

Article ID: 1006-8775(2025)01-0064-11

Deep Learning-based Bias Correction Method for Seasonal Prediction of Summer Rainfall in China

QU An-kang (瞿安康)^{1,2}, BAO Qing (包庆)¹, ZHU Tao (朱涛)¹, LUO Zhao-ming (罗昭明)^{1,2}

(1. State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029 China; 2. School of Emergency Management Science and Engineering, University of Chinese Academy of Sciences, Beijing 101408 China)

Abstract: Seasonal prediction of summer rainfall in China plays a crucial role in decision-making, environmental protection, and socio-economic development, while it currently has a low prediction skill. We developed a deep learning-based seasonal prediction bias correction method for summer rainfall in China. Based on prediction fields from the flexible Global Ocean-Atmosphere-Land System Model finite volume version 2 (FGOALS-f2), we optimized the loss function of U-Net, trained with different hyperparameters, and selected the optimum model. U-Net model can extract multi-scale feature information and preserve spatial information, making it suitable for processing meteorological data. With this end-to-end model, the precipitation distribution can be obtained directly without using the traditional method of data dimensionality reduction (e.g., Empirical Orthogonal Function), which could maximize the retention of spatio-temporal information of the input data. Optimization of the loss function enhances the prediction results and mitigates model overfitting. The independent prediction shows a significant skill improvement measured by the anomalous correlation coefficient score. The skill has an average value of 0.679 in China (0°–63°N, 73°–133°E) and 0.691 in the region of the Chinese mainland, which significantly improves the dynamical prediction skill by 1357% and 4836%. This study suggests that the deep learning (U-Net)-based seasonal prediction bias correction method is a promising approach for improving rainfall prediction of the dynamical model.

Key words: seasonal prediction; rainfall; statistical-dynamical model; deep learning

CLC number: P409

Document code: A

Citation: Qu, A. K., Q. Bao, T. Zhu, and Z. M. Luo, 2025: Deep Learning-based Bias Correction Method for Seasonal Prediction of Summer Rainfall in China. *J. Trop. Meteor.*, **31**, 64–74, <https://doi.org/10.3724/j.1006-8775.2024.040>.

1 INTRODUCTION

Seasonal rainfall prediction is foundational for rational planning and coping with meteorological changes in various sectors of society (Ye et al. 2015). It is significant in promoting sustainable development and enhancing societal disaster resilience (Li et al. 2016). Climate dynamical models are widely used in seasonal prediction by simulating the physical processes of the atmosphere, land, ocean, and sea ice on high-performance computers (Yang et al. 2008; Alessandri et al. 2018). However, since the inherent errors in meteorological observation fields, and the simplification and approximation in atmospheric physical processes and parameters, the prediction results of the climate

dynamical model have limited skills (Wang et al. 2009). Seasonal prediction of summer rainfall shows more crucial challenges due to its long period and high randomness (Alessandri et al. 2018). As the lead time increases, the forecasting accuracy decreases quickly (Randall et al. 2007). Bias correction is a common way to improve the prediction skill of dynamical models (Badr et al. 2014; Bett et al. 2018).

There are two main categories of seasonal prediction bias correction methods: dynamical and statistical methods. Dynamical methods reduce prediction bias based on the physical framework of regional climate models, and the results are physically explainable (Giorgi et al. 2009; Colette et al. 2012). However, dynamical methods are sensitive to the parameterization schemes and boundary conditions (Gholami et al. 2021). Traditional statistical methods establish mathematical correction models by leveraging the relationships between observation data and historical predictions from a dynamical model (Lee et al. 2013; Lim et al. 2015). Initially, linear statistical models were favored for their simplicity and low computational demands (Barreiro et al. 2004; Lealdi et al. 2018). Nonetheless, the inherent nonlinearity of meteorological fields imposes a limitation on the effectiveness of these linear models (Goddard et al. 2001). Recently, many nonlinear statistical models (Feng et al. 2020; Finnis et al. 2012), as well as machine

Received 2024-06-12; **Revised** 2024-08-29; **Accepted** 2024-10-24

Funding: Guangdong Major Project of Basic and Applied Basic Research (2020B0301030004); Postdoctoral Fellowship Program of CPSF (GZC20232598); China Postdoctoral Science Foundation (2024M753168); National Key Scientific and Technological Infrastructure Project “Earth System Numerical Simulation Facility” (EarthLab)

Biography: QU An-kang, M.S., primarily undertaking research on the application of machine learning algorithms in atmospheric science.

Corresponding author: ZHU Tao, e-mail: zhutao@mail.iap.ac.cn

learning models (Wang et al. 2021; Xu et al. 2019), have been used to post-process numerical seasonal climate outputs. Nonlinear models can fit complex relationships well, while these methods need manual feature engineering (Duboue et al. 2020) to select relevant features from high-dimensional data and represent these features properly.

With the popularity of deep learning (Krizhevsky et al. 2017; He et al. 2016), it has made a breakthrough in atmospheric science. Deep learning models have end-to-end structures that train a model from raw data to predictands without feature engineering (LeCun et al. 2015). Recently, deep learning methods have been applied to nowcasting (Zhang et al. 2023) and medium-range weather forecasting (Bi et al. 2023; Lam et al. 2023) and obtained significant improvements compared to the dynamical integrated forecasting system of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Bougeault et al. 2001). Many studies also apply deep learning methods to correct the bias of short-lead prediction ranging from weather to sub-seasonal time scales (Han et al. 2021; Kim et al. 2021; Liu et al. 2021). However, fewer deep learning methods are used for bias correction of seasonal prediction due to the lack of sufficient data (Ling et al. 2022). Since the high intensity and uneven spatiotemporal distribution of summer rainfall in East Asia, more advanced methods for the bias correction of seasonal prediction are urgently required. Therefore, in this study, we proposed a deep learning-based bias correction method for the seasonal prediction of summer rainfall in China, reducing the overfitting issue by modifying the loss function and data augmentation with a slide-window strategy. Although there have been studies using the U-Net model for correcting model prediction results (Lyu et al. 2023; Tong et al. 2024), this work aimed to directly apply the trained U-Net model to the backend of our FGOALS-f2 model operationally, so we trained the MU-Net model with hindcast data from operational FGOALS-f2.

The remainder of this paper is organized as follows. Section 3 details the dynamical model, observation data, data preprocessing, deep learning methods (specifically the U-Net model), validation, and loss function optimization techniques employed in this study. Section 4 presents a comparison of the seasonal prediction results obtained from the deep learning-based bias correction method with those from the raw FGOALS-f2 model, focusing on key rainfall characteristics using various metrics. Finally, Section 5 provides conclusions and a discussion of the findings.

2 MODELS, DATASETS AND METHODS

2.1 Dynamical model

The dynamical prediction data used in this study are retrieved from an operational subseasonal-to-seasonal system using the flexible Global Ocean–Atmosphere–Land System Model finite volume version 2 (FGOALS-

f2), which is a coupled global climate dynamical model developed by the Institute of Atmospheric Physics, Chinese Academy of Sciences (IAP-CAS) (He et al. 2019). The FGOALS-f2 model has been proven to be highly skilled in forecasting El Niño events and can provide skillful seasonal forecast results (Bao et al. 2018). The seasonal prediction used in this study initializes on March 20 and predicts the following six months of each year, covering the period from 1981 to 2020, with a $1^\circ \times 1^\circ$ global grid.

