Hydrology and
Earth System
Sciences

Open Access

Discussions

EGU

# *Interactive comment on* "Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM, an evaluation at a pan-India scale?" *by* Harsh Beria et al.

**Anonymous Referee #1**

Received and published: 30 November 2016

Review of "Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM, an evaluation at the pan-India scale?" by Beria et al.

General Comments: This paper evaluates the performance of the new GPM satellite rainfall product compared to that of TRMM, at the basin scale for 86 basins across India. While other studies have conducted such evaluations at the grid scale, this study aims to be more relevant for hydrology, and includes a rainfall-runoff modelling exercise.

I have read the paper with interest and feel that the results provide a useful comparison

of the two precipitation products. However, I am concerned that comparisons of rainfall between the two products have been previously published for the pan-India scale, as mentioned in the introduction. Overall, I am happy with the methodology and approach, but I would have liked to see the rainfall-runoff modelling exercise completed for more than one river basin in order to set this study apart from previous similar studies, and to strengthen the conclusion that any improvements in the rainfall did not translate into improved runoff simulations. I find some descriptions within the paper confusing and feel that in almost all sections, the amount of text should be considerably reduced as this paper is very long, indeed with 9 pages before the results are presented. The results are detailed, although the authors should provide some indication of how their results compare to the studies cited in the introduction. Most of the conclusions are clear and justified, but some require clarification - I like the clear way of presenting the conclusions. The large number of figures needs to be reduced, and I have several issues with the presentation of the figures that need to be addressed. Please see the detailed comments and suggestions below.

Specific Comments & Corrections:

Text:

Title: Slight misplacement of punctuation, I believe this should read: "Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at the pan-India scale."

Lines 47-69: Interesting, and I see why this has been included, but this much detail is maybe not required as not all of these example are directly relevant to this study; this paragraph could easily be shortened.

Lines 75-77: This is almost a repeat of lines 44-46.

Lines 99-104: I would like to see more justification of the choice to focus on the basin level, to make it clear what the benefit of this study is over the previous studies the

authors have mentioned. The authors state "most" of the previous studies - what about the remaining? How does this study improve on this? Why is the basin scale more useful for water resources and policy makers? It is not clear at the moment why this would be much more useful than the grid-scale analyses.

Lines 120 & 142: I would suggest replacing the word "scanty" with "scarce", which is much more widely used and less colloquial.

Section 2.1: While background information (and especially the maps) on the study area is always appreciated, I would recommend condensing section 2.1 - not all of the information is relevant or referred to later in the paper.

Lines 201-202: This is a repeat of lines 78-79.

Line 262: Could the authors clarify this statement?

Section 3: Throughout the results section, there are a lot of statements along the lines of "IMERG outperforms TRMM in x out of y basins, but they are similar in z basins" - the authors may be able to reduce the text and number of figures by constructing a table of the number of basins in which IMERG outperforms TRMM, the number in which they are similar, and vice versa, for each skill measure evaluated in the paper. This would also be interesting for the reader to give a quick overview of these numbers without needing to read the entire text and pick them out. Of course, it is still worth discussing these and the regional differences etc. as the authors have done, but this text could be reduced.

Lines 270-275: Some of this explanation should be included in the datasets section. It is not clear why this is done like this - why were the TRMM statistics obtained for 2 periods? Also implied here is that IMERG data is only available for March - December 2014, but later in the conclusions the authors state that a longer timeseries is available. This is confusing and should be clarified. If a longer timeseries of IMERG is available, why did the authors choose to use only 2014?

Lines 279-280: The authors state that the two datasets show similar skills, and immediately then state that IMERG is better in 70% of the basins - this is somewhat of a contradiction.

Lines 309-310: Could the authors expand on what the implications of this result are; why is it worth noting?

Section 3.3: Throughout the section on basin-wise bias, the results are difficult to follow. Typically in the literature, a positive bias indicates over-estimation, and a negative bias indicates under-estimation. I would recommend that the authors amend the presentation of the results here to also use this convention, making it more intuitive for the reader and more consistent with the literature. This is simply a case of reversing the sign in the results, i.e. using bias = simulated - observed, instead of bias = observed - simulated. Some of the language chosen in this section is also confusing, see specific comments following:

Line 352: The authors use the term "increased" bias - it is not clear if this refers to a larger negative or positive bias.

Line 354: Surely, in the 20 basins that now exhibit a positive bias which did not before, this is indeed a decay in skill for these basins? Please clarify this statement.

Line 356: What do the authors mean by an increase in the variability of the bias? This is not clear.

Lines 354-365: The terms "lower" and "higher" when referring to bias are ambiguous; it would be better to refer to "smaller" and "larger" biases. Again, it is not clear in this paragraph whether the authors refer to positive or negative biases. Please also check the rest of the section / paper for further use of these terms.

Line 408: Does section e refer to section 3.5?

Lines 474-475: What is the reason behind this part of the evaluation? What do the authors aim to gain from this analysis? This may have been mentioned earlier in the

paper but is not completely clear and it would be good to see clarification at the start of section 3.5.

Line 488: Again, the use of "high/low" when referring to bias is confusing.

