[Response to the referees' comments]

We would like to thank all the referees for the constructive and valuable comments. Overall, we agree with the referees' suggestions, and have revised the manuscript based on them. Here, we provide responses to each comment by the reviewers. Please note that we numbered the comments from the referees in a unified manner for clarification. In addition, we provide a supplemental PDF file that shows the comment numbers which correspond to each revision.

In addition to the revisions based on the referees' comments, we modified Fig. 4, because the value of Kendall's rank correlation coefficient listed in Fig. 4 (c) was mistaken. We would be grateful if the referees would compare the Figures 4 before and after the modification which are depicted below:

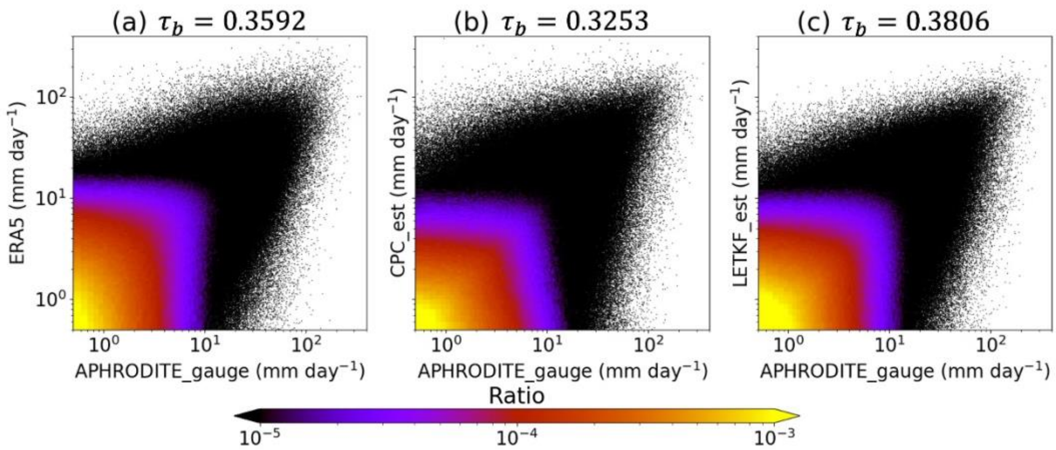


Fig. 4 (before modification)

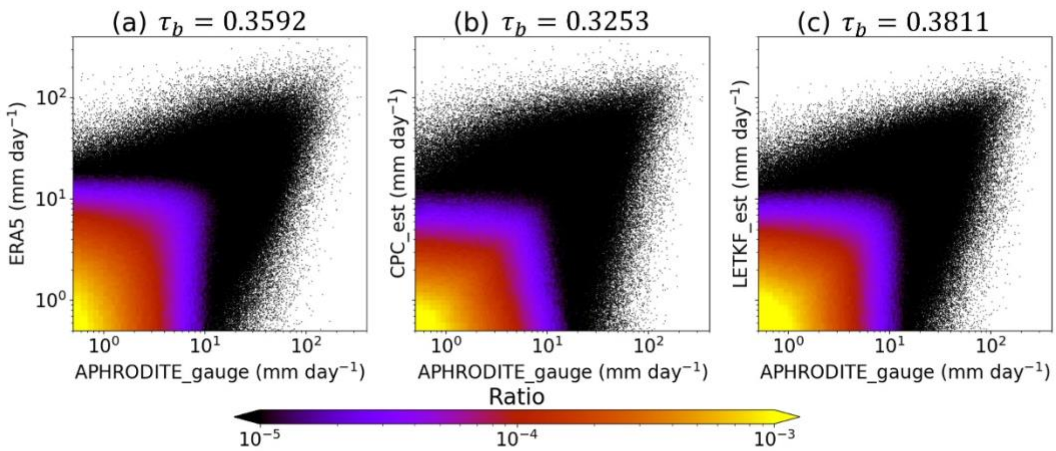


Fig. 4 (after modification)

Furthermore, based on the comments from Referee 1 and 2, we added appendices (Appendix 1-5) to enhance the readers' understanding on our study. We would be grateful if you would also check these additional materials.

The comments from the referees are written in black in the gray boxes, followed by the authors’ response written in blue.

Referee: 1

General comments

R1-G1. The product quality of this paper may rely not only for the LETKF method but daily base ERA5 data that have archived various satellite base observational information with physical model to assimilate them. I wonder that usage of ERA5 is critical, or LETKF could also act better on other reanalysis data, such as NCEP or JRA. Case studies in Fig. 2 and Fig. 5 were used to confirm the improvements, however, why they were a single day/month in old ages such as 1988 and 1985? Are they chosen to avoid recent improvements of ERA5 quality on purpose? Did you try other cases with different seasons to derive the same tendency? Please show the representativity of case studies, and discuss that how much of your improvements were rely on ERA5 quality.
Response: Thank you for the valuable comments. Firstly, we would like to clarify that we have performed a 10-year experiment (from 1981 to 1990) global precipitation fields, and this is not a case study (Please find the last paragraph in Section 2.1.2). Although Fig. 2, 3 and 7 depicts examples on a single date or month to provide detailed images to the readers, the results in Fig. 4, 5, 6, 8, 9 are all based on validations using the data samples of the whole estimation period. We apologize for the lack of explanation, and we added descriptions about the period used for the validations in Sections 2.2.1 and 2.2.2. Secondly, the estimation period of this study was selected from more than 30 years ago, considering the fact that the proposed methodology would be especially beneficial for periods when there were few satellite observations. Since this point was not clarified in the introductions, we added explanation in Section 1. Thirdly, we agree with the reviewer that the accuracy of our precipitation estimates may also differ if a different type of reanalysis data is used to construct the first guess and its error covariance. However, since the spatial resolution of various reanalysis data differs, it is difficult to directly compare how the quality of each reanalysis effects the accuracy of the precipitation estimates. We added discussion in Section 4 on this issue.
R1-G2. Reasons of the improvement in this paper is in the statistical base. The statistics changes depending on the samples derived from the areas and periods. However, the treatment of target areas or periods changes depending

