Comment:

This study combines altimetry data that measure lake levels directly with shoreline positions from optical data to create extended and denser lake level time series for the largest lakes of the TP. In that sense, the resulting dataset differs from existing lake level time series and seems thus a valuable resource for the scientific community as well as other users. The study is relevant for ESSD and worth publishing. To properly document the data and methods and to comply with ESSD's guidelines, the manuscript needs to be improved - in particular to better describe important parts of the methods, include/consider uncertainties, and properly validate the time series against existing data sets.

Response:

We really appreciate the overall evaluation, insightful comments, and recommendation by this reviewer. Our point-by-point responses to the reviewer's comments are given as follows.

General comments:

1) The study would benefit from a clearer story line and justification how this work/data fills a current knowledge gap. I only understood the plot halfway through the methods. What are the shortcomings of the existing studies/datasets, and how do you overcome these with your study? This is especially important for the introduction, but also the abstract and conclusion would benefit from an easier to understand quick summary. See also comment paragraph P8, L11ff below.

Response:

Thanks for this constructive comment. As suggested by this Reviewer in specific comments, we have reorganized several paragraphs and enhanced how our study and developed data set fill a current knowledge gap in the introduction, abstract and conclusion sections. Abstract and conclusion sections have also been improved by reducing all redundant information. Details can be found in the attached modified manuscript.

2) Method: the important novelty of your approach is the use of shoreline positions from optical data to increase the temporal resolution and extend the length of the water level time series. To do so, you relate shoreline positions to lake level elevations from spaceborne altimetry data, using a statistical relationship between the two. Currently, the statistics part is not well enough described, and uncertainties from the found relationship do not seem to be propagated to your final "optical water levels". I suggest you extend this part to provide more transparency and include also a discussion of the uncertainties, considering in particular the assumption of a linear relationship (?) and whether it is appropriate to extrapolate beyond the range of measured lake levels.

Response:

Thanks for this very insightful comment. As suggested by this Reviewer, we have extended section 3.2 (optical water level) and provided a discussion in section 4.2 to better evaluate the uncertainty in the regression relationship and how it propagates into optical water levels. The extrapolation problem is discussed in section 4.2 as an interpretation of the propagated regression uncertainty and in this response letter too (specific comment 8 of the method section). We believe that the impact of extrapolation of optical water levels possibly occurring in the time gap between two altimetry time series has been well addressed in this response letter (specific comment 8 of the method section) and will be added to the supplementary file. However, we acknowledge that little information is available to quantify the effect of extrapolation during the time window from 2000–2002, as little altimetry information is available due to either poor quality or limited observations, and available DEM is too coarse to describe the micro topography of the lake bank. We have informed potential readers/users of such a risk in the validation and conclusion sections of the revised manuscript.

3) Dataset: I'm missing a detailed description of the final dataset and its attributes and uncertainties, e.g. after the validation section.

Response:

Thanks for this comment. The description of the dataset is combined with the data availability following the validation part now.

4) Validation (and uncertainties):

   a) What is the accuracy/uncertainty of the altimetry products, and how does this propagate to your optical water levels? Consider also the uncertainty of the statistical relationship (s) you compute to derive the optical water levels.

Response:

Thanks for this insightful comment. We used the standard deviation of water levels from valid footprints in a cycle to represent the uncertainty in the altimetry product. The valid footprints are referred to as the footprints selected with the histogram method as illustrated in the manuscript. For most cases they comprise more than 80% of all available footprints in a cycle. As suggested by this Reviewer, a thorough discussion of the error propagation from the altimetry data to the optical water level through the statistical relationship has been added in section 4.2.

   b) The theoretical computation of an uncertainty (most of 4.2) based on a single UAV image is not convincing to me as it is based on a single image pair only with unknown coregistration accuracy (see comment below). The lack of hands-on data basis and the extensive length of the theoretical part makes this off-topic. Maybe this could fit as supplementary information in a separate document.

Response:

Thanks for this insightful comment. We have redone the uncertainty analysis based on high-resolution optical images from GF-2 (i.e., China's high spatial resolution satellite) and investigated a total of 4128 Landsat shoreline pixels after performing co-registration (the co-registration error was estimated to be ~2 m). Based on the new experiments and results, we have modified part of section 4.2, making it more convinced. Considering the excessive content of section 4, we will move part of the theoretical derivation to a supplementary file as suggested by this Reviewer.

c) Rather than treating the comparison to the LEGOS Hydroweb data as an application case this should be part of the validation section. How do your time series compare to the other datasets listed in table 1?

For data description, uncertainties and validation see ESSD's guidelines at https://www.earth-syst-sci-data.net/10/2275/2018/, in particular sections 3.3, 3.5 and 3.6.

Response:

Thanks for this comment. We have moved part of the comparison with Hydroweb data to the validation section. We chose to make a comparison with the Hydroweb because Hydroweb data have exploited most altimetry missions and provided densest altimetry water levels among all listed studies (also for most lakes the systematic biases between altimetry missions seem to have been well removed), very typical for altimetry-based lake studies. Other altimetry-based lake studies may include more lakes, but based on the published results they are subject to some systematic biases. Therefore, we have taken the Hydroweb data as the benchmark to see if there are improvements or advantages in our generated product.

We did compare our lake data with that of Yao et al. (2018) and show the importance of temporal resolution, as we are not comparable with the lake quantity of these kind of studies based only on Landsat images and DEM. Studies that primarily use Landsat images and DEM are able to cover a larger number of lakes and are not subject to systematic biases as those using various altimetry data sources. However, most of those studies have a low temporal resolution (e.g., annually or even lower) due to the difficulty of acquiring quality optical images covering entire lake areas at a high temporal resolution, as opposed to our study that needs optical images covering a small portion of the lake shore.

Specific Comments:

A simpler title might make it easier to understand what the study is about. Especially the rather unclear terms "densified" and "developed optical water levels" should be replaced. Focus on the data and not the application cases.

Response:

Thanks for this constructive comment. The tentative title of this study has been revised as: "Generation of long-term and high-temporal-resolution water level and storage datasets for lakes on the Tibetan Plateau using multiple altimetry missions and Landsat-derived lake shore positions and areas" for your kind suggestion.

Abstract

1) The abstract could be more to the point. Add some information on the performance of your data (uncertainties and validation). Consider removing already published findings (applications).

Response:

Thanks for this constructive comment. We have removed numerical results similar to some published work such as lake storage trends and lake overflow amount. More information on the validation and uncertainty has been added, as we performed additional experiments with high-spatial-resolution images. Details can be found in the revised manuscript.

2) L12: which altimetric missions? If there are too many to list all, specify how many and which types (e.g. Lidar altimetry, interferometric SAR altimetry...)

Response:

Thanks for this comment. All altimetric sensors used in this study have been listed in the abstract now.

