Response for Reviewer #1

High-quality global precipitation product with finer spatiotemporal resolutions and long-term temporal coverage is really critical for a variety of science communities. However, after carefully reading this manuscript, there are various aspects confusing me a lot. Most critically, the writing and organization are really too weak to understand the key ideas of this study, as well as lacking scientific innovative contributions for the community, which seems to be just mixing several global precipitation datasets without any clear new thoughts. Some more serious scientific issues could be seen as follows. Considering the high standards of the big journal, ESSD, I think this study have great limitations and too far distances from the standards.

Response: We are very grateful to the reviewer for careful reviews and valuable comments. Here, we would like to make the following clarifications, explanations, and modifications for the above comments.

(1) This study aims to generate a high-quality global land precipitation dataset. The new products (i.e., MGP-3P and MGP-6P) were compared with other five popular global products (i.e., MSWEP, IMERG-Final, GSMaP-Gauge, ERA5, and CPCU) in global land areas, and were evaluated and compared in mainland China using two high-density ground precipitation observations as the benchmark. The evaluation results show that the MGP-3P product substantially performed better than the other five research-quality products (i.e., MGP-6P, MSWEP, IMERG-Final, GSMaP-Gauge, and ERA5) in terms of most error metrics. This indicates that the MGP-3P

can provide a high-accuracy precipitation dataset for the community and has been available to the public (https://zenodo.org/record/7386441#.ZBkpxcJBxD-), which is the main purpose of this work, as well as our contribution to the community.

(2) The novel multi-source precipitation data fusion (MPDF) algorithm uses CC as the fusion weights for satellite and reanalysis, which is consistent with that of beck et al., 2017, 2019. It should be pointed out that, however, the MPDF algorithm also further considers the dependency of satellite and reanalysis precipitation errors on seasonality, which is conducive to fully taking advantage of the complementary strengths from satellite and reanalysis data. Furthermore, we used the number of rain gauges per 0.5° grid cell as the fusion weights for ground precipitation observations (i.e., CPCU), which can effectively consider the advantage of gauge observations and avoid precipitation uncertainties because the precipitation of the grid cells with no gauge observations was not considered in the merged procedure. More importantly, the spatiotemporal downscaling module used in the MPDF algorithm is different from that of beck et al., 2017, 2019. Our spatiotemporal downscaling module not only considered the advantage of satellite and reanalysis but also took into account the contributions of hourly and 0.25° grid merged precipitation to corresponding daily and 0.5° grid merged precipitation. The results of MGP-3P demonstrated this point, although it is not as good as the spatiotemporal scaling module of IMERG, please see Table 1 (Table 4 in the manuscript). Admittedly, the MPDF algorithm is very simple, but that also is its advantage that it is easy for the readers to understand and use it. In fact, the verification results also

demonstrated the MPDF algorithm is effective in considering the advantages from satellite, reanalysis, and gauge data and in improving the quality of precipitation. The MGP-3P substantially performed better than the other five research-quality products (i.e., MGP-6P, MSWEP, IMERG-Final, GSMaP-Gauge, and ERA5) in terms of most error metrics. In particular, the accuracy of the MGP-3P product is obviously better than that of another multi-source data fusion product MSWEP in the evaluation for mainland China. In addition, we found that the quality of the input products is critical for that of the merged precipitation estimates, rather than the number of the input products. This finding can give some valuable information for researchers to customize the multi-source precipitation data fusion algorithms. Given that the MPDF algorithm has the advantages of the above four points and this finding, we believe that this paper gives some new thoughts to the community.

(3) We will carefully revise the writing and organization in the next version.

In summary, this paper is to generate a high-quality global land precipitation product by using the novel MPDF algorithm to provide a high-accuracy precipitation dataset for the community. Currently, the MGP-3P product has been available to the public (https://zenodo.org/record/7386441#.ZBkpxcJBxD-). In addition, we give some new thoughts to the community. The writing issue existing in the previous manuscript might mislead the reviewer. We will significantly improve the quality of the writing in the next version. The authors hope that the above clarifications and explanations can solve the concerns of the reviewer.

Table 1 Decay rates of five products (i.e., MGP-6P, MGP-3P, IMERG, GSMaP, and ERA5) in terms of POD, FAR, and CC when from a 3 hourly resolution scaling down to an hourly resolution.