on the chapters, or even they are not clearly explained. It looks like that the samples may prepared as author's subjective convenient. Consistent data process, such as the same area with same duration, is required for daily and monthly analysis.
Response: Thank you for your comment. Firstly, please see the response to R1-G1 about the target period for the validations. Secondly, we would also like to clarify that the target areas for the validations are unified according to the type of reference data. For the validations against the AHPRODITE product, we used the samples in Monsoon Asia (MA) (Results of Fig. 4 and 8). This is because of the data coverage of the MA area APHRODITE product, and we did not intend to limit the target area subjectively. We used the APHRODITE product for MA area, because this product contains particularly dense rain gauge data independent from those in the CPC product. We rephrased some descriptions in the first paragraph of Section 2.2.1 to clarify this point. For the validations against the GPCC product, we used the samples in the global area (Results of Fig. 5, 6 and 9). However, in Fig. 9, we only showed the results over Asia and Africa to make it easier for the readers to pay attention on the areas which are focused in the discussion. To prevent the readers' misunderstanding, we rephrased the explanation about the computation of the temporal MAD (Equation (16)) and added the results of the temporal MAD for the global area in Appendix 6.
R1-G3. Most of the gauge observation has been conducted in the valley or basin where people live even in the mountainous regions, and interpolation of those gauge-based network is hard to provide unique signals in high-elevations. It is the same situation for gauge space areas, because interpolation can not produce no data areas's information. The APHRODITE is the same condition. Besides, numerical model providing the reanalysis data is expected to reproduce precipitations (not as interpolation). Direct measurements by satellite-based radar observation, such as TRMM or GPM-PR, are also expected to provide the signals, however, gauge adjusted micro-wave satellite products (such as GSMaP-Gauge) intentionally filter out the important remote signals. If the paper would like to insist that new products are beneficial for mountainous or rain-gauge-sparse regions, please show the results of comparisons with gauge data locating in the high-mountains or remote areas not included in APHRODITE or CPC networks. Also, you intended focusing on specific areas, such as Himalayas,

Zagrous mountiains, South-east Asia and central part of Africa, but they are not "mountainous or rain-gauge-sparse regions" of all. I would like to advice to exclude the sentences of beneficial/improvements of new data in "mountainous or rain-gauge-sparse regions" from the abstract and conclusion. Or you should mention that as "algorism worked better especially in Himalayas, Zagrous mountiains, South-east Asia and central part of Africa" with adequate reasons.

Response: Thank you for your valuable suggestion. We would like to clarify that we are using only the rain gauge observations included in the APHRODITE or GPCC product for the validations in our study. Furthermore, we excluded the reference samples when a CPC observation input was located at the same pixel to make the reference data independent from the observations used for estimation. Thus, as far as we understand from the samples' location in Fig. 8 and 9, we have independent reference data even in mountainous or rain-gauge-sparse areas.

Moreover, we agree that it is strange to discuss generally about mountainous or rain-gauge-sparse areas despite only showing the results in specific areas. Therefore, we added the results same as in Fig. 9 for the global area in Figure A5 (Appendix 5). In this figure, large improvement (i.e., red colors in Fig. A5 c) in the estimates can be seen also for around the Andes Mountains (40S-60S), although the entire mountain region is not covered in the validation. Moreover, despite some degradations seen in the surrounding areas, improvements can be seen around the Alps and Scandinavian mountains (especially in the highest region around 60N, between Belgen and Oslo) as well.

R1-G4. Composition of the chapters need to be revised again. In the Section 1, reviews need to derive the issues to be challenged, and clear objectives should follow. If the "Estimation" is your objective as in the title, you need to specify not only for the target periods/areas but also describe for "which purpose". As there are already so many precipitation products (Sun et al., 2018), you may want to demonstrate the efficiency of LETKF with ERA5 data. Then it is better to modify the title, and add physical explanation of why the LETKF could improve the biases in the conclusion, not only showing the statistical number. In that sense, analysis procedure explained in the Fig.1 caption should be done in the contents with more polite ways. Discussion of chapter 4 need to be done with challenges described in chapter 5. Besides, important results such as the performance of LETKF is needed in the conclusion more.

