Responses to the comments on the Manuscript “ A dataset for lake level changes in the Tibetan Plateau from 2002 or 2010 to 2021 using multi-altimeter data ”

Dear editor,

The authors would like to express thanks to the anonymous reviewers for their voluntary work and the constructive comments to improve this manuscript. All of the comments are of great benefit to us. During the past few days, we did much work to revise the manuscript according to the reviewer's comments. All of the comments have been addressed. Our revisions are as follows.

Editor comments::

Thank you very much for your new contributions. Many thanks to the reviewer for the deep-going reviews and thanks for the authors for the detailed replies and the edits on the manuscript and dataset on lake level time series on the Tibetan plateau.
The manuscript is now in minor revision and we would like to thank the authors, who revised manuscript and dataset publication and our referees for these constructive efforts.
Dear authors, please consider the reviewers comments and technically discuss all issues. Please also consider the critical final comment from the open discussion.
There is also the editorial minor request for a minor addition to the PANGAEA data publication that will not change the status of the DOI-referenced data publication in 2024. Thank you for providing the detailed information per lake in form of the kml and the md files that is a very user-friendly format. As your data set is very complex please provide also this overview in form of a short technical product guide on the file structures (and optimally on the processing scheme also including figures and maps) in pdf format that could be attached by PANGAEA to the abstract text, i.e. this would not be part of the downloaded dataset but is part of the abstract text and could be downloaded from there.

Reply:

Thank you for your constructive feedback and the opportunity to improve our manuscript and dataset. We sincerely appreciate the reviewers' insightful comments and have carefully addressed all technical and editorial suggestions. As requested, we have prepared a Technical Product Guide (PDF) summarizing: File structures: Organization of TXT files, KML, and MD files. And region maps, processing workflow.

This guide has been submitted to PANGAEA for attachment to the abstract page (separate from the dataset download) to improve user accessibility, which will not alter the DOI-referenced data publication (2024 status).

Reviewer1:

Authors have satisfactorily addressed my main concerns and I thank them for their careful answers and edits to the manuscript. I think the paper could published in ESSD, provided the authors address my specific comments below. There is no need to have a new round of review on my side.

Reply:

Thank you for your positive feedback and for acknowledging our revisions. We sincerely appreciate the time and effort you have dedicated to reviewing our manuscript.

We have carefully addressed all of your specific comments (point-by-point responses are provided below).

Specific comments:

Abstract: 'The period for the lake level change series, which affords high accuracy, can be much longer for many lake systems.' → this sentence is not clear and the term 'high accuracy' is quite subjective. Here you should provide the RMSE (and correlation) when your database is compared to few available in situ gauges (and recall the number of gauges) and the cross-validation with Dahiti, Hydroweb.next and G-REALM. It will be more informative.

Reply:

Thank you for highlighting the need for clearer validation metrics. We have revised the abstract to incorporate specific accuracy statistics and cross-validation results. Below are the details. "The lake level change series shows good consistency with in situ measurements, demonstrating a median RMSE of 0.19 m across 8 validation gauges. The dataset further exhibits robust agreement with established satellite altimetry products (DAHITI, Hydroweb.next, and G-REALM), with median RMSE values below 0.30 m in all cross-validation comparisons."

In Table 1 the time period and number of lakes of each dataset might need to be updated. This is especially true for Dahiti and Hydroweb (which is now hydroweb.next), which have near real time time series for some lakes (I don't know if this is the case for lakes in the Tibetan Plateau) and new lakes might have been added to the database (here again, I do not know if this is the case for the TP). If it has changed, then the text in section 1 should be updated accordingly. In Table 1, you should also add G-REALM (https://ipad.fas.usda.gov/cropexplorer/global_reservoir/), especially as you are comparing your results with this database in section 4.2.

Reply:

Thank you for your suggestion, Hydroweb is updated the number of lakes in TP, now it has 46 lakes, Dahiti is not changing during the period. The G-REALM is also added in the table and text.

In section 2.2.1, 'ICESat-1 is a lidar altimeter, distinct from above radar altimeters. Its technique provides high spatial resolution and small footprint, but results in less measurements over time.' This sentence should be expanded with few words like: 'because of its orbit repeat period of 91 days and impact of clouds'. Otherwise, some people might think that less measurements over time are only due to the sensor itself.

Reply:

Thank you for your suggestion, we revised the sentence into "ICESat-1 is a lidar altimeter, distinct from the above radar altimeters. Its technique provides high spatial resolution and small footprint, but results in fewer measurements over time due to its 91-day orbit repeat period and frequent data gaps caused by cloud obstruction."

