant effect on the East Asian Monsoon (You QL et al., 2020). The snow cover in the Northern Hemisphere mainly triggers the mid- to high-latitude wave train (Zhang RN et al., 2017b). Therefore, the index of snow cover on the Tibetan Plateau had a closer relationship with the precipitation on the east side of the Tibetan Plateau.

4. Conclusions

In this study, we established a deep learning model to predict S2S precipitation in the Sichuan Province. This model inputted the time-decomposed historical precipitation and multiple climate factors into the RNN neurons to fit the nonlinear relationship between input features and precipitation. Through dual validations in test datasets and updated 2021–2022 datasets, the results showed that the RNN model could accurately predict the precipitation trend and MPM in most years in the Plateau, Valley, and Basin and was able to provide a stable trend prediction for summer precipitation at different lead times.

A comparison of the optimal loss values of the three deep learning models showed the RNN had better prediction performance than the LSTM and GRU models, which increased with the depth of the hidden layers. When the depth was more than five layers, the prediction performance of the three models dropped off. This result implies that deeper hidden layers did not necessarily mean better performance. The activation function was considered important to explain the superiority of the RNN. The tanh function could properly transfer the climate indices, such as the ENSO signal, whereas the sigmoid function of the LSTM ignored the La Niña events of the ENSO cycle.

The sensitivity analysis showed that the RNN model could learn some nonlinear relationships between climate factors and precipitation. The RNN model considered that the SST anomalies represented by the Niño 4 and Niño 3.4 indexes were equally important to the precipitation prediction in the wet season. The sensitivity to snow cover on the Tibetan Plateau was higher than that in the Northern Hemisphere. An opposite sensitivity pattern appeared in the two pairs of similar factors, namely the TIOD and SIOD and the sea ice concentrations in the Arctic and Barents Sea. A sensitivity value with the same sign appeared in the AO index and the sea ice concentration in the Arctic in the model of the dry season.

This study shows that the physical basis of input features plays a key role in deep learning. Previous research has also indicated that a physically constrained model has better performance than a data-driven-only model (Raissi et al., 2019; Gao GL et al., 2021; Zhou L and Zhang RH, 2022). Therefore, a possible approach to improving the prediction of deep learning is by coupling physical processes with deep learning. It was quite interesting to find that deep learning could learn some knowledge through data fitting. In future work, we will further explore the physical connection between deep learning and input features. In addition, the Coupled Model Intercomparison Project Phase 5/6 (CMIP5/6) was able to provide the deep learning model with extensive trainable datasets because the filled observation gap and expanded climate elements improved the performance of deep learning models in the long-term prediction of precipitation (Taylor et al., 2012; Eyring et al., 2016). Although there were some obvious biases between the observation and climate models, it is worth considering how to use the CMIP5/6 datasets to improve the S2S prediction and further reduce the uncertainty in different emission scenarios.

Acknowledgments

We acknowledge the language support from Yulong Tan, and the helpful suggestions from anonymous reviewers and editors. This work was supported by the National Natural Science Foundation of China (Nos. U20A2097, 42175042), the Natural Science Foundation of Sichuan (Nos. 2022NSFSC1056, 2023NSFSC0246), the China Scholarship Council (No. 201908510031), the Plateau and Basin Rainstorm, Drought and Flood Key Laboratory of Sichuan Province (Nos. SCQXKJZD202102-6, SCQXKJYJXMS202102), the Innovation Team Fund of Southwest Regional Meteorological Center, China Meteorological Administration (No. XNQYCXTD202201), and the Sichuan Science and Technology Program (No. 2022YFS0544).

References

Brunet, G., Shapiro, M., Hoskins, B., Moncrieff, M., Dole, R., Kiladis, G. N., Kirtman, B., Lorenc, A., Mills, B., ... Shukla, J. (2010). Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction. *Bull. Amer. Meteor. Soc.*, *91*(10), 1397–1406. https://doi.org/10.1175/2010bams3013.1

Chen, Y. R., Li, Y. Q., Wang, C. G., and Deng, R. Y. (2009). Study on the relationship between South Asia high and rainfall of Sichuan–Chongqing regions in summer. *Plateau. Meteor.* (*in Chinese*), *28*(3), 539–548.

