Response to the referee comments (RCs)

Anonymous Referee #1

R: The manuscript provides a long-term datasets of reservoir storage in Mainland Southeast Asia. This is meaningful for studies about the reservoir operations and further studies on hydrological processes. However, there are still some comments needed to be illustrated as listed below.

A: We thank the reviewer for the positive feedback. We will carefully address all your comments to strengthen the manuscript.

R: Line 43, Steyaert et al., 2022 and Steyaert and Condon, 2024, these two references are not found in the reference list. Please check.

A: At line 43, we referred to:

Steyaert, J. C., Condon, L. E., WD Turner, S., & Voisin, N. (2022). ResOpsUS, a dataset of historical reservoir operations in the contiguous United States. Scientific Data, 9(1), 34.

Steyaert, J. C., & Condon, L. E. (2024). Synthesis of historical reservoir operations from 1980 to 2020 for the evaluation of reservoir representation in large-scale hydrologic models. Hydrology and Earth System Sciences, 28(4), 1071-1088.

We will ensure that these two references are cited correctly in the revised version of the manuscript.

R: In table 3, for two attributes 'Water surface area (empty reservoir)' and 'Absolute storage (empty reservoir)', why they are marked as empty reservoir?

A: The water surface area and absolute storage start from a value equal to zero (i.e., representing an empty reservoir) in the Area-Elevation-Volume curves, which is why we used the expression "(empty reservoir)". We understand that this expression can be misleading, so we will remove it from the revised paper.

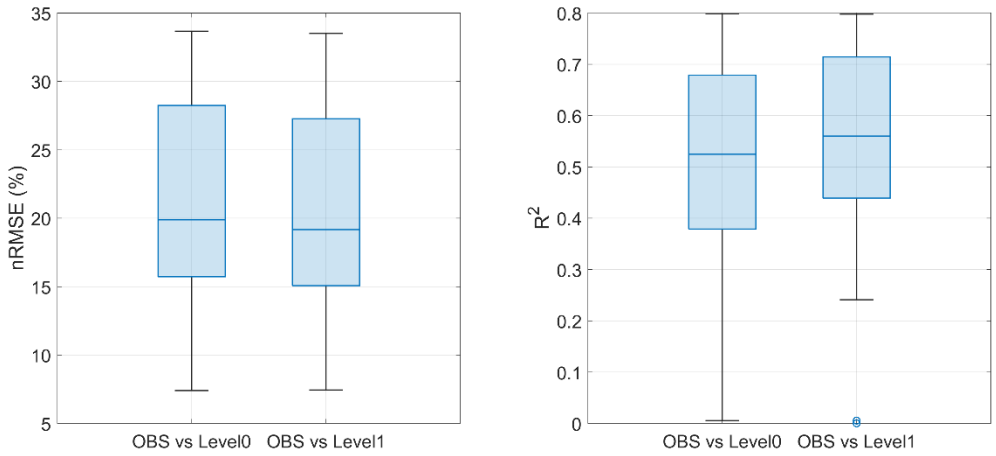
R: I've downloaded the dataset and am a bit confused about the water area extraction. There are three attributes 'Before_area', 'After_area' and 'Final_area', how are they derived respectively? How was the 'Final_area' determined? some of them equal to 'Before_area' and some equal to 'After_area'. Please add more explanations in the manuscript.

A: 'Before_area' refers to the water surface area of a given reservoir, calculated using the k-means classification technique applied to NDWI images. However, in binary classified images (water and non-water) with data gaps due to cloud masking, the estimated water surface area ('Before_area') is likely to be smaller than the actual water surface area. To address this, the binary images were enhanced to fill these data gaps, resulting in a revised water surface area estimate called 'After_area.' The 'Final_area' represents the area obtained after adjusting the boundary water pixels. Notably, if no adjustments were detected by the algorithm, the 'Final_area' remains equal to the 'After_area.'

To clarify these elements further, we will revise Section 3, providing a more detailed explanation of the dataset components.

R: Is the estimation improved in level 1 comparing to level 0? If so, how much?

A: Level-1 data are obtained after removing the outlier from the Level-0 data. The improvement in Level-1 compared to Level-0 varies between the reservoirs. To quantify it, we calculated the R^2 and nRMSE for level-0 and level-1 of the 20 reservoirs for which we have observed storage. We found that the nRMSE decreased and R^2 increased from Level-0 to Level-1, suggesting an improvement in our results. We will add the following supplementary figure in the revised manuscript.



R: Section 4.4, I understand that direct validation is limited by observations, but I assumed indirect validation can be applied to most of the reservoirs, why only 20 are presented? Can authors present more results?

A: For indirect validation, we need data from Altimeters, which have limited passes over the 185 reservoirs we studied here. There are 29 reservoirs for which altimeter passes are available; however, approximately 2/3rd of the reservoirs only have data points available for a considerable amount of time, i.e., at least ten years. Therefore, we had to limit the indirect validation to a total of 20 reservoirs. Overall, the limited availability of observations from altimeters reinforces the need to work with satellite images if one is interested in studying all / several reservoirs within a region of interest.

R: Figure 9, there are some discrepancies between the spatial distribution of precipitation deficit and storage deficit in reservoirs, especially in 2020, the lower part are not suffered from precipitation deficit but with less water stored. Please explain why.

A: The discrepancy between the spatial distribution of the precipitation and water storage anomalies is likely due to the topology of the cascading reservoir system. In other words, some reservoirs may be located in regions characterized by positive precipitation anomalies, but may

receive limited inflow from upstream reservoirs located in regions affected by droughts. We will further analyze this aspect and improve Section 4.5 to better explain Figure 9.

R: Section 4.5, drought is supposed to be a prolonged disaster that can affect the long-term water availability. It would be interesting to look into the time series of water storage in reservoirs to explore how reservoirs are affected and how they can alleviate the influence of droughts.

A: Thanks for the comment. We will include a supplementary figure showing the time series of water storage of some major reservoirs to discuss how reservoirs were affected and how they alleviated the influence of droughts in 2019-2020.

R: Line 424-426, are there any evidences that China held back water in its dams from any data or references? If not, please remove this sentence.

A: We will remove this sentence from the revised manuscript.