We chose variables including the total precipitation, the 10-m U and V wind fields, 2-m temperature, and the 500-hPa geopotential height, based on previous research on summer rainfall in China (Gong and Ho 2002; Shen et al. 2011; Zhu et al. 2014) and its seasonal prediction (Nan and Li 2003; Fan 2006; Yin and Wang 2016).

2.2 Observation data

The observation data of rainfall used in this study are from the Global Precipitation Climatology Project (GPCP) version 2.3, which provides global accumulated daily precipitation on 1° latitude-longitude grids from October 1996 to the present. This dataset offers the most comprehensive analysis of global oceanic rainfall to date and adds necessary spatial details to precipitation analysis on land (Adler et al. 2003).

2.3 Data preprocessing

This study aims to postprocess climate dynamical model outputs using deep learning methods to improve the seasonal prediction skill of summer rainfall in China. Therefore, only a square region covering China (0° – 63° N, 73° – 154° E) is selected. Additionally, to address potential overfitting issues caused by different dimensionalities of variables, the values of each variable are normalized to the range from 0 to 1.

Deep learning methods often involve complex network structures and a huge number of parameters, requiring substantial data for training (LeCun et al. 2015). However, the available data for summer rainfall on seasonal scales are limited. To solve the contradiction between the scarcity of seasonal prediction data and the high demand for data from deep learning models, this study employs a sliding window strategy for data augmentation to mitigate the potential overfitting problem. We generated artificial monthly data by summing thirty continuous daily data starting several days before or after the first day of the month (Fig. 1). It is verified that the artificially processed monthly data has similar probability distribution characteristics to the

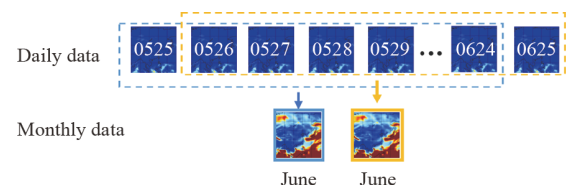


Figure 1. Sliding-window data augmentation strategy. Artificial monthly data consists of thirty continuous daily data starting several days before or after the first day of the month.

original data, which can be used as an effective method for data augmentation (figure ignored).

After data augmentation, the dataset is expanded. Then, we divide the dataset into three types of subsets: the data from 2008 and 2018 is allocated as the validation set; the data from 2000, 2005, 2010, and 2015 is designated as the test set; and the data from 1997 to 2019, excluding the validation and test sets, is utilized for the training set.

2.4 Deep learning methods

The U-Net model was initially applied in Biomedical Image Segmentation (Ronneberger et al. 2015) and has been widely employed in pixel-level regression tasks. In this study, we modified the U-Net model according to the characteristics of input data and trained the modified U-Net(MU-Net) model for seasonal prediction bias correction of China summer rainfall, as illustrated in Fig. 2. The input data size is $64 \times 64 \times 5$, covering a region of $63^\circ \times 63^\circ$ including China. The third dimension represents the five variables, including total precipitation, 2-m temperature, the 10-m U wind, the 10-m V wind, and the 500-hPa geopotential height, respectively. After encoding and decoding, the MU-Net model outputs the corrected rainfall over the same region. Solid lines in the figure represent skip connections, where the connected features simultaneously possess spatial information from low-dimensional features and semantic information from high-dimensional features.

Incorporating the total precipitation from the climate dynamical model could enable the correction of systematic errors between model forecast data and observational data. Training, including the 2-m temperature, 10-m U and V wind fields, and 500-hPa geopotential height, allows for learning the mutual constraints between physical quantities, thereby correcting the precipitation forecast from the dynamical model. Since the inputs of our model are contemporaneous data from model predictions, this has alleviated the issue of deep learning-based climate prediction methods significantly deviating from

observational results over long time scales due to the lack of physical constraint. Additionally, compared to traditional machine learning models that correct station-by-station or grid-by-grid, the U-Net model, owing to convolutional operations, can effectively utilize the local correlations of features. It considers the influence of surrounding areas on a grid point, thus improving the rationality of the correction results.

2.5 Validation

In addition to evaluation metrics commonly used in the machine learning fields (e.g., root mean squared error), the anomaly correlation coefficient (ACC) is also applied in this work. This method mainly evaluates the consistency of the forecast results in the spatial distribution of climate anomalies, which is a main metric used to measure the ability to predict seasonal rainfall. The calculation formula is as follows:

$$ACC = \frac{\sum_{i=1}^N (\Delta R_f - \overline{\Delta R_f}) \sum_{i=1}^N (\Delta R_0 - \overline{\Delta R_0})}{\sqrt{\sum_{i=1}^N (\Delta R_f - \overline{\Delta R_f})^2 \sum_{i=1}^N (\Delta R_0 - \overline{\Delta R_0})^2}} \quad (1)$$

where ΔR_f and $\overline{\Delta R_f}$ represent the forecast values and their multi-year averages of precipitation anomaly percentages, respectively, ΔR_0 and $\overline{\Delta R_0}$ represent the corresponding observed values, and N represents the total number of grid points for scoring.

We also used the consistency of drought and flood states to evaluate our model. In this work, the drought state and flood state are defined relative to their own climatology. For example, considering observed precipitation, if the summer precipitation at a specific grid point is above its own climatological average, we define the observed precipitation in this grid point as a flood state; otherwise, it is considered a drought state. Then, if the observed and predicted precipitation at a specific grid point both fall into a drought state or both fall into a flood state, we define that grid point as consistency; otherwise, it is defined as inconsistency (Table 1).

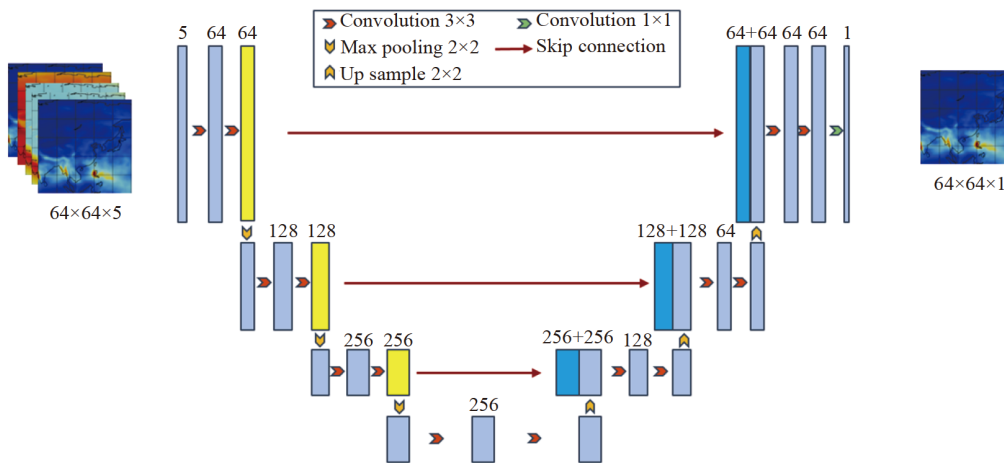


Figure 2. The architecture of the U-Net model used in this study. The input data size is $64 \times 64 \times 4$. Horizontal red arrow represents convolution with a 3×3 kernel; horizontal green arrow represents 1×1 convolution; downward arrow represents 2×2 max-pooling operation, upward arrow represents linear upsampling, and solid line represents skip connection. Blue feature maps are copied from the corresponding yellow feature maps on the left. The output size is $64 \times 64 \times 1$, representing the corrected precipitation values.