Response: Thank you for the important comments.
Firstly, we clarified the objective of our study and added explanations to support them. We mentioned that ”we focus on proposing an improved interpolation method to obtain rain-gauge-based historical global precipitation fields”, and also added description that “rain-gauge-based precipitation data would be especially beneficial for historical periods with few satellite observations available”. Moreover, we added description about the limitation of the rain-gauge-based product from CPC (CPC_est) based on a reference to emphasize the significance of our objective.
Secondly, we believe that we have explained the benefit of utilizing the NWP-based ERA5 data with the use of ensemble data assimilation (EnDA) in the third paragraph in Section 1. We have also explained the benefit of using the LETKF algorithm among various (EnDA) methods in the same paragraph. On the other hand, we agree that we did not sufficiently discuss the computational efficiency of the LETKF algorithm, and therefore added description on it in Section 4 (discussion) and 5 (conclusions).
Thirdly, we have also already explained about the caption of Fig.1 in the main text Section 2.1.2 in detail. Since we relocated Fig. 1 at the beginning of Section 2.1.2 as the reviewer suggested in R1-S5, we hope the readers could understand that the contents of the caption in Fig. 1 and the main text Section 2.1.2 matches well.
Finally, we agree that the limitations addressed in Section 5 should be discussed in Section 4. We have added discussions accordingly, and instead simplified the description on the limitations in Section 5.

Specific comments

R1-S1. Title: Better to mention the appeal terms, such as LETKF, improvements, assessments,,

Response: Thank you for your comment. The authors think that the biggest originality of this study is utilizing information from reanalysis precipitation for interpolation by applying a local ensemble data assimilation method. Therefore, we changed the title as follows:
“Estimating global precipitation fields by interpolating rain gauge observations using the local ensemble transform Kalman filter and reanalysis precipitation”

R1-S2. Abstract: Better to mention the reason of why the LETKF could improve the products according to the comparison of CPC_est.
Response: We added the explanation that the LETKF is a computationally efficient ensemble data assimilation method in the second sentence of the abstract. The reason for the improvement in the estimates due to the use of reanalysis data is described in the second last sentence (the advantage of constructing a physically guaranteed first guess and its error variance).
R1-S3. L56 Please rise the issues that previous study did not archive. Then, describe why you need new methods, for which areas/period for your target of estimation.
Response: Please see the response to R1-G1, 2 and 4.
R1-S4. L57 Clear objectives are missing. You need to set them according to the key conclusion (Chapter 5).
Response: Please see the response to R1-G4.
R1-S5. L66 Readers can not handle why the Fig.1 appeared suddenly without explanations. Move the Fig. 1 in Section 2.1.2.
Response: We moved Fig. 1 to the beginning of Section 2.1.2.
R1-S6. L68-74 Fig.1 caption includes study methods to be written in the main contents.
Response: We agree that the caption of Fig. 1 describes the methodology of this study, which is also repeated in the main text in 2.1.2. However, we would like to keep the description in the caption of Fig. 1 to enhance the readers' understanding of this figure.
R1-S7. L76 Need to explain why the CPC_est is the target of comparison.
Response: Thank you for highlighting this point. We used the same rain gauge observations as used in CPC_est in our study (and proposed an interpolation method different from the one used for CPC_est), since we wanted to compare the interpolation methods of CPC_est and our study. We added a sentence at the end of this paragraph to clarify this point.
R1-S8. L77 "Daily" mean 24 hours from 0UTC ? The original daily CPC data were not local time coordinate?
Response: Although CPC_est defines the daily precipitation by local time, we assume that the daily precipitation in CPC_est represents the 24-hour precipitation from 00:00 UTC, provided that open information on the local time used for each pixel is limited and inaccurate. We added this explanation in the first paragraph of Section 2.1.1. We also added explanation that the daily

precipitation of ERA5 is computed from the 24-hour precipitation from 00:00 UTC.
R1-S9.L78 CPC archive does not limit to the US. Please clarify the target areas/periods of your estimation here. Maps of Fig.1 includes north/south America and Australia, but you omitted them later.
Response: We have the same understanding that CPC archives do not limit to the U.S., but we intended to list some examples of the data source of the rain gauge data in the CPC product. We rephrased the commented sentence to prevent the readers' misunderstanding. Moreover, we modified the explanation on the number of stations used for the CPC product based on the latest information based on NCARS (2022). As we responded to R1-G1, we did not omit North/South America and Australia from the target of our study.
R1-S10. L80 You did not estimate the grids without gauge sites, then mask the grids in the following maps. If the multiple gauge station existed in a pixel, did you assume them in the same location in the 0.5 degree scale?
Response: We intended to explain that we only used the rain gauge observations used in CPC_est (and not the whole interpolated precipitation field of CPC_est) as the observation inputs in our estimation. We added the term "for the observation inputs in our estimation" in the commented sentence to clarify this point. Furthermore, as the referee mentioned, we assumed that all the rain gauge stations existing in a same pixel to be located at the center of that pixel. This is one limitation owing to the data availability. We added a description about this issue in the second paragraph of Section 4.
R1-S11. L85-88 I can not understand ",, over land, where rain gauge observation are available". How did you adjust 0.5 interval CPC_est with 0.25 interval ERA5 data? The same expression at L166 " converted".
Response: Response to the 1st sentence: In the data assimilation algorithm, we map the first guess to the observation space using the observation operator (Please see Equation (1) and the following paragraph). Therefore, the spatial resolution of the observation inputs do not need to be equivalent with that of the first guess. Since we assume that the observation sites are located in the center of the 0.5-degree pixels, each observation site exactly corresponds to one 0.25-degree grid point of the first guess. Hence, the observation operator $H_t()$ is simply a linear

