Dear Prof. van den Hurk,

Please accept our gratitude for your careful review, constructive criticism, and your patience in handling this manuscript and its companion paper.

Based on your earlier remark, we extended the revisions of the manuscript to better reflect on all the comments made by you and the anonymous reviewers. Below please find the details on the corrections that were made.

A marked-up PDF of the manuscript is included as a supplement file for your review.

We hope you find the corrections satisfactory.

Kind regards,
Poolad Karimi & Wim Bastiaanssen

Response to the reviewer's comments:

Reviewer #1:

We thank the anonymous referee for the thorough review, positive comments, and constructive remarks on this manuscript. We believe that, by addressing the concerns in this review, the manuscript has improved. Our reply to the numbered General and Specific Comments are given below:

General comments:

With delight I read through this paper. The authors made great efforts to summarize the current situation of Reliability and errors of remote sensing data in ET, rainfall and land use. This study will benefit the science community and will be appreciated by the peers. The subject is within the scope of HESS, the paper is clearly structured and the story line is straight forward, and the tables are well designed and informative. At the same time, to make this paper more ambitious and more profitable for the community, further concerns need to be addressed before consideration of publication at HESS. I recommend acceptance after moderate revision.

We are we very glad that the referee finds our paper to be of benefit for the science community and peers. We also appreciate the reviewer's positive feedback on the manuscripts structure and design.

Specific comments:

1. There is a lack of section presenting the directions or recommendations of future work in this field. This section is a necessary part of a review paper, to guide the peers or young generation to better advance science.

Response: Great point. We agree that the paper will improve by providing some directions/recommendations for future research in this field. We will revise the paper to include such recommendations.

Update: Paper was revised to provide recommendations for future research. This includes high resolution images for land use mapping and an example of recent work of Tsinghua University is refered to. It also includes an ensemble ET product for which a paper is currently under preparation using 6 global scale and quasi-operational ET models. For rainfall we believe that downscaling and simultaneous calibration with radar and telecommunication signals is the way forward. This was integrated in the conclusions section.

2. To further enrich the merits of this study, a summary of recommended remote sensing data and methods on ET, rainfall and land use for different regions (e.g. continental scale or small region) and different research tasks (e.g. focus on hydrological processes or on human-land-atmosphere interactions) is desirable. Although it is apparent that there is huge heterogeneity across the globe, it is still possible to give some recommendations or directions. Sometimes the researchers are overwhelmed by a long list of data and methods, and might struggle for choosing the appropriate one, and there is a big space for the authors to fill. I leave this to the authors on how to benefit a broader range of readers.

   Response: We very much appreciate and understand the value of providing such recommendations that may help users to choose appropriate methods and data products for their work. However, as it's been rightly pointed out by the reviewer, there is a great deal of heterogeneity regarding with the accuracy and reliability of remotely sensing data and methods. Oftentimes reliability of RS methods outputs, especially in case of public available processed data products is rather case and location specific. Therefore providing such recommendations and endorsing or dismissing a RS based product requires further research that is linked to a specific application which falls beyond scope of this paper. We will revise the paper to be more explicit on this issue.

Update: Revisions was done to reflect on the remark. There is no single preferred model. While scientists have the tendency to make the schematization more complete and in agreement with physical processes (e.g. 2 and 4 layer ET models), other research has indicated that this does not always necessarily yield to a better model performance. This can be partially explained by the high data demands of complex models (which then appear to be an over-parameterization) and the solution to deal with cloudy atmospheres (that has in the end great impact of the final results). The following text has been insert into the manuscript: "There is no single preferred ET model. The selection of the algorithm depends on the application, the required spatial resolution, the period for which the ET fluxes should be estimated for, the size of the study area, the land use classes present etc. A useful distinction is to discern global scale models (few) and local scale models (many). Also the level of validation and application of these models widely differ. Whereas certain models are tested with a single experimental flux site, other models have been applied in more than 30 countries. "

3. Interpretation of PDF is problematic. The distribution of figure 1 and figure 2 tends to be an exponential distribution rather than a skewed normal distribution, but the authors imposed a skewed normal distribution to fit the histogram. Do you have sufficient evidences to support this imposition? You should be very cautious to make that kind of imposition, as it would be very easily complained by statisticians. Consequently, the interpretation of the PDF should also be careful.

   Response: As stated in the paper many authors of the reviewed papers are both the developer and the tester of the algorithms, thus we believe there is a natural bias towards reporting low errors in literature. This is evident from the number of reported error that are exceptionally low. To correct this the data points were fitted by means of a skewed normal distribution so that less weight is given to the class with very low errors.

4. The measure variables of errors or accuracy should be widely accepted, or specifically defined, and consistent throughout the paper. The term “error”, “absolute error”, “deviation” show up multiple times in the paper, it is not very clear for their statistical meaning as they are kind of vague, please define these terms clearly in the paper or use widely accepted statistical terms such as mean percentage error (MPE), mean absolute percentage error (MAPE), standard error (SE), and keep these terms consistent throughout the paper.

   Response: Good point. The paper will be revised to use mean absolute percentage error (MAPE) throughout the paper to make sure the use of consistent terminology.

   Update: Revisions was done and MAPE was used throughout.

5. There are a bunch of studies using triangle method for estimating ET (e.g. review paper by Toby Carlson (2007, Sensors)), the authors might want to lump it with trapezoid method.

   Response: Paper will be revised to reflect on triangle method.