Chhetri, M., Kumar, S., Pratim Roy, P., and Kim, B. G. (2020). Deep BLSTM-GRU model for monthly rainfall prediction: A case study of Simtokha, Bhutan. *Remote Sens.*, *12*(19), 3174. https://doi.org/10.3390/rs12193174

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1724–1734). Doha: Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1179

Cleveland, W. S., and Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *J. Am. Stat. Assoc.*, *83*(403), 596–610. https://doi.org/10.1080/01621459.1988.10478639

De Andrade, F. M., Coelho, C. A. S., and Cavalcanti, I. F. A. (2019). Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models. *Climate Dyn.*, *52*(9), 5451–5475. https://doi.org/10.1007/s00382-018-4457-z

Dey, R., and Salem, F. M. (2017). Gate-variants of Gated Recurrent Unit (GRU) neural networks. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)* (pp. 1597–1600). Boston: IEEE. https://doi.org/10.1109/MWSCAS.2017.8053243

Dimopoulos, Y., Bourret, P., and Lek, S. (1995). Use of some sensitivity criteria for choosing networks with good generalization ability. *Neur. Process. Lett.*, *2*(6), 1–4. https://doi.org/10.1007/BF02309007

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*,

9(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Gao, G. L., Li, Y., Li, J. Q., Zhou, X. Y., and Zhou, Z. Q. (2021). A hybrid model for short-term rainstorm forecasting based on a back-propagation neural network and synoptic diagnosis. *Atmos. Oceanic Sci. Lett.*, *14*(5), 100053. https://doi.org/10.1016/j.aosl.2021.100053

Garson, G. D. (1991). Interpreting neural-network connection weights. *AI Exp.*, *6*(4), 47–51.

Guo, Q., Liu, X. W., Tang, H. Y., and Li, Y. H. (2021). Prediction and correction of in situ summer precipitation in Southwest China based on a downscaling method with the BCC_CSM. *Theor. Appl. Climatol.*, *145*(3), 1145–1159. https://doi.org/10.1007/s00704-021-03687-w

Guo, Y., Li, J. P., and Li, Y. (2012). A time-scale decomposition approach to statistically downscale summer rainfall over North China. *J. Climate*, *25*(2), 572–591. https://doi.org/10.1175/jcli-d-11-00014.1

He, S. P., Gao, Y. Q., Li, F., Wang, H. J., and He, Y. C. (2017). Impact of Arctic Oscillation on the East Asian climate: A review. *Earth Sci. Rev.*, *164*, 48–62. https://doi.org/10.1016/j.earscirev.2016.10.014

Jia, X. L., and Li, C. Y. (2005). Dipole oscillation in the Southern Indian Ocean and its impacts on climate. *Chin. J. Geophys.* (*in Chinese*), *48*(6), 1238–1249. https://doi.org/10.3321/j.issn:0001-5733.2005.06.004

Kingma, D., and Ba, J. (2014). Adam: A method for stochastic optimization. [arXiv preprint]. arXiv: 1412.6980. https://doi.org/10.48550/arXiv.1412.6980

Koh, P. W., and Liang, P. (2017). Understanding black-box predictions via influence functions. In *Proceedings of the 34th International Conference on Machine Learning* (pp. 1885–1894). Sydney: JMLR.org.

