Dear Bramha,

Thank you very much for your valuable comments. Your comments and suggestions were considered in revising the manuscript.

Please see our answers below in red.

With kind regards,

Eva Boergens (on behalf of the authors)

Summary: The manuscript uses GRACE(-FO) along with Altimetry, precipitation, and Evaporation datasets to analyse the spatiotemporal behaviour of the East African rift region. In terms of tools, STL and clustering algorithms were used first and then comparisons were made between several variables (lake storage, SPEI, and TWS) to draw conclusions.

General comments: the application of a clustering algorithm to identify regions with similar behaviour is one of the most interesting part of the manuscript, but this is not fully explored. The article has numerous language and grammar errors (from spelling mistakes to redundant and incorrect sentence formations). Authors indicate that they have investigated human vs climate signals, but the analysis in that direction is also weak. They found a good agreement between Altimetry and GRACE & GRACE-FO in general and that remains the most convincing part.

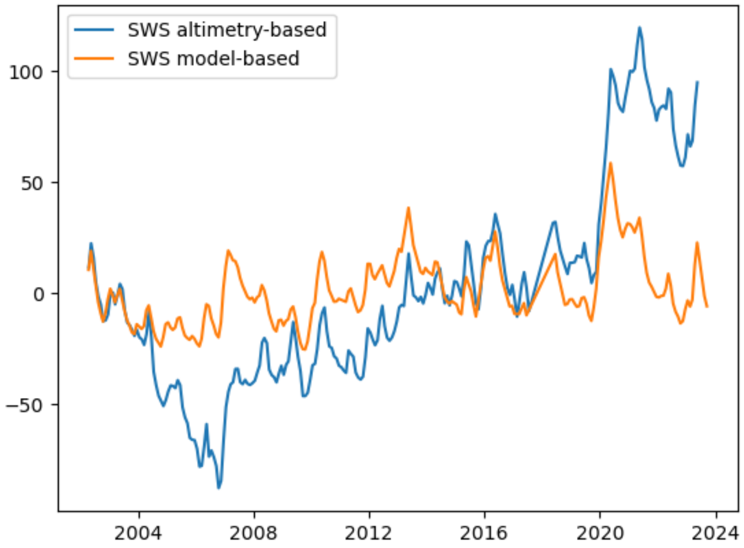
Thank you for this comment. We decided to shift the focus of the manuscript more towards the (geodetic) observations of TWS and of the water storage compartments in this region and their interpretation rather than the analysis of human vs climate signals.

Here are some recommendations/concerns/suggestions:

1. Line 5: here the study claims that it will characterize and analyze the interannual TWS variations over the East African rift region to provide a categorical classification: natural or human. Several important hydrological aspects that represent human and climate have been missed in the analysis: for example, groundwater is not accounted for in the whole analysis. The African Monsoon system has a huge impact on the decadal water resource availability in Eastern Africa, which has not been included in the discussion. Even the Monsson system is evolving with climate (see https://www.nature.com/articles/s43017-023-00397-x). Nevertheless, in the conclusions section, the characterization and analysis is not clearly written: how much of interannual variation can be explained by precipitation (or P-E) and how much of it is due to human decisions on lake outflow. It is appreciated that lake release data is not available, but some quantitative insights based on remote sensing data would add a lot of value and increase the impact of this work on our current state of understanding.

Thank you for this remark. In the revised manuscript, we focus more on the (geodetic) observations of the water storage compartments in this region and their interpretation rather than the analysis of human versus climate signals. The latter is difficult to disentangle by the integrative observations at hand and may also require a regional hydrological model to simulate the hydrological dynamics with and without human interference, which is beyond the scope of this study. Nevertheless, we used additional observations of other water storage compartments, namely root zone soil moisture and groundwater, to comprehensively analyse the contributions to the TWS changes (Section 5.3 Comparison between TWS Signals and Water Storage Compartments).

