Reply on comments on Liechti et al. manuscript for HESS-2011-257, "Comparison and Evaluation of Satellite Derived Precipitation Products for Hydrological Modeling of the Zambezi River Basin"

Reviewer #1

General Comments

The authors address the important question of which of several openly available precipitation data sets are most appropriate for hydrological modeling in a river basin that lacks sufficient raingauge data. They take a very reasonable approach, including being very careful about handling the point gauge-vs-area satellite data problem. I think a few questions need to be resolved before this paper is released. In general, the English grammar and usage are very good.

Thank you for the positive feedback.

Specific Comments

1. Choice of data sets. The manuscript adequately informs the reader of the welter of contradictory results in different locations with different datasets and different statistics. I'm less clear on precisely why the datasets chosen for evaluation were picked. In particular:
   The satellite derived datasets were chosen as they are available online without restriction and they cover the region of interest. Moreover, they had to be available at least at daily time step and on a grid not coarser than 0.25° latitude/longitude. The research aims to calibrate and validate a hydrological model for the Zambezi basin. Therefore, products with long time series were preferred.
   The PERSIANN dataset was tested in a preliminary study but gave disappointing results. (Adapted in the text)

2. Ultimate use of data sets. Is it envisioned that the datasets being examined are candidates for post-real-time research, or for near-real-time operations? The paragraph in the middle of p.8187 is not explicit. That is, what are the requirements for timeliness? As examples; the 3B42 Version 6 came out about two weeks after the end of the month, while in its new Version 7 form it will be two months after the month; PERSIANN is three days; FEWS is "next day"; CMORPH is about 18 hours after observation time; and 3B42RT is about 8 hours.
   The datasets are candidates for post-real-time research. The objective of the thesis is to assess different scenarios of water use over the Zambezi basin using a calibrated hydrological model. Therefore, the quality of data was preferred over the short timeliness for TRMM 3B42RT. The new version 7 is not yet available so the version 6 was used as the best daily data of the TRMM products. (Adapted in the text)

3. Versions In general, the authors could be more precise about the versions of data sets being used. For example (p.8180), the GPCC Full Analysis described is probably Version 4, compared to the recently released Version 5 that runs to 2009.

The GPCC Full Analysis product used is indeed version 4. The versions of the products were partially listed in Table 1 but the text has been adapted and the table completed. (Adapted in the text)

4. Original grid size (p.8182,l.9) Working at the original grid size might have implications, just as the time averaging does; I think the authors need to be a little more careful about explaining when possible scale mismatches between estimates might affect the results.
   The original grid size has been kept for the analysis as the products will be used for the hydrological modeling at their optimal spatial resolution. It may affect the results as the smaller grid size (FEWS RFE 2.0) could be a reason for better results. An explanation has been added in the text on this matter. The time averaging is the same for all the estimates in order to obtain comparable results.

5. TRMM uses GSOD (p.8186, l.2) Do the authors know this for sure? All we're told about the gauge analyses used in TRMM is that they use the GTS "first-order" stations. Are all the GSOD stations in this collective?
   After a careful literature research, it is not sure that the GSOD stations are included in the GTS first order stations. Moreover, the TRMM uses monthly gridded rainfall ground data to rescale the estimate which is quite different from the raw ground data. Therefore, the text has been adapted.

6. p.8186,l.10 I'd suggest "… the cloud of scatterplot points is on the left …"
   Done

7. P.8187,l.12 I'd suggest "… CMORPH being less precise …".
   Done

Reviewer #2

General comments

This is a nice piece of work inter-comparing and evaluating satellite precipitation products (TRMM 3B42, FEWS RFE 2.0 and CMORPH) for the purpose of hydrological modeling of the Zambezi Basin. The relative performance of the satellite products was also assessed by comparisons with surface rain gauge measurements. I enjoyed reading and reviewing this manuscript, and would like to read the revised version.