Kumar, D., Singh, A., Samui, P., and Jha, R. K. (2019). Forecasting monthly precipitation using sequential modelling. *Hydrolog. Sci. J.*, *64*(6), 690–700. https://doi.org/10.1080/02626667.2019.1595624

Li, C. C., Ren, H. L., Zhou, F., Li, S. L., Fu, J. X., and Li, G. P. (2018). Multi-pentad prediction of precipitation variability over Southeast Asia during boreal summer using BCC_CSM1.2. *Dyn. Atmos. Oceans*, *82*, 20–36. https://doi.org/10.1016/j.dynatmoce.2018.02.002

Li, C. Y., Mu, M., Zhou, G. Q., and Yang, H. (2008). Mechanism and prediction studies of the ENSO. *Chin. J. Atmos. Sci.* (*in Chinese*), *32*(4), 761–781. https://doi.org/10.3878/j.issn.1006-9895.2008.04.06

Li, Y., Li, J. P., Zhang, W. J., Chen, Q. L., Feng, J., Zheng, F., Wang, W., and Zhou, X. (2017). Impacts of the tropical pacific cold tongue mode on ENSO diversity under global warming. *J. Geophys. Res.: Oceans*, *122*(11), 8524–8542. https://doi.org/10.1002/2017JC013052

Li, Y., Chen, Q. L., Liu, X. R., Xing, N., Cheng, Z. G., Cai, H. K., Zhou, X., Chen, D., Wu, X. F., and Li, M. G. (2019). The first two leading modes of the tropical Pacific and their linkage without global warming. *Earth Planet. Phys.*, *3*(2), 157–165. https://doi.org/10.26464/epp2019019

Lu, R. Y., and Ye, H. (2011). Decreasing trend in summer precipitation over the western Sichuan Basin since the 1950s. *Atmos. Oceanic Sci. Lett.*, *4*(2), 114–117. https://doi.org/10.1080/16742834.2011.11446915

Ma, Z. F., Liu, J., Zhang, S. Q., Chen, W. X., and Yang, S. Q. (2013). Observed climate changes in Southwest China during 1961–2010. *Adv. Climate Change Res.*, *4*(1), 30–40. https://doi.org/10.3724/SP.J.1248.2013.030

Mayer, K. J., and Barnes, E. A. (2021). Subseasonal forecasts of opportunity identified by an explainable neural network. *Geophys. Res. Lett.*, *48*(10), e2020GL092092. https://doi.org/10.1029/2020GL092092

Misra, S., Sarkar, S., and Mitra, P. (2018). Statistical downscaling of precipitation using long short-term memory recurrent neural networks. *Theor. Appl. Climatol.*, *134*(3), 1179–1196. https://doi.org/10.1007/s00704-017-2307-2

Nie, Y. B., and Sun, J. Q. (2021). Synoptic-scale circulation precursors of extreme precipitation events over Southwest China during the rainy season. *J. Geophys. Res.: Atmos.*, *126*(13), e2021JD035134. https://doi.org/10.1029/2021JD035134

Olden, J. D., and Jackson, D. A. (2002). Illuminating the "black box": A randomization approach for understanding variable contributions in artificial neural networks. *Ecol. Model.*, *154*(1–2), 135–150. https://doi.org/10.1016/S0304-3800(02)00064-9

Pascanu, R., Mikolov, T., and Bengio, Y. (2013). On the difficulty of training recurrent neural networks. In *Proceedings of the 30th International Conference on Machine Learning* (pp. III-1310–III-1318). Atlanta: JMLR.org.

Peng, J. B., Zheng, F., Fan, F. X., Chen, H., Lang, X. M., Zhan, Y. L., Lin, C. H., Zhang, Q. Y., Lin, R. P., ... Zhou, T. J. (2022). Climate prediction and outlook in China for the flood season 2022.. *Climatic Environ. Res.* (*in Chinese*), *27*(4), 547–558. https://doi.org/10.3878/j.issn.1006-9585.2022.22059

Raissi, M., Perdikaris, P., and Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.*, *378*, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

Ren, H. L., Wu, J., Zhao, C. B., Liu, Y., Jia, X. L., and Zhang, P. Q. (2015). Progresses of MJO prediction researches and developments. *J. Appl. Meteor. Sci.* (*in Chinese*), *26*(6), 658–668. https://doi.org/10.11898/1001-7313.20150602