Section 3.1.1, 'potential interference or submergence of water signals by those from adjacent land areas' → I don't think 'submergence' is the correct term here. Would 'interference of adjacent land areas signals with signal from water body' fit authors' message?

Reply:

Thank you for your suggestion, we revised the sentence into "due to the interference of adjacent land areas signals with signal from water body."

The threshold level variable name is slightly different between Figure 2 (DistanceThres) and the text (e.g. DistanceThresh, p.6 lines 164 and 166). Please, use the same variable name in the figures and in the text.

Reply:

Thank you for your suggestion, we revised the figures to make sure the same variable name in figures and in the text.

Line 148 page 6, HDEM should be defined here (currently, it is explained in Figure 2 and in line 162 page 6)

Reply:

Thank you for your suggestion, we revised the sentence into "the optimal retracked levels should be within the range of the Digital Elevation Model (DEM)-based reference elevation H_DEM±20 m."

Figure 3 is a good figure to show the impact of the two-step retracker. In the legend, could you give the mean longitude and latitude of the measurements and the name of the lake to help reader to locate this lake? The water level time series provided in figure 3 has a 'zig-zag' shape for many cycles. Have you been able to validate this time series against in situ water level? If yes, it would be good to mention it in the text.

Reply:

Thank you for your suggestion, we added the lake name and which satellite (Ayakkum Lake using SARAL measurement) been used here in figure title to make the reader locating this lake.

Unfortunately, this lake does not have an insitu data, but by cross validate with other satellite in the same lake, the time series are fitting good with time series from different satellite. The final fusion time series shows the same changing during this period.

In section 3.1.2, line 179 page 7, 'Waveform classification is an effective method for identifying the noise observations' → please reformulate, as 'noise observations' is not clear. I guess you meant 'waveforms that are too noisy and difficult to process', do not you?

Reply:

Thank you for your suggestion, the sentence has been revised into "Waveform classification is an effective method for identifying highly noisy waveforms that are challenging to process accurately."

Line 189 page 7, 'tracks with fewer than five valid observations were excluded from further analysis' → provide few words to justify why below 5 valid observations the track is discarded. Besides, the term 'valid' might not be the most appropriate. You are mentioning waveforms that does not fulfill the methodology shown in Figure 2 flow chart. It does not mean that kept waveforms will provide actual lake water level (see the case of the Mattenalpsee and Oberaarsee lakes with different value of the OLTC that you provided in your answer to one of my previous review comment).

Reply:

Thank you for your suggestion, we revised the sentence into "After removing noisy observations, tracks containing fewer than five quality-controlled observations (post filtering via DEM-based elevation thresholds and waveform classification criteria) were excluded to ensure statistical reliability."

Here is the quality-controlled by H_DEM and waveform classy, we would like to keep more points to make the two-step retracker more stable, so decide when smaller than 5 observations, the tracks would be excluded.

OLTC shifts may introduce systematic elevation biases when two proximate lakes exhibit significant elevation differences ($\Delta H > 50$ m). In such cases, the tracking window could erroneously

lock onto the adjacent lake's surface, leading to implausible water level estimates for one of the pairs. This is not been processed in here, but later when do the cross validation with other satellite.

Figure 7, Change 'RMSE' with 'RMSE (in m)'

Reply:

Thank you for your suggestion, we update the figures in revised version.

In section 4.3 concerning 'Potential source of error', you should also discuss the impact of ice and snow on lake WSE estimate, which impacts a lot the radar altimeter signal. See the important literature on this issue.

Reply:

Thank you for your comment, the impact of ice and snow is big, we add several sentence to describe it. "Additionally, ice and snow cover introduce significant uncertainties in radar altimeter measurements. During frozen periods, radar pulses may penetrate snow/ice layers, measuring subsurface features rather than the true water surface (Guerreiro et al., 2017). Ice formation alters surface reflectivity, causing peak retracking misidentification (e.g., false peaks from ice-water interfaces; Chen et al., 2021, Song et al., 2020)."

In section 5, you should remove the lake without statistically significant trends in figures 8 to 10 and the numbers provided in section 5. You should write in the text that you computed the p-value and explain which threshold you set on p-value to assess if the trend is statistically significant or not. In appendix B, there are multiple lakes with p-value>0.01 and even some with p-value>0.05. You should also write in the text the number lakes without a statistically significant trend.

Reply:

Thank you for your suggestion, we have removed the lake without statistically significant trends through setting at the level 0f 0.05 ($p<0.05$) to assess if the trend is statistically significant or not, and also wrote in the text the number lakes with a statistically significant or not. Please see the revised manuscript.