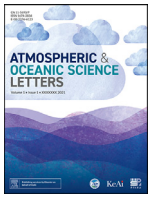
Atmospheric and Oceanic Science Letters xxx (xxxx) xxx

KeAi CHINESE ROOTS GLOBAL IMPACT

Contents lists available at ScienceDirect

Atmospheric and Oceanic Science Letters

journal homepage: http://www.keaipublishing.com/en/journals/atmospheric-and-oceanic-science-letters/



U-Net: A deep-learning method for improving summer precipitation forecasts in China

Qimin Deng [a,*], Peirong Lu [a], Shuyun Zhao [a], Naiming Yuan [b,*]

[a] Department of Atmospheric Science, CMA-CUG Joint Centre for Severe Weather and Climate and Hydro-geological Hazards, China University of Geosciences, Wuhan, China
[b] School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, China

ARTICLE INFO

Keywords:
Summer precipitation
U-Net
Subseasonal forecast
Deep learning
关键词:
汛期降水
U-Net
次季节预报
深度学习

ABSTRACT

A deep-learning method named U-Net was applied to improve the skill in forecasting summer (June–August) precipitation for at a one-month lead during the period 1981–2020 in China. The variables of geopotential height, soil moisture, sea level pressure, sea surface temperature, ocean salinity, and snow were considered as the model input to revise the seasonal prediction of the Climate Forecast System, version 2 (CFSv2). Results showed that on average U-Net reduced the root-mean-square error of the original CFSv2 prediction by 49.7% and 42.7% for the validation and testing set, respectively. The most improved areas were Northwest, Southwest, and Southeast China. The anomaly same sign percentages and temporal and spatial correlation coefficients did not present significant improvement but maintained the comparable performances of CFSv2. Sensitivity experiments showed that soil moisture is the most crucial factor in predicting summer rainfall in China, followed by geopotential height. Due to its advantages in handling small training dataset sizes, U-Net is a promising deep-learning method for seasonal rainfall prediction.

CHINESE ABSTRACT

本研究应用了名为U-Net的深度学习方法来提高中国夏季 (6–8月) 降水的预报技能, 预报时段为1981–2020年, 预报提前期为一个月. 将位势高度场, 土壤湿度, 海平面气压, 海表面温度, 海洋盐度和青藏高原积雪等变量作为模型输入, 本文对美国NCAR气候预报系统第2版 (CFSv2) 的季节性预报结果进行了修正. 结果显示, 在验证集和测试集上, U-Net平均将原CFSv2预测的均方根误差分别减少了49.7%和42.7%. 预报结果改善最大的地区是中国的西北,西南和东南地区. 然而, 同号率和时空相关系数没有得到明显改善, 但仍与CFSv2的预测技巧持平. 敏感性实验表明, 土壤湿度是预测中国夏季降雨的最关键因素, 其次是位势高度场. 本研究显示了U-Net模型在训练小样本数据集方面的优势, 为我国汛期季节性降雨预测提供了一种有效的深度学习方法.

1. Introduction

As pointed out in the IPCC's 6th Assessment Report (IPCC, 2021), global warming and climate change have increased the frequency and intensity of extreme events such as heavy rainfall and drought (Westra et al., 2013; Donat et al., 2016). Flood and drought disasters not only lead to huge socioeconomic and agricultural losses but also pose a risk to human survival (e.g., Zhang et al., 2022). Hence, forecasting summer rainfall in advance is important for disaster prevention and reduction.

Current summer precipitation prediction methods can be broadly classified into statistical, dynamic, and dynamic–statistical methods. Statistical prediction methods use seasonal climate predictors such as sea surface temperature (SST) and large-scale atmospheric and oceanic oscillation indices to build a regression function (Wu et al., 2013; Wu and Yu, 2016; Xing et al., 2016). Powered by advanced computing technology, a variety of coupled general circulation models (CGCMs) have become the most important and practical tool for precipitation prediction. However, on account of the low potential predictability of the East Asian summer monsoon and the East Asian winter monsoon in China, precipitation prediction is less skillful than in other places and

* Corresponding authors.
E-mail addresses: dengqimin@cug.edu.cn (Q. Deng), yuannm@mail.sysu.edu.cn (N. Yuan).

https://doi.org/10.1016/j.aosl.2022.100322
Received 1 November 2022; Revised 23 November 2022; Accepted 15 December 2022
Available online xxx

Please cite this article as: Q. Deng, P. Lu, S. Zhao et al., U-Net: A deep-learning method for improving summer precipitation forecasts in China, Atmospheric and Oceanic Science Letters, https://doi.org/10.1016/j.aosl.2022.100322

for other variables (Lang et al., 2014; Liu et al., 2019). Therefore, some studies have used statistical methods to revise the prediction output of CGCMs and achieve better results (Tung et al., 2013; Fu et al., 2022). Deep learning, as a nonlinear model, can offer a potential approach for combining dynamic model and statistical methods and revising the predictions of CGCMs (Reichstein et al., 2019; Jin et al., 2022; Pan et al., 2021).