Ruan, C. Q., Li, J. P., and Feng, J. (2015). Statistical downscaling model for late-winter rainfall over Southwest China. *Sci. China Earth Sci.*, *58*(10), 1827–1839. https://doi.org/10.1007/s11430-015-5104-8

Saji, N. H., Goswami, B. N., Vinayachandran, P. N., and Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, *401*(6751), 360–363. https://doi.org/10.1038/43854

Sak, H., Senior, A., and Beaufays, F. (2014). Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition [arXiv preprint]. arXiv: 1402.1128. https://doi.org/10.48550/arXiv.1402.1128

Schneider, R., Bonavita, M., Geer, A., Arcucci, R., Dueben, P., Vitolo, C., Le Saux, B., Demir, B., and Mathieu, P. P. (2022). ESA-ECMWF report on recent progress and research directions in machine learning for earth system observation and prediction. *NPJ Clim. Atmos. Sci.*, *5*, 51. https://doi.org/10.1038/s41612-022-00269-z

Shen, H. J., Luo, Y., Zhao, Z. C., and Wang, H. J. (2020). Prediction of summer precipitation in China based on LSTM network. *Climate Change. Res.* (*in Chinese*), *16*(3), 263–275. https://doi.org/10.12006/j.issn.1673-1719.2019.067

Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D Nonlin. Phenom.*, *404*, 132306. https://doi.org/10.1016/j.physd.2019.132306

Sun, J. Q., and Wang, H. J. (2012). Changes of the connection between the summer North Atlantic Oscillation and the East Asian summer rainfall. *J. Geophys. Res.: Atmos.*, *117*(D8), D08110. https://doi.org/10.1029/2012JD017482

Sun, S. L., Chen, H. S., Wang, G. J., Li, J. J., Mu, M. Y., Yan, G. X., Xu, B., Huang, J., Wang, J., … Zhu, S. G. (2016). Shift in potential evapotranspiration and its implications for dryness/wetness over Southwest China. *J. Geophys. Res.: Atmos.*, *121*(16), 9342–9355. https://doi.org/10.1002/2016JD025276

Sun, Z. H., Sandoval, L., Crystal-Ornelas, R., Mousavi, S. M., Wang, J. B. , Lin, C., Cristea, N., Tong, D., Carande, W. H., … John, A. (2022). A review of earth artificial intelligence. *Comput. Geosci.*, *159*, 105034. https://doi.org/10.1016/j.cageo.2022.105034

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Vihma, T. (2014). Effects of Arctic sea ice decline on weather and climate: A review. *Surv. Geophys.*, *35*(4), 1175–1214. https://doi.org/10.1007/s10712-014-9284-0

Vitart, F., Robertson, A. W., and Anderson, D. L. T. (2012). Subseasonal to seasonal prediction project: Bridging the gap between weather and climate. *WMO Bull.*, *61*(2), 23–28.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., ... Zhang, L. (2017). The subseasonal to seasonal (S2S) prediction project database. *Bull. Amer. Meteor. Soc.*, *98*(1), 163–173. https://doi.org/10.1175/bams-d-16-0017.1

Wang, P., Wu, X. Q., Hao, Y. R., Wu, C. H., and Zhang, J. (2020). Is Southwest China drying or wetting? Spatiotemporal patterns and potential causes. *Theor. Appl. Climatol.*, *139*(1), 1–15. https://doi.org/10.1007/s00704-019-02935-4

Wang, R., and Ren, H. L. (2020). Understanding key roles of two ENSO modes in spatiotemporal diversity of ENSO. *J. Climate*, *33*(15), 6453–6469. https://doi.org/10.1175/jcli-d-19-0770.1

Weyn, J. A., Durran, D. R., Caruana, R., and Cresswell-Clay, N. (2021). Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *J. Adv. Model. Earth Syst.*, *13*(7), e2021MS002502. https://doi.org/10.1029/2021ms002502

