GPROF V7 and beyond: Assessment of current and potential future versions of the GPROF passive microwave precipitation retrievals against ground radar measurements over the continental US and the Pacific Ocean

Response to reviewers

1 Response to reviewer 1

We thank the reviewer for investing their time to read our manuscript and provide constructive feedback to improve it. The principal changes that we have implemented based on comments from both reviewers are the following:

1. We repeated our analysis excluding all precipitation classified as frozen by MRMS, i.e., excluding not only snow but also hail. Significant effects of excluding hail were only observed in the SW region and we have updated the manuscript accordingly.

2. We have added a short section that analyzes the impact of excluding both frozen precipitation as well as precipitation of snow-covered and mountain surfaces and the behavior of the different retrievals over those surfaces.

3. We have revised all references to the a priori database and the MRMS validation data to make the distinction of the two clearer.

Finally, we have also updated that precision-recall curves to show the precision over the recall instead of the recall over the precision, which is the more common way to present these curves. During this, we also realized that the PR curves previously included samples over ocean, which we have now excluded. This lead to a minor change in the relative skill of the GPROF-NN HR retrieval.

In what follows, line and figure numbers are given with respect to the revised manuscript.

1.1 Principal comments

Reviewer comment 1

The whole validation is based on 'validation measurements' and 'reference measurements', them being ground-based radars and GPM CMB precipitation retrievals respectively. In section 3.1.1 sometimes there is a bit of confusion on how the different databases are addressed. Line 226 has 'reference precipitation', line 234 has 'retrieval database', line 270 has 'database', line 281 has 'a-priori database', line 284 has 'database precipitation' etc. I suggest to get a bit of consistency since the whole validation is based on different but very similar databases (GPM CMB 2019 is the a-priori database, GPM CMB for other years is just a comparison database etc.).

Author response:

We thank the reviewer for pointing this out. Although we did make an effort to apply language that makes the different between the two datasets clear already in the initial

version of the manuscript, we apparently fell short of this goal. We have revised the references to the two datasets in section 3.1.1 and the rest of the manuscript in an effort to make the distinction clearer.

Changes in manuscript:

1. We have revised Sec. 3.1 and the rest of the manuscript taking care to always clearly state whether we are referring to the a priori database or the MRMS validation data.

2. We have revised all figures to consistently use 'A priori' or 'MRMS' to refer to data from these two sources.

Reviewer comment 2

In section 3.1.2 you compare different regions and some of the explanations for high biases are attributed to winter precipitation (see line 331 for example). I am a bit confused on how you are dealing with winter precipitation since snow covered surfaces and MRMS frozen precipitation are excluded from the analysis. Please provide more context on how you analyze winter precipitation in the different regions.

Author response:

Since our analysis excludes all samples over snow-covered and mountain surfaces as well as precipitation identified as frozen according to MRMS also our analysis of winter precipitation only applies to liquid precipitation over snow-free, non-mountain surfaces.

Although, in general, the contribution of the exluded samples to the total precipitation is rather small, it can have a significant impact on seasonal and regional scales, as is certainly the case for the winter in the NW region. Nonetheless, even if the excluded precipitation estimates can make up for the observed underestimation, our analysis still points to differences between the retrievals and the MRMS validation data.

However, we agree with the reviewer that this must be clearly communicated and discussed in the manuscript. We therefore revised the discussion of the regional seasonal cycles and also added a sub-section that discusses the behavior of the retrieval for different surface and hydrometeor types. Furthermore, we revised relevant statements in the abstract and conclusions to clearly state that our validation focused on liquid precipitation over these surfaces.

Changes in manuscript:

1. We have extended the paragraph discussing the regional seasonal precipitation cycles:

   > **Changes starting in line 364:**
   >
   > GPM CMB and GPROF generally capture the seasonal cycles of the regions ac-

> curately. Notable deviations from the MRMS ~~reference precipitation~~ validation data are an underestimation of stratiform winter precipitation in the North-West as well as an underestimation of convective summer precipitation in the South-East. It should be noted that this analysis excludes snow and precipitation over mountains and snow-covered surfaces and thus the reported biases, especially in the NW region and during winter time, may not be representative of total accumulations. However, even if the overestimation of the excluded cases would make up for this underesimtation, these findings still point towards a notable disagreement between MRMS and the satellite retrievals for retrievals of liquid precipitation over snow-free and non-mountainous surfaces.