The authors of this manuscript have been part of the international consortium of the “Global Gravity-based Groundwater Product” (G3P) project funded by the EU as a Horizon 2020 project (https://www.g3p.eu/), joining several leading experts in Europe for satellite-based remote sensing of soil moisture (W. Dorigo, TU Wien), glaciers (M. Zemp, Uni Zurich), snow (K. Luojus, FMI) and mass changes with GRACE (A. Güntner, F. Flechtner, GFZ, T. Mayer-Gürr, TU Graz, A. Jäggi, Uni Bern). G3P provides groundwater storage changes as the difference between TWS and surface water storage (SWS), root zone soil moisture (RZSM), snow, and ice. The latest data set version, including all individual storage compartments, is available until 09/2023 (Güntner et al., 2024). While RZSM is satellite-based, the SWS variations are based on simulation results of the hydrological model LISFLOOD (Van der Knijff et al., 2008). However, LISFLOOD simulations of surface water storage changes are considered unreliable in the study region (cf. Prudhomme et al., 2024). In particular, despite similar dynamics and shorter time scales, the modelled SWS does not show the distinct and strong interannual variability we see in the altimetry-derived SWS of the study (see the following figure, not included in the manuscript).



Thus, we computed groundwater storage variations for the present study based on our altimetry-based SWS results as GWS = TWS - RZSM - SWS(altimetry). Snow and ice can be neglected in the study region.

Side note: Our altimetry-based SWS does not include river storage variations. Based on the model's different SWS components of rivers, lakes and reservoirs, we estimate that river SWS explains roughly 10% of the seasonal SWS variations in the study area and does not show large interannual trends.

In the manuscript, we now quantify and discuss the amount of TWS change explained by the different storage compartments and P-E.

After revising the manuscript, we thoroughly revised the abstract and conclusion accordingly.

Güntner, Andreas; Sharifi, Ehsan; Haas, Julian; Boergens, Eva; Dahle, Christoph; Dobslaw, Henryk; Dorigo, Wouter; Dussailant, Inés; Flechtner, Frank; Jäggi, Adrian; Kosmale, Miriam; Luojus, Kari; Mayer-Gürr, Torsten; Meyer, Ulrich; Preimesberger, Wolfgang; Ruz Vargas, Claudia; Zemp, Michael (2024): Global Gravity-based Groundwater Product (G3P). V. 1.12. GFZ Data Services. https://doi.org/10.5880/g3p.2024.001

Van Der Knijff, J. M., Younis, J., & De Roo, A. P. J. (2008). LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. International Journal of Geographical Information Science, 24(2), 189–212. https://doi.org/10.1080/13658810802549154

Prudhomme, Christel, et al. "Global hydrological reanalyses: The value of river discharge information for world-wide downstream applications–The example of the Global Flood Awareness System GloFAS." Meteorological Applications 31.2 (2024): e2192.

Looking into the monsoon system and its temporal changes in East Africa did provide further insights into our study. Thus, we did not elaborate on it.

2. Line 8: “separate the TWS signal” -- > “decompose the TWS signal”

Changed

3. Line 10: “study’s region” --> study region. This also raises the question if the study region chosen here is the same as East African Rift (EAR)? There are maps of the EAR that differ from the study region obtained via clustering. For example, the Lake Kariba and Lake Malawi (https://www.sciencedirect.com/science/article/pii/S1464343X05001251) are also part of the rift system but outside the study region here. If authors are choosing this name because it is already existing in literature, citing the source would help.

True, we only investigate the northern part of the East African Rift. We chose the name not because of existing names in the literature but because it appeared to be the best name for the region identified by the clustering in this study. We decided to change the name of the study region to the Northern East African Rift (NEAR) region (but not in the title).

4. Line 11: The sentence would read better if written as: We observe a decline in TWS un till 2006, followed by a steady increase till 2016, and a sharp increase in 2019 and 2020.

Changed

5. Line 13: “ large lakes of the region explain large parts” --> “lakes explain large parts”

Changed

6. Line 14: “alone contribute up to” --> “alone contributes up to”

Changed

7. Line 14: “Satellite altimetry reveals the anthropogenically altered discharge downstream of the dam”: This sentence hurts the coherence of the text. This may be moved to the first or second line in the paragraph.

Rephrased

8. It is already well known that lake water levels and the discharge from the Nile River are Anthropogenic. Authors have cited several papers that also find the same. Hence the last line of the abstract should contain a novel insight from this study.