Table 1. Definition of consistency or inconsistency.

Observation	Prediction	
	Dry anomaly	Wet anomaly
Dry anomaly	Consistency	Inconsistency
Wet anomaly	Inconsistency	Consistency

2.6 Loss function optimization

Mapping model outputs directly to the observation is unstable, and the model often exhibited slow convergence and tended to experience overfitting if the mean squared error (MSE) is used as the model's loss function. Considering that this study aims to correct the climate model forecast results and enable the corrected results to fluctuate around the raw model outputs, a penalty term is added to the loss function (i.e., MSE) to penalize the deviations between the corrected results of the deep learning model and the prediction from the climate dynamical model. The modified loss function, named ModifiedMSE, is calculated as follows:

$$\begin{aligned} \text{ModifiedMSE} &= \frac{1}{N} \sum_{i,j=1}^n (Y_{ij} - f(x)_{ij})^2 \\ &+ w \times \frac{1}{N} \sum_{i,j=1}^n (x_{ij}^{\text{tp}} - f(x)_{ij})^2 \quad (2) \\ &= \text{MSE}(Y, f(x)) + w \times \text{MSE}(x^{\text{tp}}, f(x)) \end{aligned}$$

where x^{tp} represents the total precipitation variable in the input of the deep learning model, the w is used to adjust the weight of the penalty term. The formula for MSE is as follows:

$$\text{MSE} = \frac{1}{N} \sum_{i,j=1}^n (Y_{ij} - f(x)_{ij})^2 \quad (3)$$

where Y_{ij} represents the observed total precipitation value at grid point (i, j) , x represents the data input to the deep learning model, and $f(x)_{ij}$ represents the total precipitation forecast value at grid point (i, j) after correction by the deep learning model.

Training with the modified loss function (i.e., ModifiedMSE), the performance on the validation set is shown in Fig. 3. When training with MSE as the loss function, the model achieves high ACC scores on the training set but poor ACC scores on the validation set, indicating overfitting. However, when training with ModifiedMSE as the loss function, the model maintains high ACC scores on both the training and validation sets without overfitting. The optimized loss function helps alleviate overfitting in the training step.

3 RESULTS

3.1 Hyperparameter determination

We trained multiple models with different combinations of hyperparameters, including model depth,

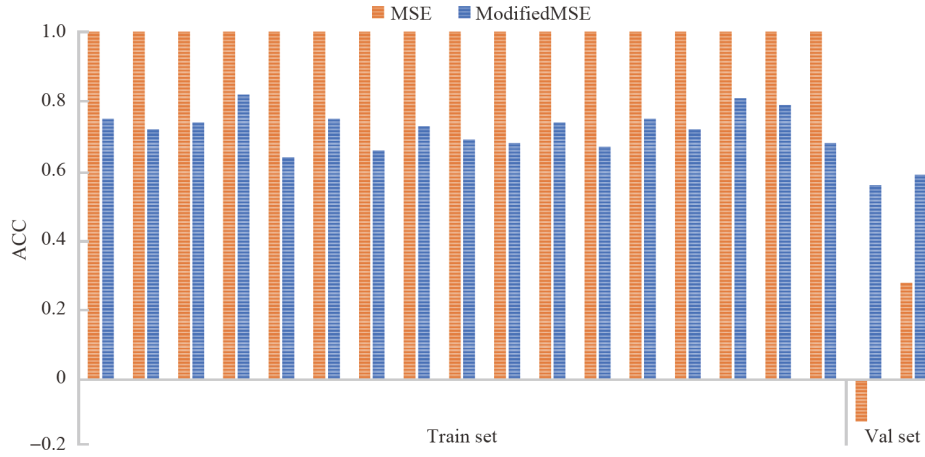


Figure 3. Comparison of ACC scores of precipitation correction models trained with MSE and ModifiedMSE as loss functions on the training and validation sets. The orange bars represent the ACC score of the model trained with MSE as the loss function on the training and validation set, while the blue bars represent the ACC score of the model trained with ModifiedMSE as the loss function on the training and validation set.

Table 2. RMSE scores of models with different hyperparameter combinations on the validation set for various months. Apart from CAS-FGOALS, “L” represents the number of layers in the MU-Net model, “P” represents the penalty coefficient, and “E” represents the number of training epochs.

Month	CAS-FGOALS	4L-0.01P-100E	4L-0.01P-150E	4L-0P-150E	4L-0.01P-300E	5L-0.01P-150E
June	124.6	121.3	71.2	158.0	79.2	82.4
July	144.0	110.1	80.3	166.6	81.9	89.4
August	132.4	121.3	75.3	162.0	76.6	81.8
JJA	133.9	111.4	75.6	162.2	79.2	84.6

number of training epochs, and penalty weight (Table 2). The results of training a 4-layer MU-Net for 100 epochs did not converge. Models without penalty exhibited poor performance on the validation set, suggesting potential overfitting and limited generalization ability.

The hyperparameter combination that yielded the best results on the validation set consisted of training a 4-layer MU-Net for 150 epochs with a penalty weight of 0.01. With this model, we evaluate the monthly total precipitation, anomaly percentage, and drought-flood mode on the test set, which includes data from 2000, 2005, 2010, and 2015, spanning four years.

3.2 Predictors determination

When the data volume is sufficient, incorporating more predictors can provide additional information, enhancing the model's ability to fit nonlinear predictions. However, insufficient data is a common challenge for deep learning applications in seasonal forecasting. Therefore, we need to select a limited set of predictors that can provide sufficient information to the deep learning model.

From a physical perspective, both SST and 850-hPa wind speed undoubtedly have significant impacts on summer precipitation in China. However, when we incorporated SST and 850-hPa U and V into the MU-Net model, its performance on the independent test sets was not satisfactory (Fig. 4). This was primarily due to overfitting caused by the limited number of training samples.

3.3 Study area determination

Expanding the study area can similarly provide the model with more information. However, like with predictors, when the data volume is insufficient, using a larger study area as input to the model can lead to severe overfitting.

In order to select an appropriately sized study area, we conducted extensive sensitivity tests regarding the selection of the study region, including using global circulation fields as inputs (Table 3). However, the MU-Net model trained with global circulation did not yield the highest predictive skill of summer precipitation on the independent test sets. Therefore, in order to strike a balance between fitting ability and generalization capability, we select a square region covering China (0° – 63° N, 73° – 154° E).

