Response to Review

Reviewer #3:

Review on "Statistical post-processing of precipitation forecasts using circulation classifications and spatiotemporal deep neural networks"

In this manuscript, the authors have proposed a statistical post-processing method that can simultaneously take into account the effects of large-scale circulation patterns and local spatiotemporal information to calibrate the ECMWF forecast dataset for the Huaihe River basin. The study is well developed and the expected results have been achieved. The new model proposed by the authors has the best calibration capability for different seasons, lead times and precipitation intensities. Overall, the study is innovative and has a high degree of completion which deserves to be published, but some issues still need to be corrected or further clarified.

Thanks for your comprehensive review and recognition of the study contribution. The constructive comments will help us improve our manuscript after revision. We provide detailed responses to your comments and our proposed manuscript revisions in the subsequent sections.

Major comments:

1. In the construction of the SOM-CNN-LSTM post-processing methodology, the SOM model was used to identify and classify different large-scale circulation patterns. In selecting of the SOM node, the authors have tested that the 2×3 configuration is physically interpretable. It should be explained what the node here refers to in the SOM model and what their role is. Also explain why 2×3 is interpretable.

Thank you for your suggestion. The SOM nodes are the clustered large-scale circulation patterns, which need to be determined before implementing the SOM model. A fewer number of nodes in the SOM array cannot capture specific circulation patterns while a greater number of nodes would produce redundant circulation patterns that are similar. Therefore, choosing the optimal SOM node is critical. In this study, we have tested several SOM arrays by quantization and topological errors, including 2×2, 2×3, 2×4, 3×4 nodes, and found that 6 distinctive circulation patterns with 2×3 configuration can provide enough details for physical interpretation and satisfactorily describe the variations of the synoptic situations in Huaihe River basin. For physical interpretability, there is a problem in our statement. What we want to express here is that the 2×3 configuration can provide enough details for physical interpretation, because fewer nodes will yield too general patterns while too many nodes may cause its interpretation to be cumbersome or impractical. We will make these points clearer in the resubmitted manuscript.

2. The authors used three statistical metrics in their study to evaluate the prediction skill

and the ability of the correction, but only one of them was used to evaluate and present the results in each of the relevant experiments shown in Figures 7 to 9, respectively. Consideration could be given to including the results of all evaluation metrics from the relevant experiments in the supporting material to more fully demonstrate the features and advantages of the SOM-CNN-LSTM method.

We fully agree. Adding all evaluation metrics to the corresponding experiments will fully demonstrate the features and advantages of the SOM-CNN-LSTM method. We have made supplementary calculations on the corresponding evaluation indicators based on the original data. The main conclusions of this study have not changed and new findings will be added to the resubmitted manuscript. For example, because small deviations may lead to large relative errors (RB), winter precipitation has a larger RB although it has the highest CC and lowest RMSE compared with other seasons(Figures 7(a), 7(c), and 7(e)).