function that extracts the first guess data at grid points where the observation exists, and $\mathbf{H}_t$ is equivalent to $H_t(\,)$. We added this explanation to the corresponding paragraph. Response to the 2nd sentence: We added the explanation on the method for converting ERA5 and LETKF_est precipitation prior to the validations in Appendix 4.
R1-S12. L100 Why the (2) was classified at 1mm/d ? No precipitation (0mm/d) is always log(2)?
Response: Thank you for your comment. As we describe in Equations (7), (9) and (10), the LETKF algorithm uses the inverse of the observation error covariance matrix $\mathbf{R}_t$ with localization to compute the precipitation estimates. Therefore, it is necessary to set a minimum limit to the error variance so that the computation would not diverge when we take its inverse. We added a sentence after Equation (2) to explain this issue.
R1-S13. L103-104 Why the "the data of the 10 years before and after the date", "surrounding 7 days for ,," ? Again, your target of study period is not clear, so I can not understand why you intended to do so.
Response: We extracted the ERA precipitation 10 years before and after the estimation date, considering that CPC_est uses the 20-year average daily precipitation as the first guess of estimation (Xie et al., 2007). Moreover, we further extracted the data of the surrounding 7 days, so that we can construct an ensemble that represents the daily climatology of the estimation date. Although we also tried constructing an ensemble from the surrounding 15 days, we eventually used the surrounding-7-days ensemble, because it had small difference with the surrounding-15-days ensemble. We added some descriptions to explain the reasons for "10 years" and "7 days" in the paragraph before Equation (3).
R1-S14. L105 It is better to divide the Fig. 1 in two, and lower part should be cited here as Fig. 1b.
Response: Revised.
R1-S15. L106 Section 2.2.1 is about the comparison for case study day. Did you perform the comparison only in the case day or multiple years? Reader can not understand the detail evaluation methods.
Response: Please see the response to R1-G1.
R1-S16. L115 Formula (5) is your original?

Response: Sorry for the lack of explanation. We added two references which Equation (5) is based on.
R1-S17. L117 "Observation site" is the location of CPC observation site used to make CPC_est? You mentioned that location of the gauge is set at a pixel (L80), so it is not clear the meaning of d (distance). Meaning of "analysis grid point" is also unclear. Is this about the ERA5 grid? Please also revise English sentence.
Response: We added explanation that we are using the assumption that all the observation sites are located at the center of the 0.5-degree pixels. Additionally, we rephrased the term "analysis grid point".
R1-S18. L124 "author's preliminary experiments" need additional explanation or citation. Some constants, such as 1000km or 10, many have meaning according to the study target.
Response: We added explanation about the results of our preliminary experiments related to the localization parameters in Appendix 2.
R1-S19. L151 Why the Fig. 2a and 2b are different areas? Still not clear the target areas of your estimation. Are you interested in Asia for daily base and global scale for monthly base? Better to unify the map (and analysis) areas. As the precipitation intensity distributions are depending on the climate (areas), following statistic (such as shown in Fig.4, 6) may change depending on the target areas.
Response: Fig. 2 (a) only covers the Asian countries, because the Monsoon Asia (MA) APHRODITE product is limited to this area. On the other hand, we would like to evaluate the precipitation estimates globally in a daily basis. Considering that the Monsoon Asia (MA) APHRODITE product has a limitation in area, and the GPCC product has a limitation in the temporal resolution, we performed validation against both data. We clarified the limitations of these products used for validation and the reason for using both data in the third paragraph of Section 2.2.1.
R1-S20. L144-149 APHODITE and GPCC were utilized in different concepts. Former data is very dense and used for hydrometeorological sense, and latter data is long and used to evaluate historical climate change. The daily biases are evaluated in local time base, and monthly biases are evaluated by subgrid scale spatial average. Such background should be referred in Section 1. Then, please clarify which kind of time scales you want to "estimate" ?
Response: We would like to clarify that only the rain gauge observations from the APHRODITE and GPCC products are used as references for validation.

(Please find the last paragraph in Sections 2.2.1) We do not intend to evaluate our precipitation estimates against the whole interpolated data of these products. We rephrased the first sentence in Section 2.2.1 to make this point clear.
R1-S21. L148 APHRODITE and GPCC may include the data by GTS, so they are not "independent".
Response: We intended to explain that we would be able to use independent rain gauge observations by extracting rain gauge observations which are not provided in the CPC product (as we describe the methods in the last paragraphs in Sections 2.2.2 and 2.2.3). We rephrased the sentences with the term "independent" not to cause misunderstanding.
R1-S22. L149 I do no think "dense rain gauge" in any MA regions. Again are you interested in the estimation of monsoon Asia? The Gauge observation is much dense in UA, Europa, Japan,, (Fig. 2b). Why you did not avoid those areas? Also, which periods of comparison with APHRODITE?
Response: We intended to explain that the rain gauge observations (which are independent from those in the CPC product) are especially dense in the MA area for the APHRODITE product. We rephrased this sentence to avoid misunderstanding.
R1-S23. L165 Again, you assumed the location of gauge at the center of pixel (0.25 or 0.5 grids), but considered the distance (d) between the grid point and observation site (gauge location) at L114, making confusion. In the later chapters, orographic effects are discussed, but such assumption (location of gauge = the center of pixel) do not affect for your interpretation?
Response: First, please see the response to R1-S17. Second, as the reviewer suggested, there is a possibility that the assumption on the gauge location may drown out the orographic effects in scales finer than 0.5 degrees. We added this description in the discussions section. On the other hand, provided the results in Fig. 7, 8 and 9, we believe that it is clear that the estimation of our study improves the estimates around large-scale mountains such as the Himalayas of the Zagros mountains.
R1-S24. L172 "to be biased" which kind of biases? Then, why you choose rank correlation coeffect? Do you want to improve the identification of extreme events, not the absolute amount?
Response: We revised the commented sentence so that the possible bias in the APHRODITE product is explained more in detail. We use the rank correlation

efficient as an evaluation index, because it is a measure based on the rank of the samples rather than the exact magnitude of them (Please see the former part of the commented sentence), which means it is less affected by the bias of the reference data. We rephrased some terms in the explanation of Kendall’s rank correlation coefficient to enhance the readers’ understanding.