3) L13: avoid putting important information in brackets. Monthly to weekly time series? L16: "partial altimetry data" and "optical water levels" are unclear terms

Response:

Thanks for this constructive comment. This sentience has been modified. Brackets in L13 have been removed and a brief explanation to optical water levels has been added.

4) L19: "densified" is unclear

Response:

Thanks for this comment. It has been replaced with "merged".

5) L20ff: Are these groundbreaking new numbers/findings? Consider removing them and focus on the dataset.

Response:

Thanks for raising this comment. These numbers are actually not that different from published studies, but they can serve as an independent source of information from relevant studies, as we have generated a new dataset with temporal resolution being greatly improved and systematic biases being well removed. We have removed these numbers and placed more emphasis on the dataset itself.

Introduction

1) P2, L3: A strong motivation for TP lake studies not mentioned here is to find out why they are expanding, i.e. a good data set will contribute to a better understanding of climate and circulation patterns and changes thereof. This is important as the TP has a strong influence on both regional climate.

Response:

Thank you so much for this comment. We have added this to the first paragraph of introduction to clearly state the motivation of TP lake studies that a good data set should contribute to a better understanding of climate and circulation patterns and changes.

2) P2, L6: source of that number?

Response:

The source is (Messager et al., 2016).

3) P2, L8: I wonder why you selected exactly these references? There are many more lake studies on the TP. References for the method (general) and local application should be separated.

Response:

Thanks for this comment. We agree that more general studies instead of local applications may be cited. Now we have cited the earliest one that we can find to represent this kind of studies using remotely sensed water surface height and extent performed by Frappart et al. (2005).

4) P2, L11: It is better to introduce radar and lidar separately as the systems and data are quite different. Also, these data are not meant for ice berg height - you probably mean ice sheet surface elevation or sea ice freeboard?

Response:

Yes, this makes sense. We have separately introduced laser and radar altimeters and added a supplementary description of the two types of altimeters to underscore the differences between them in this paragraph. We agree that the altimetry data are not meant for ice berg height. It has now been corrected in the revised manuscript.

5) P2, L16: The satellite is called ICESat, not ICESat-1. Change everywhere.

Response:

Done.

6) P2, L25: it seems you mainly mean (and in your study only use) optical data. Do

you have an example for a sensor and study that used SAR data?

Response:

Yes, SAR images from Sentinel-1 were used by Huang et al. (2018) from our group to derive the effective river width, which is calculated with the river surface area divided by the river length. The automatic extraction of the river surface area is similar to that of the lake surface area or lake shoreline changes. We may take advantage of SAR data in future studies.

7) P2, L26: why exactly these references? These are not the only or first such studies.

Response:

It is true there are many published studies on water classification/extraction. We chose these two references mainly because they are similar in study area, data source, and publishing time, showing a good comparison between methods. We would like to show a change in this kind of studies and to stress the point that manual extraction of lake boundary could be labor-intensive and low-efficiency.

8) P2, L33: references for the water index and Otsu algorithm?

Response:

Done.

9) P3, L10f: remaining bias: is this not true for your study, too? Or how do you avoid/remove such bias?

Response:

We have done our best to remove the systematic bias between different altimetry missions by using optical water levels as reference data, which is rarely seen in the literature. Hwang et al. (2019) showed that the systematic bias among different altimeters is hard to remove unless in situ water level measurements or Jason-1/2/3 data are available. Our method could provide a better solution to this problem. We would not say there is no remaining systematic bias in our data, but we are confident that the biases have largely been reduced. Even though there might be some concern about the accuracy of the optical water levels because altimetry information is involved in the generation, they are currently the best available long-term reference data for ungauged lakes.

10) P3, table1: Does this table include all studies, or how did you select? Either remove all that do not compute lake levels, otherwise consider also including "complete" TP water studies for a larger number of lakes than the ones you are listing (e.g. Pekel et al (2016) to whom you refer to earlier, or Yang et al. 2019, doi:10.5194/tc-2018-238; Treichler et al. 2018, doi:10.5194/tc-2018-238...)

Response:

Thanks for this constructive comment. We consider it is quite reasonable to exclude those references without water levels, as our study focuses on improving the quality of merged water levels and subsequently improving lake storage change estimation.

11) P4, L4: the meaning of "hypsometric curve" is unclear to me in this context.

Response:

We noticed that in some studies hypsometric curves represent the total area above a certain elevation, which means that at the lowest elevation the hypsometric curve reaches its maximum value. However, in this study, hypsometric curves represent the lake surface area at a given water level, which means that the curve reaches its maximum value when the water level is maximized. We adopted this denotation as same as the LEGOS Hydroweb. To make it clear, we have added an explanation in brackets in the context.

Study area and data

1) Parts of this (e.g. from P5, L24, or P6, L1ff) rather belongs to the method section.

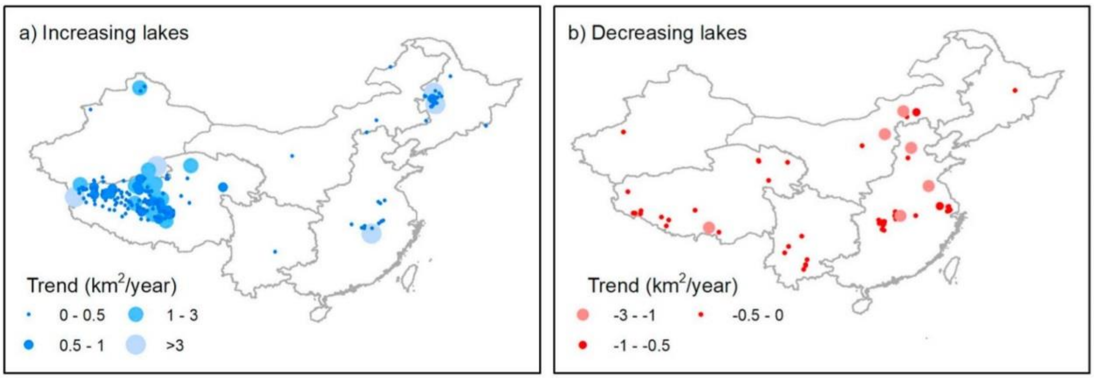
Response:

Thanks for this comment. We have moved partial content to the method section. For instance, we have moved P5 L24–L26 to the second paragraph of section 3.2.

2) P4, L16: "as opposed to many other places..." - I tend to disagree, as nearly all seem to have expanded. Can you justify or explain more clearly?

Response:

We only studied 12 lakes outside the endorheic basin for the recent twenty years, which possibly caused such an impression that all lakes have experienced expansion. Exorheic lake shrinkage in the TP in the past 50 years can be seen from (Zhang et al., 2019) as shown in the figure below.



