

Dear Editor-in-Chief,

We wish to thank you for offering us another chance to revise our manuscript (**hess-2022-185**). We detail below all of the revisions that we have undertaken in response to the comments of each reviewer.

In line with changes that we have now made to the manuscript as a whole, we also wish to propose a revision to the title to instead read as: “Evaluating the accuracy of gridded water resources reanalysis and evapotranspiration products for assessing water security in ungauged basins”. We hope that these revisions are now acceptable.

With kind regards

Elias Nkiaka (on behalf of the co-authors).

Response to reviewer 1 comments

1. Design of the study: The authors need to provide a better explanation of why they decided to evaluate runoff and evaporation from completely different sets of models. All GHMs and LSMs provide estimates of all water balance components, especially when the authors consider GHMs and LSMs as a reanalysis product. Currently, these two parts of the paper are totally distinct from each with no connection to each other. If the objective is to assess water security, I would imagine the end-user would be interested in using estimates of all water balance components from one model or a specific ensemble of models.

Response: Thanks for highlighting this flaw in our study. We have now included the evaluation of evapotranspiration estimates from GHMs and LSMs separately from the results of remote sensing-based evapotranspiration estimates. The results are presented as distinct figures and discussed separately in the revised manuscript. Furthermore, we have highlighted in the manuscript the fact the users’ needs for the application of ET estimates may vary. L121-124, revised manuscript.

2. Related to the above comment, if water security is the main intention, would not subsurface water availability be an important variable as well? The authors need to justify only evaluating evaporation and runoff. I am sure most GHMs and LSMs provide data of water storage change.

Response: Thanks for this remark. Yes, most GHMs and LSMs provide data for subsurface water. However, we did not evaluate subsurface water availability in this study because of a lack of in situ data that can be used to validate model simulations. We are aware that other studies such as (Koukoulou, M., Nikolopoulos, E. I., Dokou, Z., and Anagnostou, E. N.: Evaluation of global water resources reanalysis products in the upper Blue Nile River Basin, Journal of Hydrometeorology, 21, 935-952) used data assimilation methods to estimate changes in terrestrial water storage from WRR. In this study, we limited our evaluation to

discharge and evapotranspiration estimates. In addition, water storage change was used as a variable in estimating basin-scale evapotranspiration.

3. The authors claim that the utility of gridded datasets have not been sufficiently explored in Africa. I do not agree with the claim - authors have ignored the innumerable studies which have used gridded datasets for model calibration, forcings and validation. In fact, gridded evaporation products are routinely used for improving large scale models for African watersheds (Dile et al. 2020, Dembele et al. 2020). The authors themselves have cited many studies which evaluate these datasets over African basins.

Response: Thanks for this remark. I beg to differ with this claim. I said and I quote “Whilst the use of outputs from WRR in water management has gained significant attention in many ungauged areas such as Asia and Latin America, they remain largely under-utilized in Africa. For example, there are only a few case studies reporting on the use of these products in the Upper Blue Nile River basin and the Zambezi River basin. On the other hand, several studies evaluating the performance of gridded hydrometeorological variables in Africa have focused mostly on precipitation while a few studies that have evaluated gridded ET products focused on large basins”. See L97 – 111, revised manuscript.

Nevertheless, we believe that evaluating the different datasets across several basins of varying sizes will contribute to the contemporary debate on the performances of the different products across Africa.

4. Methodology: The authors do not make a convincing case for comparing the evaporation datasets with water balance-based evaporation estimates, especially (according to the results) when the uncertainties are large. In fact, achieving water balance closure with different sources of P, ET, and TWS is not a trivial task (Lorenz et al. 2015, Koppa et al. 2021, Pan et al. 2012) and is definitely not robust if only one source of data is used for each component.