Deep learning is a multilayer model that can automatically extract and learn the nonlinear features by transforming the lower-level features into higher ones (Lecun et al., 2015). In recent years, a growing number of studies have applied deep neural networks to climate prediction (Ham et al., 2019; Zhou et al., 2019 Pan et al., 2021), nowcasting precipitation prediction (Shi et al., 2015), recognition of extreme events (Racah et al., 2017), and replacing the traditional sub-grid parameterizations in numerical models (Irrgang et al., 2021). On the seasonal timescale, however, the observational data are normally not long enough to train a deep-learning network (Jin et al., 2022). Therefore, many researchers have utilized model simulations to pre-train the model, followed by fine-tuning with observations (Ham et al., 2019; Jin et al., 2022). Among the deep-learning models, the Convolutional Neural Network (CNN) model, benefiting from its image-convolutional structure, is highly suitable for inputting gridded climate fields and learning spatial correlations in climate variables.

In this study, we applied one of the CNN models named U-Net (Ronncberger et al., 2015) to predict summer rainfall in China. U-Net is well-known for its powerful ability in image segmentation and small-sample-size training (Du et al., 2020). The structure of U-Net is composed of down-sampling and up-sampling layers. The images are dimensionally reduced to a fixed tiny size to extract deeper-layer features. Then, up-sampling is used to enlarge the image, and each same layer feature of down-sampling and up-sampling is merged. Hence, the final convolutional layer can learn to assemble a more accurate output from this information. Compared with the original CNN, the large number of feature channels in U-Net allows for the propagation from shallow information to the higher layers through a relatively small size of the training set. Therefore, we chose the U-Net model as a post-processing method to develop a combined dynamic–statistical method. More specifically, we revised the subseasonal forecast output of the National Centers for Environmental Prediction (NCEP) Climate Forecast System, version 2 (CFSv2; Saha et al., 2014), a high-performance forecast model, and tried to achieve improvements.

The paper is organized as follows: Section 2 introduces the data needed for model training and forecasting; Section 3 describes the U-Net model design, training details, and model evaluation criteria; Section 4 summarizes the forecasting results; and Section 5 presents our conclusions and some further discussion.

2. Data and methods

2.1. Data

(1) Monthly mean gridded precipitation observation data from the CN05.1 dataset (Wu and Gao, 2013) provided by the National Climate Center (http://data.cma.cn/). This dataset is interpolated from the daily observation data of more than 2400 national stations of the National Meteorological Information Center (excluding data from Taiwan) with a horizontal resolution of 0.25°. It served as the training target and, consequently, the prediction output had the same resolution.

(2) Precipitation prediction data from NCEP CFSv2 (Saha et al., 2014).

(3) ERA5 monthly reanalysis data on a single level from ECMWF (https://cds.climate.copernicus.eu/): geopotential height; soil moisture; SST;, sea level pressure; snow depth; snow temperature; and snow albedo (Hersbach et al., 2020).

(4) UK Met Office Hadley Centre EN4.2.2 monthly sea surface salinity data (Good et al., 2013).

Table 1 Variables used as the model input.

Area	Variable	Lead time
China	Model precipitation	2 months
	500/850/1000 hPa geopotential	1–3 months
	0–1 m depth soil moisture	1–3 months
North Pacific and	Sea level pressure	Pre-winter, spring
North Indian	Sea surface temperature	Pre-winter, spring
Ocean	Ocean salinity at 5 m depth	Spring
Tibetan Plateau	Snow depth	Pre-winter, spring
	Snow temperature	Pre-winter, spring
	Snow albedo	Pre-winter, spring

A discussion regarding the selection of predictors is provided in the supplementary material.

2.2. Data pre-processing

The CN05.1 observation data during June to August (summer) were employed in this study. The rainfall prediction of CFSv2 at a lead time of two months was used to forecast the summer precipitation, so data from April to June 1982–2020 were adopted. The one-month predictions of CFSv2 were ensemble-averaged to serve as one predictor of the inputs. The monthly average precipitation data of CN05.1 for 30 years from 1991 to 2020 were used as the climatological mean precipitation, and the observation and dynamic model output data were processed as the anomaly percentage. The precipitation anomaly percentage (PP) was calculated as follows:

$$\mathrm{PP} = \frac{p_{1,2} - m}{m} \times 100(\%), \tag{1}$$

where $p_{1,2}$ denotes the observed precipitation and model output and m is the climatological mean precipitation. Using a different climatological mean in Eq. (1) barely changed the results.

The time period, area, and lead time of the variables are listed in Table 1. All the predictors were linearly normalized before inputting them into the model, and the null value areas were padded with 0.

2.3. Methods

2.3.1. Model design

The main structure of the model was the U-Net model (Ronncberger et al., 2015), which is a deep neural network with a fully symmetric encoding–decoding structure. The encoder is used for feature compression and the decoder is used to reconstruct features. Both model inputs and outputs are from a single time point, so there is no time memory in the prediction. The original U-Net has four down-sampling and four up-sampling blocks. To reduce the model complexity and overfitting, in this study, the U-Net model contained three levels (Fig. 1). Details are provided in the supplementary material.

2.3.2. Training details

We divided the coupled dataset into three parts: June–August 1982–2006 (75 months) as the training set; June–August 2007–2013 (21 months) as the validation set; and June–August 2014–2020 (21 months) as the testing set. The training set was used to iteratively optimize the parameters of the model. The validation set was used to verify the effectiveness of the completed training parameters and was not used in the training process. The testing set was used to test the real performance in the future, and was not used or checked in any previous steps.

