

Reply to the Comments from Referee #1 for AMT-2018-395

We would like to thank the referee for the review of our paper and the constructive comments. In the following, we have provided an item-by-item reply to the comments.

Major Comments:

- 1) The paper is fairly well written, and the topic falls within the scope of the journal. I feel, however, that experiment is not well designed for two main reasons. First, the accumulation time selected for the raingauges is too short. If I well understood, the sensitivity of the TB is 0.1 mm, so that any measure X has to be intended as $X \pm 0.1$ mm. By considering such short time, most of the time (given the low rainrates measured in the area) the data should be (looking at figures 6 to 9) for a single gauge 0.1 mm, resulting in a value of 1.2 ± 0.1 mm/h, that means a relative error of 100% for the most frequent value. Tokay et al., (2003, J. Atmos. Oceanic Technol., 20, 1460-1477) shows that for very light rain amount, the correlation between co-located tipping bucket raingauges can be very low. I suggest to use longer raingauge cumulation intervals (see also Porcù et al., 2014, Atmospheric Research) and to discuss the error associated to the ground measure.

Considering the accumulation time, which easily introduces errors to the evaluation, we refer to Tan et al. (2018) (<https://journals.ametsoc.org/doi/full/10.1175/JHM-D-17-0174.1>): “To extract the precipitation rate from the gauges, we will need to select an accumulation time. On one hand, since the satellite retrievals are considered instantaneous, the gauge accumulation time should be as short as possible. On the other hand, gauges have measurement uncertainty, so the gauge accumulation time should be sufficiently long to ensure a “stable” gauge measurement. After some trial and error, we determine 5 min to be a reasonable balance between these two factors.”

Furthermore, we show the DPR performance with lagged (± 30 min) gauge data (Fig. 10 – which will be updated to ± 15 min) based on the assumption that Level 2 DPR is supposed to provide only a snapshot of rainfall data. Thus, we selected the shortest accumulation data.

As a further example Amitai et al. (2012)

(<https://journals.ametsoc.org/doi/full/10.1175/JHM-D-12-016.1>) also follow the approach of using 1 minute gauge data to compare with satellite radar data including a time lag instead of using accumulated gauge data.

- 2) The second issue is on the use of binary indicators to assess the quality of the DPR estimates. Two indicators (PC and POFD) includes the number of correct negatives as input. This number should not be considered in the evaluation, since can be arbitrarily larger or smaller by changing the selection of cases, and thus the results are not general. Moreover, I suggest to use other indicators (ETS, HSS, HK) to synthesize POD and FAR information

We agree, the PC and POFD cannot be generalised since they are heavily dependent on the sample selection. We had chosen to incorporate them because the number of footprints showing precipitation almost equals the number without precipitation. However, as you pointed out the dependency on the correct negatives leads to misinterpretations. We will therefore cancel the PC and the POFD and move to ETS and CSI as well as HSS. We started the computations and will include them in the revised manuscript.

Specific Comments:

- 1) Page 2, Lines 18-19. The imager is only one (GPM Microwave Imager) and it is designed to provide a radiometric standard for the other radiometers of the GPM constellation.

Indeed, this sentence referred erroneously to the microwave instruments (microwave-radiometer, -imager, -sounder) of the constellation satellites. We change it to “The GMI completes the core satellite, enables a high temporal resolution for global precipitation maps and is used as a calibrator for the other radiometers in the constellation.”. Sorry for the confusion.

- 2) Page 3, Line 1. There are two recent papers (Speirs et al, 2017; Petracca et al., 2018) performing a similar analysis in the same region (Switzerland and Italy): they are reported in the reference list but not mentioned in the paper. Also the important paper Seto and Iguchi (2014) is in the reference but not mentioned in the text. The Authors should carefully read these papers and discuss their results.

Thank you for those remarks, we will incorporate them correctly.

- 3) Page 4, Lines 12-15. No solid precipitation in summer (i.e. hail)?