All the cited literature was focused on the drought conditions prior to 2006. Here, we found evidence in support of these findings. However, the floods of 2019/2020 have not been investigated in the same way before, and here, we do not find any clear indication of the anthropogenic influence of the Nalubaale dam on TWS variability. Now, the last sentences of the abstract are:

The Nalubaale Dam regulates Lake Victoria's outflow. Water level observations from satellite altimetry reveal the impact of dam operation on downstream discharge and on TWS decrease in the drought years before 2006. On the other hand, we do not find evidence for an impact of the Nalubaale Dam regulations on the the strong TWS increase after 2019.

9. Line 21: delete: “cover equally surface and subsurface water storage compartments, i.e., they” (this info is redundant please remove)

Removed

10. Line 24: Please rephrase. Either it has to be complementary data or delete “and invaluable complement to all other”.

Removed

11. Line 25: “tiny” please use a more quantitative adjective such as (micrometer level).

Changed

12. Line 25: please rephrase : two twin satellites: language wise it appears that there are 4 of them.

Changed

13. Line 25: instead of “trailing each other” it should be “one following the other”.

Changed

14. Line 26-27: “From collecting these .... derived” --> These intersatellite range measurements over a month are then processed to obtain monthly mean gravity field of the Earth.

Changed

15. Line 27: “by computing and comparing .....investigated” --> Changes in the Earth’s gravity field are then represented in terms of mass changes near the Earth’s surface.

Changed

16. Line 35: Rewrite as: Quantifying continental scale terrestrial water storage (TWS) variations has been possible only with GRACE.

Not changed as it would alter the meaning of the sentence.

17. Line 46: “These region’s lakes ... ecosystems ” --> “These lakes have been named in the Global 200 eco ... ecosystems”

Changed

18. Line 52: “monitoring, standardised indices got well established, namely the .. . For example” --> monitoring, well known indices such as SPI and SPEI have been used extensively.”

Changed

19. Line 57: “storage variations are by now commonly .. “ --> Storage variations are now also monitored.. “

Changed

Similar changes are recommended for the rest of the manuscript. A thorough proof reading is essential. I will now only point out spelling mistakes for the manuscript after line 60.

20. Line 72: It is true that the region experienced drought and more water was released. However, then an independent Hydrologic engineer broke this news (https://archive.internationalrivers.org/resources/dams-draining-africa-s-lake-victoria-4117) and the treaty’s terms and conditions were enforced which led to a swift recovery. It would be nice to acknowledge that the dam water release was disproportionate and when ensured they were within agreed limits, conditions improved.

Thank you very much for pointing this out. We acknowledged this in the text.

21. Line 77: The 2018 Rodell paper has termed the TWS increase as “probable natural variability”, however, there are studies that also investigate the severity and cause for the trend. In Vishwakarma et al., 2021 (https://iopscience.iop.org/article/10.1088/1748-9326/abd4a9/pdf) the trend observed over the region is found to be “extreme gain” in comparison to the long-term hydrological natural variability. Then by Zhong etal., 2023, these trends have further been attribute to precipitation driven and non-precipitation driven (https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2023WR035817). They found that the trends observed are mostly non-precipitation driven events. This puts the region into “anthropogenic” category, not the “natural variability”. Discussion must be added to improve the attribution, or the lack of it.

Thank you very much for pointing out these two publications. We have added them to the introduction. However, as stated above, we will reduce the focus on discussing anthropogenic vs natural causes in the manuscript.

22. Line 93: Rephrase.

Rephrased.

23. Figure 2: this map is not of the full East African Rift region, but a part of it. See the map in (https://www.sciencedirect.com/science/article/pii/S1464343X05001251).

See our answer above. We changed the name to Northern East African Rift (NEAR).

24. Line 139: continues --> continuous.

Changed

25. Line 140 – 145: How does this method compares to that given by Wang et al., 2011 for converting lake height to storage? (https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2011WR010534)

The method used by Wang et al. is based on a parametric description between water level and volume. Thus, it does not allow for more complicated shore profiles. Further, data on both quantities is needed to estimate the power law relationship between water level and volume, which is unavailable in our study region.

We added to the text:

Here, we assume a monotonic but non-parametric relationship between WL and WSA. Thus, the ECDF method is more flexible for complicated terrains than methods fitting a parametric curve through the WL -- WSA relationship (e.g. Wang et al, 2011).

26. Line 149 –151: Please mention the interpolation technique used

Added that we use a linear interpolation.

27. Line 158: I am not sure what is meant by “multi-year interannual” Maybe I am wrong, but multi-year variations and interannual variations are synonyms.