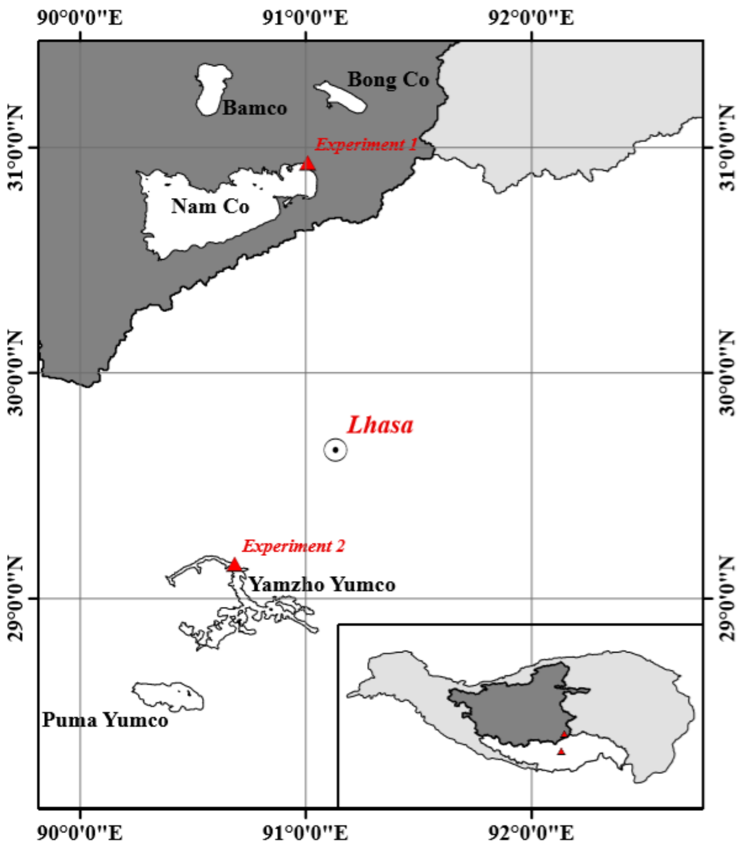
In addition, most global endorheic basins have experienced water loss in recent years, whereas the endorheic region in the TP has gained water (Wang et al., 2018). This phenomenon has also drawn a lot of attention for the endorheic basin in the TP.

3) P5, L4ff: why did you choose these lakes in particular? And where is Lake Yamzhog Yumco? An overview map might be useful.

Response:

The reasons why we chose Yamzhog Yumco and Nam Co are threefold: (1) they are close to the city, making it easier for logistics and transportation; (2) they are both large lakes, typical in our study; and (3) one of them is located in the endorheic basin (Nam Co), and the other is from the exorheic basin (Yamzhog Yumco), increasing the representativeness of the experiment.

Following figure will be added into the manuscript to clearly show the two experiment locations:



4) P5, L17: "moderate set of orbital parameters" is unclear

Response:

Thanks for this comment. We have made it clear. We meant to show that Envisat has a lower orbit than Jason-1/2/3 but higher than ICESat, thereby for sampling frequency: ICESat<Envisat<Jason-1/2/3, and for spatial coverage: ICESat>Envisat>Jason-1/2/3.

5) P5, L30: when were the drone data acquisitions?

Response:

The drone images were acquired in the morning on May 19 and 21, 2018, for Yamzhog Yumco and Nam Co respectively. The Landsat images used for validation purposes were both acquired on May 19, 2018.

6) P5, L31: "similar" in what sense? What have Huang et al done?

Response:

Huang et al. (2018) used UAV images to evaluate the performance of water auto-extraction with four water indices based on Landsat 8 images. The accurate water surface boundary was extracted manually from the UAV images using ArcGIS, and then water extraction results from Landsat using different water indices were compared with the accurate water surface area from the UAV images. Our data source and method are similar, but focused on different targets. On the other hand, we have performed a systematic analysis to link the uncertainty in water surface area extraction to the uncertainty in optical water levels.

7) P6, table 2: Some of the missions included many instruments (e.g. ENVISAT: 10 sensors). You need to specify which sensor and data you used. Here, you distinguish between "radar" and "interferometer", which is also based on radar (SAR/interferometric radar altimeter). This is confusing, and it would be useful to explain the technologies/differences either in the introduction or in a separate paragraph in the data or methods section

Response:

Thanks for this constructive comment. It is important to clarify the sensors and data we used, and they have been added to the table now. The classification of different radar altimeters in the original manuscript might be confusing as indicated by the reviewer. Therefore, we have provided a brief explanation after the first paragraph of section 2.2 on the mechanism of different altimeters including SIRAL onboard CryoSat-2.

Methods

1) The first paragraph seems to explain what this study is about and would thus fit (better?) to the introduction (it is missing there!).

Response:

Thanks for this comment. They have been moved to the introduction section.

2) P6, L8: "comparing the mean water level of the overlap period" is vague. Explain better.

Response:

Thanks for raising this issue. It has been explained in detail in our response to short comment 1 (the first question in short comment 1). We have also added a separate paragraph at the end of section 3.1 to better explain this part.

3) P7, figure 1: refer to the figure in the text, e.g. when you introduce the data and where you are talking about overlap periods. Consider adding the optical data to the figure to show the overlap periods you use to create the optical lake level-lake surface elevation relationship.

Response:

Yes, we have added references to Figure 1 in three places where we think it is necessary. In addition, the optical data are presented in Figure 1. However, it is not easy to show the time period we used to derive optical water levels from altimetry data, because for different altimetry missions may be used to derive optical water levels for different lakes. For instance, if Jason-1/2/3 data are available, optical water levels are generated by fitting with the merged Jason-1/2/3 water levels. If ICESat and CryoSat-2 data are available for a lake, optical water levels are generated first by fitting with CryoSat-2 data. After the extended CryoSat-2 data are merged with the ICESat data, the optical water levels generated throughout the entire study period are checked again by fitting with the merged altimetry water levels to see if there is an extrapolation problem. We will discuss this issue in detail in response to the specific comment 8 below in this section.

4) P7, L18: It is very unclear what "ENVISAT product" you used.

Response:

Thanks for this comment. It has been changed to Envisat/RA-2.

5) P7, L23: "highest bucket" is an unclear term. What elevation bin spacing did you choose for your frequency histograms? It seems you are losing information by binning your surface elevation measurements. How does that affect the accuracy of the extracted lake level elevations? I assume you have t-distributed data, i.e. roughly bell-shaped elevation distributions with long tails. It might be more appropriate to use the median elevation measurement, maybe in combination with a threshold to remove biased measurements in the tails. From reference DEMs, you should know the true surface elevation (of the lake shore).

Response:

Thanks for this insightful comment. We used a 0.6-meter bin space to generate a histogram and the 'highest bucket' represents the histogram bin with the highest frequency. It has now been clarified in the revised manuscript. We do not think much information is lost, as for most cycles (>70%) there are more than 80% measurements falling into the highest bin. We first used the median value of each cycle to represent the lake water level, which is noisier/less smoother than that using the histogram. It turns out that a 0.6 m bin space is large enough to capture valid measurements in a cycle.