Lines 533-537: This reads as though it should be part of the methodology of the paper, rather than results.

Line 543: The term "slightly" is ambiguous - how much worse are they? How much better is the NSE? How much larger the bias?

Section 3.6: This section is presented in the introduction as a major part of the novelty of this study, but in comparison to the proportion of the paper spent discussing the rainfall results, very little discussion is offered in terms of the hydrology. The implications of the findings are not discussed, and with only one basin used in this experiment, it is not possible to say whether the results would be similar for other basins in India or elsewhere. The aim of this experiment is left unclear and while I think it could be a very interesting part of the study, it seems somewhat unfinished. I would like to see, as the authors state would indeed be interesting, a comparison of these results for other basins in different regions in the study area.

Conclusion 1: To which parameter do the authors refer to with the quoted values?

Conclusion 3 (line 582): Do the authors refer to spatial variability? Please clarify.

Conclusion 5: Use of "higher" bias, as before.

Conclusion 7: If a longer timeseries of IMERG is available - why was this not used? This should be clarified / justified.

Lines 601-604: These statements are somewhat contradictory. The authors state throughout that IMERG outperforms TRMM in various aspects, and here state that there is a reasonable improvement, and also that the improvement is only incremental and not ground-breaking, but also that IMERG is a worthy successor of TRMM. These

statements leave the reader somewhat confused as to what the overall conclusion of the study is.

Line 611: "post forecast data assimilation scheme" - do the authors refer to post-processing?

Figures:

I'm afraid there are too many figures included in this manuscript, particularly considering the majority of figures are multi-panel. I would suggest moving a number of the figures into supplementary material. Please see below my detailed comments and recommendations for each figure:

Figure 1: Thank you for including this map, this is incredibly helpful for those readers who are not as familiar with the geography of the region. I would recommend splitting Figure 1 into two figures, one containing the two geographical maps (a) and (d), and the second comprising of (b) and (c). Also, the colour scales used for (b) and (c) are confusing - please modify these; the best way to present these would be a colour bar with just one colour for each map, ranging from light to dark with increasing values.

Figures 2.1 and 2.2. Firstly, it is strange to label two separate figures as 2.1 and 2.2 - surely these should be figures 2 and 3. Secondly, what exactly is the precipitation shown here? Is it daily precipitation? Is it averaged over a time period? This should be clarified and included on the axes / in the caption, for both 2.1 and 2.2.

Figures 3 and 5: These figures are very difficult and confusing to interpret - this data is not continuous (it represents the independent basins, rather than e.g. a continuous time period), and this is not the best way to present it. I would in fact recommend removing figures 3 and 5, and just discussing their results in the text as you have done. Figures 4 and 6, the corresponding spatial maps, are a much clearer way of presenting the data.

Figure 4: I like these maps, it is clear what they show and intuitive to interpret. However,

the colours used are very confusing - please amend the colour scale to use just one colour from 0 to 1 (light to dark), and avoid rainbow colours. In the case of (j), (k) and (l) it is not immediately obvious that there is a negative correlation in one or more of the basins and it is hard to spot. So on these maps, two colours should be used - the same as (a-i) for 0 to +1, with white at 0, and a different colour for the negative values. For example, the colour scale the authors have chosen for figure 8 would be perfect for figure 4, with white at 0.

Figure 6: Again, I like this figure, but the colour scale should be improved. I would recommend again a scale such as that used in figure 8, where 0 is white and the darker the colour, the larger the value. Please note that the colour scale has a big impact on the way the reader interprets the data, and incorrectly used colour scales can be misleading.

Figure 7: Again, the colour scale here is not the best option. For this data, the best would be to use one single colour, from light at 0 to dark at 1. For example, in figures (j-l), at first glance it seems that the blue basins have an opposite result to the red basins, but this is not the case.

Figure 8: While this scale would be perfect for the results shown in figure 4(j-l) and figure 6, it is not the good choice for the data presented in figure 8. As with figure 7, the best option would be one colour from light at 0 to dark at 1.

Figure 9: This graph could be removed and just discussed in the text.

figures 10 - 17: While I can see why the authors have presented the data in this way, again, there is the issue that this data is not continuous so this type of graph is not really correct, and also this is confusing for the reader. There are also a large number of similar plots here, I would suggest to pick one or two which show the most interesting results to present in the main body of the paper, and move the rest to supplementary material. Most readers would not analyse all the information in all of these figures and would appreciate the highlights, but the interested reader could easily find all the

graphs in the supplementary material. This would solve the problem of the overwhelming number of figures included in this paper. Also, I would recommend that the authors display all of these graphs (whether in the main body or the supplementary material) instead as scatter plots of the rainfall/elevation vs. bias/correlation. This would be a much more accurate and easy-to-understand way of displaying the data.

Figure 18: What is "data points"? Is it time on the x axis? Please change this label, and if it is indeed time as I suspect, please display the dates.

Recommendation: I believe that the work is of interest / useful, and warrants publication, but the manuscript indeed requires some work in terms of the descriptions and presentation of the study and its results, and clarification of some confusing aspects of the paper. Ideally, I would like to see the rainfall-runoff exercise extended.