Major remarks

The authors are evaluating different satellite-based rainfall estimates over Ethiopia. Such satellite estimates are very helpful over data sparse regions so that studies evaluating their accuracy over specific regions are important. In order to understand biases in the satellite estimates, these studies should also include analyses and speculations on why biases occur in the precipitation estimates. In my opinion the current manuscript is good in its descriptive part, but it has several shortages (see also several minor remarks) in the background and speculation part (too much number crunching without providing explanatory details leading to better understanding).

> [Author's Response] The authors would like to thank the reviewer for the comments. Responses are provided below for each of the comments.

A good result is that the authors found a relation of bias and elevation. But no explanation or speculation is provided what is causing this relation. Here, also the knowledge obtained in previous studies/publications may be taken into account.

> [Author's Response] We have added the following paragraphs within Section 3.3 that provide an explanation for the physical mechanisms that could possibly cause the relationship between bias and elevation.
>
> "The physical mechanisms or processes that could potentially cause a relationship between elevation and precipitation could be explained by the fact that the precipitation in the highlands of Ethiopia (Northwest) is strongly influenced by the ITCZ, whereas the precipitation in the lowlands of Ethiopia (Southeast) is influenced more by the southerly winds than the ITCZ.
>
> In the highlands of Ethiopia (Northwest), one hypothesis for the overestimation exhibited by CMORPH and TMPA 3B42RT could be associated with the deep convection of the ITCZ in the lower elevations leading to an increase in ice aloft, which is perceived by the MW sensors to be precipitation. Segele et al (2008) showed that the ITCZ, which is the major rain producing mechanism during the Kiremt, is centered to the North of Ethiopia and extends into the Blue Nile River basin. The overestimation of precipitation by CMORPH within deep convective systems has been shown by Nesbitt et al. (2008). The observance that PERSIANN is closer to the rain gauge data at these lower elevations could possibly be that the ice that is aloft is sufficient to bring the cloud top temperature within a range for the IR sensor to associate it with a precipitation intensity closer to the rain gauge data. The decrease in the bias ratio when moving from the lower elevations to the higher elevations for CMORPH and TMPA 3B42RT could be attributed to shallow convective systems. Nesbitt et al. (2008) showed that there is an underestimation of precipitation by MW products, such as CMORPH and TMPA 3B42RT, within shallow convective system because there is not enough ice aloft for the MW sensors to detect. In the case of PERSIANN, the underestimation could be due to poor detection of light rain events, consistent with the findings of Hong et al. (2007).

In the lowlands of Ethiopia (Southeast) the underestimation over regions such as the Rift Valley, Genale Dawa and Wabi Shebele River basins could be attributed to the southerly winds (Segele et al. 2008) not producing sufficient ice aloft (MW sensors) or having warmer cloud top temperatures (IR sensors)."

Partially it is unclear, which reference data are used in the data comparisons. Partially the SREs itself are used as reference data, even though this paper aims to evaluate them (e.g. Table 3). You cannot use data as reference in a study where you want to evaluate these data.

[Author's Response] Table 3 was used to understand the distribution of rainfall within the different regions. We have used intercomparison to understand the differences and similarities in the three SREs. We then used the rain gauge data to evaluate the errors.

In the conclusions section, for example, it is written: "This study has found that all three products capture the varied spatial precipitation pattern over the mountainous northwest, and the homogeneous mean annual precipitation and rainy days in the lower elevation of the southeast."But no rainfall pattern based on rain gauge data is shown, which would be necessary to support this statement.

[Author's Response] The authors acknowledge that a statement about the SREs ability to capture the number of rainy days is not possible given the absence of a comparison to the rain gauge data. A spatial distribution for comparison to the SRE was not possible given that the rain gauge data was provided in monthly totals, with some months missing, and that the missing months were not filled. This statement has been replaced by the following statement "All three products exhibit a Northwest to Southeast gradient in total precipitation, with higher precipitation totals found in the Northwest and lower precipitation totals found in the Southeast."

The question on the limitations/uncertainty in the results of the present study is omitted up to now. For example, are there any shortcomings or limitations of the study that are induced by the fact that the rain gauge data are only monthly and that they also comprise missing data in the rather short time period of 5 years?

[Author's Response] We have pointed out this limitation in the last paragraph of Section 2.3.

In summary, I suggest that the paper may be accepted after major revisions are conducted.

[Author's Response] Thank you.

Minor remarks

Abstract – par. 1 - p. 7670 – line 16-20

This part is very descriptive ("performs better than"), but no indication on why one estimate performs better than another is given.

[Author's Response] We have removed this phrase and have added an explanation for the performance difference in Section 3.3.

Introduction – par. 3 - p. 7671 – line 15

... and *western mountains* during ...

[Author's Response] This change has been made to the manuscript.

Introduction – par. 3 - p. 7671 – line 17

... in the *western mountains* and...

[Author's Response] This change has been made to the manuscript.