It is true that a bell distribution is quite common for most lakes. But setting constant thresholds to remove outliers for each lake does not seem to work well in our study. We did try this method before but it always ends up in how to choose an appropriate threshold. If the threshold is too large, invalid measurements will be involved in a lake. Otherwise, certain amount of information would be lost. For instance, Lake Kusai experienced a water level jump up to ~10 m in 2011. If we do not know this information before, then a threshold must be larger than ± 10 m from the mean water level/DEM to capture the water level jump, which will definitely introduce a number of inaccurate measurements in normal cycles.

6) P8, L4ff: How large are the biases you found? Are they constant over time and in space? I assume you compute this per lake?

Response:

Thanks for this comment. The spatial distribution of systematic biases seems quite random to us, varying from place to place, even the sign of the systematic biases is not stable between two certain altimeters (except for Jason-1/2/3). The range of biases is within ± 5 m. Fortunately, the systematic bias is quite stable in time, as we compared the merged altimetry data with the optical water levels. If the bias is not stable in time/elevation, which means that the additive correction is not effective enough, the multiplicative correction may be needed. Overall, we did not see the necessity of using the multiplicative correction nor did we find any relative research reporting such corrections.

7) P8, L13: it is unclear what you mean with "merging using optical water levels"

Response:

It should be clear now as we have provided a separate paragraph at the end of section 3.1 to summarize the merging process. Thanks for this comment.

8) paragraph P8, L11ff: Only after reading this paragraph I think I finally understood the purpose of this study: You want to generate continuous lake level (volume?) series for as many lakes as possible. This requires elevation (and areal?) data from different sensors, as missions only last for a few years. As an additional challenge, the satellites in question have different orbits that only cover some lakes each, so not all elevation datasets can be used for each lake. For each lake, you therefore combine lake level elevation time series from the different sensors with data for that lake, using the overlap periods to correctly align the records, i.e. you remove potential elevation bias between the time series and make sure they are consistent. Where there is no sufficient overlap, you use optical data as a proxy: you create a statistical relationship between lake levels (from altimetry data) and corresponding shoreline position (from optical data acquired at the same time), and then apply (extrapolate?) the relationship to (optical) shoreline positions for

time periods where you lack surface elevation data, but do have optical data. I propose you add something like this to the introduction. Secondly, this paragraph would be much easier to understand if you first introduce optical water levels and refer to Figures 2 and 3 in the text. Given the importance of the relationship for your results you might want to explain your method in more detail. An important missing detail is whether you only interpolate or also extrapolate beyond the available data range?

Response:

We really appreciate these accurate and comprehensive summary and highlights on our work. As the referee suggested earlier, we have enhanced the introduction section to clarify the purpose and underscore the contributions of this study. We have added references to Figure 2 and Figure 3 in this paragraph and we have moved part of it to the end of the optical water levels section (section 3.2). The interpolation and extrapolation may be the most concerned issue here. Below we provide a few examples to justify our methodology.

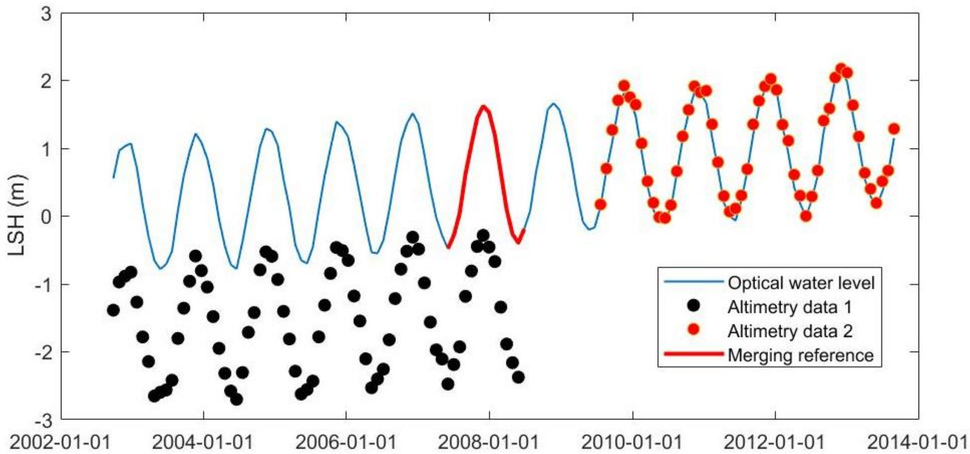
Note that we have performed two regressions to generate the optical water levels. For the first regression, we only used one altimetry data product and optical images-derived lake shoreline positions. After merging the altimetry water levels, we performed the second regression using the merged altimetry water levels and the optical water levels temporally close to the altimetry water levels throughout the entire study period. This information is missing in the original manuscript and we will add it in the revised manuscript/supplementary file. Here we show that part of the extrapolation problem is evitable in nature with the second regression:

a) When and where does extrapolation exist?

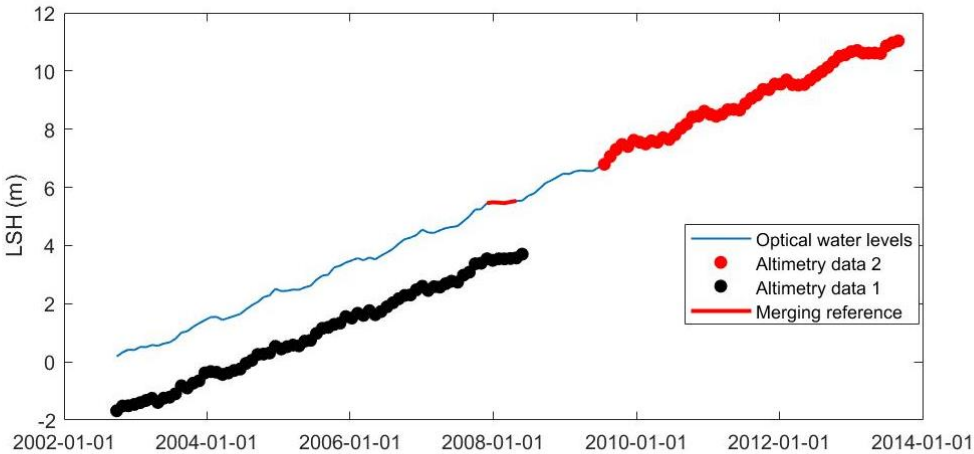
   First, extrapolation here means the extrapolation of the linear relationship developed from the regression analysis between altimetry water levels and lake shoreline changes. For instance, if the altimetry water levels used for the regression analysis have a range of 4500–4502 m, then the generated optical water levels beyond/below this range are regarded as extrapolated values. On the other hand, if an optical water level H_1 acquired in 2003 is within 4500–4502 m, though the altimetry water levels used for such a regression were from 2010 to 2017, H_1 is still regarded as an interpolated value because it is within the elevation range of the linear regression.