3.4 Comparison of monthly total precipitation

This study evaluates the monthly total precipitation corrected by the MU-Net model. In the independent prediction stage, the corrected monthly total precipitation is spatially and numerically closer to the observation than the original model prediction. Taking 2000 as an example (Fig. 5), the original model showed less precipitation in South China in July. After correction, the intensity of July precipitation in South China increased, aligning more closely with the observed results. Using RMSE as the evaluation metric, the RMSE for June decreased from 124.6 in the original prediction to 71.2 after correction, for

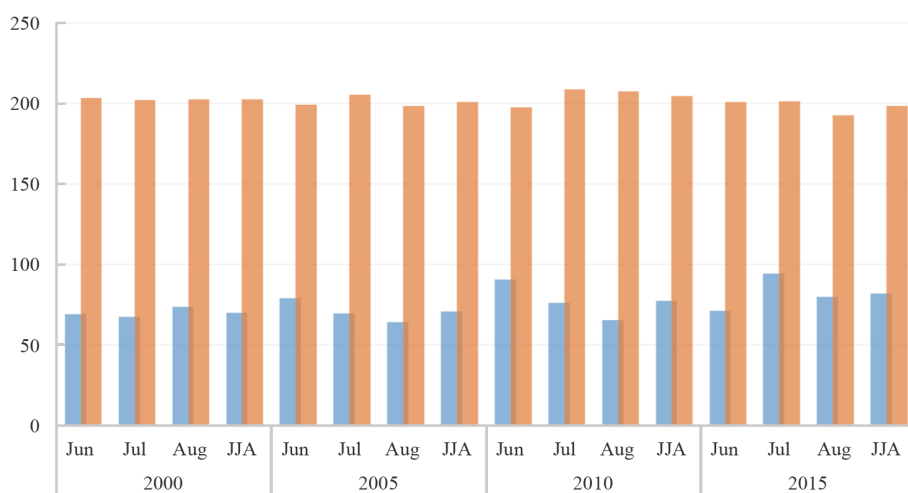


Figure 4. Comparison of the RMSE of the MU-Net model with and without the inclusion of SST and 850-hPa U & V wind components on the independent test sets (2000, 2005, 2010, and 2015). The blue bars represent the RMSE values (units: mm month^{-1}) for the MU-Net model without SST and 850-hPa wind components, while the orange bars represent the RMSE values when SST and 850-hPa wind components are included. Results are shown for June, July, August, and the whole summer (June–July–August, JJA) as a whole.

Table 3. RMSE scores (units: mm month^{-1}) for MU-Net models across different study regions on the independent test sets. The italicized text indicates the study region selected in this work.

Month	CAS-FGOALS	0° – 63° N, 73° – 154° E	3° – 54° N, 73° – 134° E	Global
June	124.6	<i>71.2</i>	74.2	148.2
July	144.0	<i>80.3</i>	78.7	156.8
August	132.4	<i>75.3</i>	72.5	146.3
JJA	133.9	<i>75.6</i>	75.1	150.4

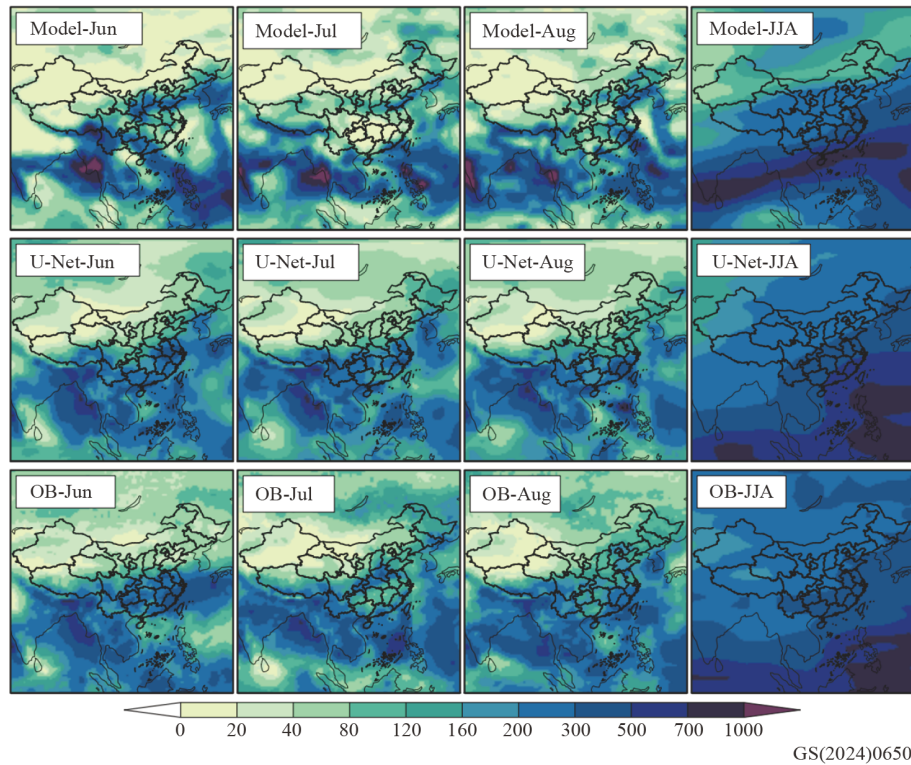


Figure 5. Comparison of monthly total precipitation (units: mm) between model forecast results and corrected results in the rectangular region for the year 2000 (JJA represents monthly average results). The first row shows the model output results, the second row shows the corrected results, and the third row shows the observed results.

July from 144.0 to 80.3, for August from 132.4 to 75.3, and for the entire summer season from 133.9 to 75.6. Additionally, the average results show that the results after correction by the deep learning model are closer to the observed results in terms of spatial distribution.

3.5 Comparison of anomalous precipitation percentage

In the independent prediction stage, the anomalous precipitation percentage is also compared between the original model prediction and the MU-Net corrected results. It can be found that the corrected anomalous precipitation percentage is closer to the observation than the original model prediction. Similarly, we use the data from 2000 as an example (Fig. 6), a wrong dry anomaly is predicted in the eastern region of China in June by the original model, while it has been corrected to the wet anomalous mode in the MU-Net corrected results. Moreover, the MU-Net outputs successfully predict the wet anomaly in Xinjiang during the entire summer of 2000, which is in agreement with the observation. In contrast, the original model shows a wrong drought anomaly in that region. Similar improvements by the MU-Net corrected model could be also observed in the Yangtze River Basin and the northeast region.

Further, the improvement of the anomalous precipitation percentage by the MU-Net corrected model was quantified by ACC scores (Fig. 7). It can be found that the ACC in anomalous precipitation percentage obtained from the MU-Net correction model has been significantly increased both in monthly scale (June, July, and August)

and in the average of the whole summer. Specifically, the multi-year averaged ACC is improved from -0.124 in the original model prediction to 0.579 in the MU-Net corrected model in June, from 0.107 to 0.493 in July, from -0.032 to 0.630 in August, and from -0.054 to 0.679 in the whole summer.

For China, accurate seasonal prediction of summer land precipitation has even more practical significance. Therefore, this work further focuses on the anomalous precipitation percentage on the mainland of China (Fig. 8), and calculates the ACC scores of the original model prediction and the MU-Net corrected results on the Chinese mainland, respectively (Fig. 9). It's found that the ACC scores in Chinese mainland are also significantly increased after the MU-Net correction, in which the multi-year averaged ACC is increased from -0.102 in the original model prediction to 0.547 in the MU-Net corrected model in June, from -0.104 to 0.572 in July, from 0.078 to 0.639 in August, and from 0.014 to 0.691 in the whole summer.

