

Geosci. Model Dev. Discuss., author comment AC3  
<https://doi.org/10.5194/gmd-2022-213-AC3>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC3

Fang Wang et al.

---

Author comment on "Customized deep learning for precipitation bias correction and downscaling" by Fang Wang et al., Geosci. Model Dev. Discuss.,  
<https://doi.org/10.5194/gmd-2022-213-AC3>, 2022

---

The paper entitled "Customized Deep Learning for Precipitation Bias Correction and Downscaling" developed a customized deep learning model for precipitation downscaling and tested the advantages of incorporating a weighted loss function, multitask learning, and accounting for physically relevant covariates. The manuscript is generally well structured and clearly presented. I have a few minor comments regarding the explanations of the training processes.

- Line 88: (Chen, 2020)

*Response: Thank you for pointing it out. We have fixed it.*

- Section 2.2.4 How many training steps (iterations) do you have for each scenario? It would be nice to see the learning curve for each scenario.

*Response: Thank you for your comments. The total iterations for each scenario are about  $2.5 \times 10^5$  and the learning curves for the 6 scenarios are plotted in Figure S1 in the Supplement document. We have added the information in Section 4.1 of the updated manuscript.*

- Line 172: Please explain why you specifically developed the weighted mean absolute error (MAE) loss function rather than using the regular MAE.

*Response: Thank you for your comments. We have added the following explanation in Section 3.1.3 of the updated manuscript: "Precipitation data is highly skewed and unbalanced especially at hourly time scale, which could cause deep learning algorithm to focus more on no rain events and ignore heavy rain events if using regular loss functions."*

- Line 211: How are these covariates chosen? I wonder if the authors have tested if all covariates are necessary for precipitation downscaling or if other covariates are necessary (for example, convective available potential energy).

*Response: Thank you for your comments. These covariates are chosen based on precipitation formation theory (cloud mass movements and thermodynamics) and other studies on estimating precipitation. We have included the following explanation in Section*

3.1.4 of the updated manuscript "We chose these variables based on precipitation formation theory (cloud mass movements and thermodynamics) and other studies on estimating precipitation as listed above." We did not test whether all covariates or other covariates are necessary for this study. We have made a note in the discussion section 5 of the updated manuscript as follows: "Note that we did not explore the importance rank among these covariates in improving the model performance in this study, which could be a potential avenue for future work."

- Line 240: Can you show the data distribution during the training period (2002-2015) and the test period (2019-2021)? It is hard to tell if the methods have extrapolation capability without seeing the differences in precipitation distribution. Are the increases in annual precipitations caused by a systematic increase each timestep or an increase in extreme precipitations?

*Response: Thank you for your comments. Since there are too many no-rain hours, the data distributions during training and testing periods are hard to distinguish. Therefore, we calculated the probability for each precipitation bin with a range of 0.5 mm/h during the training and testing periods and calculated their difference of probability for each bin (see Table S1 in the Supplement). The table suggests that the climatology differences between testing and training datasets (Figure 2) are because there is a higher percentage of rains greater than 0.5 mm/h in the testing data, while there is a higher percentage of no rain or drizzling (0 to 0.5 mm/hr) in the training dataset. We have added the following explanation in Section 3.1.4 of the updated manuscript: "Wetter conditions are observed in most of the study area in the testing period (average 0.03 mm/h) than the training period, which is due to a higher percentage of rains (with values greater than 0.5mm/h) during the testing period than during the training period based on analyzing the Stage IV data (Table S1 in Supplement)."*

*We train the model using hourly data and evaluate that at different scales including hourly, daily, and monthly scales. The increase in skill in aggregated time scales (e.g., daily and monthly) is due to the smoothing of the hourly data. We did not aggregate the data into an annual time scale, but it is expected that the skill for the annual time scale would be better than the monthly scale due to further smoothing.*

- Line 244: Briefly describe how QDM is applied in this study

*Response: Thank you for your comments. We have included more explanations in Section 3.2 of the updated manuscript as follows "The QDM method corrects systematic biases at each grid cell in quantiles of a modeled series with respect to observed values. Compared to the regular quantile mapping method (Panofsky and Brier, 1968; Thrasher et al., 2012; Wood et al., 2002), QDM also applies a relative difference between historical and future climate data (here, training and testing periods), thus it is capable of preserving trend of the future climate (Cannon et al., 2015), which is critical for this study since there are substantial differences between the precipitation during the training (2002 to 2015) and testing (2019 to 2021) periods (see Figure 2)".*

- Line 302: There is no section 3.1

*Response: Thank you for pointing it out. We have fixed it.*