The overall presentation of this manuscript is well structured and clear. However, some details of the description are not very clear and concise. The statistics used needs to be justified. The language, in general, is pretty good, but I did notice a few grammar errors and typos. In summary, this interesting research is conducted reasonably well, and results are helpful to partially mitigate the lack of adequate evaluation of satellite products over Africa. Therefore, I recommend that this manuscript be accepted for publication in Hydrology and Earth System Sciences, after the following comments are carefully addressed.

Thank you for the positive feedback. We will clarify the points addressed below and consult a native English speaker for the grammar corrections.

Specific comments

1. Page 8176, Line 24, " : : :.bias is null", clarify please.
   In the study cited (Nicholson et al., 2003), the bias is calculated as the difference between the satellite mean and the gauge mean and in the case of the TRMM-merged product. In this study there was no significant bias compared to gauge data.

2. More detailed description about ground data is required. As a reader when I read through the entire manuscript, I'm still not sure whether these 32 ARA and 48 GSOD stations are referred to rain gauges. Since gauge measurements are discussed in Introduction and Conclusions, I guess they must be rain gauges. I'd like to know more info about these stations.
   Are they weighted gauges, tipping bucket gauges or other types?
   Unfortunately, the description of the GSOD dataset is quite evasive. The data are referred as station but seems to be in reality rain gauge (one per station) and the type of rain gauge is not specified.
   The ARA-Zambeze data are collected by standard gauges of 5 inches diameter without automatic spillage. The reading is manual.

   What are the sampling solution and temporal interval for raw gauge data?
   For the GSOD data, the sampling solution is not specified and the temporal interval of the rain report varies between 6h and 24h.
   For the ARA-Zambeze data, the most important gauges report daily via radio the remaining ones are collected from time to time by a team.

   How many gauges at each station? If there is only one gauge at each station, I'd like to suggest using word "gauge" to replace "station".
   For the GSOD data, the description of the dataset suggests that there is only one gauge at each station. For the ARA-Zambeze data, there is only one gauge at each station. Therefore, the word "gauge" has been used in the text instead of the word "station".

   Accurately measuring rainfall from gauges poses difficult challenges. Gauges data are subject to many possible error sources such as mechanical and electrical problems due to the harsh environment, inadequate calibration before and after deployment. In addition, gauge data error sources can result from the sampling mechanism, wind effects, off-level gauge placement, funnel surface wetting and evaporation, or animal and human interference, etc. The gauge data used in this study are not exempted from these problems.
   The gauge data used in this study are not considered as a perfect measure of the rainfall. They are used for comparison with the satellite estimates in order to evaluate their precision but the assessment is also based on other analysis.

   Therefore, "an extensive automated quality control is applied to correctly decode as much of the synoptic data as possible, and to eliminate the random errors" in this study. I'd like to learn more about the "extensive automated quality control". Relevant references might be also helpful.
   Unfortunately, there is no reference available for the GSOD dataset except the readme file available on the website ftp://ftp.ncdc.noaa.gov/pub/data/gsod/readme.txt. The only explanation given is the following: "In deriving the summary of day data, a minimum of 4 observations for the day must be present (allows for stations which report 4 synoptic observations/day). […] As for quality control (QC), the input data

undergo extensive automated QC to correctly 'decode' as much of the synoptic data as possible, and to eliminate many of the random errors found in the original data. Then, these data are QC'ed further as the summary of day data are derived. However, we expect that a very small % of the errors will remain in the summary of day data."

ARA is listed in Table 1, but why GSOD is not listed there?
GSOD is also listed in the table.

