Response to anonymous referee 1

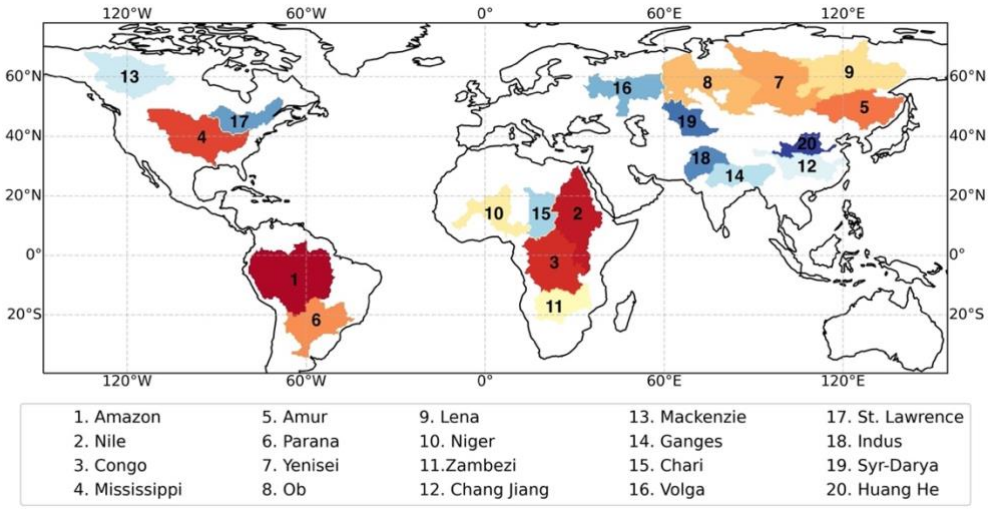
General comment

Overall, I believe this is an interesting work. The authors studied the water and energy budget in several larger rivers on both short and long-time scales. The manuscript could be accepted after major revisions.

Thanks for your helpful feedback, we really appreciate all your comments and have tried our best to address each of your concerns.

1.Please label the river basins in the Figure 1.

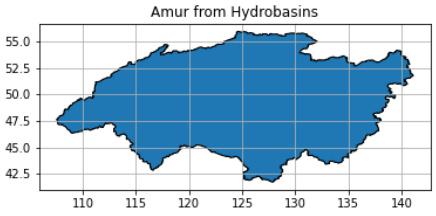
We have produced an updated version of Figure 1 (below). This includes a key to label the basins.



Colours are used to make the black labels easier to read. The colour scale is associated with basin size.

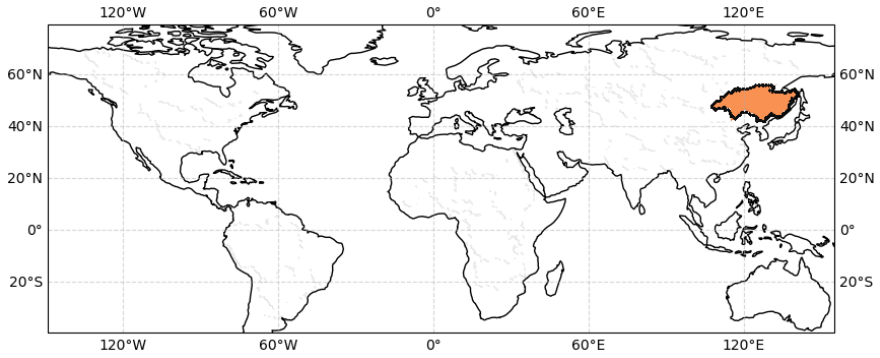
2. The boundary of the Amur River is not correct, which would make following results not right. Please check different maps to use the correct boundary.

We used a mask taken from GFZ Terrestrial Water Storage product (Boergens et al., 2020). However, we agree the Amur basin was odd compared to other maps, e.g., Hydrobasin https://www.hydrosheds.org/products/hydrobasins Lehner, B., Grill G. (2013). We have contacted GFZ authors to enquire.