You are right; the formulation was not precise enough. Fortunately, the events in this study do not incorporate a hail event. In general, hail events affect only very few stations, which are then treated by the quality control of the WegenerNet like those that are clogged (e.g. by leaves). Their data are not taken into account in the calculation of the gridded product. In the entire WegenerNet time series there is just one event, where three stations have been simultaneously affected by hail.

- 4) Page 4. Line 27. Inverse distance interpolation does not add information to the gridded data, since the only information is in the raingauge measurements. The increased resolution can be good or bad, depending on how the precipitation pattern agrees with the inverse distance model.

This is true, but the interpolated data can include rainfall information from gauges that are located outside of DPR footprints but still within a radius of influence. Additionally, in most cases there is very little difference between gauge and gridded WegenerNet data (see Figure A1/A2). We think that the better coverage of the gridded data delivers more reliable results (see event no. 9).

- 5) Page 5, Line 1. Frozen hydrometeors attenuate radiation as well.

Thank you for careful reading, we will add that.

- 6) Page 5, Lines 19-20. It has to be considered that the radar measures precipitation in a volume at a given altitude (near surface bin, I guess), and it takes time to the raindrops to reach the ground. In case the Authors claim for a very precise temporal matching, this issue should be discussed. It should also be mentioned if the “near surface bin” value is used to evaluate the vertical distance between the DPR estimate and the ground.

We wanted to point out the challenge here. However, we used the precipitation rate for the estimated surface, which definitely needs to be mentioned. This will be done in section 2.3 “Selected data”. The effect of potential time lags is addressed in Fig. 11 (within +- 30 min [Note that this figure will be updated to show just +- 15 min]).

- 7) Page 7, Line 12. It is “8 to 10” or 8-12 (Page 6, Line 1) stations for footprint?

It is “8-12 stations”. Thank you for the hint.

- 8) Page 9, Lines 15-16. Bias and Normalized Bias give different information and should be computed both. The fact that the NB gives huge values simply tells that the error is much larger than the measurement, this is often the case when a too short cumulation time is used (5 min).

Indeed, in case of a very small mean, the normalised bias gives information that it is hardly possible to detect anything in this event. We follow your suggestion and will show both the bias and the normalised bias in the revised manuscript.

- 9) Page 9, Lines 19-22. To build any contingency table the threshold to define rain and no-rain sample has to be carefully defined and reported here.

The threshold was set to 0 mm/5min. We will report that. A change of the threshold to 0.1 mm/5min has little impact. The contingency table for KuPR is then:

<i>92 hits</i>	<i>-</i>	<i>47 false al.</i>	<i>instead of</i>	<i>95 hits</i>	<i>-</i>	<i>44 false al.</i>
<i>28 misses</i>	<i>-</i>	<i>86 corr. neg.</i>		<i>31 misses</i>	<i>-</i>	<i>83 corr. neg.</i>

- 10) Page 9, Lines 23-24. As mentioned, the direct use of correct negatives should be avoided in any validation study (see Nurmi, 2003, Recommendations on the verification of local weather forecasts. ECMWF Technical Memoranda. Technical Memorandum No. 430 for reference). What is the precise meaning of the “careful choice” of the events? For this reason the indicator PC and POFD should be removed by the analysis.

Starting from the end of your comment: The “careful choice” means that we found a balanced environment of rain/no-rain events. However, we will remove PC and POFD, to avoid the usage of correct negatives. Consequently, the “careful” choice can and will be omitted as well.

- 11) Page 10, Line 21. A key indicator is the Equitable Threat Score (see Nurmi, 2003) that summarizes both POD and FAR, and gives the skill with respect to the random assignment of rainy footprint.

We will include the ETS and the Critical Skill Score (CSI) to score the rain events.

- 12) Page 10, Line 25-26. This sentence is useless and should be cancelled.

Cancelled.

- 13) Page 10, Line 27. What is the “general structure”?