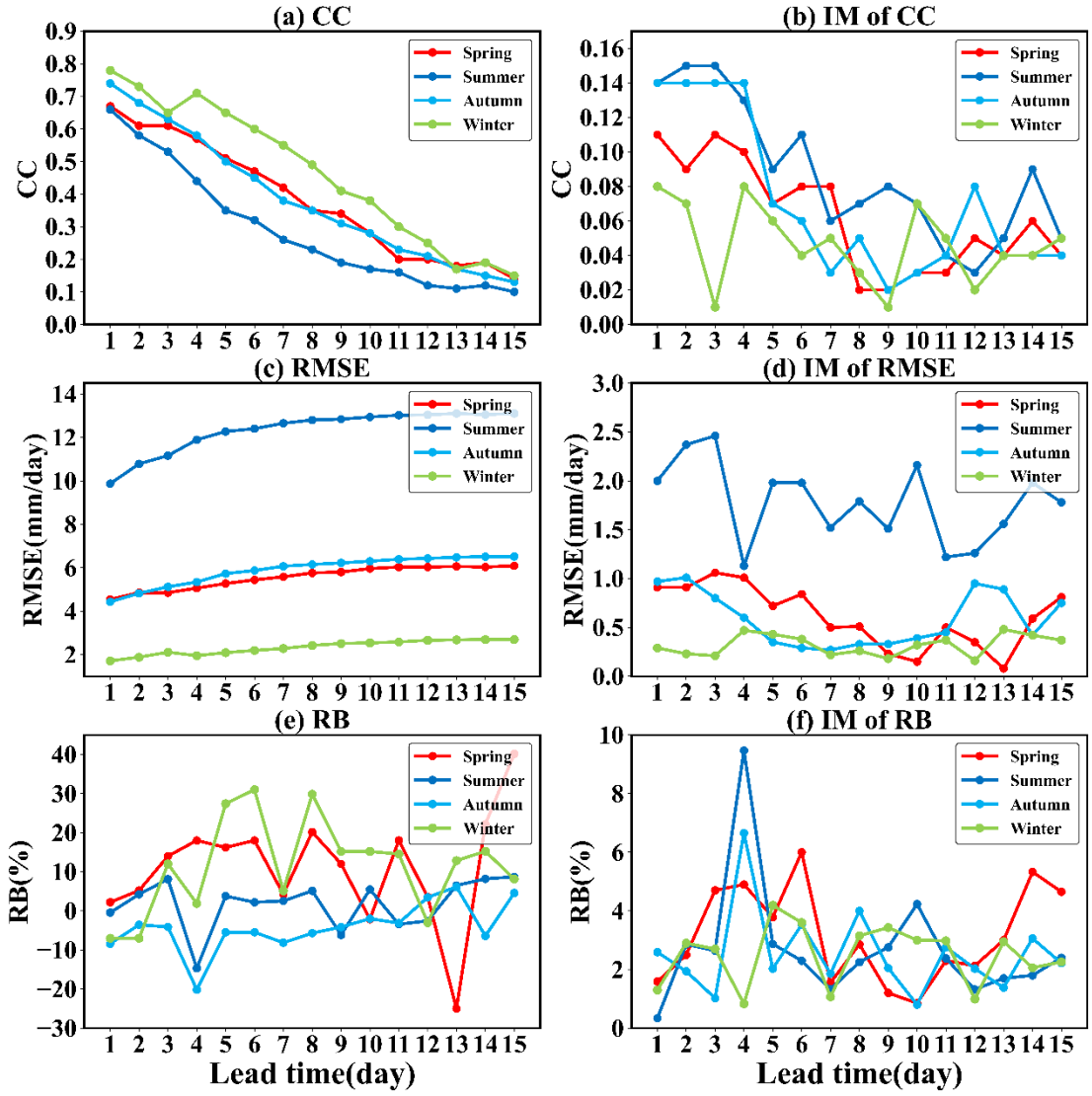

Figure 7 (a) CC, (c) RMSE, and (e) RB of SOM-CNN-LSTM method over 1-15 lead days during spring, summer, autumn, and winter. The second column is the (b)improvement (IM) of CC, (d) RMSE, and (f) RB relative to raw forecasts.

For Figures 8-10, we also add the corresponding statistical metrics and the new figures will be placed in the supporting material.

3. The study focuses on the Huaihe River basin in China. The application and development of similar research in the region should be described in the manuscript to further highlight the main purpose and innovation of this study.

We fully agree. We have found some similar research in the region. Particularly, for Huaihe River basin, Tao et al. (2014) adopted the ensemble pre-processor (EPP) method to calibrate the TIGGE multimodel ensemble forecast precipitation and Li et al. (2022b) adopted the CNN model to correct raw forecast precipitation by considering multi-spatial information. Although the above results show that post-processed precipitation forecasts have substantial improvement over the raw forecasts, these traditional post-processing methods overlook the influence of large-scale circulations and spatiotemporal information on precipitation. To overcome the problem, we propose the SOM-CNN-LSTM post-processing method. we compare the method with other benchmarks, including CNN, LSTM, and CNN-LSTM methods. We will add these sentences to the discussion section in the resubmitted manuscript.

Reference:

Tao, Y., Duan, Q., Ye, A., Gong, W., Di, Z., Xiao, M., and Hsu, K.: An evaluation of post-processed TIGGE multimodel ensemble precipitation forecast in the Huai river basin, Journal of hydrology, 519, 2890-2905, https://doi.org/10.1016/j.jhydrol.2014.04.040, 2014.

Li, W., Pan, B., Xia, J., and Duan, Q.: Convolutional neural network-based statistical post-processing of ensemble precipitation forecasts, Journal of Hydrology, 605, https://doi.org/10.1016/j.jhydrol.2021.127301, 2022b.

Minor comments:

1. L82,305 ‘we’ => ‘We’.

We are very sorry for our incorrect writing. We will replace “we” by “We”.

2. L95 ‘contains’ => ‘contain’.

We are very sorry for our incorrect writing. We will replace “contains” by “contain”.

3. L102 The title of section 3 is wrong.

We are very sorry for our incorrect writing. We will replace “Study area and datasets” by “Methodology”.

4. L123 The formula is incomplete.

We are very sorry for our incorrect writing. We will replace "$\langle Z\rangle = \frac{Z-Z_{mean}}{\sigma_Z}\cos$" by "$\langle Z\rangle = \frac{Z-Z_{mean}}{\sigma_Z}\cos\phi$".

5. Change the use of color table in Figure 8. The authors use only one color table in Figure 8 to represent two types of data, correlations and changes in correlations, which can be confusing. Also, this color table is more appropriate to represent the variation between positive and negative values, which is not the case for the two variables in this figure.

