Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-233-RC2, 2019

Hydrology and Earth System Sciences Discussions

Open Access

HESSD

Interactive comment

Interactive comment on "Can we trust remote sensing ET products over Africa?" *by* Imeshi Weerasinghe et al.

Anonymous Referee #2

Received and published: 10 August 2019

The article "Can we trust remote sensing ET products over Africa?" by Imeshi Weerasinghe et al. presents an evaluation analysis of the eight satellite - based evapotranspiration (ET) products over selected African river basins against the ET estimates derived from the water balance equation. The main conclusion of the study ranks the selected ET products in accordance with the results of the comparison analyses.

The topicality and scientific relevance of the research question addressed in this study is high considering the sparseness of the in situ ET data in the region as well as the urgency of having a high quality ET estimates for the climate related problems in Africa.

However, at this point I cannot recommend publication of this article as it (i) – contains a number of significant methodological inaccuracies and (ii) – gives poor explanation and presentation of the performed analyses and graphics. Also, stylistically and structurally the manuscript needs a substantial improvement. I highly recommend a major revision of the manuscript followed up by an internal review prior to the re-submission. Details follow.

Major comments:

================

Generally, the presentation style of the paper makes it often hard to understand the correctness and hence the added value of the illustrated results. The lack of accompanying relevant information in a well-written form along with the multitude of presented data combinations in a variety of forms and at different scales in many cases confuses an understanding of (i) which data sets were used for this concrete calculation, (ii) in which form the data went into the following graphic, (iii) what the estimates were compared to, (iv) how many and which basins were used this time, (v) when data mean was used and over which scales the averaging was done? In my opinion, the paper did not succeed in wrapping up the results in a clear manner. Usage of multiple data levels, i.e. 3 reference rain datasets, 8 ET products, 27/20 basins with/without trends, different temporal resolutions (from one value to time-series), two spatial levels (from basin integral to pixel-basis) comprises a fairly large number of levels of information which the authors should unwrap and present in a very simple, consequent and logical manner. In the present version of the paper this have not been archived.

The presentation style needs a thorough improvement, including restructuring of the manuscript, improvement of English grammar and scientific wiring style itself. One of the major remarks is that the whole manuscript text is written in a very intermittent and superficial manner. The explanations throughout the whole paper are significantly lacking concreteness. Also, confusion and replacement of some terms used throughout the study (e.g. trends as trends or trends as tendency to show certain value, among others) together with the multitude of abbreviations used in the text makes it very hard to follow the presentation (details given below).

Finally, some data descriptions and methodological assumptions (in particular, downscaling to the smallest grid and usage of different time periods) raises a number of questions related to their correctness and validity. The details are given further below.

More specifically:

==============

* Scientific relevance:

1. From the abstract, introduction and methodology sections it remains unclear how new is the water balance (WB) method, how accurate is the method and which other studies already used it for similar tasks. The abstract even makes a false impression that the authors developed a method and not used the existing one (p1, L6). The introduction in turn makes an impression that the focus of the study is the methodology (p3, L9) and not the evaluation of the existing products.

In this view, the authors must provide an overarching literature review of the studies that already used the WB ET estimates for satellite products evaluation, and also studies which evaluated the same satellite products over Africa using the same or other techniques. One of such study examples is the Miralles et al. (2016), which is also referenced in the present manuscript. Note, that Miralles' study also involved the African river basins. This has to be explicitly mentioned in the introduction. The authors should also then place their results into the findings of others. This is not done at the present state of the paper.

2. The authors should be more careful in formulating their scientific conclusions. The following sentence in the abstract: "However our recommendation would be the three highest ranked products being CMRSET, SSEBop and WaPOR." sounds rather subjective and needs further motivation (especially considering the huge differences in spatial-temporal scales between the products, as well as the manipulations on interpolation, then vice versa - integration - done within the study). Why first three? The same

Printer-friendly version

Discussion paper

remark applies for the conclusions.