Products	POD (%)	FAR (%)	CC (%)
MGP-6P	13.25	0.00	13.04
MGP-3P	10.59	1.81	12.68
IMERG	7.58	4.44	10.00
GSMaP	10.39	5.77	13.04
ERA5	15.00	5.00	20.00

Serious scientific concerns including but are not limited to:

1. what's your basic assumptions for fusing these global dataset? If only consider the CC as the fusion weights, it is too simple and too weak. Beck et al., 2017, 2019 have already investigated such explorations. Please carefully reading such critical references.

Response: we used CC as the fusion weights for satellite and reanalysis, as it can objectively indicate the performance of satellite and reanalysis products. Also, Beck et al. (2017, 2019) verified the effectiveness of this method in merging multiple precipitation estimates. However, the MPDF algorithm did not only consider the CC as the fusion weights and is different from the algorithm of Beck et al. (2017, 2019). The advantages of the MPDF algorithm are reflected in the following aspects:

(1) The MPDF algorithm considers the dependency of precipitation errors on seasonality because the performance of satellite and reanalysis has a significant dependency on seasonality. The global land maps of the weights designed to satellite, reanalysis, and CPCU in four seasons (i.e., MAM, JJA, SON, and DJF) demonstrated this point, please see Figure 1 (Figure 13 in the manuscript). Thus, the dependency of satellite and reanalysis precipitation errors on seasonality considered in the MDFP algorithm is necessary.

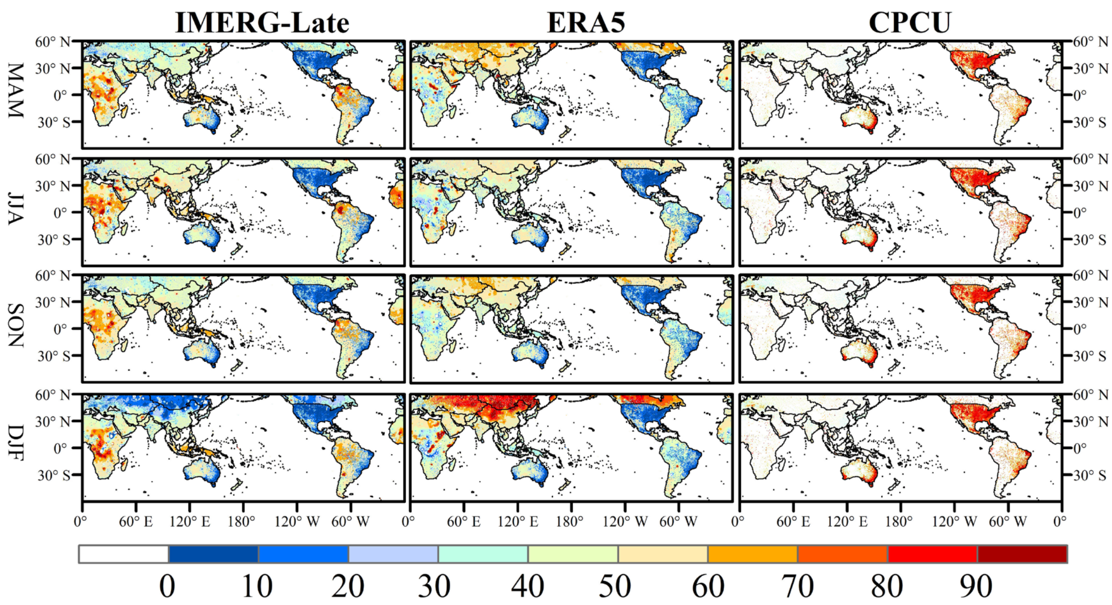

Figure 1 Global land maps of the weights designed to satellite, reanalysis, and CPCU in four seasons (i.e., MAM, JJA, SON, and DJF) for scheme 2.

(2) The number of rain gauges for each 0.5° grid cell was used as the fusion weights for ground precipitation observations (i.e., CPCU). This method can fully take advantage of the strengths of gauge observations as the quality of the ground precipitation observations depends on the spatial density of the rain gauges (Krajewski et al., 2000; Villarini and Krajewski, 2007; Roca et al. 2010; Prakash et

al., 2019; Chen et al., 2020). Additionally, this weighting method avoids resulting in precipitation uncertainties caused by gauge observations.