Sect. 2.1 - p. 7672 – 7674

This section is partially redundant with Table 2. Please focus on important characteristics of the basins, and leave out information on ranking or the mere repetition (without any specific comment) of numbers that can be easily obtained from the table.

[Author's Response] We have removed the redundant portion of this section and reworded the introductory paragraphs to point the reader to Table 2.

Sect. 2.1 – par. 2 - p. 7673 – line 8

... and *provide* six ...

[Author's Response] This change has been made to the manuscript.

Sect. 2.2

Provides numbers of annual precipitation, thereby pretending that these are the truth observed values. But looking at Table 3, it becomes obvious that these number are derived by averaging the three satellite estimates, despite the fact that the paper aims at evaluating these estimates in the following sections and despite of any biases of specific estimates that are not known in this part of the paper. This does seem adequate. It would be better to base these estimates on rain gauge measurement, and later on compare the satellite estimates.

Again, ranking is rather unimportant as long as no specific comments are made for this, e.g. with regard to climatic situation.

[Author's Response] We are unable to use rain gauge measurements because of the scarcity of the data. We use the average SRE for background information on basin wide totals and seasonal contributions.

Sect. 2.3 – par. 2 - p. 7675 – line 26

... for the *six* basins ...

[Author's Response] This change has been made to the manuscript.

Sect. 3.1 – par. 1 - p. 7677 – line 21

It is written:

... to the other SREs (correlation of 0.912) …

Correlation of what with what? You refer to two other SREs, but only give one value.

[Author's Response] The wording in the manuscript has been changed to "…to CMORPH and TMPA 3B42RT (correlation of 0.912 and 0.910, respectively)".

Same remark applies to Sect. 3.1 – par. 2 – p. 7678 – line 28.

[Author's Response] The author's believe that the reviewer was pointing out p. 7678 – line 2, rather than line 28, which states "…to the other SREs (correlation of 0.936)." The wording in the manuscript has been changed to "…to CMORPH and PERSIANN (correlation of 0.936 and 0.931, respectively)."

Sect. 3.1 – par. 1 - p. 7677 – line 22

It is written:

…, whereas PERSIANN underestimates …

This implies that PERSIANN has a low bias, so you pretend that the other two estimates are the truth. This is not justified in the text. Thus, you can only state that PERISANN estimate is lower than the other two SREs.

> [Author's Response] The reviewer brings up a very good point and the wording suggested by the reviewer "…lower than the other two SREs" is a much better choice. The manuscript has been revised as such.

Sect. 3.2 – par. 1 - p. 7678 – line 14

The *Rift* Valley …

> [Author's Response] This change has been made to the manuscript.

Sect. 3.3 – par. 1 - p. 7679 – line 4-6

In this section, the three SRE estimates are compared to rain gauge data. But correlation numbers seem to be given for the SREs to each other, and not to the rain gauge data. The reason for this is not clear.

> [Author's Response] The authors have added to this section by providing the correlation of the SRE data used to generate the 5 year average to that of the data used to generate the Rain Gauge 5 year average.

Sect. 3.4 – par. 1 - p. 7679 – line 24-26

It is written: A bias ratio greater than one indicates overestimation by SRE, a bias ratio less than one indicates underestimation by SRE, and a bias ratio of one indicates no bias in the SRE. Why there is no range around 1 that considers that the bias is within the uncertainty of the rain gauge measurements? In the current statement, it is assumed that a) rain gauge data are perfect, and b) no acceptable range of bias is defined around 1 where the estimate can still be considered as good.

This is certainly not adequate!

> [Author's Response] We have modified the statement as follows "…A bias ratio greater than one indicates overestimation by the SRE, a bias ratio less than one indicates underestimation by the SRE, and a bias ratio of one indicates no bias in the SRE (all with respect to the rain gauge data)."

Sect. 3.4 – par. 2 - p. 7680 – line 11

… basins *dominated by* less …

> [Author's Response] This change has been made to the manuscript.

Fig. 1

Use discrete colour steps! For black/white figures, not more than 5 steps are recommended.

> [Author's Response] The author's made an attempt to incorporate the reviewer's comment using only 5 discrete color steps; however, the complex topography was lost. After several iterations it was determined that 7 could be used without losing too much of the detail. The author's would like to thank the reviewer for this comment as it has improved the figure for the manuscript.

Fig. 2

Legend is too small. Please use 5 instead of 6 or 7 colour steps.

In addition I don't understand the colour scales ranging from 4 to 1 to 1.6 to 0.4 (upper panels) and 4 to 1 to 16 to 4 (lower panels).

> [Author's Response]: The upper scale ranges from values < 400 to 2,800 in 400 mm / year increments. The lower scale ranges from values < 40 to 240 in 40 days / year increments.

Fig. 3, 4, 5, 6

Legends are too small.

> [Author's Response] The author's appreciate the comment and have tried to balance out the size of the legend with the content of the figure.

Fig. Caption 5, 6

It is written: Based on river basins with at least 10 or more rain gauges. This statement is confusing as in the whole paper, only 6 river basins are studied.

> [Author's Response] This statement has been deleted from the Figure captions.