Dear reviewer,

Many thanks for carefully reading our manuscript and for your helpful suggestions for improving it. Our responses to your comments are given below in reddish color.

The authors have introduced the Global Gravity-based Groundwater Product (G3P), a satellite observation-derived dataset of monthly groundwater storage variations from 2002 to 2023 at 0.5° of spatial resolution. The authors underlined the challenge of monitoring global groundwater changes due to the limited groundwater in situ data and the need to have such a product. They used observations from the Gravity Recovery and Climate Experiment satellite mission which provide total water storage anomalies (TWSA), unique means of quantifying groundwater storage anomalies (GWSA) at large spatial scales. To derive the changes in GWSA, they subtract the storage anomalies of surface water storage anomalies (SWSA), soil moisture in the unsaturated zone (UZS), snow water equivalent (SWE), the mass of glaciers (GM) and ice sheets (ISM) from TWSA, following a water budget approach. For these different storage anomalies, they have estimated the uncertainties which they propagated to get the one of GWSA. The findings cover the development and the assessment of the global GWSA dynamics which revealed regional variations indicating both depletion and recovery in different areas.

This is an important paper in the sense that the developed G3P data set which relies mostly on the observed datasets of different water storage compartments serves as a prototype for an operational global groundwater service, crucial for sustainable water management. The paper is impressive in its scope and thorough in its analysis. While the results look convincing, well-written, and nicely presented, I have suggested a few minor changes that might be helpful for the authors to further improve the manuscript. Other than these small typos, there is a lot to recommend publishing this paper.

Line 32 - 35, could you consider inserting a brief statement of the results itself of the evaluation. Abstract should briefly provide the main results (e.g., number on statistics). The way this sentence is presented, it just describes the step followed in the paper without highlighting the main results.

We add a sentence on the evaluation results in the abstract.

Line 40, in the sentence below, there is a typo error. The verb "are" should be replaced by "is" I think. The authors could also consider changing the verb "weigh" to an usual verb such as "measure" for fluidity in the comprehension. Here is the sentence: "A main source of the data are observations with satellites 40 that weigh the changing amount of water by its mass attraction effect".

We rephrased the sentence to: An important part of the data comes from satellites that track changes in water by its gravity effect.

Line 44 - 49, there is no reference in these sentences. Please could the authors consider inserting some references at different sentences.

These introductory sentences describe basic textbook knowledge on the water cycle and on groundwater. We suggest not to include references for it.

Line 113 - 127, this part did not reference some recent relevant papers which also point out the origine of uncertainties related to GWSA estimation using the mass balance equation: 1) Getirana et al. 2025 (https://doi.org/10.1029/2025GL119197) and Arifin et al. 2025 (https://doi.org/10.1029/2025GL118580). I suggest the authors include these references as this part is dedicated to a kind of literature review.

Also following the second reviewer's suggestion, we add the reference of Getirana et al. 2025 to this paragraph. Also, Arifin et al. 2025 is additionally added here, while we referred to it already in the original manuscript with respect to G3P uncertainties and evaluation (line 777).

Line 137, many acronyms have been already defined in the upper part of the text. Why are the authors continuing using the expanded form and keep putting their acronym in brackets? Just use the acronym provided above. e.g., "…. Monthly ground-water storage anomalies (GWSA)……".

Thank you for this hint. We carefully checked the manuscript and removed full variable names with their acronyms, except for sub-chapter titles and for one instance in the conclusions (CDR) to keep clarity.

Line 168 - 170, For the GWSA data set that covers the period 04/2002 to 09/2023, the anomalies are calculated with the same long-term mean over 04/2002 –12/2020? or the long-term mean has been updated to the period 04/2002 to 09/2023? I will suggest to the authors to make it clear in the text in this paragraph as done further in the text (Level-2B Processing section)

Yes, GWSA is calculated relative to the long-term mean over 04/2002 –12/2020, similar to the GRACE Level 2B data, even though the data coverage is 04/2002 to 09/2023. We resorted and rephrased these lines to clarify this point.

Line 180, for consistency and uniformity in text, I will suggest the authors change " The TWS anomaly" to "The TWSA".

Done.

Line 255, I guess this acronym "RMSE" stands for root mean square error, but its extended form is not provided anywhere in the text. Please, the authors should consider providing this.

Correct, we add the full expression to the revised manuscript.

Line 302 – 304, in the equation (4), SWI and K stand for what? They are not defined. Please, could the authors consider defining these variables.

Thanks for pointing to this issue. In fact, SWI was falsely used instead of RZSM in equation 4. In the revised version, we corrected the equation, inserted an explanation of the gain factor K, and rearranged the paragraph to improve the description of the method.