3.6 Comparison of drought and flood states

In addition to total precipitation and anomalous precipitation percentage, the drought and flood states of precipitation may have more practical application significance, especially in the field of agricultural production. Therefore, in this subsection, we calculate the state consistency for the original model prediction and the MU-Net corrected results, respectively in the independent prediction stage. Fig. 10 shows the 2020

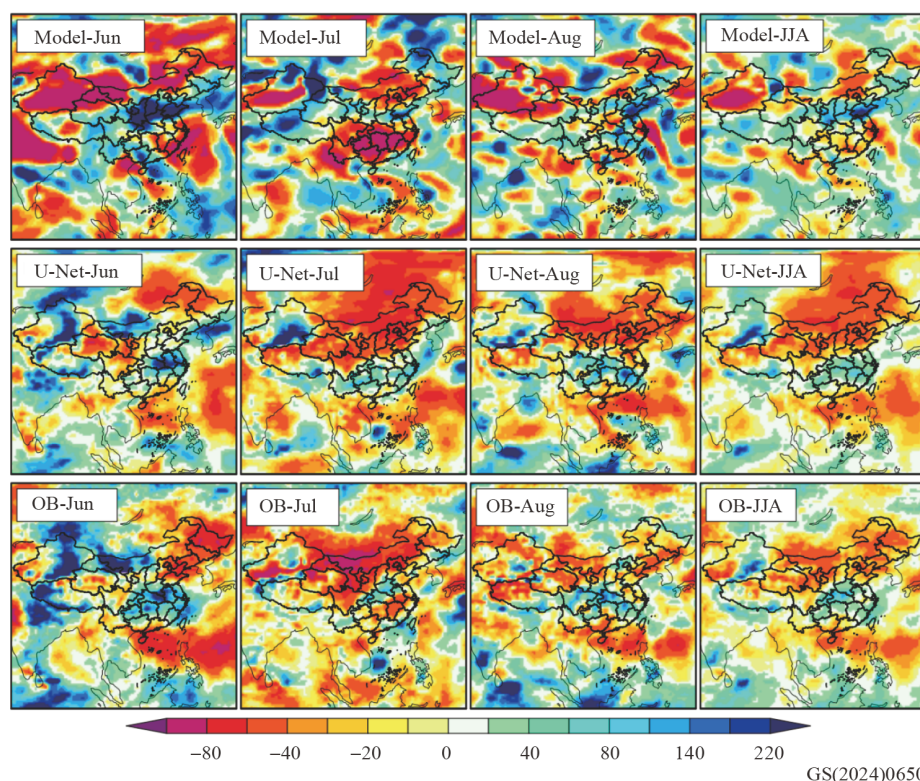


Figure 6. Comparison of anomalous precipitation percentage results (units: %) between the model forecast and corrected results for the rectangular region in 2000. The first row represents the model output results, the second row represents the corrected results, and the third row represents the observed results.

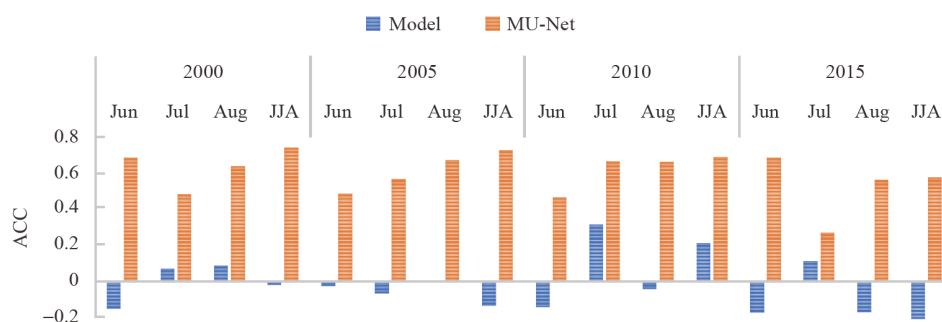


Figure 7. Comparison of the correlation coefficients of anomalous precipitation percentage for June, July, August, and JJA between the model forecast results and the corrected results by the MU-Net model for the years 2000, 2005, 2010, and 2015 in the rectangular region. The blue bars represent the ACC scores of the model forecast results, while the orange bars represent the ACC scores of the results after correction by the MU-Net model.

case, in which the shading regions represent the state consistency between the observation and the prediction. It's found that MU-Net corrected precipitation shows more shading regions, indicating better consistency with the observed drought and flood state regarding spatial distributions. Especially in the Northeast, Inner Mongolia, and Xinjiang regions, the consistency of the drought and flood states in the original model prediction was poor, while the precipitation state showed better consistency with the observation after correction by the MU-Net. Quantificationally, the shading regions accounted for an average of approximately 33.7% in the original model prediction, while the shading regions increase up to 76.2% in the MU-Net corrected model in

average.

3.7 Comparison of dominant EOF modes

We also conducted an evaluation of the dominant EOF modes of precipitation variability, originally predicted by FGOALS and corrected by MU-Net, respectively. We focused on the first four EOF modes, as these are traditionally considered to account for the most variance and are more predictable (Wang et al. 2021). Specifically, we projected the observed, original predicted, and MU-Net-corrected summer precipitation anomalies onto the first four observed EOF modes. This approach allowed us to obtain the summer-to-summer principal components for each dataset, where the closer the predicted principal component values are to the observed

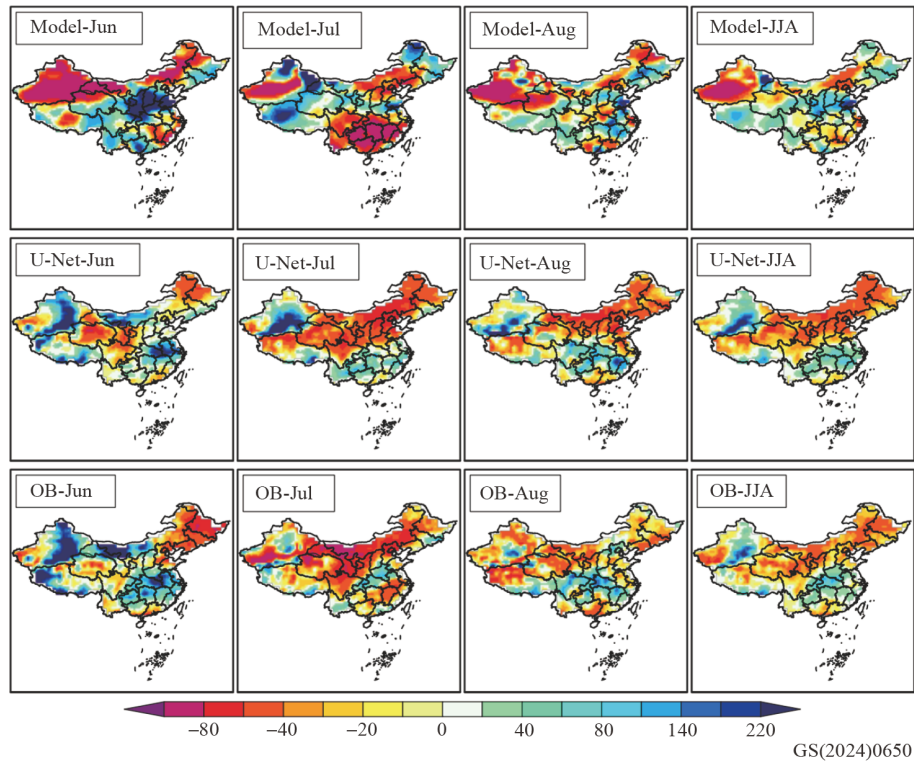


Figure 8. A comparison of the anomalous precipitation percentage results (units: %) between forecast and corrected results over the Chinese mainland in 2000. The first row shows the model output results, the second row shows the corrected results, and the third row shows the observational results.

