Review of "GRACE water storage estimates for the Middle East and other regions with significant reservoir and lake storage" by Longuevergne et al.

This paper addresses an interesting research question and provides good analysis, including a synthetic analysis and illustrative examples. It is generally well organized and clearly written. The only major comment I have is for the synthetic study. The authors have done a nice job examining different cases, but I wish the authors provide a more solid and in-depth analysis on the synthetic study. Besides this, I will provide below some minor suggestions to improve this manuscript. Overall, I would recommend accept this manuscript if the following can be addressed:

Thank you very much for the constructive comments. For the sake of clarity, we better introduced the "point mass" problem with a methodological figure and associated discussion. Finally, we strengthened the "case studies" part with evaluation of uncertainties. Concerning the analysis of the synthetic study, the authors believe that the three points highlighted are the most critical and should not be lost in a longer discussion.

Major comments:

Numerical study:

1. The authors investigated two scenarios but the percentage of the reservoir mass (change) over the total basin mass (change) in the synthetic basin is not mentioned. Additional examination/discussion with this could further strengthen this paper. Please at least provide discussion for this.

The synthetic study considers a 1 km³ mass change within a basin, in both cases. This is indicated in the text, and above each experiment. Moreover, the equivalent mass change over the basin is indicated as TBS (true basin storage). The caption now includes this information.

2. The authors quickly reach three major findings in this section. I feel there are more interesting results in Figure 1 and the relative figures in the supplement, e.g, the change of the curves.

The authors believe that the three points highlighted are the most critical. In particular, this experiment clearly shows the impact of mass distribution on GRACE basin-wide average storage change. However, we extended the discussion section in the synthetic experiment (see text).

3. How possibly the GRACE processing could infect these results/findings? The authors mentioned that smoothing, does the results hold true for different smoothing methods? Please provide discussion.

Beyond the synthetic experiment, examples illustrate the impact of point masses for regular processing (i) truncation at degree 60 plus 300-km Gaussian smoother (unconstrained solutions) and (ii) truncation at degree 50 with no further filtering (available constrained solution). As shown, smoothing affects the way point masses are captured. More filtering options are shown in figure in supplementary material (figure S1 and S2). Some more words are added on point (3). Therefore, filtering (truncation and smoothing) affects the way GRACE captures the impact of point masses. Generally speaking, the main conclusion holds true, the impact on GRACE of a point mass for basin averages rarely equals the reservoir mass, whatever filtering is used. Reducing filtering makes GRACE basin-wide averages more sensitive to point masses

4. Overall, I enjoy reading the numerical study and it is nicely done. Please list/discuss any limitations of such a design, if possible. What is the limitation of such a numerical study?

There is no real limitation to the numerical study, it is here to illustrate the importance of mass distribution on GRACE basin-wide averages. Then the computation of the impact of point masses is conditioned by the actual knowledge you have, so we underline the information required for a complete interpretation of GRACE data: beyond gathering LSM output, a complete image of storage compartments (amount and distribution) is required.

In the second illustrative study, the authors mentioned that not all surface storage is considered. Could such a numerical study possible include similar situation?

Yes indeed, synthetic experiments allow the testing of multiple hypotheses and a best understanding of observation data. The article is quite long by itself and we wanted to focus on the hypothesis "quantifying mass changes within a basin depends on the mass distribution inside and outside the area of interest", which has never been investigated before. It is quite important that people interested in GRACE data could understand this point.

As GRACE processing is linear, it is quite straightforward to compute the potential impact of masses not considered in a study, but you need basic information such as total mass and distribution.

Minor comments:

1. In abstract, L 15: seasonal variations of what?

Seasonal variations in water storage, updated in the revised manuscript

2. P11134, L11: what does "the disaggregation equation" refer to? Please re-arrange this sentence to avoid confusion.

This has been updated. The objective of GRACE studies generally consists in extracting the individual contribution of groundwater storage changes, which requires disaggregating Δ TWS as follows: $\Delta GWS = \Delta TWS - \Delta SWES - \Delta SWS - \Delta SMS$.

3. P11136, L 10: Consider "Method and dataset" as the section title. Also, please reorganize the subsections in section 2. How about putting current 2.3, 2.4 and 2.5 under 2.2, since these are the datasets used for the case studies but not for the numerical study.

ОК

4. P11139, section 2.4: since many acronyms are used and this makes this section hard to read. I suggest the authors to provide a simple schematic figure which can clearly convey the main message. It helps the readers to catch the text.

This part is indeed very descriptive but required to understand the following points. The paragraph has been slightly reorganized to be easier to understand

5. P11139 L20: please put "Oki and Sud (1998)" at the end of the sentence.

6. P11139 L23: Pease justify why these two GRACE resolutions are examined. Is it because these are usually examined in the literature? Please provide information for readers to follow.

Added after the description of the products. CSR and GRGS solutions use the same GRACE range-rate data and similar geophysical background models, but employ independent processing strategies. They have been chosen because CSR solutions are considered as one of the least constrained solution, while GRGS is considered as one of the most constrained.

7. P11140 L14: change "to simulate GRGS" to "to simulate the data from GRGS". Similar change applies to the other.

Right. Updated.

8. P11140, L21: Please provide simple description of the difference/relation between ABS and TBS.

Added: In other words, ABS is what GRACE "sees" from the true mass TBS.

9. P11140, L20: Please consider to replace "the sum of" and re-write this sentence. It is misleading. If you claim it is the sum of different items, you should be able to decompose the total effects.

This is true. Replaced with "the combined effects"

10. P11144, L26: should it be "not all surface bodies..." instead "all surface bodies..."?

Corrected.