Thank you for your suggestion. We have changed the figure with two color tables. The revised figure is as follows:

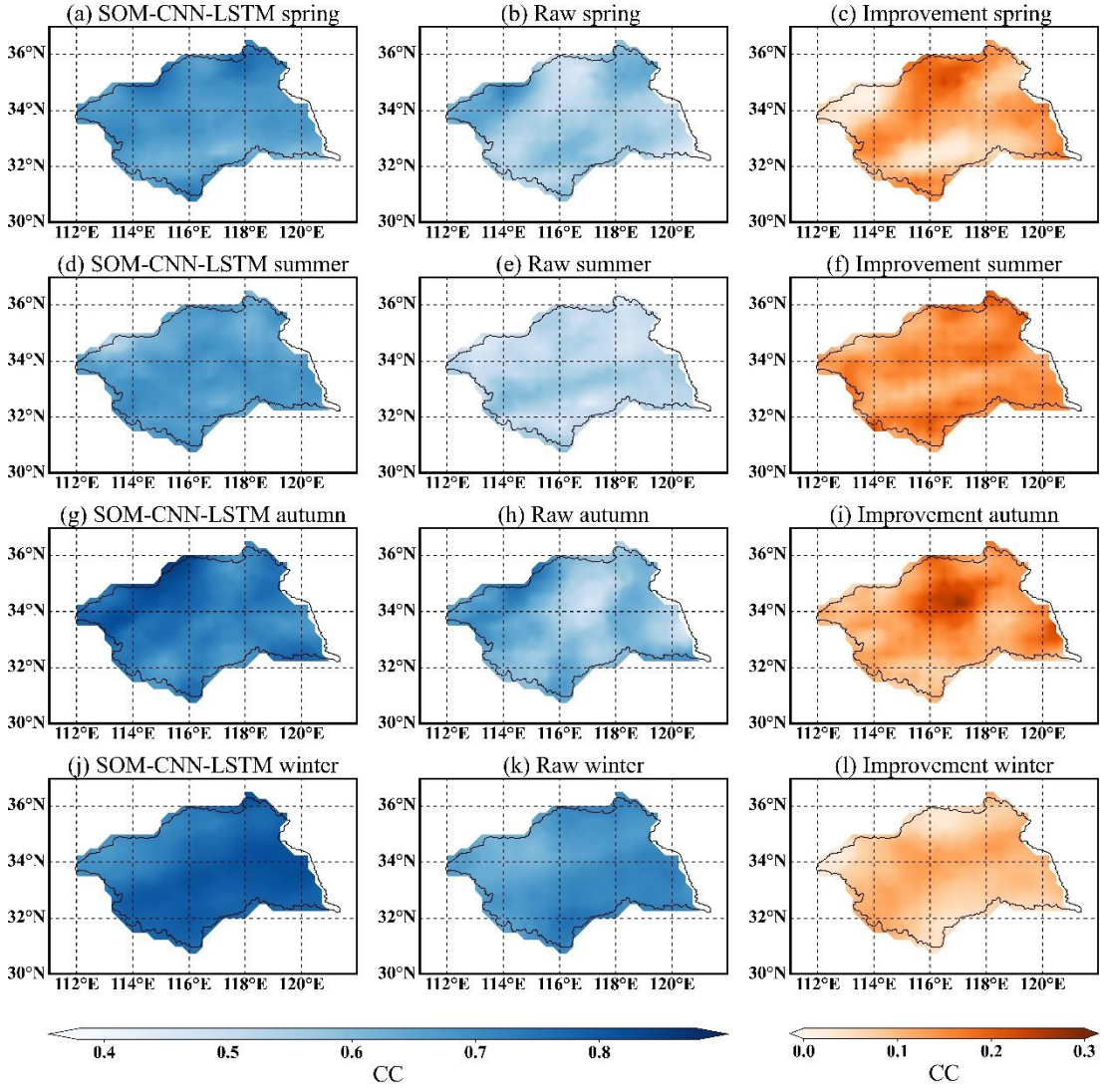

Figure 8 Spatial distributions of the CC for SOM-CNN-LSTM method and raw forecasts at the lead time of 1 day. The third column is the improvement of CC in

spring, summer, autumn and winter.

6. In Figure 9, the conclusion the authors most wanted to express would have been the difference between the precipitation predictions for different years, but at the same time they also point out that the SOM-CNN-LSTM method performs the best. However, the color table used and the type of Figure 9 make the latter conclusion very unclear, at least compared to the other figures in the paper. Also, the correspondence between color table and value is not fixed.Therefore,the author should consider a more appropriate way of presenting the relevant conclusions.