R1-S25. L183-186, L200-203 These parts should be explained before.

Response: Thank you for your comment. However, we would like to not change the orders of the commented sentences, because we consider that it is appropriate to describe the characteristics of each evaluation index after their definitions.

R1-S26. L142 “2.2 Validation” This chapter would be in the “3. Result” of your analysis.

Response: We consider that it is appropriate to explain the validation methods in the methods section before the results section. On the other hand, we changed the title of section 2.2 to “Validation methods” to clarify that we only explain the methods for validation here.

R1-S27. L185 “APHRODITE < 0.5 mm/d is excluded”. Your statistics exclude the non-rain days. Please mention clearly in advance. This is not the matter of rage accuracy.

Response: Thank you for pointing out this issue. We agree that we exclude the no-rain days by excluding samples where APHRODITE_gauge is < 0.5 mm day^{-1}. On the other hand, we consider that no-rain days are not included in the computation of the rank correlation coefficient regardless of this exclusion (Because the computation of the rank correlation coefficient neglects the samples with the completely same values in APHRODITE_gauge (or in the precipitation estimates) (please see the sentence after Equation (12), and because there are more than one no-rain cases in APHRODITE_gauge or the precipitation estimates). Hence, we clarified that the computation of the rank correlation coefficient itself excludes no-rain days.

R1-S28. L215 Why you chose the old post-monsoon month in both hemisphere (1988, Nov. 15th) that may also miss heavy precipitation events? As you evaluate the difference as ranking correlation (L174) and would like to discuss the orographic enhancement (Fig. 7), the day should be in summer. Why the legend is exponential without color?

Response: The date Nov. 15th, 1988 was chosen randomly from the estimation period of our study (from Jan. 1st, 1981 to Dec. 31st, 1990). The validations are

performed using not only the data of this single time step but of the whole estimation period (Please see the response to R1-G1). We used the exponential color bar because it is easier to understand the difference in the order of the precipitation amount. Additionally, we selected this color bar which allows universal readability following the instructions from the Copernicus manuscript templates.
R1-S29. L211 I could not understand "broader precipitation areas". Where is the Himalayas and Zaguroud mountains? Please mention in the map.
Response: Although we considered mentioning the Himalayas and Zagros mountains in the map as the referee suggested, we decided not to, because adding signs on the map in Fig. 3 would hide the precipitation patterns in some areas, making it difficult to capture the overall precipitation field.
R1-S30. L220 The sample of the Fig.4 is not clear. Is this from one day distribution on Fig. 3 or from certain periods. Why you limit the areas in Fig. 2a? You will discuss the signals in Africa later on (Fig. 9).
Response: We used the samples of the whole estimation period in the MA area (Please see the response to R1-G1, R1-S19). To enhance the readers' understanding, we added the explanation that the validation method using the Kendall's rank correlation coefficient is described in Section 2.2.2. Fig. 9 is the results of the validation against the GPCC product, so we performed it for the whole global area.
R1-S31. L224 Although the correlation coefficient is highest, is it significant? Please show the statistical significance.
Response: Thank you for your comment. The τ_b value of LETKF_est has statistically significant differences between those values of ERA5 or CPC_est with the P-value of 0.01. This description was added to the manuscript.
R1-S32. L234 Monthly comparison in Fig. 5&6 was done in which areas, in MA or global scale? If it is in global scale, North/South America is included and why it was different from the comparison area of APHRODITE?
Response: The validation results in Fig. 5 and 6 are performed in a global scale. To enhance the readers' understanding, we added the explanation that the validation methods against the GPCC product are described in Section 2.2.3. Please also see the response to R1-S19 explaining the difference in the target area for validation.
R1-S33. L257 There are many kind of dynamics of orography affecting precipitation system (Houze, 2012). Please explain why you assume the first