To reduce the influence of extreme values in the precipitation forecasts of CFSv2 on the model, those values exceeding the 95% quantile were converted to the 95% quantile. Considering the regional differences in precipitation features over China, the entire space was split into four regions (region 1: <200 mm; region 2: 200–400 mm; region 3: 400–800 mm; region 4: > 800 mm) according to the climatological mean state's total annual precipitation. The forecast models for the four

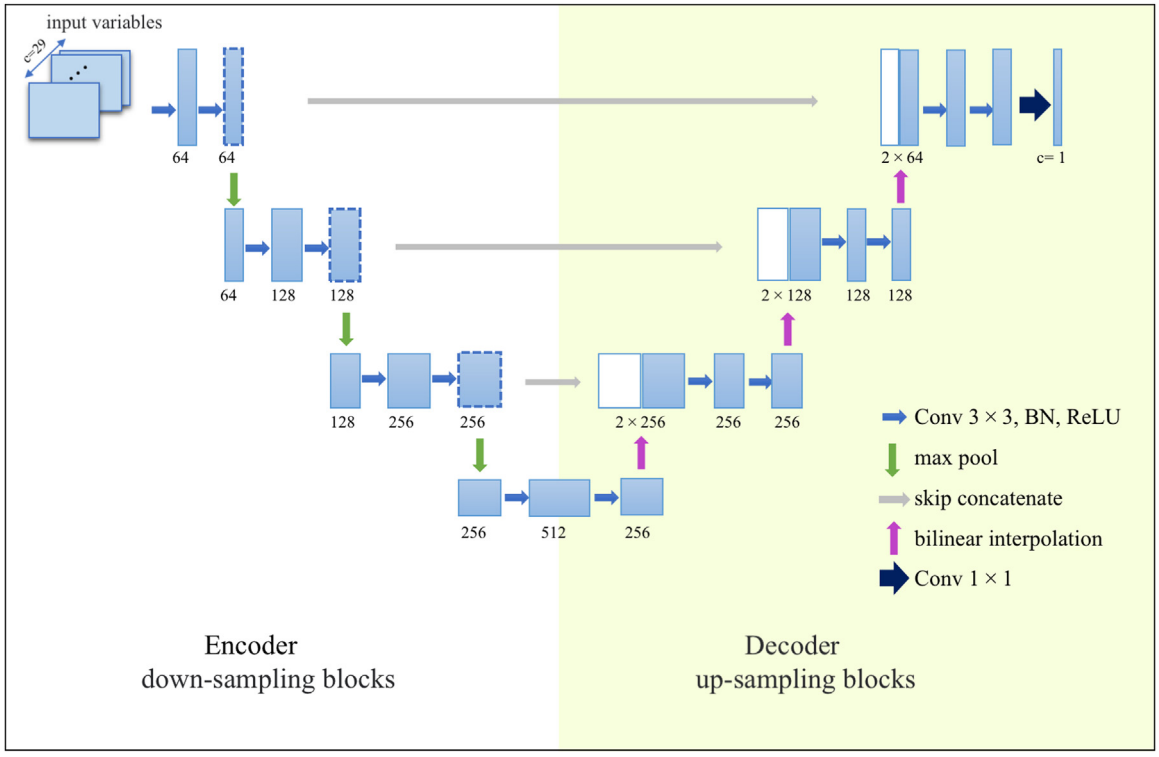


Fig. 1. U-Net model structure. The numbers in the figure denote the number of channels in each step, and the dashed boxes denote replication of the one to the left. As the model goes deeper, the resolution of the image reduces, and the features extracted by convolution double.

regions were trained separately. To expand the sample size, image mirroring (left–right and up–down mirroring) and 180° rotation were conducted as data augmentation. The sample size of the training set was hence increased to 75 × 4 = 300.

2.3.3. Evaluation indices

To assess the forecast skill as comprehensively as possible, four indices were employed: root-mean-square error (RMSE); anomaly same sign (AS); temporal correlation coefficient (TCC); and anomaly correlation coefficient (ACC). The RMSE was used to evaluate the Eulerian distance between predictions and observations:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Y_i - f(x_i)\right)^2}, \tag{2}$$

where N is the total number of samples, and Y_i and $f(x_i)$ denote the observed and predicted values, respectively. AS was calculated to judge the accuracy of the prediction in terms of whether it was positive or negative precipitation compared with the climatological mean:

$$\mathrm{AS} = \frac{N_t}{N} \times 100\%, \tag{3}$$

in which N_t denotes the number of samples where the predicted and observation have the same sign or one of the two is zero. TCC and ACC are designed to evaluate the temporal and spatial correlation, respectively:

$$\mathrm{TCC}_i = \frac{\sum_{j=1}^{M}\left(Y_{i,j} - \overline{Y_i}\right)\left(f(x_{i,j}) - \overline{f(x_i)}\right)}{\sqrt{\sum_{j=1}^{M}\left(Y_{i,j} - \overline{Y_i}\right)^2 \sum_{j=1}^{M}\left(f(x_{i,j}) - \overline{f(x_i)}\right)^2}}, \tag{4}$$

$$\mathrm{ACC}_j = \frac{\sum_{i=1}^{N}\left(Y_{i,j} - \overline{Y_j}\right)\left(f(x_{i,j}) - \overline{f(x_j)}\right)}{\sqrt{\sum_{i=1}^{N}\left(Y_{i,j} - \overline{Y_j}\right)^2 \sum_{i=1}^{N}\left(f(x_{i,j}) - \overline{f(x_j)}\right)^2}}, \tag{5}$$

Table 2 Indices measured during 1982–2006 (training set), 2007–2013 (validation set), and 2014–2020 (testing set).