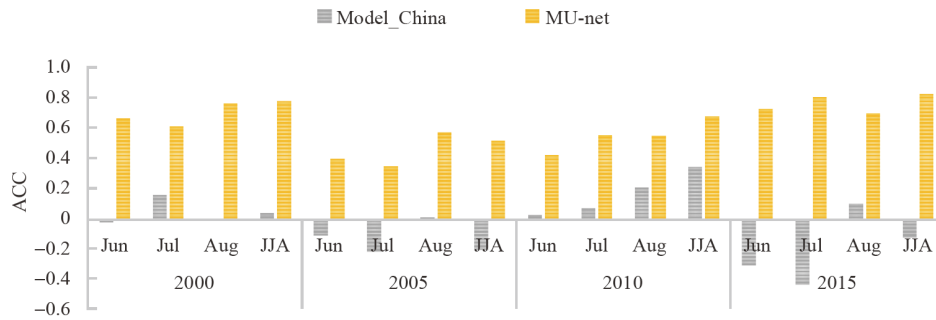


Figure 9. Comparison of anomalous precipitation percentage correlation coefficients between model forecast results and corrected results for June, July, August, and JJA over the Chinese mainland for 2000, 2005, 2010, and 2015. There is a significant increase in the precipitation anomaly percentage correlation coefficients after the correction.

values, the more accurate the prediction of precipitation variability.

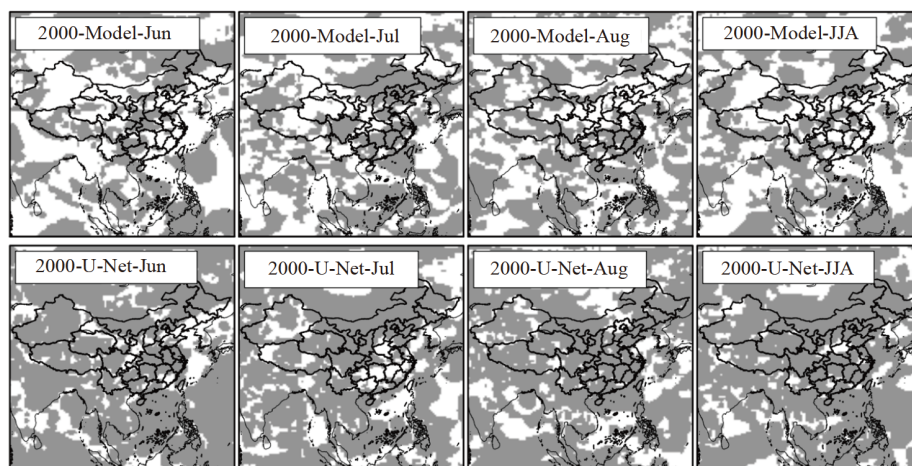
As shown in Fig. 11, on the independent test set, 6 of the original FGOALS-predicted principal components were completely opposite to the observed ones (i.e., PC1 in 2000 and 2015, PC2 in 2005, PC3 in 2000, and PC4 in 2000 and 2005). However, after correction with the MU-Net model, the principal components of the predicted precipitation variability were not only correct but also closely matched the magnitude of the observations. This result indicates that the U-Net method can effectively predict the dominant modes of precipitation variability.

3.8 Comparison of different correction methods

Although our MU-Net method significantly improves the seasonal precipitation prediction skill of the FGOALS-

f2 model, other correction methods can also yield relatively good results when the climate model's predictions are already accurate. To demonstrate the superiority of the proposed method, we compared the results of the MU-Net method introduced in this work, the classic U-Net method and traditional correction methods.

As shown in Table 4, the optimized MU-Net model consistently produces lower RMSE values compared to the classical U-Net model on the independent test sets, both in monthly prediction (June, July, and August) and for the entire summer season. This demonstrates that the enhancements in the MU-Net model result in more accurate predictions. We believe the primary reason is the addition of a penalty term in the loss function for deviations from climate model predictions, which



GS(2024)0650

Figure 10. Comparison of drought and flood states between model forecast results and corrected results for 2000 in the rectangular region. The shading regions indicate consistency with the observed drought and flood states, while the white regions indicate inconsistency with the observed drought and flood states.

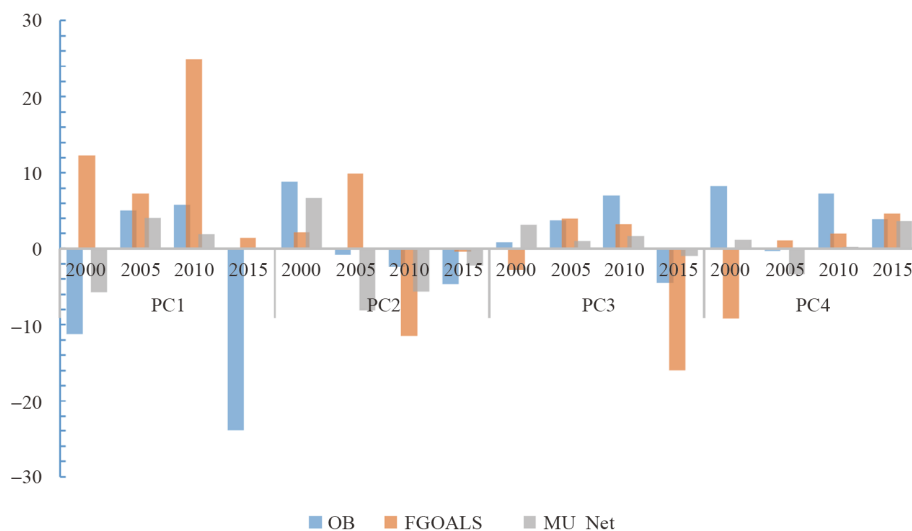


Figure 11. Comparison of the first four principal components of summer precipitation anomalies from observations, original FGOALS predictions, and MU-Net corrected predictions on the independent test sets. The blue bars represent the observed principal components, the orange bars represent the original FGOALS-predicted principal components, and the gray bars represent the MU-Net-corrected principal components.

Table 4. RMSE scores (units: mm month^{-1}) of different methods on the independent test sets. MU-net represents the method used in this work, U-Net represents the classic U-Net architecture, and SVD represents a traditional statistical correction method.