Response: Thanks for highlighting this issue. The use of water balance-based evapotranspiration estimates for validating global evapotranspiration estimates is a well-established technique in hydrology including in gauged and ungauged basins. Few examples of such studies that have applied the water balance concept to evaluate evapotranspiration estimates at basin-scale include:

1. Weerasinghe, I., Bastiaanssen, W., Mul, M., Jia, L., and Van Griensven, A.: Can we trust remote sensing evapotranspiration products over Africa? *Hydrology and Earth System Sciences*, 24, 1565-1586.
2. Baker, J. C., Garcia-Carreras, L., Gloor, M., Marsham, J. H., Buermann, W., da Rocha, H. R., Nobre, A. D., de Araujo, A. C., and Spracklen, D. V.: Evapotranspiration in the Amazon: spatial patterns, seasonality, and recent trends in observations, reanalysis, and climate models, *Hydrology and Earth System Sciences*, 25, 2279-2300
3. Blatchford, M. L., Mannaerts, C. M., Njuki, S. M., Nouri, H., Zeng, Y., Pelgrum, H., Wonink, S., and Karimi, P.: Evaluation of WaPOR V2 evapotranspiration products across Africa, *Hydrological processes*, 34, 3200-3221
4. Liu, W.: Evaluating remotely sensed monthly evapotranspiration against water balance estimates at basin scale in the Tibetan Plateau, *Hydrology Research*, 49, 1977-1990

We acknowledge that the uncertainties are large and that is why we decided to identify the dominant sources of uncertainties in this study which is not the case in several studies that have used basin-scale water balance estimates to validate global evapotranspiration datasets. We

believe that identifying the sources of uncertainties is a first step towards reducing them and also to inform policy decisions.

We applied only CHIRPS precipitation estimates in this study because a recent study has provided an average of the uncertainty estimates inherent in monthly CHIRPS estimates across the world including the African continent. However, we are aware that there may be regional differences in the uncertainty estimates across Africa. Nevertheless, we believe that our approach is robust and there is no method that is free from uncertainties.

1. Shen, Z., Yong, B., Gourley, J. J., Qi, W., Lu, D., Liu, J., Ren, L., Hong, Y., and Zhang, J.: Recent global performance of the Climate Hazards group Infrared Precipitation (CHIRP) with Stations (CHIRPS), *Journal of Hydrology*, 591, 125284

We are also aware that GRACE data is processed and made available by three different research centres. We decided to use estimates from Jet Propulsion Laboratory as it is one of the most commonly used GRACE datasets. Moreover, GRACE estimates from Jet Propulsion Laboratory are provided with uncertainty estimates which for each grid point. For each basin, we averaged the uncertainty estimates for all grid points located within the basin to estimate the GRACE uncertainty for that basin. We also wish to highlight to the author that every study adopts different methods, and we believe we have provided sufficient justifications and clarifications on our approach and methods.

The author may also wish to refer to the following article on GRACE estimates produced by JPL.

Wiese, D. N., Landerer, F. W., and Watkins, M. M. (2016). Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution, *Water Resources Research*, 52, 7490-7502, <https://doi.org/10.1002/2016WR019344>.

5. Despite previous studies using GRACE at higher resolution, I have serious doubts about the applicability of TWS estimates for basins as small as 9,620 sq.km (an order of magnitude smaller than intended GRACE footprint).

Response: Thanks for highlighting this issue. I believe we raised this issue in the manuscript and provided a few examples where GRACE data was used in catchments smaller than the size highlighted by the reviewer e.g.,

1. Liu, W.: Evaluating remotely sensed monthly evapotranspiration against water balance estimates at basin scale in the Tibetan Plateau, *Hydrology Research*, 49, 1977-1990.

Moreover, in each of our basins, there was at least one GRACE grid point located within each basin which was used to represent the TWSC for the whole basin. Where there were two or more GRACE grid points, we calculated the average of all the grid points located within the basin.

6. In summary, the above two points casts serious doubts on the robustness of the ET_{wb} estimates and its use as a reference dataset for evaluating other datasets.

Response: We believe that we have provided sufficient justifications on the use of ET_{wb} as a reference data for evaluating ET estimates derived from different sources. We wish to

reiterate to the reviewer that this is not the first study to use this concept to evaluate ET estimates. In addition, we went further to identify the dominant sources of uncertainties when using this method which is a novelty compared to most other studies that have used this method for evaluating ET estimates. However, inherent uncertainties in the data cannot be a basis for disqualifying the use of this method as hydrologists have to deal with the challenge of uncertainty in data in every study.