3. Page 8179, Line 10, Huffman et al 2007 is about TMPA, instead of GPCC. You may need to move the reference a few lines ahead to TRMM 3B42.
Done

4. Any particular considerations to use R2 instead of r? In statistics, R2 is usually referred to as "coefficient of multiple determination" which is a measure of the fit of a multiple linear regression (y=a0+a1x1+ : : :+anxn). R2 is not the square of the Pearson correlation coefficient between y and any of xi (i=1,2, : : :,n). For a simple linear regression (Y=a+bx), R2=r2, r=R or r=–r. So r provides more information than R2. In this study, as described (Lines 3-4, Page 8181), R2 is square of the Pearson correlation coefficient between two time series at the same pixel. In this case, I think the Pearson correlation coefficient r, instead of R2, should be used. R2 is used in Figs 3,6,7 whereas r is shown in Fig 4. They had better be consistent.
Thank you for noticing this imprecision. The correlation coefficient used in the study is the following: $r(x,y) = \frac{\sum_{i=1}^{N}(x_i-\overline{x})\cdot(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\overline{x})^2}\cdot\sqrt{\sum_{i=1}^{N}(y_i-\overline{y})^2}}$ which corresponds to the Pearson correlation coefficient and not to the square. Therefore, R^2 has been replaced by R in the text and the figures.

I don't understand Eq. 2. Double check it and make sure it's precise.
Eq 2 has been adapted as following:

$$\overline{COR}_{p(i,j)r} = \overline{\sum_{m=i-r}^{i+r}\sum_{n=j-r}^{j+r} COR_{p(i,j)}(m,n)} - \overline{\sum_{s=i-(r-1)}^{i+(r-1)}\sum_{t=j-(r-1)}^{j+(r-1)} COR_{p(i,j)}(s,t)}$$

It describes the mean correlation on a squared ring of 1 pixel width at a distance of r pixel from the pixel p(i,j). It is calculated by subtracting the mean correlation at a distance of r-1 pixels to the mean correlation at a distance of r pixels.
The text has been adapted to clarify the equation.

5. About Table 2. The gauge numbers in the table is less than maximum possible (32/48 for ARA/GSOD). The explanation seems to have been provided at Lines 9-13 on Page 8182. Suggest moving it to the 1st paragraph of Page 8182 right after Table 2 was first mentioned.
The explanation has been moved.

What's the threshold unit in Table 2? The unit should be provided in both text and Table 2.
The unit is mm. It has been added in both text and table.

6. Page 8187, 1st paragraph. Proper references may be required here when discussing the gauge's point measurements. A number of studies (e.g., Fisher 2007, J. Appl. Meteor. Climatol.; Wang and Wolff 2010, J. Appl. Meteor. Climatol.) have reported that cautions must be taken when the gauge measurements are used as the "ground truth" reference for the area-averaged rainfall due to the fact that gauges lack areal representativeness.
   The text has been adapted and includes references about the reliability of gauge measurements for area-averaged rainfall.

7. Figs 5, 8 would be more informative if the correlation, bias and sample size were shown in the figures.
   Tables with characteristics of the samples were added.

8. Version number for each rain product should be provided, as many studies have shown different comparison results using different versions of the same product.
   Version number has been added where missing in both text (paragraph 2.2) and Table 1.

9. A few minor issues:
   Table 1, 3B42 resolution 0.25o, not 0.25"
   Done
   Fig 4, bottom panel, monthly is labeled as "10-daily"
   Done
   Figs 6,7 need to label y-axis.
   The title has been moved to the axis.

A bit of grammar check and proofreading would be helpful, e.g. Page 8179, Line 27, "precipitations : : :. on the afternoon, : : :"

Done

References

Nicholson, S. E., Some, B., McCollum, J., Nelkin, E., Klotter, D., Berte, Y., Diallo, B. M., Gaye, I., Kpabeba, G., Ndiaye, O., Noukpozounkou, J. N., Tanu, M. M., Thiam, A., Toure, A. A., and Traore, A. K.: Validation of trmm and other rainfall estimates with a high-density gauge dataset for west africa. Part ii: Validation of trmm rainfall products, Journal of Applied Meteorology, 42, 1355-1368, 2003.