Comment on hess-2021-350
Anonymous Referee #3

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.

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.

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.

Line comments:
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.

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

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

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?

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.

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