Dataset	Model	RMSE	TCC	ACC	AS
Training	U-Net	\	\	0.948	0.914
	CFSv2	\	\	0.238	0.477
Validation	U-Net	49.298	0.181	0.108	0.544
	CFSv2	212.008	0.166	0.258	0.476
Testing	U-Net	54.233	0.196	0.181	0.479
	CFSv2	210.035	0.202	0.250	0.493

where i and j denote different spatial grids and time, respectively; $Y_{i,j}$, $\overline{Y_i}$, and $\overline{Y_j}$ are their observed values, their time mean, and spatial mean, respectively; and $f(x_{i,j})$, $\overline{f(x_i)}$, and $\overline{f(x_j)}$ are the predicted values, their time mean, and spatial mean, respectively.

3. Results

3.1. Mean forecast skill assessment

The spatial and temporal mean prediction performances of the original CFSv2 and U-Net on the training, validation, and testing sets are given in Table 2. The samples of the training sets were augmented, so their RMSE and TCC are not shown here. The sample mean ACC and AS of CFSv2 were 0.238 and 0.477, respectively. In contrast, the sample mean ACC and AS of U-Net both improved to above 0.9. However, these two indices of U-Net declined in the validation and testing sets, especially the ACC. Compared to the original CFSv2 prediction, the TCC of the U-Net model in the validation period improved slightly from 0.166 to 0.181, while AS improved markedly from 0.476 to 0.544. As for the testing set, U-Net maintained its superiority over CFSv2 but without significant improvement.

In terms of the RMSE, U-Net exhibited more powerful prediction skill than CFSv2. The values were 49.298 and 54.233 in the validation

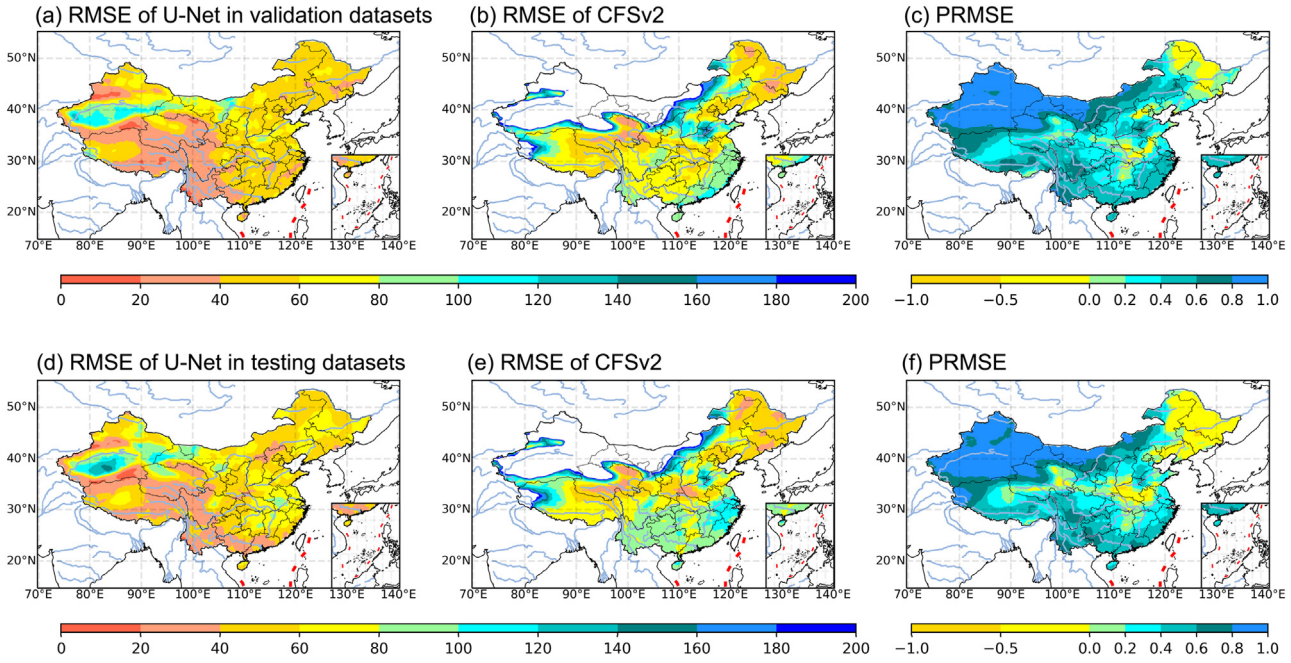


Fig. 2. RMSE and PRMSE of the predicted PP: (a–c) validation set; (d–f) testing set; (a,d) U-Net; (b,e) CFSv2; (c,f) PRMSE. The areas with RMSE exceeding 200 are not colored.

and testing sets, respectively, which were much smaller than the values (212.008 and 210.035) of CFSv2. This is because we used a mean-square error (MSE) loss function (Eq. (6)) to optimize our parameters in the model:

$$L_{\mathrm{MSE}} = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2, \tag{6}$$

where x_i is the predicted value and y_i is the target value. Therefore, the RMSE will be reduced preferably and the results proved the effectiveness of MSE loss.