guess could take into account the orographic effect? Please explain somewhere in the paper.
Response: As described in the first paragraph of Section 4, we presume that we could take into the account the orographic effects owing to the interpolation method that uses the dynamically guaranteed first guess and background error covariance constructed from the ERA5 data, which is based on numerical weather predictions.
R1-S34. L260 Why you choose the day of Jun. 7th, 1985 in MA? The date is old and different from Fig. 3. The feature of "reproducing the orographic changes in precipitation" was also confirmed by other days?
Response: We chose Jun. 27th, 1985 to show an example on a date in the Monsoon season within the estimation period. The results in Fig. 8 and 9 were computed using samples from the whole estimation period (from Jan. 1st, 1981 to Dec. 31st, 1990), confirming the reflection of orographic effects for other days as well. To enhance the readers' understanding of the investigation of Fig. 7 and 8, we added explanation to the following sentences 1. Last sentence in the first paragraph in Section 4. 2. First sentence in the third paragraph in Section 4.
R1-S35. L261 Monsoon rain along the Himalayas dominates in the night (e.g. Sugimoto et al., 2021). So, you mean that your algorism work for the nocturnal rain? Orographic ascending type precipitation along the Ghats mountain range was reproduced in both products (Fig. 7e&f). Is this consistent with your idea? Please explain the consistency if you would like to mention "LETKF succeeded in reproducing orographic changes in precipitation".
Response: Thank you for the important comment. As the reviewer suggests, the precipitation along the Ghats mountains are reproduced in both CPC_est and LETKF_est (Fig. 7 e, f). However, we presume that CPC_est succeeded in reproducing the precipitation along the Ghats mountains because there are some observations showing high precipitation around this region in CPC_gauge (Fig. 7 c). On the other hand, LETKF_est also succeed in reproducing the precipitation along the Himalayas despite the lack of observation inputs around this region. We added explanation to the commented paragraph to clarify this point.
R1-S36. L270 Is this June 7th or 27th? This map is also different from MA (Fig. 2a). Your comparison changes areas/periods according to your interests. I hope the

analysis in the same areas because your results are depending on statistical evidences.
Response: June 27th is correct. Sorry for the mistake. We corrected the main text. Although we focused on the precipitation around South Asia in Fig. 7 as an example, the target area for the validations against the APHRODITE and GPCC products are unified to the MA area and the global area, respectively. We added explanation to clarify the target areas.
R1-S37. L276 Only one case does not to fit to mention "as the Himalayas in general".
Response: Please see the response to R1-S34
R1-S38. L277-278 Need to mention the sample periods for Fig. 8. I could not understand how one pixel could get more than 1800 samples.
Response: We added explanation on the target period (the whole estimation period from 1981 to 1990).
R1-S39. L279 Please show the way of statistical significance if you insist "significantly".
Response: We rephrased "significantly" to "largely".
R1-S40. L299 I can not see grey pixel.
Response: The authors have re-checked that the pixels where the validation weren't performed were colored in gray appropriately in Fig. 8. We may be misunderstanding something, so could the referee comment again in detail if there is any further problem?
R1-S41. L282 Please explain the meaning of "samples". Is this months, then which period?
Response: We rephrased "time steps" to "monthly time steps", and also added explanation on the target period (the whole estimation period from 1981 to 1990).
R1-S42. L285 Here you mentioned "temporal MAD", but formula (14) defined the spatial MAD. Please explain the difference.
Response: We added explanation on the difference between "spatial MAD" and "temporal MAD" after Equation (16).
R1-S43. L286 Figure 9d-f covers Africa. Do you also want to estimate the precipitation in Africa? Please explain the reason of area extension.
Response: We added the results of the temporal MAD for the global area in the Appendix 6. Please also see the response to R1-G1 and R1-S36.

R1-S44. L287 "methods is beneficial for those areas in general" mean your methods works especially in the Himalayas and Zagrous Mountains or mountainous areas in general? Why in general?
Response: We intended to describe that our method is beneficial for the Himalayas and Zagros Mountains throughout the estimation period, provided the validation results using the whole period. We rephrased "in general" to "throughout the estimation period".
R1-S45. L288 "gauge stations are especially sparse, such as South-east Asia and central part of Africa" Such crude descriptions should be avoided. Where is the central part of Africa? There are dense gauge networks even in Asian countries.
Response: We rephrased "central part of Africa" to 0˚S–20˚S of Africa" for clarification. Additionally, we added the term "some regions in" before "South-east Asia" to prevent the readers' misunderstanding that rain gauge networks are sparse in whole South-east Asia and Central Africa.
R1-S46. After L253 Chapter 4 and 5 must be revised carefully after the revision according to the former comments. Exclusion of North America, Australia and Arabian Peninsula is excused in the ending part of the conclusion; however, this way is very strange. You need to mention the target areas in the beginning with reasons.
Response: Please see the response to R1-G1.