Month	CAS-FGOALS	MU-Net	U-Net	SVD
June	124.6	71.2	88.9	105.2
July	144.0	80.3	95.4	112.3
August	132.4	75.3	92.5	110.0
JJA	133.9	75.6	92.3	109.2

effectively alleviated the overfitting issue. This penalty term restricts the model's predictions from deviating from known climate patterns, thereby enhancing the model's generalization ability. Additionally, the optimized MU-Net model demonstrates a more stable training process, converging faster under the same parameter settings, whereas the classic U-Net model experiences unstable

oscillations during training.

Regarding the comparison with traditional statistical correction methods such as SVD, the results of SVD were significantly worse. The primary reason for this is that traditional statistical correction methods rely mainly on correcting systematic biases in the predicted variable itself (e.g., precipitation) without incorporating additional

physical constraint information. These methods do not leverage the higher predictive skill of atmospheric circulation fields or account for the nonlinear relationships between circulation and precipitation. In contrast, the MU-Net-based correction approach capitalizes on the higher predictive skill of circulation fields by learning the nonlinear constraints between circulation and precipitation. This allows the high predictive skill of circulation to be transferred to precipitation prediction, resulting in better overall performance. Moreover, this physically constrained correction method not only yields superior results but also has greater potential for further improving precipitation prediction accuracy.

4 CONCLUSIONS AND DISCUSSION

In this study, we developed a MU-Net-based method to correct the bias of seasonal prediction of summer rainfall in China. With this end-to-end model, the precipitation distribution can be obtained directly without using the traditional method of data dimensionality reduction, which could maximize the retention of spatiotemporal information of the input data. To address the issue of insufficient data for training seasonal-scale models, we employed a slide-window strategy for data augmentation, increasing the dataset size to alleviate potential overfitting problems due to data scarcity. Furthermore, we optimized the loss function by incorporating the difference between the corrected results from the MU-Net model and the original model prediction as a penalty term, which helped improve the convergence speed of the model and mitigate overfitting issues (Table 2).

We trained multiple MU-Net models with different sets of hyperparameters and selected the model with the best performance on the validation set for evaluation on the test set. The evaluation included assessments of monthly total precipitation, anomalous precipitation percentages, and precipitation drought-flood states. The results reveal that using the MU-Net model for correcting original model prediction could efficiently improve the seasonal prediction skill over the China summer rainfall.

The MU-Net model corrects bias by capturing the relationship between model outputs and observed precipitation, the constraint relationship between other physical quantities and total rainfall, and the influence of surrounding grid points on the target grid point. Further research is needed to determine how the MU-Net model improves model seasonal prediction skills and what systematic errors exist in the model. Seasonal forecasts are often influenced by global weather and climate factors. Therefore, it is essential to consider the global situation when studying the summer rainfall of China. Traditional computer vision algorithms cannot effectively capture the local correlations on Earth and the interactions between various atmospheric pressure

layers. Models based on transformer can address this issue effectively; further research will explore this aspect in more detail.

REFERENCES

- Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–Present). *J. Hydrometeorology*, **4**, 1147–1167, [http://doi.org/10.1175/1525-7541\(2003\)004<1147:tvvgpcp>2.0.co;2](http://doi.org/10.1175/1525-7541(2003)004<1147:tvvgpcp>2.0.co;2).
- Alessandri, A., M. D. Felice, F. Catalano, J.-Y. Lee, B. Wang, D. Y. Lee, J.-H. Yoo, and A. Weisheimer, 2017: Grand European and Asian-Pacific multi-model seasonal forecasts: maximization of skill and of potential economical value to end-users. *Climate Dyn.*, **50**, 2719–2738, <http://doi.org/10.1007/s00382-017-3766-y>.
- Badr, H. S., B. F. Zaitchik, and S. D. Guikema, 2014: Application of statistical models to the prediction of seasonal rainfall anomalies over the Sahel. *J. Appl. Meteor. Climatol.*, **53**, 614–636, <http://doi.org/10.1175/jamc-d-13-0181.1>.
- Bao, Q., X. Wu, J. Li, L. Wang, B. He, X. Wang, Y. Liu, and G. Wu, 2018: Outlook for El Niño and the Indian Ocean Dipole in autumn-winter 2018–2019. *Chin. Sci. Bull.*, **64**, 73–78, <http://doi.org/10.1360/n972018-00913>. (in Chinese)
- Barreiro, M., and P. Chang, 2004: A linear tendency correction technique for improving seasonal prediction of SST. *Geophys. Res. Lett.*, **31**, L23209, <http://doi.org/10.1029/2004gl021148>.
- Bett, P. E., and Coauthors, 2018: Seasonal forecasts of the summer 2016 Yangtze River Basin rainfall. *Adv. Atmos. Sci.*, **35**, 918–926, <http://doi.org/10.1007/s00376-018-7210-y>.
- Bi, K., L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, 2023: Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, **619**, 533–538, <http://doi.org/10.1038/s41586-023-06185-3>.
- Bougeault, P., and Coauthors, 2001: The MAP special observing period. *Bull. Amer. Meteor. Soc.*, **82**, 433–462, [http://doi.org/10.1175/1520-0477\(2001\)082<0433:tmsop>2.3.co;2](http://doi.org/10.1175/1520-0477(2001)082<0433:tmsop>2.3.co;2).
- Colette, A., R. Vautard, and M. Vrac, 2012: Regional climate downscaling with prior statistical correction of the global climate forcing. *Geophys. Res. Lett.*, **39**, L13707, <http://doi.org/10.1029/2012gl052258>.
- Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. Climate change 2007: The physical science basis. *Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR)*, Cambridge University Press, 589–662.
- Duboue, P., 2020: *The Art of Feature Engineering: Essentials for Machine Learning*, Cambridge University Press, 1–274.
- Fan, K., 2006: Atmospheric circulation in Southern Hemisphere and summer rainfall over Yangtze River Valley. *Chin. J. Geophys.*, **49**, 599–606, <http://doi.org/10.1002/cjg2.873>. (in Chinese)
- Feng, G.-L., and Coauthors, 2020: Improved prediction model for flood-season rainfall based on a nonlinear dynamics-statistic combined method. *Chaos, Solitons & Fractals*, **140**, 110160, <http://doi.org/10.1016/j.chaos.2020.110160>.
- Finnis, J., W. W. Hsieh, H. Lin, and W. J. Merryfield, 2012: Non-linear post-processing of numerical seasonal climate forecasts. *Atmos.-Ocean*, **50**, 207–218, <http://doi.org/10.1080/07055900.2012.667388>.
- Gholami, S., S. Ghader, H. Khaleghi-Zavareh, and P. Ghafarian, 2021: Sensitivity of WRF-simulated 10 m wind over the Persian Gulf to different boundary conditions and PBL parameterization