2. We added the information that our validation focuses on liquid precipitation over snow-free and non-mountainous surfaces to the abstract and the conclusions.

3. We have included a new section that analyzes the general retrieval characteristics for the excluded regimes and analyzes the impact of excluding those samples. This section includes the new figure shown in Fig. 1.1.

4. We have extended the discussion of the limitation of our study, which now also discusses the potential effect of the left-out samples as well as the potential effect the spatial sampling of the validation data in this.

1.2 Other suggestions

Reviewer comment 1

Line 37: 'resolution of 10 km' - given the global nature of IMERG maybe 0.1x0.1deg is more appropriate?

Author response:

This is correct. We will change this in the manuscript

Changes is manuscript:

> **Changes starting in line 38:**
>
> .... satellites and rain gauges to produce the level 3 GPM IMERG product (Huffman et al., 2020), which provides global precipitation estimates at a nominal spatial resolution of ~~10 km~~ 0.1° and a temporal resolution of 30 min.

Reviewer comment 2

Line 171: 'neighboring pixels' - is this the distance from the centers of neighboring pixels?

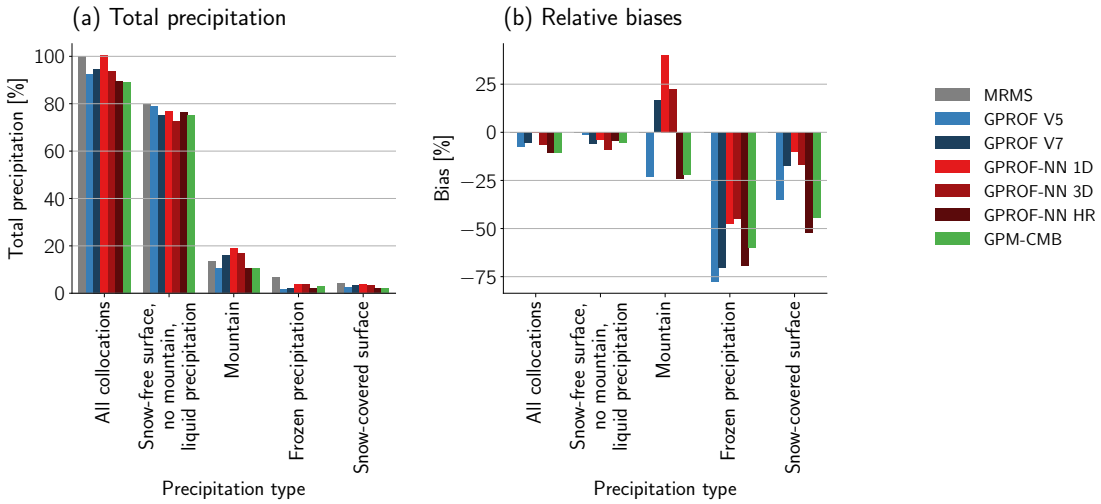

Figure 1.1: Contribution of frozen precipitation and precipitation over snow-covered and mountain surfaces to the total precipitation in the MRMS validation data. Panel (a) displays the total volume of precipitation for different retrieval scenarios relative to the total MRMS precipitation for all validation collocations from the water years 2021 and 2022. Panel (b) shows the corresponding retrieval biases relative to corresponding mean MRMS precipitation.

Author response:

Yes, that is the distance between pixel centers.

Changes in manuscript:

We have included this information in the revised manuscript.

> **Changes starting in line 184:**
>
> The distance between the centers of neighboring pixels of the GMI swath is approximately 13.5 km in along-track direction and 5 km in across-track directions at the center of the swath.

Reviewer comment 3

Line 217: 'conditioned on the validation precipitation' - do you mean the analysis is made only on pixels where it is precipitating according to the validation (MRMS) dataset?