Line 310, Considering the different T-values of the EF corresponding to the four depth layers, and given that this depth layers depend on plant species (Martens et al., 2017), how were the T-values implemented? Did you use the mean value of the reported range, or did you differentiate the T-values according to plant or vegetation species for example?

We add a sentence to the revised manuscript to explain how the T values for each depth layer were derived: “The optimised T-value for each depth layer is the one that maximises the Pearson correlation between satellite soil moisture, after applying the EF to the SSM estimates, and field measurements taken within that depth interval, sourced from the International Soil Moisture Network (ISMN).”

Line 324, there is a typo error in this sentence " The here developed RZSM data…", please, kindly remove " here" in the sentence.

Done.

Line 406, typo error " …geodetic surveys.,", kindly remove the comma from this sentence

Done.

Line 417, Why the mean annual mass change time series uncertainty is multiplied by "1.5" only. How did this number come in? could the authors insert a bit of explanation on how "1.5" was chosen to get the relative monthly uncertainty.

Due to logistic reasons, there are not enough monthly glaciological in-situ (!) measurements that would allow to assess the monthly uncertainty from observations at global scale. Hence, we decided to increase the error bar of the annual product, which is based on annual observations, by a somewhat arbitrary factor of 1.5. This is added to the revised manuscript: “This multi-plication factor is somewhat arbitrary, but it ensures that the monthly downscaled product has a larger uncertainty than the annual product."

Line 636 – 639, For comparison purpose and to check if current results align with existing knowledge, I suggest the authors to compare their results for the major basins (e.g., Amazon, Congo, Mississippi, Nile, etc) with the results of Figure 2 of Getirana et al. 2017 (https://doi.org/10.1002/2017GL074684). This could be inserted in the text as an additional table reinforcing the results of an observed GWSA compared to a modeled GWSA.

As also stated in our responses to comments by reviewer 2, we found that model-based GWSA can vary substantially between hydrological models (in our case in the ISMIP model ensemble) and that a comparison to one modeled GWSA output may not be conclusive in evaluating the G3P performance. Nevertheless, we computed the climatology (mean seasonal cycle) of

TWSA, GWSA and the other storage compartments for the major river basins, similar to Figure 2 in Getirana et al. (2017), as suggested in the reviewer's comment above. The results (see Figure I below) show similarities of the GWSA seasonality of G3P with the Noah-MP model-based GWSA seasonality of Getirana et al. (2017) for many river basins, but also notable differences for some basins, in particular Nile, Niger, Yangtze and Tigris-Euphrates. These differences occur in particular in river basins where the seasonal variations of soil moisture (RZSM) in G3P show strong amplitudes, similar in size or even larger than the TWS amplitudes. It should be noted, though, that RZSM uncertainties have the largest contribution to the final GWSA uncertainties in the G3P data, see Figure 15 in the original manuscript.