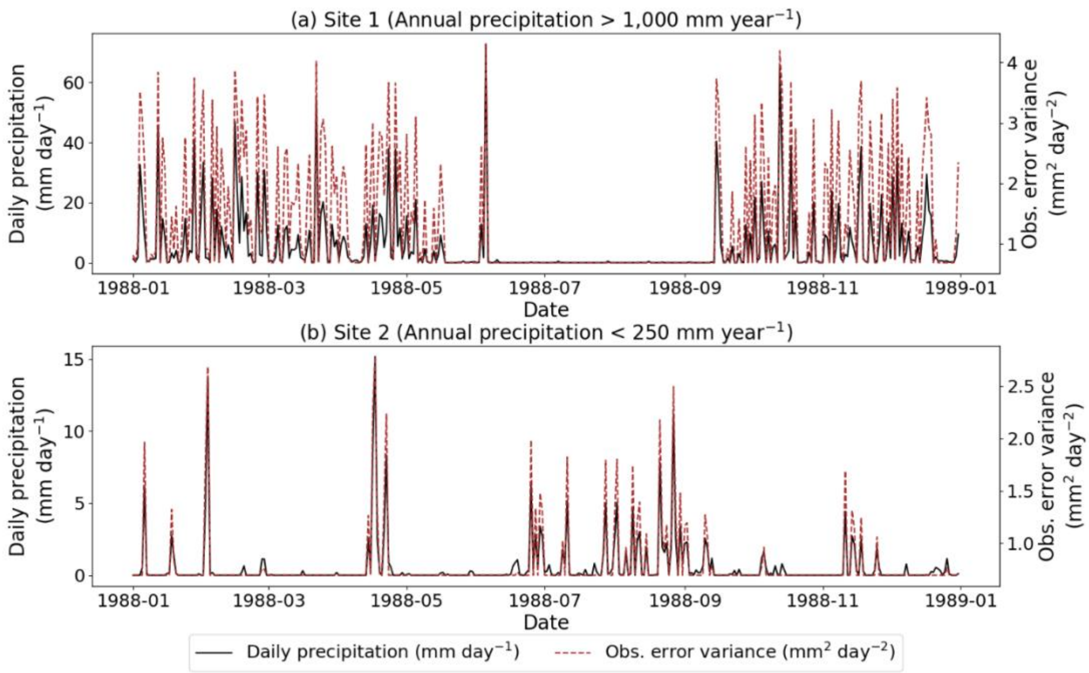
===

Referee: 2

Specific comments

R2-S1.The title should be revised to incorporate the ERA5 dataset, given its large contribution to the improved estimate, if the reviewer understands correctly from the author. As such, the reviewer wonders if the method proposed by the author can also enhance the precipitation field on recent periods, furthermore, the other fields (e.g., soil moisture) from ERA5 datasets? Could the author provide a brief discussion on the applicability of this method in the "discussion" section?
Response: Thank you for your comment. We agree with the reviewer that the biggest originality of this study is utilizing information from reanalysis precipitation for interpolation by applying a local ensemble data assimilation method. Therefore, we changed the title as follows: "Estimating global precipitation fields by interpolating rain gauge observations using the local ensemble transform Kalman filter and reanalysis precipitation. Moreover, we agree that this estimation method could by applied for other variables, although the applicability may by largely affected by the accuracy, frequency and spatial density of observations. We added this description in the discussions section.
R2-S2. In the second paragraph of "Abstract", could the author add the results of comparisons to the APHRODITE and GPCC products to support the author's demonstrations that the method of this study is shown to be particularly beneficial for mountainous or rain-gauge-sparse regions.
Response: The description about the validations performed for each pixel is added to the last sentence in the abstract.
R2-S3. In section 2.1.2 regarding to LETFK, the author mentioned that the parameterization of error covariance of observations are based on preliminary sensitivity experiments. Could the author briefly introduce the experiments? The corresponding result can be placed in the supplementary materials. Additionally, the review also suggest making plots of spatial and temporal error distribution (may put it in the supplementary materials), so the reader can further understand the observation error better and evaluate the improved estimate.
Response: We used the logarithm transformation equation to compute the observation error variance based on Lien et al. (2016)'s equation. In addition, we

performed sensitivity experiments for a coefficient that multiplies the logarithm transformed value in Equation (2), and consequently the value 1.0 was selected as the coefficient (i.e., equivalent to placing no coefficient). Such description is added to the explanation of Equation (2) in Section 2.1.2.
Moreover, we added a figure that depicts the spatial distribution of observation errors for a specific day in Appendix 1. We also depicted the time series of the observation errors for two specific sites (Sites with high and low annual precipitation amount, respectively), as in the figure below. However, we would like skip posting this figure in the appendix, because we consider that the figure in Appendix 1 can sufficiently show the change in the observation errors depending on the observations.



R2-S4. Clarification is needed in section 2.1.2 regarding whether parameter values such as 10 days, 7 days, 2*sqrt(10/3), 1000 km and 10 are optimal for this case or are they generic values in widely-used sense? It would be valuable to discuss the sensitivities of data assimilation results to the variations in these parameters in the "Discussion" section?

Response: As for the "10 years" and "7days" used for constructing an ensemble, please see the response to R1-S13 .
Equation (5) (the localization function) was defined based on previous studies. We added two references which Equation (5) is based on.

In addition, we added explanation about the results of our preliminary experiments related to the localization parameters in Appendix 2.
R2-S5. In section 2.1.2, could the author cite the source for Equation (5) and give a more detailed explanation? The review suggest including figures depicting temporal and spatial distribution of the localization function L(d) for reference in the supplementary materials.
Response: Sorry for the lack of explanation about Equation (5). We added two references which Equation (5) is based on. Additionally, we added a figure which depicts the Localization function L(d) depending on the distance of a grid point and an observation site with different localization scales σ in Appendix 3.
R2-S6. The author declared that the orographic effects considered in the EAR5 results in the superior performance of the analyzed precipitation on the mountainous regions. It would be better if the author add a short description of the interpolation method and ancillary data (e.g., especially whether the elevation data is included) used in the interpolation of CPC, GPCC and APHRODITE products, as such, the reader can get insights on author's declaration.
Response: We added a brief explanation about the interpolation method of the CPC product in the "Discussion" section. Since we only used the rain gauge observations included in the GPCC and APHRODITE products, we would like to skip the details of the interpolation methods used for these two products.
R2-S7. The reviewer proposes integrating the 'Discussion' section into the "Results" section, as it shows the results of comparisons between LETKF_est with the existing datasets. Furthermore, the content in the third paragraph in the "Conclusions" section could be discussed in the 'Discussion' section in a more detailed way.
Response: Thank you for your suggestion. Firstly, we considered intergrating the "Results" and "Discussions" sections as the referee suggested, but we would like to leave these sections apart, because the "Results" section focuses on whether the estimates of our study could outperform the CPC estimates, where as the "Discussion" section focused more on the reasons why our estimates were evaluated to be better. Secondly, we reconstructed the contents in the "Discussion" and "Conclusion" sections based on the referee's suggestion.