White, C. J., Carlsen, H., Robertson, A. W., Klein, R. J. T., Lazo, J. K., Kumar, A.,

Vitart, F., De Perez, E. C., Ray, A. J., ... Zebiak, S. E. (2017). Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorol. Appl.*, *24*(3), 315–325. https://doi.org/10.1002/met.1654

Wu, B. Y., Su, J. Z., and D'Arrigo, R. (2015). Patterns of Asian winter climate variability and links to Arctic sea ice. *J. Climate*, *28*(17), 6841–6858. https://doi.org/10.1175/jcli-d-14-00274.1

Wu, Y., Sheng, Z., and Zuo, X. J. (2022). Application of deep learning to estimate stratospheric gravity wave potential energy. *Earth Planet. Phys.*, *6*(1), 70–82. https://doi.org/10.26464/epp2022002

Xiao, C., Wu, P. L., Zhang, L. X., and Song, L. C. (2016). Robust increase in extreme summer rainfall intensity during the past four decades observed in China. *Sci. Rep.*, *6*, 38506. https://doi.org/10.1038/srep38506

Yang, S., Li, Z. N., Yu, J. Y., Hu, X. M., Dong, W. J., and He, S. (2018). El Niño–Southern Oscillation and its impact in the changing climate. *Natl. Sci. Rev.*, *5*(6), 840–857. https://doi.org/10.1093/nsr/nwy046

You, Q. L., Wu, T., Shen, L. C., Pepin, N., Zhang, L., Jiang, Z. H., Wu, Z. W., Kang, S. C., and AghaKouchak, A. (2020). Review of snow cover variation over the Tibetan Plateau and its influence on the broad climate system. *Earth-Sci. Rev.*, *201*, 103043. https://doi.org/10.1016/j.earscirev.2019.103043

Zhai, P. M., Zhang, X. B., Wan, H., and Pan, X. H. (2005). Trends in total precipitation and frequency of daily precipitation extremes over China. *J. Climate*, *18*(7), 1096–1108. https://doi.org/10.1175/jcli-3318.1

Zhang, Q., Zheng, Y. J., Singh, V. P., Luo, M., and Xie, Z. H. (2017a). Summer extreme precipitation in Eastern China: Mechanisms and impacts. *J. Geophys. Res.: Atmos.*, *122*(5), 2766–2778. https://doi.org/10.1002/2016JD025913

Zhang, R. N., Zhang, R. H., and Zuo, Z. Y. (2017b). Impact of Eurasian spring snow decrement on East Asian summer precipitation. *J. Climate*, *30*(9), 3421–3437. https://doi.org/10.1175/jcli-d-16-0214.1

Zhao, J. H., Chen, L. J., and Zhang, D. Q. (2022). Overview of climate prediction for the summer 2021 in China and its precursors. *Meteor. Mon.* (*in Chinese*), *48*(4), 479–493. https://doi.org/10.7519/j.issn.1000-0526.2021.110501

Zhou, C. Y., Li, Y. Q., Bu, Q. L., and Peng, J. (2011). Features of drought–flood coexistence in west and east of Sichuan–Chongqing basin in midsummer and its relating background of atmospheric circulation. *Plateau. Meteor.* (*in Chinese*), *30*(3), 620–627.

Zhou, K. Q., Dong, Y. F., Wang, K. X., Lee, W. S., Hooi, B., Xu, H., and Feng, J. S. (2021). Understanding and resolving performance degradation in deep graph convolutional networks. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 2728–2737). Queensland: ACM. https://doi.org/10.1145/3459637.3482488

Zhou, L., and Zhang, R. H. (2022). A hybrid neural network model for ENSO prediction in combination with principal oscillation pattern analyses. *Adv. Atmos. Sci.*, *39*(6), 889–902. https://doi.org/10.1007/s00376-021-1368-4