3.2. Spatial differences in U-Net predictions

To clearly show the difference between the prediction skill of CFSv2 and U-Net, the RMSE difference percentage (PRMSE) was defined as follows:

$$\mathrm{PRMSE} = \frac{r_{\mathrm{CFSv2}} - r_{\mathrm{unet}}}{r_{\mathrm{CFSv2}}}, \tag{7}$$

where r is the RMSE of the predicted precipitation anomaly. If the U-Net forecast skill is better than that of CFSv2, then the PRMSE should be positive.

The RMSE for the validation and testing sets had similar spatial distributions (Fig. 2). The RMSEs of U-Net were all below 40 in Southwest China, and between 20 and 80 in most of eastern China. Only in a few areas was the RMSE above 80 but still below 100. Except for parts of Northeast and Central China, where the PRMSE values were negative, the RMSE of U-Net was reduced by 20%–80% compared with CFSv2, and by more than 80% in Northwest China (Fig. 2(c,f)), in both the validation and testing sets. The mean PRMSEs in the whole area were 49.7% and 42.7% for the validation and testing set, respectively. Generally, the prediction error of the summer precipitation forecast of U-Net was much smaller than that of CFSv2, and the most improved areas were Northwest, Southwest, and Southeast China. Although we considered the precursor signals of snow cover on the Tibetan Plateau and the SST and sea level pressure in the North Pacific and North Indian Ocean, the prediction skill in Northeast China and the areas between the Yangtze and Yellow River basins were not improved.

3.3. Predicted time series in typical regions

Considering the spatial difference and relatively low ACC (Table 2), the time series of the observed PP and that predicted by U-Net and CFSv2 in typical cities are presented and analyzed (Fig. 3).

Firstly, all CFSv2 predictions produced strong annual cycles. In contrast, there was no obvious annual cycle in the observed PP series and U-Net predictions. Since CFSv2 predicts the total precipitation directly and according to Eq. (1), the annual cycle of PP predicted by CFSv2 indicates that the climatological mean of CFSv2 has large deviations from the observed climatological mean. Hence, there are still annual components remaining in the PP of CFSv2. Besides, CFSv2 could not predict the large-scale trend of PP and, in some areas, the mean states were totally wrong compared with the observed PP. On the contrary, the U-Net predictions were closer to the observation, which is consistent with the finding that its RMSE was greatly reduced. U-Net was able to predict more accurately in Guangzhou, Lhasa, Shanghai, and Urumqi, where the intensity or trend of PP was captured very well. For instance, the U-Net predictions in Urumqi matched the observations very well, including the decreasing trend from 2016 to 2020. Although better predictions of U-Net were apparent in their time series, their TCC still presented no obvious priority (Table S1).

4. Discussion and conclusion

4.1. Model outfit and loss functions

What cannot be ignored is the marked difference between the prediction skill of U-Net in the training and validation/testing periods. It indicates that the model suffers from overfitting due to the small sample size, despite the sample extension via rotation and mirroring. Fortunately, compared with CFSv2, the RMSE of the U-Net model was steady and remarkably reduced in all periods, as we expected. The reasons for the poorer performance of the TCC, ACC, and AS can be explained in two aspects: (1) Since the model input and output samples are two-dimensional vectors, the tuning of the model parameters during the training process will be more favorable to the spatial mean accuracy of the whole region and ignore the spatial differences. (2) Since we only used the MSE