==

Referee: 3

We would like to thank Referee 3 for all the editorial remarks. We have revised all the editorial mistakes that the referee pointed out, with one exception below:

Editorial remarks

R3-E24. line 275-276: I would phrase it differently as the Authors analyse one specific example of the LETKF_est in the mountainous areas, even if it is a significant one. I would rather say: 'Using the examples of the Himalayas, we investigate whether the precipitation of LETKF_est is more accurate than that of CPC_est around mountainous areas' or something of that sort
Response: Thank you for your suggestion. We agree that the commented sentence includes an ambiguous expression, and therefore rephrases the term “in general” to “not only on a specific date but for the whole estimation period”. On the other hand, considering that we have performed additional validations using the whole area coverd in APHRODITE_gauge and GPCC_gauge, we conserved the basic structure of the commented sentence.

Here, we give response to all the specific comments.

Specific comments

R3-S1. In the introduction, possibly in line 26, right after the first sentence of this section I propose to insert an additional one emphasizing the importance of global gridded precipitation fields for validation of data assimilation (for example data assimilation of space borne lightning observations would lead to forecast precipitation fields that one may want to compare with gridded precipitation fields) as well as climate studies.
Response: Thank you for your suggestion. We added descripitons emphasizing the issue described above in the first sentence of Section 1.
R3-S2. Line 52 in the introduction: It is not clear to me how EnDA is used to obtain climatological covariances. Based on caption in Fig.1 would you say that your covariances are climatological but date-specific. Maybe 'daily climatological covariances'?
Response: We rephrased the term “climatological” to “daily climatological”. As the referee says, the first guess and its error covariance are date-specific, and we

agree that the term “climatological” may be misleading.
R3-S3. Line 115: I am puzzled by sqrt(10/3) in the formula expressing L(d)? Is there a particular reason for such a choice? I mean, I would understand 3 sigma, or possibly also 2 sigma, but I am puzzled by the value of the constant the Authors used. There is essentially nothing wrong with such a choice if it serves the purpose but I have been wondering if there was a justification for it.
Response: We apologize for the lack of explanations. Equation (5) (the localization function) was defined based on previous studies. We added two references which Equation (5) is based on.
R3-S4. Section 2.2.2. Could the Authors elaborate on how the Kendall's coefficient has been computed? What criteria do you use to rank the precipitation fields in the 3 analysed data sets and in the APHRODITE_gauge? When you talk about concordant/discordant correlations do you mean their sign?
Response: u_i and v_i (in the sentence just before Equation (11)) denotes the daily precipitation samples in the precipitation estimates and APHRODITE_gauge, respectively. In the computation of the Kendall's rank correlation coefficient, we count the number of cases where the magnitude relationship of u_j and u_k is concordant (or discordant) with that of v_j and v_k. Consequently, the Kendall's rank correlation coefficient measures how similar the magnitude relationship of the samples in the precipitation estimates are to that in APHRODITE_gauge. We rephrased some terms in the explanation of the Kendall's rank correlation coefficient to enhance the readers' understanding.
R3-S5. Line 212: regarding Zagros Mountains, looking at your plots, it seems to me that you do not really have rainfields reconstructed in this particular region, which is between the Caspian Sea and the Persian Gulf. I think that the coloured region on your map is in Kazakhstan, Uzbekistan and Turkmenistan (Turan Depression?). I also think that your differences with respect to GPCC are more significant over South America. By the way, Zagros Mountains are correctly identified in Fig.9
Response: Sorry for the misleading description. We intended to explain that precipitating areas can be seen around the Himalayas and the Zagros mountains in LETKF_est, while not in CPC_est. around the Himalayas and the Zagros mountains in LETKF_est, while not in CPC_est. Since this is a different situation from Africa, the Indochina Peninsula, and South America (as the referee suggested), we divided the sentence into two to explain each situation.

R3-S6. Section 3, line 228 in the caption of Fig. 4: What does the ratio really represent? I mean what do the bins refer to?
Response: The bins refer to 0.1 mm day^{-1} × 0.1 mm day^{-1} bins in a 2-dimensional histogram of daily precipitation of the two compared datasets. We added description in the caption of Fig. 4 to enhance the readers' understanding.
R3-S7. Figs 5 and 6: I do not want to add to your work but since GPCC is a reference, it would be better to name the plots CPC_est vs. GPCC and LETKF_est vs. GPCC. In Fig. 8 I would also rather say CPC_est versus APHRODITE_gauge and LETKF_est versus APHRODITE_gauge. Not sure if you need to replot the figures for that and how difficult it is for you at this stage.
Response: We modified the titles and legends in Figures 5, 6, 8 and 9 upon the referee's suggestion.
R3-S8. Line 250, 251: Figure 6 legend and caption. There is an inconsistency between the legend: *dark-red circles represent low latitude* and the caption: *dark-red circles represent mid- and high-latitude regions*. In addition, please also check if the statements in lines 236-241 are correct
Response: We apologize for the mistake. The legend in Fig. 6 is correct. We modified the caption.