   As shown in following figures (both are conceptualized examples, optical water levels are fitted with the second altimetry product), when seasonal signal is dominated in the time series, there is no need for extrapolation. The red line in the optical water levels (which serves as the merging reference to altimetry data 1) are within the range of the linear regression. The merging between the two altimetry water levels can subsequently be achieved by removing the difference (symmetrical bias) of the mean water levels between altimetry data

1 and altimetry data 2 during the reference period (the red solid line) from altimetry data 1 (typically ICESat data).

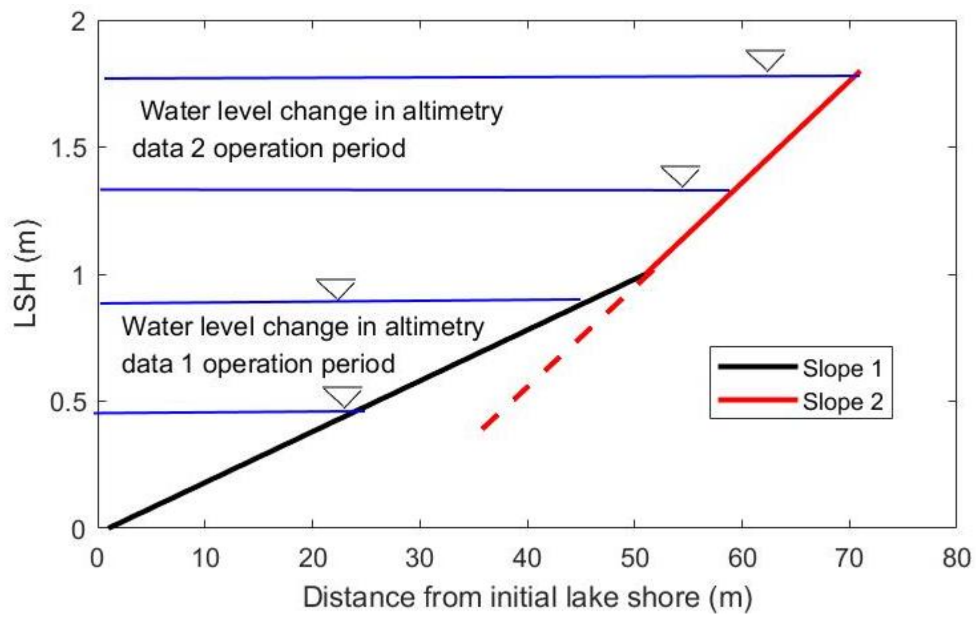


When a multiyear trend is dominated in the time series, the merging reference is out of the range of the regression relationship, and then extrapolation may occur. Both situations are common in our study. The first situation comprises 60% of all study lakes, and extrapolation can take place in ~40% lakes. The two altimetry datasets in the extrapolation case can still be merged using the similar procedure and optical water levels shown in the interpolation case above.



b) How does extrapolation become a problem?

In the merging process, extrapolation becomes a problem only if the lake bank slope experiences an abrupt change at the exact elevation where both altimetry products fail to cover, as illustrated in the following figure:



Such a situation may happen, but the possibility is relatively low. If it happens, the extrapolation will result in a remaining systematic bias in the merged altimetry water levels and consequently jeopardizing the accuracy of the optical water levels.

c) How can the problem be avoided?

By performing the regression analysis twice, it is possible to detect if there is an abrupt change in lake bank slope. If the situation in b) does happen, we can easily see from the scatterplot of the second regression analysis that the linear assumption is no longer met (i.e., the scatterplot would show two slopes/curvature). Once an obvious failure in the second linear regression occurs, we will re-choose the region of interest (ROI) and go through the entire process of generating optical water levels again. However, it only happened twice or three times in our study.

We will provide the details of generating the optical water levels discussed above in the supplementary file as they may be too detailed for general readers.

9) P9, Figure 2: refer to the figure in the text, e.g. where you introduce the data sets and in section 3.1

Response:

Yes, we have added reference to Figure 2 (now Figure 3, because we inserted a new figure after Figure 1) in the first paragraph of section 2.2 and fourth paragraph of section 3.1.

10) P9, 3.2: The optical water levels should be introduced before P8, L15ff.

Response:

The sequence has been changed.

11) P10, L4ff: the part about "shifting gaps" and the ROI is unclear. Do you mean that the Landsat 7 gaps are not always exactly at the same place? Did you choose your ROIs such that they never contain no-data pixels? How did you ensure that, given the large amount of Landsat 7 data?

Response:

Yes, the position of gaps in Landsat 7 data is various with time. But they are more like vibrating around a fixed location. So, narrowing down the width of ROI can assure higher data availability. It is true that filtering a large amount of Landsat 7 archives is really tough, but our study was primarily based on GEE and we performed an invalid-pixel detection to get rid of images with missing pixels in the ROI. The algorithm is straightforward: comparing the valid pixel number in the ROI with that from an intact image. If the missing pixels in the ROI exceed 2% then the image will be excluded. Using 2% instead of 0% is due to the consideration of the algorithm robustness, but there is not much difference in the results as the ratio of in-valid pixels is either very high (>20%) or extremely close to zero.

12) P10, L17: reference for the Otsu method?

Response:

It has been added.

13) P10, L22: How did you decide whether to use a linear or 2nd order polynomial fit?

Response:

Thanks for raising this comment. In fact, it only happened in two lakes: Zhari Namco and Chibzhang Co, where we already have Jason-1/2/3 data for altimetry data merging. For other lakes we only performed linear regression, and if the scatterplot of the regression has a clear curvature, we will re-choose the ROI (see our response to comment 8 in the method section). For Zhari Namco and Chibzhang Co, if we use linear regression, a clear discrepancy will show up at either low water levels or high water levels. Therefore, using a higher order regression is a choice.

14) P10, L25: How did you determine cloud cover?

Response:

The cloud cover was calculated in GEE based on the quality band of Landsat 5/7/8. Pixels in the quality band categorized as cloud or cloud shadow will be masked with a mask function provided in GEE. Then, the cloud/cloud shadow pixels will be regarded as invalid pixels and a corresponding rate can be calculated by dividing invalid pixels with the total pixels in the ROI. If the cloud rate is higher than 5%, the image will be discarded.

15) P11, L2: How much data pairs did you end up with per lake, and how did you

select pairs with regard to acquisition dates? I assume you did not always have altimetry and shoreline data from the same date (?)

