Reply to Reviewers' comments (Reviewer#1)

Legend

Reviewers' comments

Authors' responses

Direct quotes from the revised manuscript

We thank the reviewer for his/her time in reading our manuscript and detailed comments on our manuscript. Point-by-point replies to the comments or suggestions made can be found below.

Reviewer #1: This research provides a large ensemble of water balance ET estimations using existing global datasets. Sound comparisons with mainstream ET products have been reported. Uncertainty assessment is also performed on different scales. The manuscript is within the scope of ESSD journal. Minor revisions need to be considered before publication. Please find details below:

Response: We thank the reviewer for his/her time in reviewing our manuscript, providing comprehensive suggestions for improvement, and recognizing the potential for publication of the ET-WB dataset. Revisions for texts and figures have been made in the new version, as suggested. Please find our specific response to your comments below.

(1) ET-WB incorporates both the natural ET from various sources and human-induced part, e.g., irrigation. What about the products used for validation, particularly for GHM such as WaterGap? Can human activity be a significant reason causing difference between ET-WB and them?

Response: As rightly indicated by the reviewer, ET-WB includes the influences from both natural and anthropogenic systems because it is calculated using the physics-based water mass balance method forced by either pure observations (e.g., satellite products) or realistic simulations (e.g., GloFAS), both have incorporated natural and human-induced parts.

Please find the details about the auxiliary ET products below,
MODIS, GLEAM, and FLUXCOM products: In terms of the three auxiliary global ET databases (MODIS, GLEAM, FLUXCOM) for comparison, the governing mechanisms are divergent but consistent on the line of representing combined impacts from natural and human activities. For direct observational ET data from space (MODIS) and ground (FLUXCOM), the sources of ET are not partitioned so they also incorporate the human impacts, because the actual earth surface states (e.g., albedo) and flux (e.g., water vapour exchange) are measured. GLEAM, which uses satellite data to estimate potential ET and then convert it to actual ET by a stress module (assimilated with soil moisture observations), also shares a similar condition.
WGHM case: The standard run of WGHM that is used by our study, models major human activities like irrigation, dam construction (including their commissioning years), reservoir management, regulated lakes, and therefore the resulting ET also includes the influences by these human activities (Müller Schmied et al., 2021).

Therefore, all the selected ET products represent the total ET, and the observed differences with ET-WB primarily reflect the bias in modelling, measuring, and/or processing. Despite this, slight differences in the representation of human activities in various datasets might have led to minor discrepancies in our results and may be studied in the future as and when better datasets/models are available.

Reference:

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., Peiris, T. A., Popat, E., Portmann, F. T., Reinecke, R., Schumacher, M., Shadkam, S., Telteu, C.-E., Trautmann, T., and Döll, P.: The global water resources and use model WaterGAP v2.2d: model description and evaluation, Geosci. Model Dev., 14, 1037–1079, doi:10.5194/gmd-14-1037-2021, 2021.

(2) The authors evaluate ET-WB on several time scales longer than a month. What about the performance on the sub-monthly scale? This point should either be supplied or discussed.

Response: Thanks for bringing up this point. We see potential in the sub-monthly scale evaluation of the water balance in essence. However, there are two major challenges that continue to make such analyses and/or subsequent interpretations ambiguous.

1) *Coarse resolution of the available datasets:* Limited by the availability of directly observed input datasets (i.e., P, R, and ΔS) on a finer scale, ET-WB is evaluated on the monthly timescale and evaluated only on timescales longer than a month (i.e., month, year, and multi-year).
2) *Outweighed uncertainties at finer scales*: Some of the water balance variables, for example, GRACE-based ΔS, have been interpolated on a daily scale (Kvas et al., 2019). However, their performance over the native coarser resolution is much debated. More extreme values of, for example, precipitation, will likely be observed on shorter time scales with higher uncertainty (Kim et al., 2020; Tabari, 2020). Therefore, sub-monthly (e.g., daily) calculations of ET-WB tend to pose challenges in confident interpretations, especially in a quasi-global domain.
   Nonetheless, timely updates to our ET-WB dataset would be integrated when appropriate datasets (e.g., daily data of constituent water balance components) can be acquired. We have added this point in the Discussion Section of the revised version.