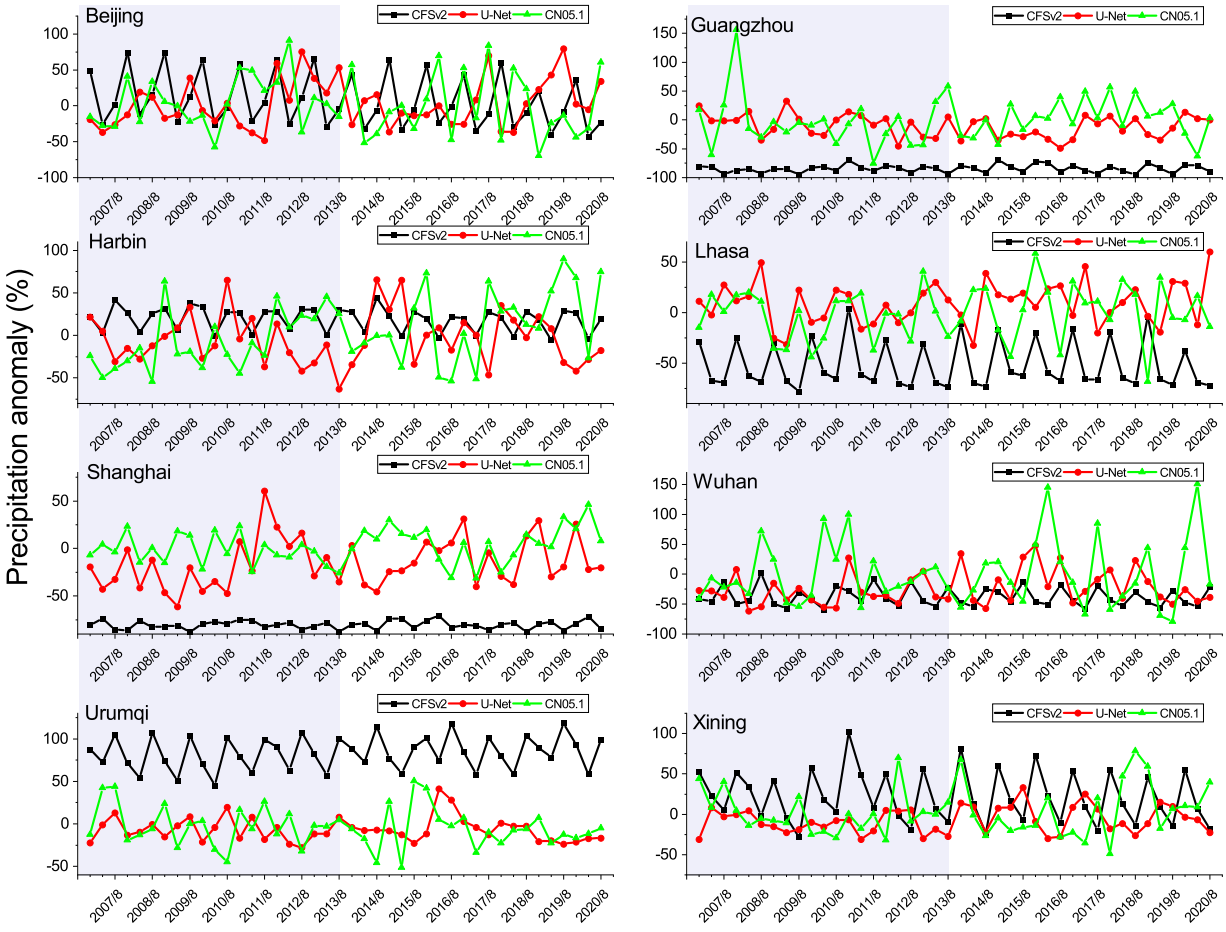


Fig. 3. The PP in typical cities from June to August 2007–2020. The grey panels mark the validation period.

loss function to reduce the Eulerian distances between predictions and observations, there are no controls on the spatial and temporal correlations. To solve this problem and further improve the predictions, we can add other loss functions. For instance, a binary cross entropy (BCE) loss function

$$L_{\mathrm{BCE}} = -\sum_{i=1}^{N}\left[y_i \ln\left(\delta\left(x_i\right)\right) + \left(1-y_i\right)\ln\left(1-\delta\left(x_i\right)\right)\right], \tag{8}$$

where x_i is the predicted value and y_i is the target value, was designed to solve a classification problem for two categories (Ruby and Yendapalli, 2020), and we can preprocess the precipitation anomaly into positive or negative states, and hence using BCE loss should improve the AS. Also, the mean absolute error (MAE) loss function,

$$L_{\mathrm{MAE}} = \frac{1}{N}\sum_{i=1}^{N}\left|x_i - y_i\right|, \tag{9}$$

can increase the TCC to some extent. How to develop a loss function concerning multiple evaluation indices will be further investigated in the future.

The other shortcoming of U-Net mainly comes from its requirement of balanced samples. Therefore, we need a balanced sampling technique to ensure equal distribution of all kinds of precipitation in the training set, which means this method is more suitable in predicting the level of precipitation rather than the precise amount. This can also explain why the U-Net model performs better in southern China and western China; in those places, precipitation samples are more balanced. The more severe the precipitation varies and more extreme events happen, the harder it is for the model to learn the mechanism.

Despite these issues, U-net still shows superiority in reducing overfitting compared with the traditional back propagation network (or so-called fully connected network) and CNN, especially in the case of small sample sizes, such as only 60 years of observation for summer rainfall in China. Under such a small training size, other machine-learning methods can barely reach the equivalent level of RMSE in U-Net. Certainly, transferring learning using model simulations can solve parts of these problems, but U-Net can greatly ease the process and save time and computing resources. In addition, the U-Net model is quite robust for the number of hidden layers and kernel sizes of convolution layers. Even so, we suggest to not use too many layers (fewer than five layers being best) when the sample size is small, and the minimum sample size suggested for a reliable forecast might be 180 (15 years). The results of control experiments that supported the above inferences are given in Table S2.

4.2. Sensitivity experiments

In order to find the key predictors, we set up six sensitivity experiments with different input variables to determine the contributions of different predictors to summer precipitation predictions. The prediction skill indices were calculated and the results are presented in Fig. 4, in which the x axis show the variables that were not included in the experiments. The control run was exactly the main experiment in this study. It shows that the RMSE increase significantly in the experiment in which soil moisture was removed. In other words, when soil moisture was not selected as one of the model inputs, the model performance declined. The next most important factor was the geopotential height. On the one hand, soil moisture can "memorize" the precipitation of the last few months through rainfall infiltration, while on the other hand it