Removed “multi-year”

28. Line 163: The guidelines for choosing STL parameters are indicative only. For regions such as EAR, where there are two wet seasons and Monsoon system exists (see climatology in Figure 1(d) in https://www.nature.com/articles/s43017-023-00397-x and Figure 5 in https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2013WR014350), you may also

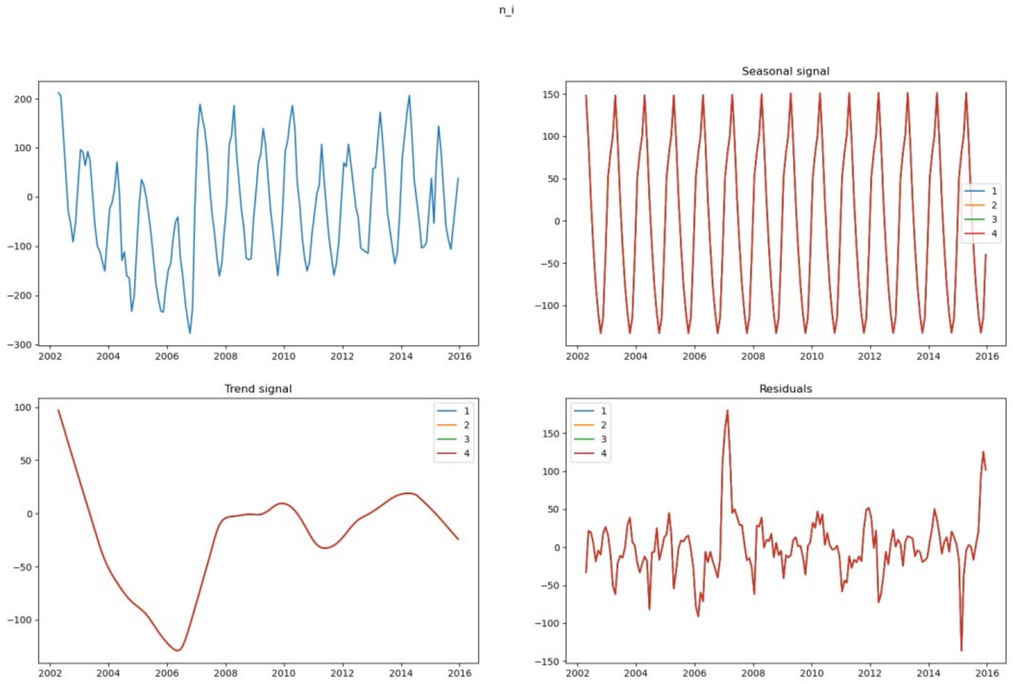
test the seasonal signal to be semiannual. Maybe more investigation is needed to show that these parameters are a good choice.

We tested the parameter. We found the semiannual signal to be less well-fitting to the observed TWS signal than an annual signal. Although precipitation follows a semiannual cycle, TWS does not.

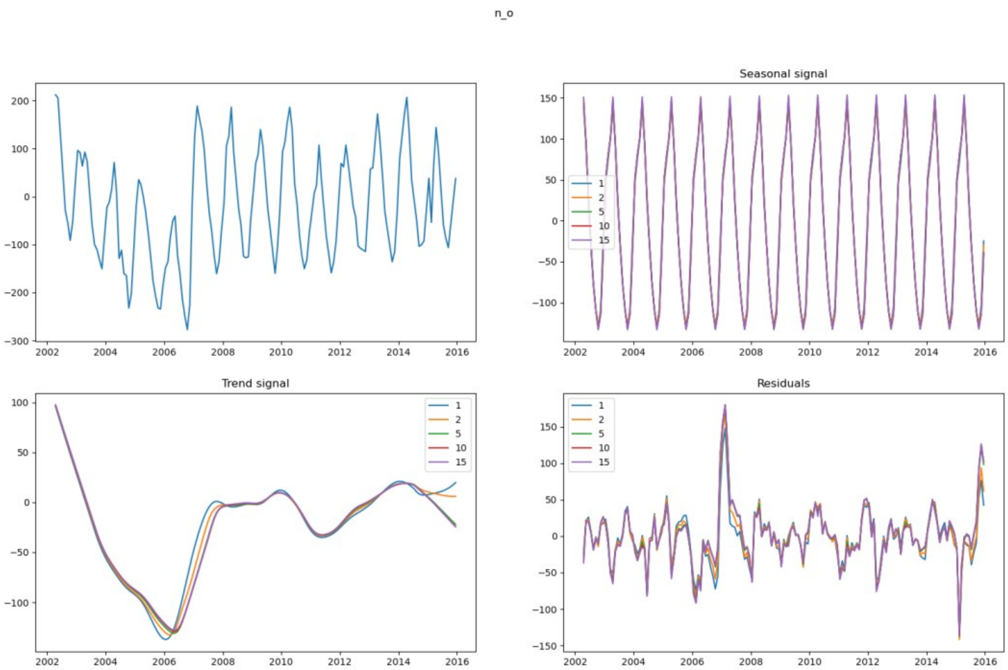
The following investigations will not be fully included in the manuscript but are included for your information, as we know you are an expert using STL. Nonetheless, we included more explanations in the text.