Thank you for your suggestion. To make the latter conclusion clearer, we add the "*" to highlight the best method with the lowest RB for each lead time. In addition, we have fixed the value for all figures to better analyze the annual forecast skills of different methods. From the revised figure, we found that the underestimation is more appropriate in 2021 than in 2013, and the overestimation is more appropriate in 2009, 2011, and 2012. Furthermore, when the lead time exceeds 12 days, forecast precipitation is overestimated in most years, especially in 2013 and 2014. We will change them in the in the resubmitted manuscript. The revised figure is as follows:

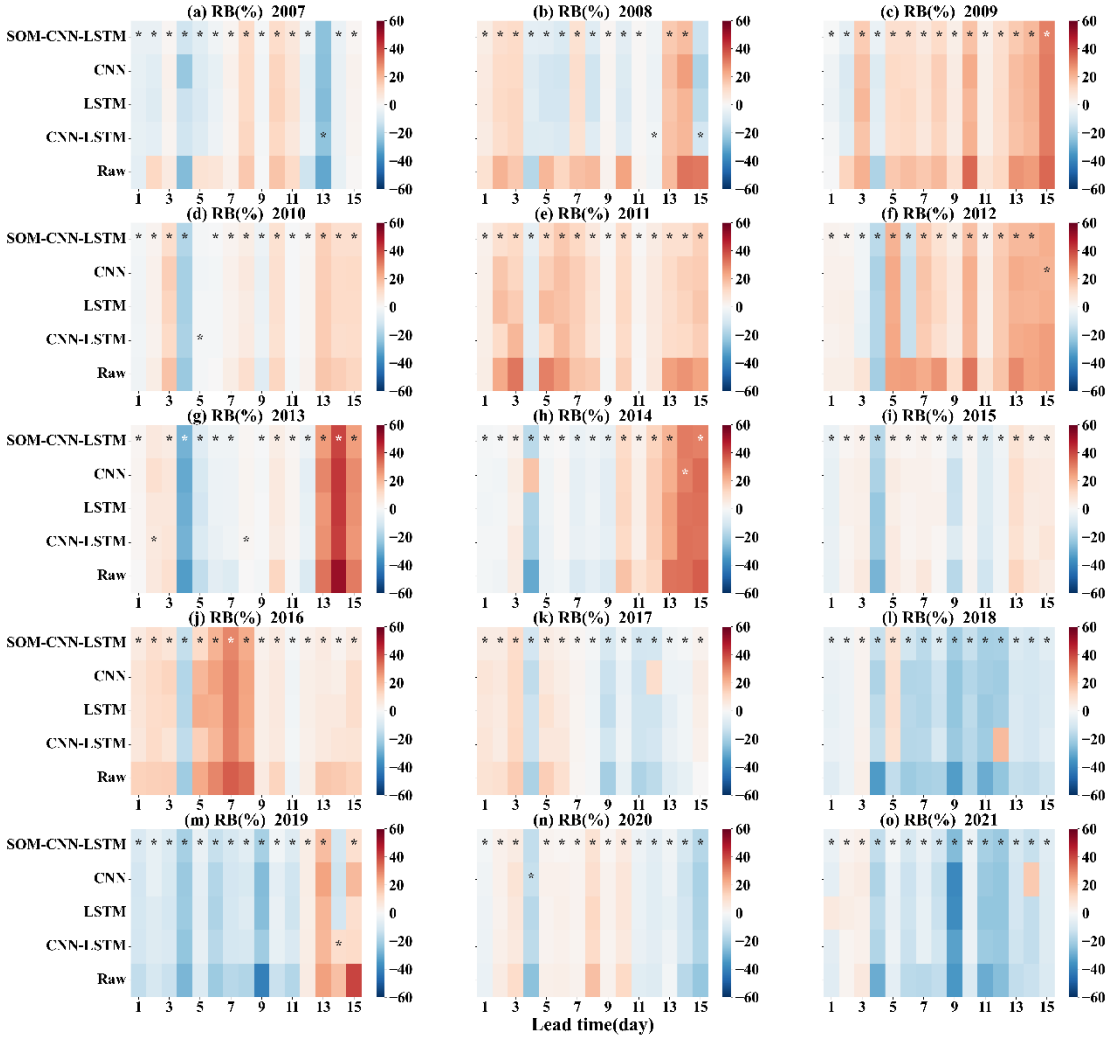

Figure 9 RB of different methods for each summer over 1-15 lead days from 2007 to 2021. The"*" indicates the best method with the lowest RB for each lead time.

7. L307 Is the 'SHAP' used here incorrectly? If not, it is needed to clarify this abbreviation.

We are very sorry for our incorrect writing. We will replace "SHAP" by "WPSH".

We appreciate for Reviewer's warm work earnestly, and hope that the correction will meet with approval.
Once again, thank you very much for your comments and suggestions.