Discussion comment 2

Although many publications already exist in the literature, errors and uncertainties associated with global satellite precipitation products are still difficult to characterize. Thus, I believe this work could be of interest to the HESS readership and worth publication. However, I have a few comments that I would like the authors to consider before accepting the manuscript for publication.

Many thanks for the important comments and suggestions from this reviewer.

1. **First off, I recommend revising the language, since there are a few grammatical mistakes**

In the revision, the manuscript will be carefully checked and revised to avoid all the grammatical mistakes.

2. **Second, I suggest clarifying the main goal of this work. There are currently 5 goals mentioned at the end of the Introduction that read more like tasks. I think that a more focused article would be more effective. In other words, is the goal to validate the SPPs? Is it to model their errors/uncertainties? Is it to investigate what are the factors causing more errors/uncertainty in one location/product/season vs. another? Is it to inform users and algorithm developers on how to use/improve such products?**

Thank you for the suggestions. We will further clarify the main goal of this work at the end of the Introduction according to reviewer's suggestion.

3. **In this regard, the abstract should be more concise and highlight the main goals**

and findings of this work

Ok, we will further refine the abstract to highlight the main goals and findings of this work.

4. **Is including both IMERG Early and Late really necessary? The algorithm is same and – as expected – their performance very similar. Same goes for GSMaP-NRT and GSMaP-MVK. This may help with my comment above of a more focused article.**

Thank you. In fact, the performance of IMERG-Early is similar to that of IMERG-Late. According to reviewer's suggestion, we will remove the IMERG-Early and GSMaP-NRT in our revised manuscript.

5. **The section on "transferability of the regional assessment results to other areas" is weak and not well justified. However, this part of the study is also one of the most interesting, since improving our knowledge of how SPPs perform in regions of the world where no ground observations are available can be extremely useful (e.g., hydrologic applications). Ground observations are mainly available in plain areas (and sparse vegetation). Thus, how can we generalize such results to densely vegetated and highly complex regions?**

Thanks. We entirely agree with the reviewer's opinion. The verification about the transferability of the evaluation results seems weak and is not well justified. However, this part of the study is rather important. Therefore, in the revision we chose two suitable areas (i.e., Chinese Fujian (FJ) and Zhejiang (ZJ) provinces) for comparing and investigating the potential for the transferability of the regional assessment results to

other areas. These two selected adjacent areas, which located in southeastern China, have similar topography and climate. The spatial distribution of the total bias for each SPP over FJ and ZJ were given in the following Fig.1. Apparently, the spatial distributions of the total biases between FJ and ZJ exhibit evident differences, suggesting that the evaluation results between similar areas could not be extended to one another. The reason might be attributed to the large performance differences existed between various satellite sensors and the differences of the satellite samples in retrieval systems over different regions. We will revise this issue in this revised version.

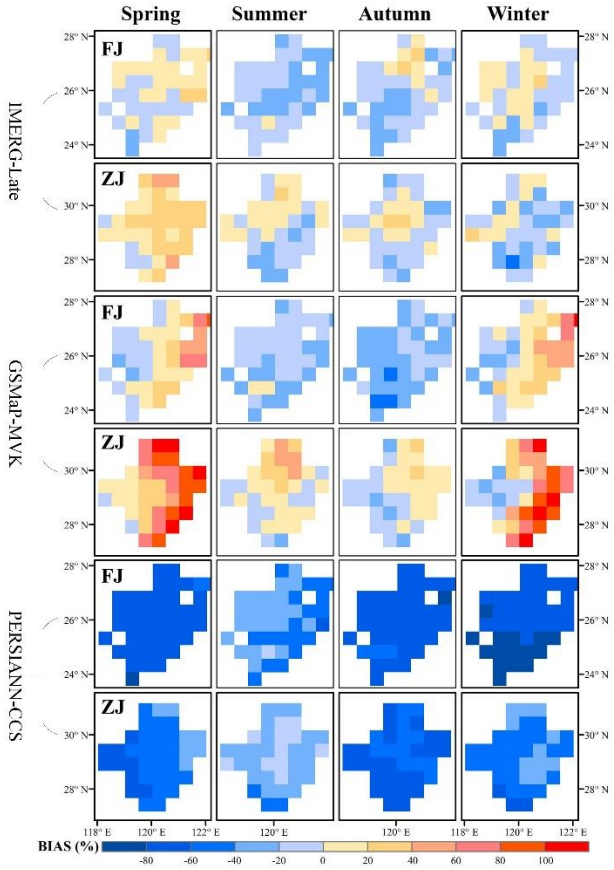

Fig. 1 The spatial maps of the total biases of three SPPs for four seasons over the Fujian (FJ) and Zhejiang (ZJ) provinces, respectively.