We tested for the same example time series one parameter after each other. Thus, all parameters were fixed to the values given in the manuscript except for one variable.

n_i: The following figure shows that the inner loop quickly converges. Thus n_i=1 is sufficient.



n_o: We cannot assume that TWS data is Gaussian. Thus, n_o has to be larger than 0. For this example, convergence is reached for n_o=5. However, to be safe for all time series, we choose n_o=10.


n_o

Seasonal signal

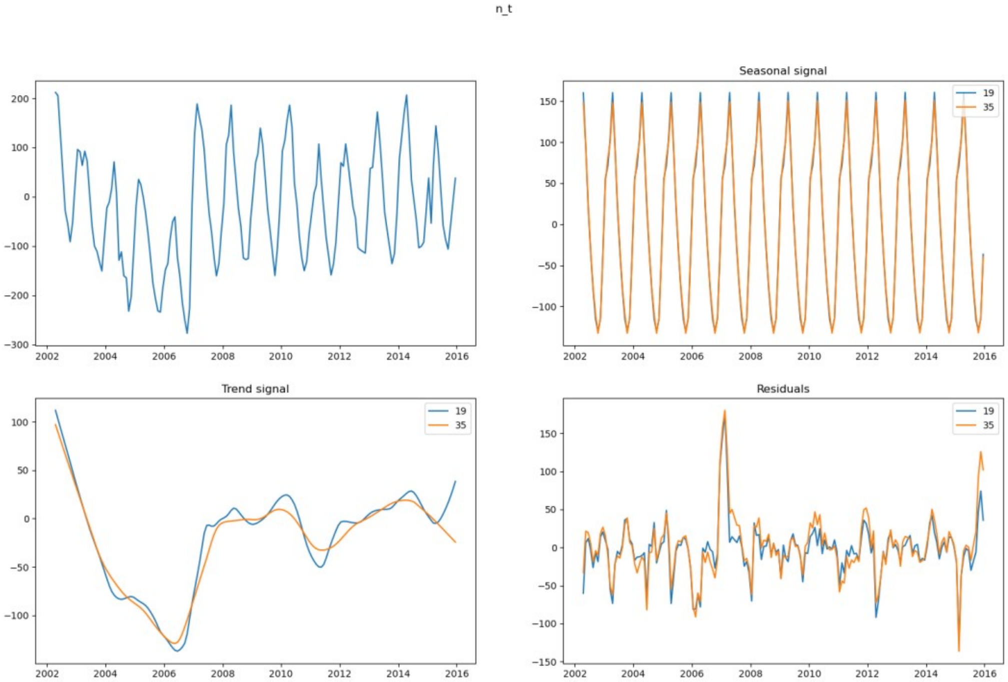
Trend signal

Residuals

n_l: Following the reasons given in Cleveland et al. (1990)

n_s: We understood Cleveland et al. (1990) that this parameter should be an integer multiple of the seasonality (here 12) minus 1. Thus, we produced the seasonal diagnostic plot for n_s=11, 23, 35, 47. Here, the seasonal diagnostic plots for n_s=11 and 23 show variations we interpret as noise rather than signal. With n_s=35, the seasonal signals no longer exhibit such residual noise. Although we expect some changes in the amplitude of the seasonal signal due to climate change, these changes are expected to be slow. Thus, our choice for n_s=35. We will not show the seasonal diagnostic plots due to their large size here.

n_t: Having fixed n_s=35, the choice of n_t would be 19 if following the reasoning given in Cleveland et al. (1990). However, we found that too much noise was still present in the resulting time series with this parameter choice. Thus, after trying out larger values of n_t, we decided to choose n_t=35 (see the figure below).