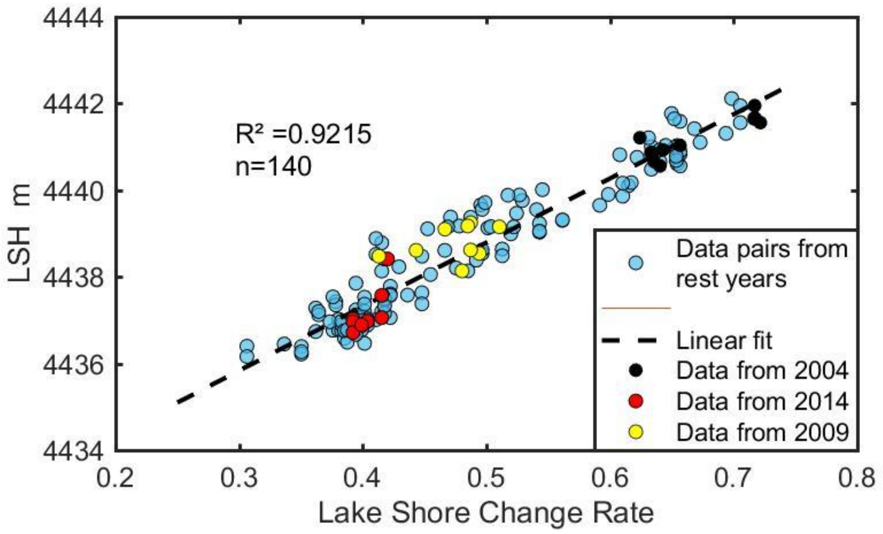
Response:

We have an average of 55 data pairs for the second regression of optical water levels. About 70% of the study lakes have more than 20 data pairs. The time difference of data pairs is within 5 days. However, if there are not enough data pairs (<10 pairs) with a time difference smaller than 5 days, we will increase the time difference to 10 days. Only for a lake named Xuru Co, where altimetry information is very limited, we increased the time difference to 30 days.

16) P11, figure 3: c) You might want to colour the dots according to time to check for (and show the readers that there is no) temporal bias. From d), it seems optical water levels are somewhat too high around 2004 and 2015, but too low around 2009?

Response:

Yes, this is a nice suggestion. Based on the following figure, it seems that in 2009 the optical water levels might be a little lower than expected. It may be caused by the uncertainty in altimetry water levels. In 2009 the main data source is Envisat, which has poorer quality than other altimetry products (except Jason-1) in our study. Overall, the impact of this problem is quite limited as the linear relationship is still strong.



17) P12, L9: How did you derive these ROIs? Are they drawn manually?

Response:

Yes, they are drawn manually. Selection criterion is illustrated in the manuscript. However, it still requires some experience.

18) P12, L15: regression between the lake area and ..?

Response:

It is between lake area and merged lake water levels, including altimetry water levels and optical water levels, but most data pairs are lake areas and optical water levels, because they usually come from the same Landsat image.

19) P13: As far as I am aware, Strahler's catchment hypsometric model is based on river catchments with a pour point, not endorheic lake catchments as it is the case for the TP. I am not entirely sure what you used this model for (to compute lake water volumes?), but I am not convinced that this is a correct approach. I am also not sure why you need that relationship at all? If you have lake area and lake level time series, you can directly compute volume changes from these?

Response:

Thanks for raising this comment. We intended to provide some justification that a parabolic relationship between the lake area and lake water level is reasonable. But it seems that such a justification is unnecessary and inappropriate because the assumption of exorheic basins is not met. We will remove this analysis from the revised manuscript.

The reason why we use the lake area-water level relationship to calculate the volume change is the lack of lake water areas with a sufficient temporal resolution. In general, we only have ~20 lake area observations for each lake, because the ROI for lake area extraction is much larger than that of lake shoreline changes, reducing the data availability. If we use the volume formula for computation, we can only get ~20 volume change values. With a lake area-water level relationship, we can derive the lake volume-water level relationship and convert all water level estimates into lake volume changes.

20) P14, Figure 6: It is unclear what the parameters y, x, z, a and d represent.

Response:

Thanks for this comment. We will remove this part from the revised manuscript.

21) P14, table 3: state nr. of data pairs (optical shoreline position + altimetric lake level) rather than optical data points

Response:

Yes, they have been added.

Validation

1) P16, L5: unclear sentence

Response:

The sentence has been reorganized.

2) P16, L25: the drone GPS tracker alone might not be very accurate, you may easily get a skewed/stretched image composite. Did you use ground control points?

Response:

We did not get ground control points. It is true that there may be skewing or stretching distortions in UAV images. So, we redid the experiment with some commercial high-resolution data such as GF-2 (China's High Resolution Satellite, GF-2, with a panchromatic resolution of 0.8 m), which has larger coverage and more ground features for co-registration with Landsat OLI image.

3) P16, L27: This seems a rather dodgy way to determine the resolution of your image composite.

Response:

Yes, it is not very rigorous, and we have abandoned it.

4) P17, figure7: which lake? images a) and e) should have the same size/spatial resolution. An overview map would be useful.

Response:

Thanks for this comment. We performed UAV scanning and water level sensor installation in both lakes. However, the water level sensor in Nam Co was broken down soon after installation and did not provide much information. Figure 7 shows pictures acquired at the Nam Co experiment spot. An overview map has been added into the study area section (section 2.1), as the referee suggested before. We decided not to use the UAV image as a validation basis, but we keep it here to show the environment at the experimental spot. In addition, the up-left image from Landsat 8 has been changed into an overview map of Nam Co and the experiment location.

5) P17, L9ff: extrapolated or interpolated? Provide the parameters and statistical relationship here, maybe even in an additional figure.

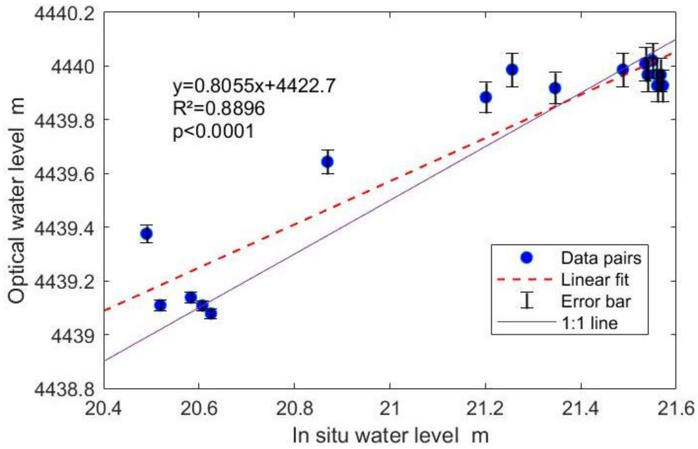
Response:

The optical water levels of Yamzhog Yumco used for validation are interpolated. The statistics of regression are already shown in Figure 3.