6. It would be useful to see how many gauges are available for each stdev class

shown in Fig. 10.

Ok, we will provide the available gauges for each topography class in the version, as shown in the following Fig.2. One can see that enough ground rain gauges, exceeding 135 rain gauges for each terrain class, were used as the references to ensure the reliability and robustness of evaluation results. We will add the corresponding discussion for this point in the section 3.3.

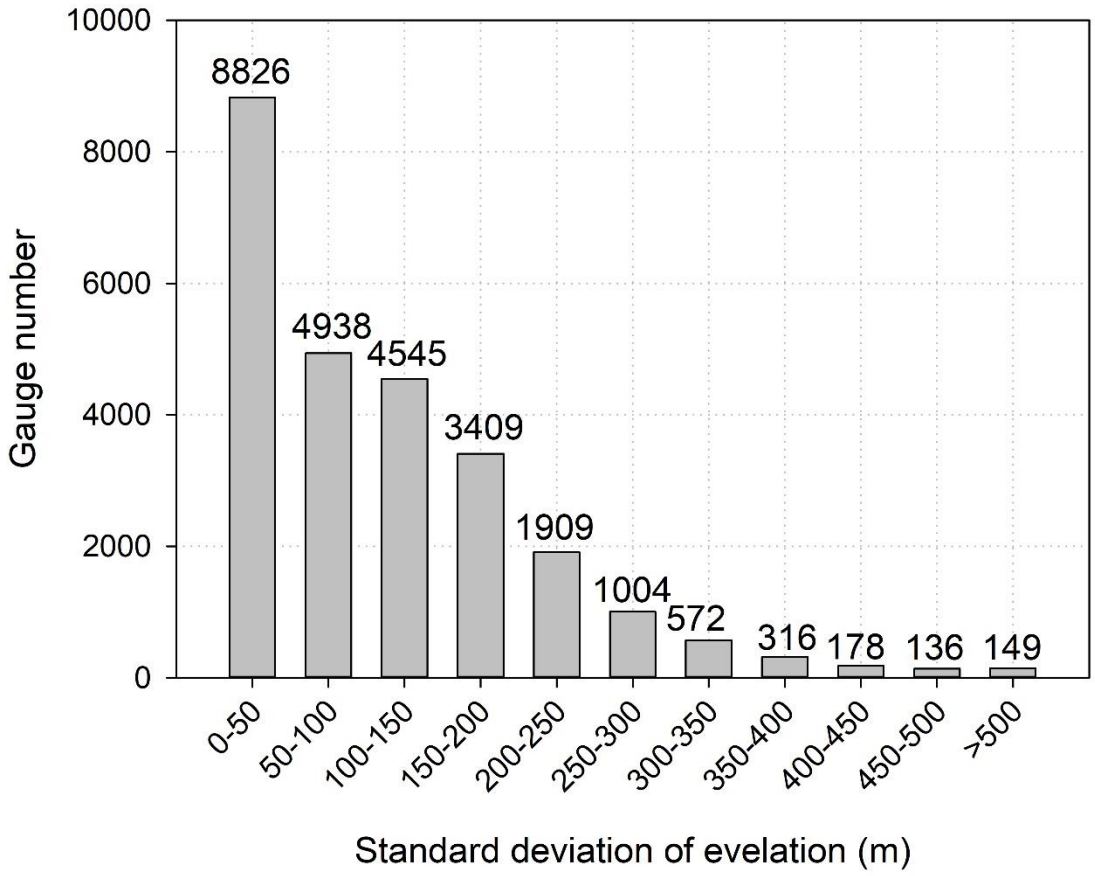

Fig. 2 The gauge number of each topography class.