We added to the text:

The results of the STL decomposition depend on several parameters that govern the smoothness of the interannual and the annual signal. Cleveland et al. (1990) provide guidelines for choosing the parameters, which we used together with empirical testing and visual inspection. That results in the following parameter values. n_p: Length of annual signal, 12 in our case; n_i and n_o: the number of passes through the inner and outer loop, set to 1 and 10, respectively; n_l: the width of the low-pass filter, to be set to the least odd integer larger than n_p, thus set to 13; and n_t and n_s: the trend and seasonal signal smoothing parameter, both set to 35. While the former four parameters are straightforward, n_s requires more considerations. We chose the value n_s=35 in such a way that we consider the interannual variability of the seasonal signal no longer governed by noise. n_t depends on the value of n_s. However, we found that the value provided by the rationale given in Cleveland et al. (1990) (n_t=19) produced a trend component containing still too many short-term variations. Finding the value n_t=35 was done with empirical testing and visual inspection.

29. Figure 4: The caption appears to be incorrect. Original TWS time-series is red while the data gap is in black. Is blue and green also correctly matched?

The colours were incorrect. Corrected in figure and caption.

30. Line 195: 700 km diameter is good enough to be resolved by GRACE (Vishwakarma et al., 2018: https://www.mdpi.com/2072-4292/10/6/852). Not sure why it was termed “not meaningful” here.

Thank you for this comment. You are entirely correct that a 700km diameter is sufficient for GRACE studies. The reason for not further increasing the number of clusters is that with only one cluster more, ring-shaped clusters will appear (see current Fig. 7, removed in the revision). Such shapes will become difficult to interpret.

We changed the text to (m is the number of final clusters):

We decided on the final value for m based on the results, especially the size and shape of the regions. We sought the largest number of clusters while keeping them reasonable for GRACE data interpretation. Here, we found m=8 to be the optimal number.

The ninth cluster would be ring-shaped and only about 100km across in the narrowest place. Thus, such a region is no longer meaningfully interpretable with GRACE data.

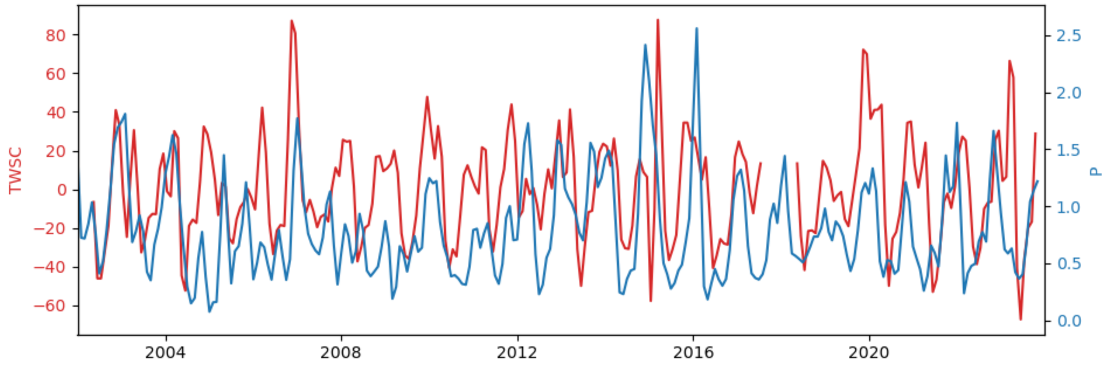
31. I believe the figure 5 has some interesting time-series. The Cluster analysis divided the region into entities that could also be explained via climatology and human intervention. This aspect was not explored further for Africa (maybe something in the future), but at least for EAR, there should be some discussion about what makes it unique in terms of climate and why this clustering makes sense.

Exploring Africa's clustering in more detail is beyond the scope of this study, but we invite you to contribute to future work on this!

We want to point out that the clustering is oriented towards TWS variations and thus independent of other spatial categorisation schemes such as river basins, precipitation patterns, or others. As the study aims to investigate the drivers for the observed TWS changes, we do not think it necessary to anticipate discussions about this already in the clustering results section.