- schemes. *Atmos. Res.*, **247**, 105147, <http://doi.org/10.1016/j.atmosres.2020.105147>.
- Giorgi, F., C. Jones, and G. Asrar, 2008: Addressing climate information needs at the regional level: the CORDEX framework. *WMO Bull.*, **53**, 175–183.
- Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal to interannual climate predictions. *Int. J. Climatol.*, **21**, 1111–1152, <http://doi.org/10.1002/joc.636>.
- Gong, D. Y., and C. H. Ho, 2002: Shift in the summer rainfall over the Yangtze River valley in the late 1970s. *Geophys. Res. Lett.*, **29**, 78–71–78–74, <http://doi.org/10.1029/2001gl014523>.
- Han, L., M. Chen, K. Chen, H. Chen, Y. Zhang, B. Lu, L. Song, and R. Qin, 2021: A deep learning method for bias correction of ECMWF 24–240 h forecasts. *Adv. Atmos. Sci.*, **38**, 1444–1459, <http://doi.org/10.1007/s00376-021-0215-y>.
- He, B., and Coauthors, 2019: CAS FGOALS-f3-L model datasets for CMIP6 historical atmospheric model intercomparison project simulation. *Adv. Atmos. Sci.*, **36**, 771–778, <http://doi.org/10.1007/s00376-019-9027-8>.
- He, K., X. Zhang, S. Ren, and J. Sun, 2016: Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, Institute of Electrical and Electronics Engineers, 770–778, <https://doi.org/10.1109/CVPR.2016.90>.
- Kim, H., Y. G. Ham, Y. S. Joo, and S. W. Son, 2021: Deep learning for bias correction of MJO prediction. *Nat. Commun.*, **12**, 3087, <http://doi.org/10.1038/s41467-021-23406-3>.
- Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2017: ImageNet classification with deep convolutional neural networks. *Commun. ACM*, **60**, 84–90, <http://doi.org/10.1145/3065386>.
- Lam, R., and Coauthors, 2023: Learning skillful medium-range global weather forecasting. *Science*, **382**, 1416–1421, <http://doi.org/10.1126/science.adi2336>.
- Lealdi, D., S. Nurdianti, and A. Sopaheluwakan, 2018: Statistical bias correction modelling for seasonal rainfall forecast for the case of Bali island. *J. Physics: Conference Series*, **1008**, 012018, <http://doi.org/10.1088/1742-6596/1008/1/012018>.
- LeCun, Y., Y. Bengio, and G. Hinton, 2015: Deep learning. *Nature*, **521**, 436–444, <http://doi.org/10.1038/nature14539>.
- Lee, J.-Y., S.-S. Lee, B. Wang, K.-J. Ha, and J.-G. Jhun, 2012: Seasonal prediction and predictability of the Asian winter temperature variability. *Climate Dyn.*, **41**, 573–587, <http://doi.org/10.1007/s00382-012-1588-5>.
- Li, C., and Coauthors, 2016: Skillful seasonal prediction of Yangtze river valley summer rainfall. *Environ. Res. Lett.*, **11**, 094002, <http://doi.org/10.1088/1748-9326/11/9/094002>.
- Lim, Y., J. Lee, H.-S. Oh, and H.-S. Kang, 2014: Independent component regression for seasonal climate prediction: an efficient way to improve multimodel ensembles. *Theor. Appl. Climatol.*, **119**, 433–441, <http://doi.org/10.1007/s00704-014-1099-x>.
- Ling, F., Y. Li, J.-J. Luo, X. Zhong, and Z. Wang, 2022: Two deep learning-based bias-correction pathways improve summer precipitation prediction over China. *Environ. Res. Lett.*, **17**, 124025, <http://doi.org/10.1088/1748-9326/aca68a>.
- Liu, Y., J. Zhang, L. Chen, H. Chu, J. Z. Wang, and L. Ma, 2022: SSAS: spatiotemporal scale adaptive selection for improving bias correction on precipitation. *IEEE Trans. Cybern.*, **52**, 12175–12188, <http://doi.org/10.1109/tcyb.2021.3072483>.
- Lyu, Y., S. Zhu, X. Zhi, Y. Ji, Y. Fan, and F. Dong, 2023: Improving subseasonal - to - seasonal prediction of summer extreme precipitation over Southern China based on a deep learning method. *Geophys. Res. Lett.*, **50**, e2023GL106245, <http://doi.org/10.1029/2023gl106245>.
- Nan, S., and J. Li, 2003: The relationship between the summer precipitation in the Yangtze River valley and the boreal spring Southern Hemisphere annular mode. *Geophys. Res. Lett.*, **30**, 2266, <http://doi.org/10.1029/2003gl018381>.
- Ronneberger, O., P. Fischer, and T. Brox, 2015: U-Net: convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th international conference*, Munich, Germany, Springer, 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.
- Shen, B., Z. Lin, R. Lu, and Y. Lian, 2011: Circulation anomalies associated with interannual variation of early- and late-summer precipitation in Northeast China. *Sci. Chin. Earth Sci.*, **54**, 1095–1104, <http://doi.org/10.1007/s11430-011-4173-6>.
- Tong, X., W. Zhou, and J. Xia, 2024: Improving boreal summer precipitation predictions from the Global NMME through Res34 - Unet. *Geophys. Res. Lett.*, **51**, e2023GL106391, <http://doi.org/10.1029/2023gl106391>.
- Wang, B., and Coauthors, 2008: Advance and prospectus of seasonal prediction: assessment of the APCC/CLIPAS 14-model ensemble retrospective seasonal prediction (1980–2004). *Climate Dyn.*, **33**, 93–117, <http://doi.org/10.1007/s00382-008-0460-0>.
- Wang, J., J. Yang, H.-L. Ren, J. Li, Q. Bao, and M. Gao, 2021: Dynamical and machine learning hybrid seasonal prediction of summer rainfall in China. *J. Meteor. Res.*, **35**, 583–593, <http://doi.org/10.1007/s13351-021-0185-0>.
- Xu, L., N. Chen, X. Zhang, Z. Chen, C. Hu, and C. Wang, 2019: Improving the North American multi-model ensemble (NMME) precipitation forecasts at local areas using wavelet and machine learning. *Climate Dyn.*, **53**, 601–615, <http://doi.org/10.1007/s00382-018-04605-z>.
- Yang, S., Z. Zhang, V. E. Kousky, R. W. Higgins, S.-H. Yoo, J. Liang, and Y. Fan, 2008: Simulations and seasonal prediction of the Asian Summer Monsoon in the NCEP climate forecast system. *J. Climate*, **21**, 3755–3775, <http://doi.org/10.1175/2008jcli1961.1>.
- Ye, K.-H., C.-Y. Tam, W. Zhou, and S.-J. Sohn, 2015: Seasonal prediction of June rainfall over South China: model assessment and statistical downscaling. *Adv. Atmos. Sci.*, **32**, 680–689, <http://doi.org/10.1007/s00376-014-4047-x>.
- Yin, Z., and H. Wang, 2016: Seasonal prediction of winter haze days in the north central North China Plain. *Atmos. Chem. Phys.*, **16**, 14843–14852, <http://doi.org/10.5194/acp-16-14843-2016>.
- Zhang, Y., M. Long, K. Chen, L. Xing, R. Jin, M. I. Jordan, and J. Wang, 2023: Skillful nowcasting of extreme precipitation with NowcastNet. *Nature*, **619**, 526–532, <http://doi.org/10.1038/s41586-023-06184-4>.
- Zhu, Z., T. Li, and J. He, 2014: Out-of-phase relationship between boreal spring and summer decadal rainfall changes in Southern China. *J. Climate*, **27**, 1083–1099, <http://doi.org/10.1175/jcli-d-13-00180.1>.