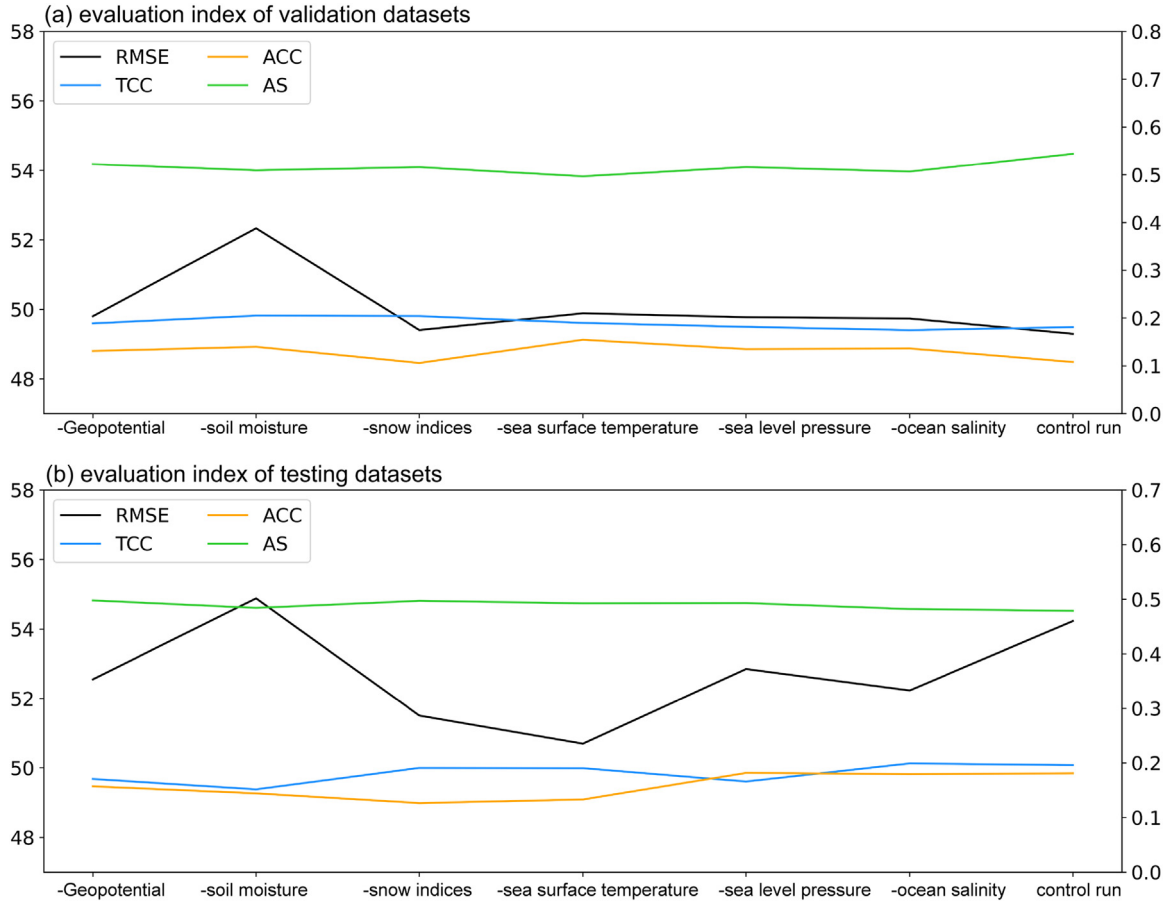


Fig. 4. Prediction skills of the six sensitivity experiments and the control group in the (a) validation and (b) testing sets.

can modulate the evaporation and other surface energy fluxes through a feedback loop (Koster et al., 2004; Asharaf et al., 2012). We suggest that using the precipitation of the last few months may result in a similar performance owing to the importance of considering sequential memory in the time series prediction. However, this needs to be further verified. It should be noted that the validation set, not the testing set, was used to determine the model. The performance of the testing set should be completely unknown before applying the finished trained model. Therefore, although the experiment in which the SST was removed performed better than the control run in the testing set, it cannot be the basis to determine the model.

5. Conclusion

In this study, we applied U-Net, a novel deep-learning model, in the forecasting of summer precipitation. The model was effective in reducing the RMSE for predicting precipitation anomalies in almost all areas of China, compared with the predictions by CFSv2. The spatial and time mean RMSEs of U-Net in the period 2007–2013 and 2014–2020 decreased from 212.008 and 210.035 to 49.298 and 54.233, with average percentage reductions of 49.7% and 42.7%. The most improved areas were Northwest, Southwest and Southeast China. The limited improvement of other indices was attributed to the design of the loss function and model outfit owing to the small sample size for training. Although the TCC of precipitation predicted by U-Net did not improve significantly, the predicted PP time series show performed better than in CFSv2. The U-Net predictions were closer to the observations and indicated the trend and intensity of the observed PP. The sensitivity experiments showed that soil moisture is the most crucial factor in predicting summer rainfall in China, followed by geopotential height. Although we specifically chose some long-memory predictors, such as SST and Tibetan Plateau snow cover, to improve the prediction skill in Northeast and Central China, these two regions showed no significant improvement. In general, with no need for pre-training and transfer learning, using U-Net for seasonal summer precipitation forecasts over China is promising and deserves more attention in the future.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the Chinese Universities Scientific Fund [Grant No. CUG 2106108] and the Natural Science Foundation of Hubei Province of China [Grant No. KZ22Z3017].

Acknowledgments

The authors wish to thank Prof. Xiaorui Niu and Dr Xiangze Jin for their useful discussions on the selection of meteorological predictors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.aosl.2022.100322.