Reviewer comment 4

Scatter plots over log-scales are quite naturally limited to precipitating pixels since they only display the behavior of the retrieval for non-zero a priori or MRMS precipitation

rates. However, the rest of our analysis always includes both raining and non-raining pixels.

To make this point clearer, we reformulate the sentence in question to stress that the conditioning refers to the distributions displayed in Fig. 2 and not the conditioning of the remaining analyses on raining pixels.

Changes in manuscript:

> **Changes starting in line 236:**
>
> ~~Distributions of retrieved and reference~~ The conditional distributions of retrieved instantaneous precipitation conditioned on the ~~validation precipitation~~ corresponding precipitation from the a priori/training database and validation measurements from MRMS are displayed in Fig 2.

Reviewer comment 5

Line 218: 'GPROF a priori database' - since this database is the same as a priori or training for NN, maybe use 'a priori/training'.

Author response:

We have adopted the suggestion in the revised version of the manuscript.

Changes in manuscript:

> **Changes starting in line 237:**
>
> The first column of panels shows the distribution of retrieved precipitation with respect to the GPROF a priori/training database.

Reviewer comment 6

Line 221: can the spread also be due to the preprocessing clustering?

Author response:

Our analysis uses the un-clustered database in order to avoid potential side effects that the clustering could have.

We would like to thank the reviewer for bringing up this point, as it should be made clear in the description of the validation data. We will modify the manuscript accordingly.

Changes is manuscript

Changes starting in line 134:

The GPROF V5 and V7 retrievals cluster the a priori database based on the similarity of the observations in order to reduce the computational complexity of performing the retrieval. However, both the training of the GPROF-NN retrievals as well as the analysis presented here use the unclustered database.

Reviewer comment 7

Line 226: 'conditioned on the reference precipitation' - is this the same as line 217, 'validation precipitation'? As mentioned in comment 1, there is a bit of confusion in the naming of the different datasets used.

Author response:

We have revised this sentence to clearly state that this is with respect to the MRMS validation precipitation.

Changes in manuscript:

Changes starting in line 251:

Apart from this, however, the accuracy of each retrieval conditioned on the MRMS validation precipitation exhibits little inter-annual variability.

Reviewer comment 8

Line 229: 'GPROF V5 is based on a different a priori database' - I suggest to specify that V5 was based on DPR over land and CMB over ocean.

Author response:

That is a very good point, we will include this in the manuscript.

Changes in manuscript:

Changes starting in line 255:

Since GPROF V5 is based on a different a priori database, ~~whose climatology may deviate from the V7~~ which, in contrast to the a priori database of GPROF V7, uses estimates from GPM DPR-Ku over land, this is expected.

Reviewer comment 9

Line 234: 'retrieval database' – which one is the retrieval database? I suppose you are referring to GPM CMB? This should be stated more clearly earlier in the section and be consistent throughout the manuscript.

Author response:

Again, thank you for pointing this out. We have revised the sentence to clearly state that the database we were referring to is the a prior/training database.

Changes in manuscript:

> **Changes starting in line 260:**
>
> Since the ~~retrieval~~ a priori/training database is derived from ~~the GPM CMBretrieval, the GPM CMB retrieval~~ GPM CMB, GPM CMB is practically bias-free compared to the retrieval database.

Reviewer comment 10

Line 244: 'introduced rain gauge correction' – replace with 'introduced by the rain gauge correction'.

Author response:

We have fixed this in the updated version of the manuscript.

Changes in manuscript:

> **Changes starting in line 270:**
>
> Given that the CMB over land is largely a radar-derived product, it is possible that the bias relative to MRMS is introduced by the rain gauge correction applied to the validation data and caused by precipitation properties that may not be resolved by the radar observations.

Reviewer comment 11

Figure 2: I see a very interesting behavior in the low values trend lines. The GPM CMB vs a-priori dataset (which is GPM CMB 2019) have overestimation of the GPM CMB 2019 compared to GPM CMB 'other years', while all the others have the opposite behaviors. Also it looks like the comparison with MRMS 2021 and 2022 shows higher bias for low values. Also the trend for higher values is worth attention. It might be nice to reference this behavior in the section and in the bias description since it provides more information on the range of precipitation that has most issues.