   Update: Update: Revisions was done to include triangle method by presenting it together with the trapezoid method. Certain standard references with longer and shorter histories are provided already in the first manuscript

6. The uncertainties of measurements should be taken into consideration when you summarize the accuracy of different methods and RS products. For example, the uncertainties of ET measurements at eddy covariance flux tower are up to 30%, and the wind induced precipitation under catch for precipitation measurements in the Northern High Latitudes is prevalent.

   Response: As stated in the paper we agree that conventional methods of measuring hydrological processes (e.g. rainfall and discharge) are not flawless

and their accuracy needs to be verified. The paper will be revised to elaborate more on this issue.

Update: Revisions was done to reflect on the remark. The following sentence was inserted: It is important to note that conventional ground measurements come with their own errors and uncertainty that should ideally be taken in consideration when used for verifying the accuracy of satellite-based estimates. This holds true for ET where the number of operational flux towers is limited, but also for rainfall that has distinct micro-scale variability, that cannot be measured by a single gauge. However, in most documented studies these ground measurements are treated as "the best available estimates" in the absence of reliable information on their accuracy. As such they are widely used to validate satellite based data.

7. Page 1085, line 2, what does "seasonal ET" exactly mean here? For each record of Table A1, do you calculate mean percentage error (MPE) for monthly ET or ET during growing season or annual ET? What does "Deviation (%)" in Table A1 exactly stand for (link to afore mentioned comment 4)? In statistics, "deviation" stands for the difference between the value of an observation and the mean of the population, it is a measure variable having unit rather than a percentage in a normal case. It is not clear for the calculation process, please specify explicitly. Besides, adding information of time step of measured ET for each record in Table A1 might be helpful.

   Response: Seasonal ET refers to the accumulated ET pertaining to a certain growing season which on average is about 5-6 months. Numbers in Table A1 represent mean absolute percentage error (MAPE) either growing season or annual ET. The term "Deviation" will be replaced with MPAE. Information on time step of measured ET will be added to the table.

   Update: Revisions was done to reflect on the remark. All time steps of Table A1 are either annual or seasonal. It has been mentioned at several places that our interests is mainly in ET fluxes integrated with time, because our work is driven by the need for operatioanl water balances and water resources assessment reports.

Technical corrections:

1. Page 1074, line 21, Vörösmarty et al (2010, Nature) is a good paper to cite.

   Response: the paper will be revised to include the suggested citation.

   Update: Reference provided.

2. Clear definition of the term "water accounting" for a broad range of readers is necessary.

   Response: Water accounting is the process of communicating water related information about a geographical domain, such as a river basin or a country, to

users such as policy makers, water authorities, basin managers, and public users. The paper will be revised to include the clear definition of water accounting.

Update: the definition of water accounting was included.

3. Page 1075, line 21, "land use" is not a commonly accepted hydrological variable, although it is an input variable for WA+, please be careful when you phrase it.

   Response: Noted. Will reflect on it.

   Update: Correction was done.

4. Page 1086, line 5, it is not clear what does the "spatial layers of ET maps" exactly refer to, please specify or rephrase it.

   Response: it should read "spatial layers of ET". Will be revised.

   Update: Correction was done.

Reviewer #2:

We thank the anonymous referee for the helpful comments, and constructive remarks on this manuscript. Our reply to the numbered General and Specific Comments are given below:

This paper reviews the accuracy of remote sensing information for the relevant components of Water Accounting: rainfall, land use and evaporation. The paper is well structured and well written. The topic is relevant for Hydrology and Earth System sciences. The content of the paper is also relevant since it puts a substantial number of important studies on satellite based estimates of hydrological components in perspective. I have two main concerns:

We are pleased to know that the referee finds our paper well written, well structured, and of relevance to HESS.

1. In the introduction reference is made to a specific tool Water Accounting Plus (WA+).
   Subsequently the title of section 2 contains the same name WA+. Does this mean that the review is limited to and geared to a specific tool? If so this should be clearly stated in the introduction and preferably also be reflected in the title of the manuscript.

   Response: The use of results of this review is not limited to WA+. However the choice of the hydrological parameters that have been investigated for their accuracy in this paper (ET, Rainfall, LULC) is based on the parameters that are used in WA+. The paper will be revised to explicitly reflect on this matter.

Update: Revisions were made to reflect on the remark. The following text is part of the paper and has been slightly modified: In addition to that, hydrological variables derived

from remote sensing can also be used for spatially distributed hydrological modeling. Studies by Houser et al. (1998), Schuurmans et al. (2003), and Immerzeel and Droogers (2008) have for instance demonstrated that such inputs have improved hydrological model performance for river basins in Australia, The Netherlands and India respectively.

2. The very small errors (1%) reported on a number of studies where SEBAL is used for evaporation estimation should be explained. Whereas any ground truth evaporation measurement will have a larger uncertainty then 1%, it is unclear what this errors of
1% actually represent.

   Response: All the inputs for this review come from published scientific papers and reports. As explained in the paper we agree that there is bias towards reporting low errors in published works in which often the authors are both developer and tester of the proposed algorithms. For this reason, the data points were fitted by means of a skewed normal distribution so that less weight is given to the class with exceptionally low errors.

Minor issues:

3. Why not keep the same sequence in section 3 as in section 2: Rainfall, Land use, evaporation?

   Response: Noted. Will revise the sequence.

   Update: Sequence was revised.

4. P1077:L22. Note that under convective daytime conditions a decrease in the wind speed, with a reduction of turbulent mixing may also increase surface temperature and this will not necessarily lead to a higher sensible heat flux.

   Response: Good point. We will reflect on this issue.