6) P18, figure 8b: add 1:1 line and error bars for the data points

Response:

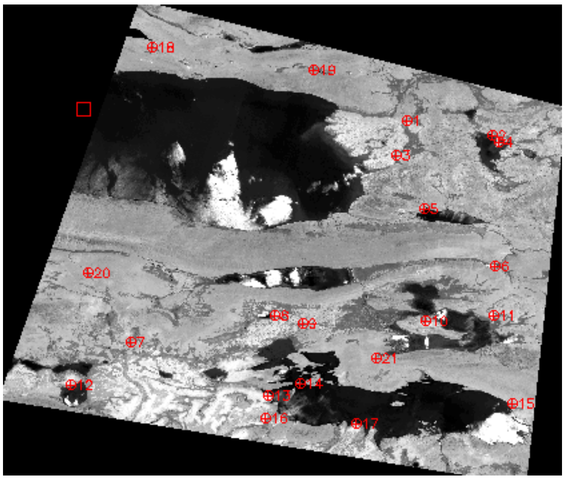
Yes, they have been added as shown in the figure below.



7) P18, 19: How did you coregistrate the UAV image composite and Landsat image? It seems a spatial shift will completely alter the (relative) shoreline position and thus the basis for your entire analysis: In Figure 9, shifting the shoreline only slightly in e.g. north-south direction will greatly change water/land (sub) pixel counts and thus the basis for the relationship in (b). In my opinion, an error analysis would require several image pairs (UAV and satellite-borne) and a solid coregistration basis, e.g. river/road crossings as clear tie points, or at least a round lake or elongated peninsula rather than a straight shore line.

Response:

This is a very constructive comment. We agree that there might be a spatial shift in the UAV image. Therefore, we no longer use the UAV image because there are very few ground features for image co-registration. Instead, we purchased some high-resolution commercial images obtained by GF-2 (0.8 m resolution at the panchromatic band) to repeat the analysis. The GF-2 images cover a much larger area and more diverse ground features, making it easier for image co-registration. The following figure shows control points that we selected for one of the GF-2 images. The co-registration error is 1.2 GF-2 pixels, say ~1 m. The other two GF-2 images have a co-registration error of 2.45 pixels and 2.72 pixels, respectively, corresponding to ~2 m.



8) P18, L7: what do you mean with "concurrent"? What dates?

Response:

It means the "same period" image. We have changed this expression.

9) P22, L1ff: Do not forget the local conditions: ice, snow, wet, dry, muddy shore conditions or also waves greatly affect the water classification result.

Response:

Thanks for this comment. Yes, the local condition is an important factor affecting the water area classification accuracy. Therefore, we chose three high resolution images acquired in different seasons and different places representing typical local conditions around the TP lakes, covering turgid water (wet season), lake ice, and dry season. As for vegetation, most of the TP lakes do not have much vegetation on the lake bank, with the Landsat images unable to detect information on vegetation.

Applications

1) P22, L10f: Are these your own numbers? How do they compare to previous estimates?

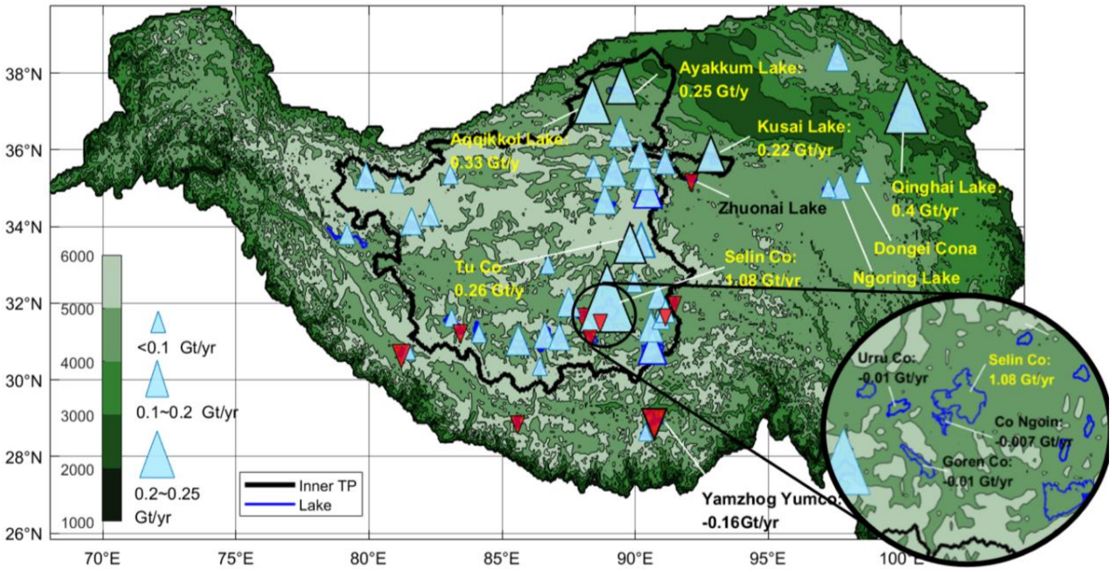
Response:

Yes, they are results generated from our product. There has not been any published study that has exactly the same study period or lakes as what we did. But for the overlap periods and lakes, the results are similar. We have made many comparisons with published studies or open source data, including the comparison between our product and Hydroweb data in Figure 14 in the original manuscript (now Figure 11).

2) P22, L14ff: mark all lakes mentioned in the text in the map. If they are very close to each other, an extra zoom-in map might be useful.

Response:

Yes, this has been done, as shown in the figure below:



3) P23f: restructure section 5 to avoid splitting the Selin Co basin analysis in two sections (5.1 and 5.3). How much of this is new, i.e. has not been published before? How does your dataset make a difference?

Response:

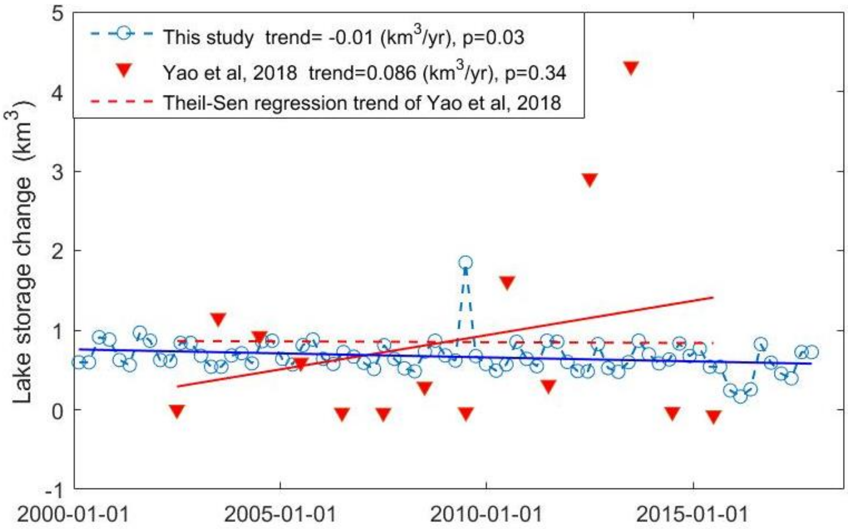
Thanks for this comment. We have only talked about Lake Kusai in section 5.3. All discussion about Selin Co is shown in section 5.1. There are some published studies that report the unusual spatial pattern of lake area/water level/storage changes in the Selin Co basin. However, there is no discussion about the reason. We proposed a possible explanation. On the other hand, given the complexity of modeling a multi-lake endorheic basin (Zhou et al., 2015), our product does provide a chance for investigating the structure of such a endorheic basin with complicated lake-river systems. For instance, the height of outlet of three upstream lakes in the Selin Co basin may be inferred from the dense time series from our product with the help of a hydrologic/hydrodynamic model.