Author response:

The opposing behavior of GPM-CMB and GPROF in the comparison against the a priori/training database, is likely due to the different resolution of the GPROF(-NN) and GPM CMB retrievals. GPM CMB having lower resolution than the precipitation in the a priori database causes and underestimation of light precipitation and overestimation of heavy precipitation.

It is true that there is a distinctive change in the conditional mean for light MRMS precipitation starting in the water year 2021. We suspect that this is connected to the change in the processing of the MRMS measurements that is also behind the change in the biases that occurrs at the same time.

We have revised the manuscript to discuss both of these points.

Changes in manuscript:

1. We have extended the discussion of the retrieval accuracy with respect to the a priori database:

   > **Changes starting in line 238:**
   >
   > The spread in these distributions is due to the limitations of the retrieval method and the ill-posed character of the retrieval and they thus represent the best-case accuracy of ~~each retrieval~~the GPROF and GPROF-NN retrievals. Some spread is observed even between GPM CMB and the a priori database~~, which~~ . This is due to the spatial smoothing applied to the precipitation in the a priori database, which causes precipitation measurements in the a priori/training database to have lower resolution than GPM CMB. The difference in the resolution between the GPROF(-NN) retrievals and GPM CMB with respect to the a priori database also explains the opposing behavior of the respective conditional means at light and heavy precipitation rates.

2. We have added the following sentence to the discussion of the inter-annual variability of the scatter plots:

   > **Changes starting in line 248:**
   >
   > The distributions corresponding to different validation periods are very similar for all retrievals. ~~This indicates that , conditioned on the reference precipitation~~An exception to this is a notable increase in the overestimation of light precipitation after the water year 2020 that affects all retrievals. As will be discussed below, this coincides with a change in the regional biases and is likely due to a change in the processing of the MRMS esimates (Anonymous Referee 2, 2023)

Reviewer comment 12

Line 260: 'conventional GPROF' – both V7 and V5?

Author response:

Yes, conventional GPROF is meant to refer to both GPROF V5 and V7. To avoid confusion, we have replaced 'conventional GPROF' with 'GPROF V5 and V7'.

Changes in manuscript:

> **Changes starting in line 290:**
>
> For the ~~conventional GPROF~~ GPROF V5 and V7 retrievals, correlations are between 0.4 and 0.6 over most of the CONUS.

Reviewer comment 13

Line 270: 'When compared to the database' – which database?

Author response:

Here we mean the a priori/training database. We have updated the manuscript to make this clear.

Changes in manuscript:

> **Changes starting in line 302:**
>
> ~~When compared to the database, all~~ All retrievals exhibit weak biases of the order of a few percent compared to the a priori/training database.

Reviewer comment 14 and 15

Line 294: 'the fraction of confirmed raining pixels among those retrieved as raining' – would this be 'the fraction of confirmed raining pixels in the a priori database among those retrieved as raining by MRMS'?

Line 295: 'the fraction of confirmed raining pixels that are detected by retrieval' – would this be 'the fraction of confirmed raining pixels in the a priori database that are detected by retrieval'? I might have interpreted these last two sentences incorrectly which suggests the importance of clarifying which datasets you are talking about.

Author response:

We have revised the corresponding paragraph to clarify the definition of precision and recall as well as the role of the a priori database and MRMS validation data.

Changes in manuscript:

> **Changes starting in line 326:**
>
> The ~~curves display the trade-off between precision , i.e.,~~ precision of the retrievals is the fraction of ~~confirmed raining pixels among those~~ truly raining pixels and the total number of pixels retrieved as raining~~, and the recall, i.e. ,~~. The recall corresponds to the fraction of ~~confirmed~~ actually raining pixels that are detected by ~~the retrieval.~~ each retrieval. Here, *actually raining* is defined with respect to the assumed ground truth, which is the a priori/training database or MRMS validation data depending on which of the two sources the retrievals are compared against.