We intended to state a broad agreement between the DPR and the WegenerNet.

- 14) Page 10, Lines 28. What does it mean “70% of the GPM-DPR precipitation rates are within the range of the respective WegenerNet gauges”? What is the “range of WegenerNet gauges”?

I. e. the interval $[min_value, max_value]$ of the WegenerNet within the respective footprint. For 70% of the DPR rain rates (r_{DPR}) $r_{DPR} \in [min_value, max_value]$ holds.

- 15) Page 11, Line 7. What does it mean “close to zero”? How are the numbers rounded? How many significant digits are considered?

One digit after the decimal point is considered, which means zero = 0.0 mm/h. The “rounded to zero” will be omitted, that was for the visual interpretation.

- 16) Page 12, Lines 8-10. Please avoid misleading numbers and cancel PC and POFD from the analysis.

We will skip them as addressed above.

- 17) Page 13, Lines 1-4. FBI gives a measure of the under-/overestimation of the wet area, not of precipitation.

Thank you, this is a neat explanation, we will include it.

- 18) Page 14, Lines 1-end. I’m not sure this analysis is well designed. If I understood well, hit is when the DPR products falls within +- 1 standard deviation of the corresponding gauges value, misses is if the DPR is lower than the gauges – 1 standard deviation, and false alarm if the DPR is higher than the gauges+1 standard deviation value. Who are the correct negatives? How large is the standard deviation (roughly)? The main problem I see is that the distribution of rainrates is strongly asymmetric (power law) so that the standard deviation is asymmetric with respect to the mean value. In case of very light rainrate the value gauge-1 standard deviation could become negative. I suggest to cancel this section or to better argument its goal and procedure.

As this analysis is confusing to everyone, we cancel it from the manuscript.

- 19) Section 4.2. In many cases, there are footprints where the rain gauges do not measure rain, while gridded value is above zero. In some cases (e.g. footprints 2, 5, 8, 11 and 13 of event 9) eight or nine gauges measure zero, but interpolation fills the footprints with not negligible amount of rain. The Authors should justify the use of the interpolated data.

We use the gridded data because the DPR represents areal rainfall within its footprint while station data are only point-data. Thus, the interpolation increases the spatial resolution and can include rainfall information from gauges that are located outside of DPR footprints but still within a radius of influence. This is one of the advantages of a dense network.

- 20) Page 16, Line 5. The convective/stratiform discrimination can be done in several way (see Bringi, et al., 2003, J. Atmos. Sci. 60, 354–365 for an example). How it is done here? It is quite strange that light precipitation belongs to convective events.

This is an error in the manuscript, it is exactly vice versa than written. Sorry for the confusion.

- 21) Page 16, Line 10. Are the DPR footprints in chronological order? In terms of milliseconds?

Yes, they are. Ordered as given in the data. However, in order not to confuse any reader we delete the reference to the chronological order, since it is not important for our analyses. "The precipitation comparison at each footprint is given in Fig. 6"

- 22) Page 19, line 5 to the end. I understand the point (see my previous comment), but if the lag is supposed to be due to the time needed to raindrops to reach the ground, it is largely overestimated here. A raindrop of 2 mm of diameter has a terminal velocity around 6 m/s, that means that in 5 minute it covers around 1800 m, and I guess the DPR near surface bin is closer to the ground. Thus to search a time lag larger than 5-10 minutes is not justified. The relatively higher correlation found at lag of 10-20 minutes are very likely due to the patchy rain pattern and to the very small rainrates.

This describes in short what we found as well. The large lag time was not only to show a potential time lag caused by the fall velocity of a raindrop but also to have a look at the patterns that might arise (especially of interest in case of light rain). However, since this obviously causes confusion, we decided to restrict the lag-window to 15 minutes.

- 23) Conclusions. Page 20, Line 13. Probably gridded data add information, but there is no guarantee that the information is correct."

Please refer to the answer in specific comment (19).