32. Section 5: Monthly precipitation and TWS are not directly comparable because of the water budget equation, where P = ET + R + d(S)/dt. To make a fair comparison its recommended that the TWS time-series is differentiated and then compared to the weighted mean Precipitation data (see equation 2 and 3 in Lehmann et al., 2022).

Thank you for this suggestion. We tested the suggested formulas to compare dS/dt with the monthly precipitation. Again, we do not have ET data. The comparison between P and TWSC shows only a weak relationship between these two (correlation = ~35%). See the following figure:



These findings do not improve the interpretations of TWS vs precipitation.

However, instead of differentiating TWSC from TWS, we can also integrate the other parts of the water budget equation. In the manuscript, we choose to do this by accumulating precipitation and using the drought index SPEI that employs accumulated data.

We added to the manuscript at the beginning of now Sec 5.2:

TWS and precipitation are not directly comparable but linked to each other via the water budget equation: d (TWS)/dt = P - ET - R, with R being the runoff. We decided against differentiating TWS for the comparison but to temporally integrate P and P-ET to evaluate similarities. We call the temporally integrated precipitation ``accumulated precipitation''.

Also the concept of accumulated precipitation is not clearly explained. After reading the first paragraph of section 5 three times, I had three interpretations. For example, is the TWS compared with Annual averages of P? Or Is there a moving window of 12 months to compute accumulated P? or the P is accumulated for 12 months and then again for 12 months, which is then added to the last sum of 12 months? It is unclear. Figure 9 has a plot with accumulated P and TWS plotted, which helps me rule out the first and third option, but still not clear.

We applied the same accumulation intervals for P as used for the published SPEI. Thus, the values for each month are the sum with a moving window.

We clarified the text:

Instead of monthly precipitation, we consider time series of accumulated precipitation. To this end, the accumulated precipitation value for each month is the sum over the preceding n months, with n taking integer values between 1 and 48.

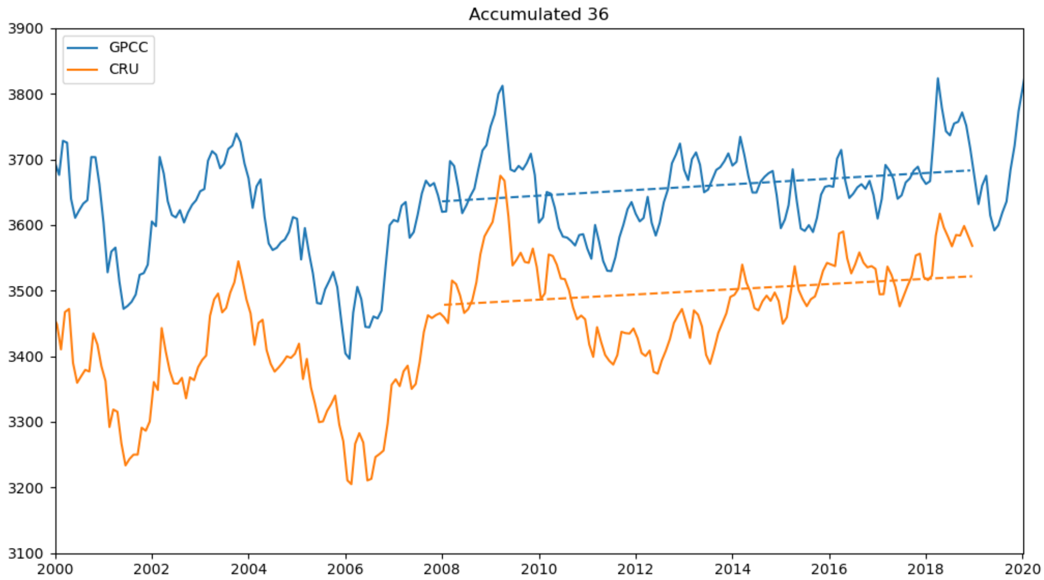
33. Figure 9 axis label says precipitation, but the caption says accumulated precipitation. Also, the units of precipitation should be (mm / [time]) and when accumulated (integrated) should become mm. Please revisit this aspect as well.

34. Figure 9: Why is the magnitude of TWS in excess of 4500 mm and there is no negative value. Is TWS also accumulated? The interannual TWS in figure 7 has a range [-200 +400].