Reviewer comment 16

Line 294-295: you talk about raining pixels. So frozen precipitation is excluded also from GPROF? I mean, it makes sense, but you mentioned it is excluded from MRMS earlier in the manuscript and never mentioned what you are doing for GPROF or CMB. I think this is a big point since it eliminates a lot of winter observations that, together with the winter precipitation mentioned in the regional analysis, needs to be clarified earlier in the manuscript.

Author response:

Samples identified as frozen precipitation are excluded from most of our analyses from both the MRMS as well as the GPROF(-NN) results. For the identification of frozen precipitation we rely on MRMS because we deem it to be more reliable than corresponding classification derived from satellite retrievals.

We hope that the changes listed in response to the reviewer's second principal comment make this point clear.

Changes in manuscript:

See principal comment 2.

Reviewer comment 17

Line 308: 'For both the database and the MRMS' – do you mean the a priori database?

Author response:

Yes, we mean tha a priori database. We have rewritten this sentence to make this clear.

Changes in manuscript:

> **Changes starting in line 341:**
>
> The error metrics for the retrievals in the six regions are shown in Fig. 8. For both the a priori database and the MRMS ~~collocations~~ validation measurements, the regional biases are generally larger in magnitude than they are for the full CONUS.

Reviewer comment 18

Figure 6 caption: Panel (a) shows the detection skill for the database collocations – I would specify a priori database.

Author response:

We have adopted this suggestion in the revised manuscript.

Changes in manuscript:

We have updated the caption of Fig. 6:

> **Changes starting in line :**
>
> Precision-recall curves for the GPROF retrievals. Panel (a) shows the detection skill for the ~~database~~ collocations with the a priori/training database. Remaining panels show the results with respect to the MRMS validation measurements for the years 2019, 2020, 2021, and 2022.

Reviewer comment 19

Figure 12: for better comparison I would suggest to add a column with the GMI results in these plots.

Author response:

We have updated both Fig. 12 and 13 in the revised version of the manuscript, which now looks as shown in Fig. 1.2 and Fig. 1.3, respectively.

Reviewer comment 20

Line 464-465: I actually see more bias for GPROF V5 and V7 than from GPROF NN, am I missing something?

Author response:

Yes, the remaning biases are larger for GPROF V5 and V7 than for GPROF-NN. What we were trying to point out was that the *the reduction in bias* is more substantial for the