(3) We used the rations between the merged precipitation from different spatiotemporal resolutions as the weights for the spatiotemporal downscaling of precipitation. This method not only considered the advantage of satellite and reanalysis but also took into account the contributions of hourly and 0.25° grid merged precipitation to corresponding daily and 0.5° grid merged precipitation. The results of MGP-3P indicated that this method is effective, although it is not as good as the spatiotemporal scaling module of IMERG, please see Table 1 (Table 4 in the manuscript). This is also one of the key modules of the MPDF algorithm.

(4) The MPDF algorithm is simple but that also is its advantage that it is easy for the readers to understand and use it. More importantly, the MPDF algorithm is effective in improving the accuracy of precipitation, the evaluation results in mainland China demonstrated this point.

The revised manuscript will clarify that Beck et al. (2017, 2019) have used the CC as the fusion weights.

References

Krajewski, W. F., Ciach, G. J., McCollum, J. R., and Bacotiu, C.: Initial validation of the Global Precipitation Climatology Project monthly rainfall over the United States. J. Appl. Meteorol. 39(7), 1071-1086, https://doi.org/10.1175/1520-0450(2000)039<1071:IVOTGP>2.0.CO;2, 2000.

Villarini, G., and Krajewski, W. F.: Evaluation of the research version TMPA three-hourly 0.25×0.25 rainfall estimates over Oklahoma. Geophys. Res. Lett. 34, L05402, https://doi.org/10.1029/2006GL029147, 2007.

Roca, R., Chambon, P., Jobard, I., Kirstetter, P. E., Gosset, M., and Bergès, J. C.: Comparing satellite and surface rainfall products over West Africa at meteorologically relevant scales during the AMMA campaign using error estimates. J. Appl. Meteorol. Clim. 49(4), 715-731, https://doi.org/10.1175/2009JAMC2318.1, 2010.

Prakash, S., Seshadri, A., Srinivasan, J., and Pai, D. S.: A new parameter to assess impact of rain gauge density on uncertainty in the estimate of monthly rainfall over India. J. Hydrometeorol. 20(5), 821-832, https://doi.org/10.1175/JHM-D-18-0161.1, 2019.

Chen, H., Yong, B., Qi, W., Wu, H., Ren, L., and Hong, Y.: Investigating the evaluation uncertainty for satellite precipitation estimates based on two different ground precipitation observation products. J. Hydrometeorol. 21(11), 2595-2606, https://doi.org/10.1175/JHM-D-20-0103.1, 2020.

2. The title tells that this study aims to public a global dataset, however, it only provides precipitation estimates over Land. It is really not rigorous. Please take care of such issues.

Response: thank you. We will modify this issue in the next version. The title will be revised to "MGP: a new 1-hourly 0.25° global land precipitation product (2000-2020) based on multi-source precipitation data fusion”.

3. How did the authors consider the negative effects from the different input precipitation estimates, especially in terms of the precipitation events? The weights based on CC could be only achieved at daily scale. So how did you consider the systematic and random errors at hourly scales from the input datasets?

Response: those are good questions. In fact, the MPDF algorithm only considers the precipitation amount, without considering precipitation events. Notably, the MPDF algorithm fully takes advantage of the complementary strengths from satellite, reanalysis, and gauge data, but also propagates negative effects from different input products into merged precipitation estimates. For instance, the FAR of the MGP products is higher than that of most global precipitation products, which is due primarily to the false precipitation of ERA5 being propagated into the MGP suite. In addition, the MGP suite is not the best in terms of the total bias, as the rainfall amount of the satellite and ERA5 reanalysis was not corrected before the weighted merging. Those issues of the MPDF algorithm were found, and we have discussed those in the manuscript, please see lines 610-622.

The weights at hourly scale are from those at daily scale, because the weights based on CC could be only achieved at daily scale. The relative size of ratio for CC between different input products could be well consistent across different time scales, as the quality of all satellite precipitation products decreases when scaled down to shorter time periods (Tan et al., 2017). However, we believe that the differences would be found in areas with complex terrain and strong spatiotemporal heterogeneity of precipitation. We will discuss this in the revised manuscript.

Reference

Tan, J., Petersen, W.A., Kirstetter, P.E., and Tian, Y.: Performance of IMERG as a function of spatiotemporal scale. J. Hydrometeorol. 18 (2), 307-319, https://doi.org/10.1175/JHM-D-16-0174.1, 2017.