Thank you for spotting the mistakes in the figure! We corrected both.

35. Figure 9: Another important observation is that the SPEI (GPCC-based) shows no rise between 2008 to 2019, while there is a rise in TWS as well as SPEI (CRU-based). If there is a lack of trust in precipitation product then how reliable are the conclusions drawn based on them?

We investigated the accumulated precipitation time series for GPCC and CRU (following figure) with an accumulation period of 36 months (the same period used for SPEI).



The amplitude and the trends since 2008 differ between the two data sets.

We cannot rule out that differences in PET used for the two SPEI data sets may also contribute to the SPEI differences. However, the PET data were unavailable for a direct comparison.

We added the figure of the accumulated precipitation to the appendix and added to the text:

The comparison of the two precipitation data sets revealed significant differences in the overall volumes of accumulated precipitation but similar interannual dynamics (see B1 in the Appendix). Increasing rainfall trends since 2008 are slightly larger for CRU (4.3 mm/year) than for GPCC (4.0 mm/year) which may partly explain the diverging patterns of the two SPEI data sets after 2008. Nevertheless, differences in the (potential) ET data used for the two SPEI times series may also contribute to the differences, but the (potential) ET data itself were not available to us.

Further, we added at the end of the section:

Further, the discussed differences between the two SPEI data sets and their input precipitation data sets limit the explanatory power of the comparison between SPEI and TWS. Hence, we will investigate the different storage compartments of TWS in the next section to identify further drivers of TWS variability.

36. Line 228: The decision to choose 48 over 36 needs more thought. The parameters chosen for STL window size could be the guiding light.

Thank you for this suggestion. After revisiting the mentioned part of the manuscript, we decided to shorten the hydrologically trivial (according to reviewer #2) discussion and instead use the STL parameter as a guideline. We found that for both the smoothing of the seasonal and trend components, the minimum length was 35 months. Thus, we also assume that the accumulation period should be at least 3 years, which is also in line with the investigations currently provided in the manuscript. We decided to change the accumulation period of precipitation and SPEI to 36 in the study and revised the part of the manuscript accordingly.

37. Line 250: govern --> governed, despise --> despite

Changed

38. Since the signal leakage due to filtering is a problem that will reduce the quality of observations, is it possible to rather use a leakage-correction method for improving GRACE data (such as the forward modelling approach by Chen et al., 2015) instead of filtering the SWS data from altimetry and Lake area?

We decided against using the leakage correction provided together with the used GravIS TWS data, as this correction is only suitable for regional means. Although large parts of the study are done on the regional mean time series, we also investigate spatial patterns. We think it better to have a common data basis for all parts of the study and not switch between leakage-corrected time series and uncorrected spatial patterns.

For applying another leakage approximation based on forward modelling (like Chen et al., 2015) or hydrological models, we decided against this as we do not trust the hydrological models sufficiently in the region (see answer above about LISFLOOD) as input to this method.

Additionally, we found only small differences between the region's leakage-corrected and uncorrected time series. Thus, we felt that leakage is not a big issue for this study.

39. Figure 10: The SWS change slowly while TWS drastically between 2010 and 2017. Precipitation is also not increasing much as seen from Figure 9. The increased lake storage

might also interact with the groundwater system. Hence a different rate of change could be attributed to groundwater-recharge (see Figure 5 in https://www.sciencedirect.com/science/article/pii/S0048969721044284)

As explained above, we introduced groundwater storage change data into the investigation. This allowed us to investigate the yearly storage changes more closely due to the different storage compartments. We also looked into the different behaviour of the storage compartments during dry and wet years but found no clear correlation.

40. Line 89: In this chapter --> in this section.

Changed

41. Section 8, conclusions: Authors claim that there are clear linear trends over Northern India. If one uses STL there is a strong interannual variability over North-west India as well. As we increase the time length, interannual (decadal also) variations start to appear, which is why a longer time-series is needed for climate analysis.

Line 310: courser --> coarser

42. Line 361: this became possible --> this was made possible

Three comments above were taken into account during the revision of the conclusion.

The manuscript is easy to read in parts and requires some effort from the readers in others. A thorough proofreading is required. This is quite an interesting problem and I wish the authors all the best. I hope these comments will be helpful.

Best wishes,

Bramha Dutt Vishwakarma