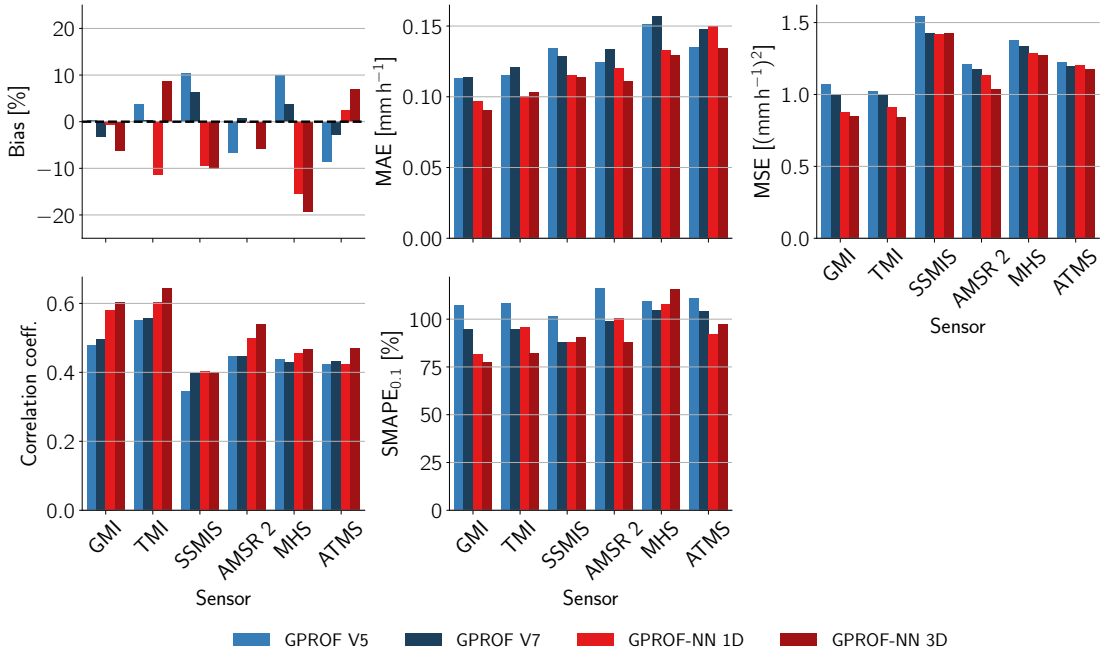

Figure 1.2: Accuracy metrics for a selection of other sensors of the GPM constellation evaluated against gauge corrected MRMS measurements of CONUS.

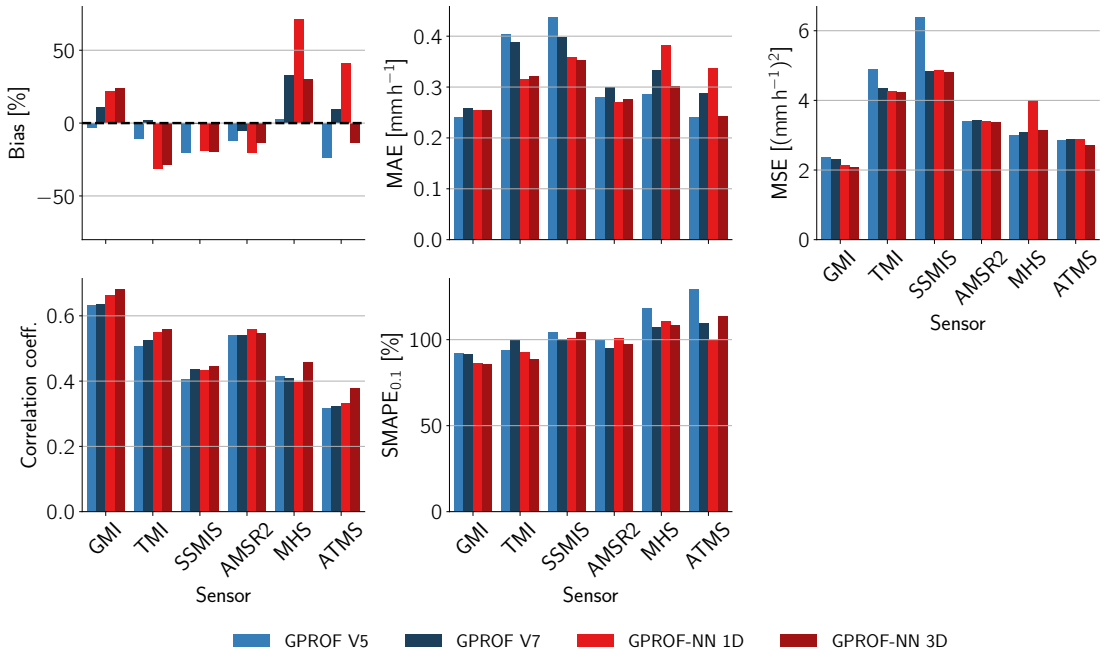

Figure 1.3: Accuracy metrics for a selection of other sensors of the GPM constellation evaluated against the K-POL radar in the tropical Pacific.

neural-network-based retrievals leading to overall lower biases after accounting for the database biases.

We have reformulated the offending section to avoid this kind of confusion.

Changes in manuscript:

> **Changes starting in line 539:**
>
> Subtracting the GPM CMB biases reduces the overall magnitude of the biases and their temporal and spatial variability. The ~~impact~~ reduction in bias is more pronounced for the GPROF-NN retrievals indicating that their higher accuracy puts increased weight on the reference data.

Bibliography

Anonymous Referee 2 (2023). Comment on egusphere-2023-1310. https://doi.org/10.5194/egusphere-2023-1310-RC2. Accessed: 2023-10-08.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Joyce, R. J., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., Wolff, D. B., and Xie, P. (2020). *Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG)*, pages 343–353. Springer International Publishing, Cham.