4. The resolutions of the MPG is very strange with 1-hourly and 0.25°. most popular satellite and reanalysis precipitation datasets have finer resolutions at 1-hourly and 0.1°, for instance, IMERG, GSMaP, and ERA5-Land. Particularly, PERSIANN-CCS is quiet finer with resolutions half-hourly and 0.04°. So what's your purpose of the resolutions at 1-hourly and 0.25°?

Response: thank you for the comments. The spatiotemporal resolution is mainly determined by the accuracy of the MGP-3P product. We noted that the spatiotemporal downscaling module can maintain the accuracy of the MGP in a favorable position at 0.25° spatial and hourly temporal resolutions. In addition, if the spatial resolution of the MGP-3P product is finer (< 0.25°), only IMERG-Late can be used as input for the spatiotemporal downscaling module, which cannot ensure the accuracy of the MGP-3P product at a finer spatial resolution. Consequently, the purpose of the resolutions at 1-hourly and 0.25° is to ensure that the accuracy of the MGP-3P product.

5. The authors seems to have not enough background information on such satellite-based and reanalysis-based precipitation datasets. For instance, the most important aim of PERSIANN-CCS is to capture the first glimpse of the possible precipitation, not the quality. The authors considered the PERSIANN-CCS to provide what information at 1-hourly and 0.25° for developing the qualified research level precipitation product?

Response: thank you for the comments. In this study, we would like to answer the question of which is crucial for the quality of merged precipitation estimates: the number of input products or the quality of input products? Therefore, we have designed six input sources for merging scheme 1 to generate MGP-6P, without considering the quality of input products, whereas we consider the quality of input products in merging scheme 2 for generating a research level precipitation product MGP-3P, merging the best satellite-only IMERG-Late and model-based ERA5 reanalysis precipitation products. A comparison between MGP-3P and MGP-6P was used to answer the above question. Finally, we found that the quality of input products is critical for that of merged precipitation estimates, rather than the number of input products. The detailed information can be seen in section 4 and lines 302-328.

6. In terms of evaluation and comparison, the results have various weak aspects and do not make me convinced, especially due to the black box merging model: (1) only evaluated at mainland China? Would it be reasonable to represent the global situations? (2) what are the reasons for improving the POD and FAR of MGP-6P and MGP-3P? just because there were merged based on CC only achieved at daily scales? and (3) why not evaluate and compare these precipitation products over CONUS where have enough ground observations for public?

Response: thank you for the comments. We only performed the evaluation and comparison in mainland China, due to lacking the high-quality ground precipitation observations for the rest of the world. We admit that the evaluation results from mainland China cannot represent the global situation, and we also indicated the limitations of evaluation in the manuscript, please see lines 631-649. We are very grateful to the reviewers for the suggestions that evaluate and compare the precipitation products over CONUS. We accept this suggestion and will do it in the revised manuscript.

In terms of the reasons for improving the POD and FAR of MGP-6P and MGP-3P, we believe the reasons mainly include the following aspects：

(1) The MPDF algorithm fully take advantage of the strengths from ERA5 which has the best detection capability compared to satellite-only precipitation products (Jiang et al., 2021; Chen et al., 2023). We believe that this is the main reason for improving

the POD of the MGP suite. This has been discussed in the manuscript, please see lines 425-426.

(2) The weighting method considered the dependency of satellite and reanalysis precipitation errors on seasonality, which fully takes advantage of the complementary strengths from satellite and reanalysis data.

(3) The weighting method of gauge observations (i.e., using the number of rain gauges as the fusion weights) is effective in considering the advantage of gauge observations, and the precipitation estimates in the grid cells with no rain gauges were not merged, avoiding precipitation uncertainties caused by gauge observations.

(4) The spatiotemporal downscaling module not only considered the advantage of satellite and reanalysis data but also took into account the contributions of hourly and 0.25° grid merged precipitation to corresponding daily and 0.5° grid merged precipitation, which improves the performance of the MGP in terms of POD.

Reference

Jiang, Q., Li, W., Fan, Z., He, X., Sun, W., Chen, S., Wen, J., Gao, J., and Wang, J.: Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland. J. Hydro. 595, 125660, https://doi.org/10.1016/j.jhydrol.2020.125660, 2021.

Chen, H., Wen, D., Du, Y., Xiong, L., and Wang, L.: Errors of five satellite precipitation products for different rainfall intensities. Atmos. Res. 285, 106622, https://doi.org/10.1016/j.atmosres.2023.106622, 2023.