However, we have now created a new mask for the Amur basin based on the Hydrobasins and reproduced results for the Amur. Results have only some slight differences and the new manuscript will be updated based on this new mask. The Amur is updated in Figures 4 and 8, values in Table 2, and Appendix Figures A and B. Discussion is also edited where necessary. Figure 1 above now includes the new Amur boundary. We compared the boundary for all other basins against Hydrobasin and all others seem to be ok.

New Amur mask:



The black outline comes from converting the hydrobasin coordinates onto 0.5 degree grids to match the other data used.

3. Please use appropriate font size and keep consistent in each figure. The font is too small to be readable.

I have reproduced all figures with larger font size 13 for axis labels and legends, and font size 12 for axis values. Previous font size used was python default size of 10.

4. Lines 374-381. It seems these are methodology, and should not be placed in the results.

Relevant parts have been moved to methodology section. However, some of these lines are describing other products that are shown in Figure 5, and we believe them relevant in results section to explain this figure.

5. Lines 419-427. This part is not well written. Each paragraph has only two or three lines. Please rearrange the text.

These lines have been rewritten to improve the quality of writing, and the paragraphs have been restructured to make sure they flow better.

6. Line 471. References are needed to support your statement.

We have adjusted the statement to "In previous budget studies longer-timescale constraints on the water budget have often not been applied Abolafia-Rosenzweig et al., 2021; Hobeichi et al., 2020)", and added references which are examples of studies where no long-term constraints have been used in optimisation.

7. When talking about optimization, we always cannot forget some popular optimization algorithms, such as SCE, DDS, GA, etc. What are the differences between your method and these popular ones?

Our optimisation involves linear budget equations only and therefore always has a unique monthly solution, so results should not depend on optimisation algorithm. The difference in our optimisation method comes from the constraints, which could also be embedded in other optimisation algorithms. We will try to be clearer about what we mean by optimisation in manuscript.

8. When reading paper, we always want to see the differences between your study and previous ones. Lines 457-470 stressed the similarities but not the differences. Please dig a little bit more to show the differences.

The similarities referred to here are only to emphasise that adjustments are of a similar size to previous budget closure studies (i.e. within Obs errors) but our results achieve an improved long term consistency of water storage changes with GRACE, wherein lies the difference in our results. We will re-emphasise this in text.

9. From the conclusion, I can see the main contribution from your study is that you introduced a sequential optimisation approach. Other than this, is there any new findings that different from other studies? There are a lot of optimization method can do the similar job. I want to see new findings that can advance our understanding of the hydrological processes.

The advance is in fully utilising the GRACE water storage observations to constrain regional water fluxes from monthly to decadal timescales. Previous optimisation studies have only used GRACE on a monthly timescale and have failed to match low frequency storage variations which are important to understand through hydrological modelling. By accounting for longer timescales, it enables us to use more information to provide stronger constraints than these other studies. These more consistent optimised products can then be more useful for model comparison studies for example. See also first response to reviewer 2.

10. A better judgment of the selections of the river basins should be given. Is it because the observation data in these rivers are better than others? Or other reasons. Some important river basins, such as the Mekong River, are not selected. No rivers in Western Europe are selected. I don't mean you have to select all the rivers, but an appropriate reason should be given.

The basins selected capture a range of imbalances in their observed budgets from the initial data, including basins with strong interannual variability, basins from a variety of latitudes, and basins that have other optimised flux products already in the literature. We are also restricted to larger basins, preferably with simple basin boundaries as these will have smaller GRACE storage errors as described in Weise et al. (2016). A more detailed discussion will be added into section 2.7 to give a clearer justification of the selection of basins. Note that a larger set of basins are included in the appendix figures.

References

Boergens, E., Dobslaw, H., Dill, R., Thomas, M., Dahle, C., Murböck, M., Flechtner, F. (2020): Modelling spatial covariances for terrestrial water storage variations verified with synthetic GRACE-FO data. International Journal on Geomathematics 11, 24, https://doi.org/10.1007/s13137-020-00160-0.

Lehner, B., Grill G. (2013). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrological Processes, 27(15): 2171–2186. https://doi.org/10.1002/hyp.9740.

Wiese, D., 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