Review of Petch et al, EGUsphere, 2023

Reviewer: Ruud van der Ent, Delft University of Technology

General comments

Petch and co-authors study the water and energy balance of 20 large river basins. Given my own expertise, I will mainly comment on the water balance part. Using GRACE-derived terrestrial water storage changes and globally available observation-based product of precipitation, evaporation and runoff they first study the imbalance between these products.

Then, they optimize their storage time series to nearly match that of GRACE and doing so update the fluxes as presented in Figure 3. Then, I failed to understand what the usefulness of the "Our Optimised Storage" and associated figures 4 and 5 are. It seems quite circular to me that if you force it to match GRACE it matches GRACE better than other products. From a hydrologist perspective I would rather see the optimized *P, Q* and *E* compared to other products. Perhaps even an independent better regional precipitation or river discharge dataset for specific basins, to see whether the optimized fluxes match that better and which would then clearly demonstrate the strength of their method. I am not sure if this is going to require major changes to the paper, or just a clearer explanation of the objectives and results.

A second major, but not difficult to solve issue, is that I find the authors somewhat sloppy regarding equations and symbology. Unfortunately, this set of guidelines has disappeared recently from the HESS manuscript preparations guidelines online: https://iahs.info/Publications-News/Other-publications/Guidelines-for-the-use-of-units-symbols-and-equations-in-hydrology.do, but I personally still appreciate it if we all try to follow this as much as possible. Of serious concern are Eq. (1) and Eq. (5), which should obviously read d*S*/d*t* instead of dS as the fluxes are per unit of time. It would also be helpful if the equations would contain the dimensions, thus, e.g. length^3 per time [L^3 T^{-1}] for Eq. (1) so this becomes obvious. Even expressed per unit area [L T^{-1}] would also be fine of course as long it clearly remains a flux and not a stock. Moreover, please use single italicized symbols, so something like S_{fi} instead of FIS, which makes it directly clear we are talking about a storage. I know many other papers invent funny acronyms as well, maybe it is even the rule rather than the exception, but in my opinion, it is simply not pleasant for any reader.

A third major question is why the authors chose the data they chose and whether it matters for the main point they are trying to make. For precipitation and evaporation, many more observation-based products exist, so did they select the 'best' according to some previous studies or did they just select 'good' data and does it not matter a lot whether it is really the 'best'. I hope the authors can explain. Moreover, the runoff data is even dependent on precipitation and evaporation from GSWP3, which is a bit of a vague product in terms of how it was constructed and I think it may even rely partly on GPCP and FLUXCOM, making the estimates of *P, E* and *Q* not completely independent. Moreover, I fail to see why spatially varying runoff is necessary at all, as on the basin scale, the actual river discharge measurements at the river mouth would suffice for which, for example, the GSIM archive (Do et al., 2018) could have also been used.

I hope to authors to give a quick response and perhaps we can even already settle some issues in the open discussion phase.

Specific comments

L1-2: "improving climate and earth system models"

I would say ‘validating’ or ‘assessing the capability of’ which is to be done first before anything can be improved.

L6: “the corresponding turbulent heat fluxes ranges between ± 10 W m^{-2}”

I suppose something should range between value x and value y, thus this sentence misses something.

L8: “This exposes mismatches in seasonal water storage”

Mismatches between what and what exactly?

L12: “The optimization also reduces formal uncertainties on individual flux components”

Sounds great, but I failed to clearly identify this result in the paper itself.

L14: “The FIS metrics”

What are ‘the FIS metrics’?

L23: “Water is a conservative quantity”

Technically speaking this statement is incorrect. Water is used by plants for photosynthesis and released by decomposition or fire. Probably it is an order of magnitude lower than the errors made in the products of *P, E* and *Q*, but not entirely negligible.

Table 1 “present” and general period statements

It is rather irrelevant whether e.g. GRUN is available until ‘present’ or that it starts in 1902, what matters is which years you used for the analysis.

L91 “evaporation”

I strongly support the use of evaporation over the ambiguous term evapotranspiration, see Miralles et al. (2020) for the arguments why that is, so perhaps you could simply use evaporation also elsewhere in the manuscript.

Equation 5

The integral is between what and what? What does the to the power 0 between brackets mean? Is this equation supposed to present a time series? Then it would be clearer if $S_{fi,w}(t)$ was explicitly written.

Technical corrections

L93: “Earths”

Earth’s

L101: “Land”

land

References

Do, H. X., Gudmundsson, L., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata, Earth Syst. Sci. Data, 10, 765–785, https://doi.org/10.5194/essd-10-765-2018, 2018.

Miralles, D. G., Brutsaert, W., Dolman, A. J., and Gash, J. H.: On the Use of the Term “Evapotranspiration,” 56, https://doi.org/10.1029/2020wr028055, 2020.