3. The study does not mention anything at all about the quality of the reference precipitation data sets, nor about the quality of the final WB- based ET product. The study should also provide or mention some quantitative assessment to the magnitude of the differences which arise only due to application of different rainfall products. This will (i) - substantiate introduction of three different rainfall products in the paper and (ii) – justify better the obtained differences in ET between the products.

* Data products:

1. The data section needs a major elaboration. Many paragraphs appear rather like a snippets of information with lack of logical sequence, and hence they often fail to deliver main message of the paragraph or peculiarity of the concrete product. My suggestion would be: i) to either extend the data product descriptions to make them more complete and understandable or vice versa, provide only a reference links to the web sources and main papers of the products, and use instead the data section to discuss / group the products by their similarities and differences, advantages and disadvantages which can further help interpreting the paper results. ii) to omit the repetition of the time period and resolution information since they are already given in the table; iii) to place all products into the tables for consistency and clarity, i.e. also precipitation products, discharge data and reference data should be summarised in the same or separate table.

2. Many product descriptions and references miss version numbers. Those must be included, since depending on the product version there might be some already known issues related to a parameter derivations.

3. Check carefully the correct citations, once you add the product version numbers. I am more familiar with the GLEAM product, and I know that for GLEAM v3 (if you used that version) the correct references are Miralles 2011 (HESS) and Martens 2016 (GMD).

4. It is also a rule of a good scientific practice to provide/cite the data source: a web-page, ftp or a personal communication. No data sources are mentioned in the current manuscript version. For GLEAM, for example it should be the web portal: www.gleam.eu; For MSWEP: http://www.gloh2o.org/ (?), etc.

5. One of my major remarks here concerns inaccurate or sometimes even false information in data set descriptions. That is unacceptable. Please, check carefully all product descriptions you are giving!

On the example of GLEAM: - GLEAM is not a physically-based model, Prietsley-Taylor, the interception loss model, the stress module, and the water-balance model in GLEAM which form the core of GLEAM are all empirical! One can call it a process-based model, as it empirically describes the process needed to estimate E from satellites; - Table 1: GLEAM does not use CMORPH at all! That is simply wrong information. - Alemohammad et al., 2017 is the reference to a paper where they describe another method of deriving E. It is not clear why this is included in this GLEAM section?

* Methodology and results:

The methodological and result parts (i) - generally raise many questions as a result of incomplete and poor description of the calculation steps, applied quality assumptions and figures, and (ii) - casts doubt on the validity of some methodological steps and hence, on the accuracy of the final study results. Details follow.

1. As it was already mentioned earlier, the presentation of the calculation steps is done in a rather superficial manner. Lots of information is not given or remains unclear. E.g.: - which concrete quality control steps were involved in the selection of the basins and, which additional analyses were done and by whom? (e.g. p8, L7-10) - how were the basin boundaries defined? - what is the time-period of available discharge data? - how the integration over the basins is exactly done? were the simple mean or the area-weighted mean of ET or P fields used when averaging over the basin area? - which manipulations were done with the precipitation data prior averaging it over the basins?

Were the data also re-scaled to the 0.0022 deg resolution and then averaged over the basins? Never mentioned. - In their paper Miralles et al., 2016 applied additional quality control check on the difference between the GRDC-reported area and the area calculated from basin boundaries. Would not it be also relevant for the present study? - How was the MPM calculated? The products have big differences in resolution. The averaging of the products to get the MPM without applying corresponding weights can be a source of errors.

2. Correlation analysis:

(i) – It remains not very clear from the text over which values the correlation analysis is performed? Over time-series of annual means? Over multi-year averages of different basins? Should be put more clear. (ii) - Units of the correlations are not common, and confuse the interpretation. Correlation should rank from -1 to 1. Besides, it is never clear from all the graphs with percentages, by what value the normalisation was done. (iii) – Correlations should always provide significance measure, or the latter should be mentioned in the text.