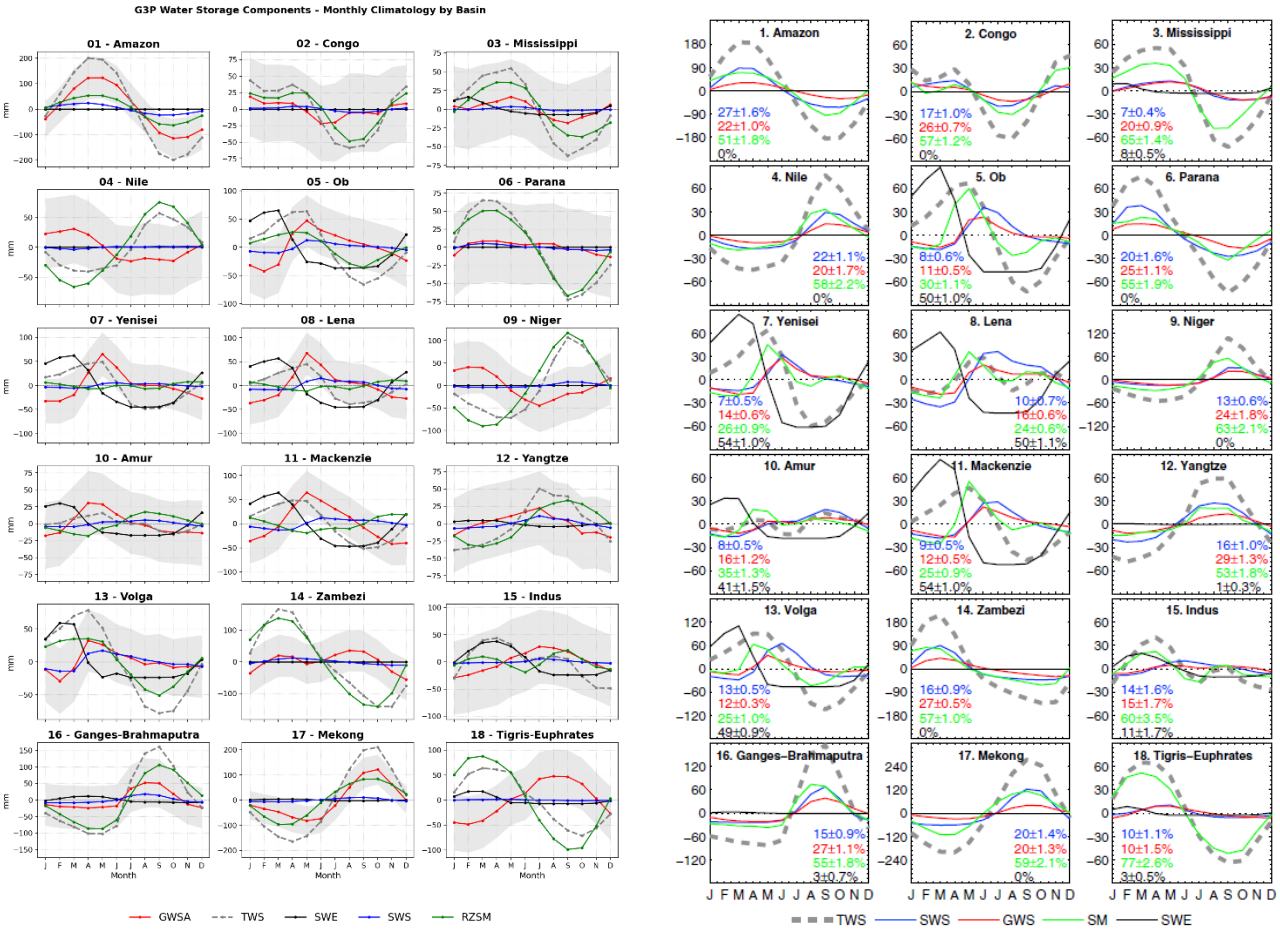

Figure I: Seasonal basin-average variations of water storage (climatology for the period 2003-2014) in different WSCs and in TWS for the G3P data set, with propagated uncertainty ranges of GWSA (grey shaded area) (left panel) and for the simulation output of the Noah-MP land surface model + HyMAP routing scheme from Figure 2 of Getirana et al. (2017) (right panel).

Line 773 – 776, I suggest that the authors refer to the recent paper by Getirana et al. 2025 (https://doi.org/10.1029/2025GL119197) in this paragraph. The paper deals with inconsistencies in GRACE-Based Groundwater Storage Estimation.

Reference to Getirana et al. (2025) added here (and in the introduction).

Line 781 – 793, I will suggest the authors to refer in this discussion paragraph to a recent study on the Congo Basin by Kitambo et al. 2025 (https://doi.org/10.1016/j.ejrh.2025.102810) where G3P groundwater dataset was compared with another groundwater dataset. The use of observation-based surface water storage in this study was the game changer which enhances the temporal dynamics of the derived groundwater. This was assessed indirectly with in situ discharge data.

Thank you for pointing to this recent study that used G3P data for comparison. Referencing Kitambo et al. (2025), we add a sentence on the difference between the model-based surface water storage data used for G3P and the observation-based SWS product used in that study for the Congo River basin. In spite of similar amplitudes, the observation-based SWS led to a different seasonal phase of the resulting groundwater storage variations. We add this sentence to the paragraph where we discuss the performance of the SWS data in the revised manuscript. (This location in the text is different from the reviewer's suggestion to add it to the paragraph where G3P evaluation against in-situ groundwater observations is discussed.)

Line 837 – 839, I will suggest the authors to mention in this paragraph the recent advances in the estimates of SWSA from satellite observations in different environments worldwide. Many have been done to measure observed-SWS changes over all types of water bodies in major river basins and there is ongoing effort for the development of such SWSA datasets at the global scale. Kindly refer to Papa and Frappart, 2021 (https://doi.org/10.3390/rs13204162). There is also this paper, Lin and Song, 2026 (https://doi.org/10.1016/j.jhydrol.2025.134458). For lakes, there is Yao et al. 2023 (DOI: 10.1126/science.abo2812).

We add the suggested references on global-scale SWS monitoring to the revised manuscript.

Line 900, there is a typo error here "… due climate variability and/or climate change…". Please, kindly add "to" between due and climate.

Corrected.