References

Kvas, A., Behzadpour, S., Ellmer, M., Klinger, B., Strasser, S., Zehentner, N., & Mayer-Gürr, T. (2019). ITSG-Grace2018: Overview and evaluation of a new GRACE-only gravity field time series. Journal of Geophysical Research: Solid Earth, 124. https://doi.org/10.1029/2019JB017415

Kim, S., Eghdamirad, S., Sharma, A., &Kim, J. H. (2020). Quantification of uncertainty in projections of extreme daily precipitation. Earth and Space Science,7, e2019EA001052. https://doi.org/10.1029/2019EA001052

Tabari, H. Climate change impact on flood and extreme precipitation increases with water availability. Sci Rep 10, 13768 (2020). https://doi.org/10.1038/s41598-020-70816-2

(3) The authors explain the spatial pattern of uncertainty in ET-WB by individually presenting the uncertainty in different variables. However, the temporal changes lack necessary justifications, e.g., the decrease of auxiliary ET products in 2016 according to Fig. 9f.

Response: The abrupt decrease of uncertainty in auxiliary ET products after 2015 is due to the limited time coverage of FLUXCOM and MODIS products. Since the number of ET products used in the uncertainty calculations based on the standard deviation method has decreased from four to two, the RMS value of uncertainty has abruptly dropped from ~5 to ~2 mm/m. This decline does not necessarily mean that the uncertainty of ET products has decreased since 2015, while more likely a mathematical result of the lesser ensemble members. We have added explicit descriptions for such behaviour in the revised manuscript.

(4) It would be good if the authors could add the uncertainty range of the selected auxiliary ET products in the comparisons with ET-WB in terms of the long-term mean and annual trends.

Response: As suggested, we have added the change range of uncertainties, denoted as the standard deviation across the ensemble members, citing the lack of available uncertainties in the individual ET products (Long et al., 2014), in both the long-term mean and annual trends in four ET products in Figs. 7-8 of the revised version.

Reference:

Long, D., Longuevergne, L., and Scanlon, B. R.: Uncertainty in evapotranspiration from land surface modelling, remote sensing, and GRACE satellites, Water Resour. Res., 50, 1131–1151, doi:10.1002/2013WR014581, 2014.

(5) Fig. 1. The pie chart representing the irrigation percentage rate should be enlarged.

Response: As suggested, we have increased the size of the pie chart for better readability in Fig.1 of the new version.

(6) You may indicate the explicit number of datasets used in your workflow in Fig.2.

Response: We have added the exact number of applied datasets adjacent to each variable in Fig.2 of the new version. We have also modified the figure caption accordingly.

(7) Table 1 may go to supplementary file as (1) it occupies too much space, (2) it is somehow duplicate with Fig.2.

Response: Thank you for the suggestion. We have moved it to the supplementary file (Table S3).

(8) For colored shading plots like Figs. 3 and 5, the authors should indicate the meaning of the central red line in the legend.

Response: For these coloured shading plots, the shading area shows the spread range among different datasets and the central solid line meaning the ensemble median value. We have added explanations in the figure captions.

(9) Line 702-704: 'The human-induced inordinate fluctuations of water balance (e.g., through reservoir management, groundwater extraction) can influence the quality of ET-WB by impacting the accuracy of the specific forcing variable (e.g., R).' The statement needs to be re-organized for better comprehension.

Response: We have re-organized this statement as follows in the revised manuscript:
The human-induced inordinate fluctuations can influence the water balance and, subsequently, the quality of ET-WB by impacting the accuracy of the specific forcing variable (e.g., impact R through reservoir management).

(10) The authors should either direct the authors to Table S2 or explain the meaning of the rank of basin ID in Fig. 10.

Response: We have revised the caption of Fig. 10 to (a) explain the meaning of the basin ID and (b) direct the readers to Table S2 for salient features of the river basins.