3. The choice of the highest resolution is one of the two major remarks that I have to the methodological part: (i) – Generally, it is not common to interpolate products to the highest resolution, especially when the difference between the highest and the coarsest resolution is that high. It would be more correct to upscale the higher resolved data to the coarser estimates to minimise the bias. (ii) – Besides, the fact that the comparison of the products is mostly done at the basin level, the downscaling does not seem to make sense at all. First you interpolate the coarse data to the very high resolution, and then, you integrate it back again over the river basin. This clearly can be a source of additional biases and errors, which also raises my doubts about the validity of the ranking results. (iii) – All the above inter alia also raises a question of what is the minimal area of the smallest basin you have, and whether it is resolved by the products with the coarsest resolution at all?

4. The second and the most major remark of mine is related to the application of the analyses at different time-periods.

(i) - First of all it never comes clear what is the time period of available discharge data for every basin; (ii) – From the Tabel 2 it appears that for most of the basins discharge data does not extend the whole period of available precipitation data at all, and the spread of data periods is huge among the basins. In this view I do not understand at all how the analyses tests were done?

(ii) – The test for the effect of temporal variability on annual means mentioned in the discussion section was done only for the four basins, while 20 basins are analyzed throughout the study. Moreover, Congo - one of the four tested basins - has data only till 2010, while remote sensing ET products span up to 2017. In this view, I would not be able to call it a fair validity test!

(iii) – Clearly, the exclusion of periods with trends does not account for the temporal variability of data which can still result in the pretty different annual means. So, the effect of temporal variability on annual means must be done for all the basins, which are used for the evaluation of the satellite ET products in order to draw a fair conclusions.

(iv) – Calculating trends only for the WB ET reference data set is not a complete analysis. If a satellite data product has a trend, this also has to be mentioned, and maybe even that product should not participate in the validation (?)

To conclude, if the tests will show that the variability indeed matters, then none of the performed analyses is valid since they will all be affected by the differences due to variability.

5. My advice would be to not use percentages for all the figure results. This only confuses the interpretation. Use -1 to 1 scale for correlation and differences.

6. The raking of products based on visual inspection is rather speculative for me. For ex on Fig 9 it was impossible to follow the text conclusions: I did not see where irrigation

area is, to which reference product other products are compared, why MPM and some other products have no data and why GLEAM is concluded to perform worst? The same for Fig 10.

6. The analyses of comparing products over one irrigation area, and one lake, where some products have no data, and others do not even resolve the region does not make much sense to me, nor it is complete enough to make a serious conclusion on which product is better or worst.

8. It maybe due to the presentation style, but I could not follow the result section presenting the crop coefficients very well. It has to be structured and presented in a more clear manner: objective and reasoning for location, crop types, etc, data used, hypothesis to prove, which products in which form are tested, and what do results show.

* Stylistic and structural comments:

I highly recommend to look into the papers of Zeng et al., 2012 (ERL) and Miralles et al., 2016 (HESS) as an example of a good presentation style, and especially of the methodological part, as well as their choice of graphs.

Few comments:

1. Usage of term throughout the paper changes and confuses the reader. For ex, the term trend is first used as trend itself, but also to indicate tendencies if I understood it correctly (p13, L5 or p16,L9)

2. Typos are also present throughout the paper (e.g. p3 L2, p7 L17, p15 L7, p20 L3). A proper internal resew is required.

3. Discussion section rather reads like methods and should be incorporated to methods. Instead, the discussion section should place the paper findings into the existing knowledge as was already mentioned earlier.

4. Nothing is mentioned about the Budyko result in Discussion or Conclusions

5. Nothing is mentioned about the differences between using three precipitation products

6. Many abbreviations are not opened. Add abbreviation table to the paper. Use less abbreviations, i.e. if possible leave it open. Very hard to follow

7. Explanation of the results in Figures are often not complete. Not clear which products were used, which reference, etc. Figures are too small.

8. Lots of sentences are too wake, “Based on the elements being analysed… p21, L30” What is meant? Be more concrete.

9 . Titles of the sections should be reconsidered

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-233, 2019.