Response to Reviewer #3 Comments:

This study presents a multi-satellite remote sensing approach to understand long term storage changes in over six thousand reservoirs around the world. The authors combine well-established remote sensing based reservoir monitoring techniques to build monthly time series of storage variations. These variations are then synthesized with streamflow to provide insight into long term trends. This is an important study that pushes the boundaries of our understanding of global reservoir storage variations and explores possible drivers of the observed changes. However, I have two major concerns and several minor concerns that should be addressed before publication.

We would like to thank the reviewer for the thoughtful comments and suggestions. The reviewer provided us with helpful comments, which will greatly improve our manuscript. Below please find our response to reviewer's comments in detail.

R3C1) First, I am unsure of the value of using long term trends to characterize reservoir storage as increasing or decreasing between 1985 and 2015 (as in lines 210-240). Figure 2 suggests that reservoirs of these sizes can go through shorter, but still multi-year periods of increased and deceased storage throughout this time period. For example, Fort Peck and Fairbairn Reservoirs show ~10 year long oscillations in storage that are not easily characterized by simply increasing or decreasing trends.

We agree. For our long-term analysis, we not only calculated whether there were increasing or decreasing trends from 1984-2015, but also tested whether these trends were significant or not using the Mann-Kendall trend test ($p<0.05$). The red and blue points in Fig.4 (L225-227) and the basins with black outlines in Fig.7 (L269-271) showed significant increasing or decreasing trends in storage. The changes in storages for Fort Peck and Fairbairn Reservoirs are non-significant trends according to the Mann-Kendall trend test ($p<0.05$). In contrast, for example, the change in storage is significant in one basin of southwestern USA, although there are 10-year-long oscillations in storage (Fig.4).

R3C2) Second, I am unconvinced of the conclusion that human intervention is an insignificant contribution to storage variability. According to equation 8, changes in storage are related to Qin and Qout (assuming small E). One could argue that any change in storage is due to human alteration of Qout, because without modification of Qout (relative to Qin) there would be no storage variation at all. Without some quantification of the drivers of Qout (hydropower demand, irrigation needs, etc.) I find it hard to make an argument for Qin to be the dominant driver with only what has been quantified in this study. Perhaps an alternate way to frame your findings is that Qin can be used as a good predictor of positive or negative reservoir storage variations.

We thank the reviewer for this constructive suggestion. We agree that the conclusions on the influence of human intervention are not as robust as desirable without direct evidence, but we would argue that our logic to deduce the role of human activity is reasonable and our conclusions sufficiently circumspect, although we will look for phrasing that reflects that better. We are also able to refer to some individual regional studies that came to a similar conclusion in the revised manuscript. Please refer to our response to R2C1 for full details.

Line comments:

R3C3) Lines 65-79: The limitations of past efforts and techniques are summarized well here, but how this study overcomes these limitations and provides something new should also be given a sentence or two here.

Thank you, we will add a few sentences to highlight the advancement of this study over previous ones.

R3C4) Line 125: This figure could use a legend describing what the colors and inner and outer rings area.

We will add a legend to this figure

R3C5) Line 150: Would reservoirs constructed during the study period have an impact on the quantified Qin for older reservoirs?

Thank you for this question. The hydrological modeling used in this study does not simulate human interventions on river flows and therefore would not reflect any such changes.

R3C6) Line 171-190: I was confused by the methods for calculating reliability, resilience, and vulnerability. How does assuming 90% reliability simplify the calculations? Why is this a reasonable assumption?

We apologize for the confusion. We will include a real reservoir as an example to introduce reliability, resilience, and vulnerability. Please see our response to R2C2 for full details.

R3C7) Line 205: The two vertical axis on Figure 2 and 3 need to be equal for each subplot. As it is now, only correlation is apparent, but it would be much more realistic to plot the observed and predicted on the same vertical scale to get a realistic sense of the errors.

Thank you for this suggestion. We will change it to use the same vertical scale

R3C8) Line 342-350: This paragraph felt a little out of place here. Maybe consider moving the content to the methods section.

Agree. We will move this paragraph to the method section.

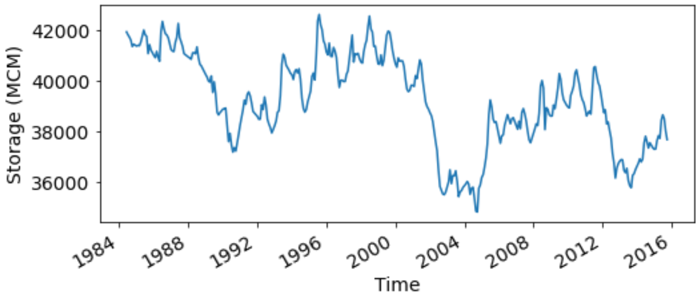

Figure 4 Total monthly storage dynamics with significant decreasing trend in one basin of southwestern USA.