References

Asharaf, S., Dobler, A., Ahrens, B., 2012. Soil moisture–precipitation feedback processes in the Indian summer monsoon season. J. Hydrometeorol. 13 (5), 1461–1474.
Donat, M.G., Lowry, A.L., Alexander, L.V., O'Gorman, P.A., Maher, N., 2016. More extreme precipitation in the world's dry and wet regions. Nat. Clim. Chang. 6, 508–513.
Du, G., Cao, X., Liang, J., Chen, X., Zhan, Y., 2020. Medical image segmentation based on U-Net: a review. J. Imaging Sci. Technol. 64, 1–12.
Fu, Z., Huang, A., Guo, Q., 2022. Improvements of the sub-seasonal precipitation predicted by the BCC_S2S forecast system over eastern China in summer using the singular value decomposition bias correction method. Front. Earth Sci. 9, 770167.
Good, S.A., Martin, M.J., Rayner, N.A., 2013. EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. J. Geophys. Res. Oceans 118 (12), 6704–6716.
Ham, Y.G., Kim, J.H., Luo, J.J., 2019. Deep learning for multi-year ENSO forecasts. Nature 573 (7775), 568–572.
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz, S. J., Nicolas, J., et al., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049.
IPCC, 2021. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1557–1563.
Irrgang, C., Boers, N., Sonnewald, M., Barnes, E.A., Kadow, C., Staneva, J., Saynisch-Wagner, J., 2021. Towards neural Earth system modelling by integrating artificial intelligence in Earth system science. Nat. Mach. Intell. 3 (8), 667–674.
Jin, W.X., Luo, Y., Wu, T.W., Huang, X.M., Xue, W., Yu, C.Q., 2022. Deep learning for seasonal precipitation prediction over China. J. Meteorol. Res. 36 (2), 271–281.
Koster, R.D., Dirmeyer, P.A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C.T., et al., 2004. Regions of strong coupling between soil moisture and precipitation. Science 305 (5687), 1138–1140.
Lang, Y., Ye, A., Gong, W., Miao, C., Di, Z., Xu, J., Liu, Y., Luo, L., Duan, Q., 2014. Evaluating skill of seasonal precipitation and temperature predictions of NCEP CFSv2 forecasts over 17 hydroclimatic regions in China. J. Hydrometeorol. 15 (4), 1546–1559.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.
Liu, Y., Ke, Z., Ding, Y., 2019. Predictability of East Asian summer monsoon in seasonal climate forecast models. Int. J. Climatol. 39 (15), 5688–5701.
Pan, X., Lu, Y., Zhao, K., Huang, H., Wang, M., Chen, H., 2021. Improving nowcasting of convective development by incorporating polarimetric radar variables into a deep-learning model. Geophys. Res. Lett. 48 (21), e2021GL095302.
Racah, E., Beckham, C., Maharaj, T., Ebrahimi Kahou, S., Prabhat, M., Pal, C., 2017. Extremeweather: a large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. In: Proceeding of the 31st International Conference on Neural Information Processing Systems. 3405–3416.
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566 (7743), 195–204.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation. In: MICCAI 2015: Medical Image Computing and Computer-Assisted Intervention, pp. 234–241.
Ruby, U., Yendapalli, V., 2020. Binary cross entropy with deep learning technique for image classification. Int. J. Adv. Trends Comput. Sci. Eng. 9 (4), 5393–5397.
Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., et al., 2014. The NCEP climate forecast system version 2. J. Clim. 27 (6), 2185–2208.
Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-k., Woo, W.C., 2015. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In: Proceedings of the 28th International Conference on Neural Information Processing Systems, Montreal, Canada. MIT Press, pp. 802–810..
Tung, Y.L., Tam, C.Y., Sohn, S.J., Chu, J.L., 2013. Improving the seasonal forecast for summertime South China rainfall using statistical downscaling. J. Geophys. Res. Atmos. 118 (11), 5147–5159.
Westra, S., Alexander, L.V., Zwiers, F.W., 2013. Global increasing trends in annual maximum daily precipitation. J. Clim. 26, 3904–3918.
Wu, J., Gao, X., 2013. A gridded daily observation dataset over China region and comparison with the other datasets. Chin. J. Geophys. Chin. Ed. 56 (4), 1102–1111.
Wu, Q.Y., Yan, Y., Chen, D.K., 2013. A linear Markov model for East Asian monsoon seasonal forecast. J. Clim. 26 (14), 5183–5195.
Wu, Z.W., Yu, L.L., 2016. Seasonal prediction of the East Asian summer monsoon with a partial-least square model. Clim. Dyn. 46 (9), 3067–3078.
Xing, W., Wang, B., Yim, S.-Y., 2016. Long-lead seasonal prediction of China summer rainfall using an EOF-PLS regression based methodology. J. Clim. 29 (5), 1783–1796.
Zhang, S., Chen, Y., Luo, Y., Liu, B., Ren, G., Zhou, T., Martinez, V.C., et al., 2022. Revealing the circulation pattern most conducive to precipitation extremes in Henan Province of North China. Geophys. Res. Lett. 49, e2022GL098034.
Zhou, K., Zheng, Y., Li, B., Dong, W., Zhang, X., 2019. Forecasting different types of convective weather: a deep learning approach. J. Meteorol. Res. 33 (5), 797–809.