4) P23, L5ff: You mention only the study of Yao et al. (2018). How about other publications? Also, from figure 12 it seems quite clear that the Yao data contains two outliers. Consider using a robust fitting method rather than regular linear regression.

Response:

Thanks for this comment. Song et al. (2013) notice the decreasing trend of the three lakes in the upstream of Selin Co during 2003 to 2009 when ICESat data are available, but there are no comments or discussion about the reason. We found that Hydroweb data do not catch the decreasing trend of Urru Co after 2000. Jiang et al. (2017) did not investigate the decreasing trend of Urru Co from 2003 to 2015 as their altimetry data from ICESat and CryoSat were not linked together but separately discussed instead. Hwang et al. (2019) reported a similar problem as Jiang et al. (2017). Other studies do not present specific statistics for the comparison nor do they cover those lakes.

With a robust linear fit method (Theil-Sen estimator), the result from Yao et al. (2018) did show a decreasing trend, consistent with our result. But they clearly did not use a robust fitting in their published paper/dataset.



5) P24, L10: Depicting intra-annual variation is a strength of your dataset that you might want to emphasize more.

Response:

Yes, we did describe the intra-annual variation in the lakes we studied.

6) P24, 5.2: Rather than treating the comparison to the LEGOS Hydroweb data as an application case this should be part of the validation section!

Response:

Yes, we have moved part of section 5.2 to the validation section (section 4.2).

7) P25, L8: "some kind of" bias removal: be more specific. The magnitude of the vertical shift between the two datasets fits to e.g. geoid/ellipsoid height confusion, but the temporal variability of the shift is worrying. Rather than speculating about the cause and assuming that the Hydroweb data is wrong you ought to find the reason for the differences - which may lie in your data processing/method.

Response:

Thanks for raising this insightful comment. The reason for the vertical shift between our product and Hydroweb data possibly lies in different geoids/reference ellipsoids, as illustrated in our response to referee comment 1 (General comment 3). However, we respectfully disagree on the point of the temporal variability in the vertical shift.

In the manuscript, we just indicate that partial Hydroweb data are not quite consistent with the optical water levels (e.g., in the three lakes shown in Figure 15), which are able to provide a straightforward answer to "in which period the lake has higher water level". As we have clarified in the revised manuscript, such a relationship on relative magnitudes reflected by the optical water levels does not change with the linear fitting parameters (unless using a negative slope, which is impossible) and that is why we regard it as robust. What we did was merging different altimetry data sources based on the reference provided by the optical water levels. Therefore, it is not likely to be a problem for this straightforward and robust scheme for merging altimetry data.

8) P25, L13ff: "reverse relationship" and the conclusion you draw (Hydroweb may "underestimate decreasing trends"): unclear what you mean

Response:

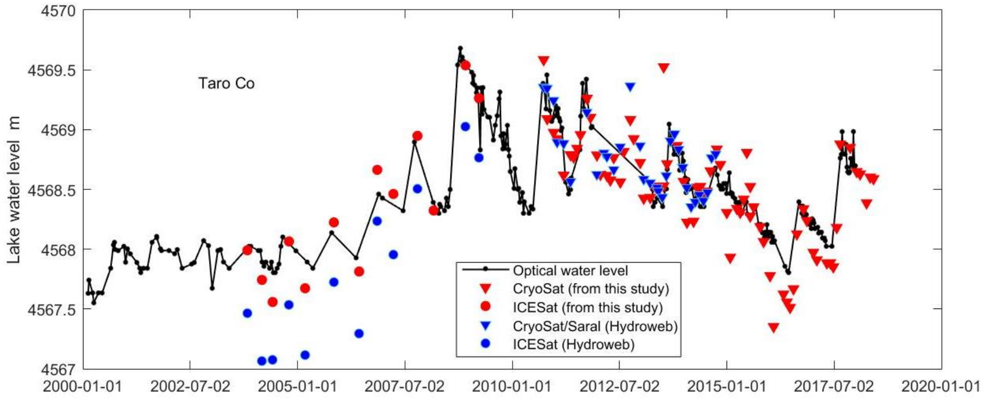
Thanks for raising this comment. We apologize for making a wrong expression in the original manuscript: the conclusion should be "…there is a possibility that Hydroweb data overestimate the increasing trend of water levels in Taro Co from 2003 to 2015".

As shown in Figure 15 (a), the last two measurements from ICESat should equal or be even larger than the first two/three measurements from CryoSat/Saral based on optical water levels, but the Hydroweb data show a reverse relationship that the last two ICESat measurements is 0.3~0.4 m smaller than the first two CryoSat/Saral measurements. This phenomenon suggests that ICESat water levels of Taro Co from Hydroweb is 0.3~0.4 m lower than the expected (in other words, CryoSat/Saral time series from Hydroweb is 0.3~0.4 m higher than the expected). It would therefore result in an overestimation of increases in lake water levels in Taro Co during the time window. In addition, the optical water levels in Taro Co were interpolated with the developed statistical relationship. Therefore, the discrepancy between Hydroweb and our product is not attributed to the extrapolation of the optical water levels.

9) P25, figure 15: What are the red/blue shaded areas? (a) compare the series after removing the shift. Sadly, the series (b) and (c) are no where discussed. The temporally varying offsets between the series from different data sources should be analysed and removed, or at least explained.

Response:

Thanks for this comment. The red and blue areas were meant for highlighting/comparing the periods when an obvious discrepancy between Hydroweb data and optical water levels from our product occurs. As for Figure 15 (a), we have removed the systematic vertical offset between our dataset and Hydroweb data of Taro Co, which is shown in the figure below:



As we suggest earlier in the response letter, there might be a remaining systematic bias between ICESat and CryoSat/Saral data from Hydroweb. Based on optical water levels, the peak water level of 2009 shall be higher than that of 2010 (again, such a relationship does not change regardless of the uncertainty in the linear fitting parameters during the generation of optical water levels), which means that the last two ICESat measurements are supposed to be higher or equal the first a few CryoSat/Saral measurements. However, this is not seen in the Hydroweb data for this specific lake.

As for Figure 15 (b) and (c), they show other examples of possible remaining systematic biases in Hydroweb data. The explanation is exactly the same as that of Figure 15 (a) and we did explain the discrepancy in the manuscript for Figure 15 (a). Thanks for your kind attention to this.

10) P26, figure 16: again, there seems to be some time-dependent offset between the optical and altimetry-based lake levels, e.g. optical levels are too high around 2005 in the top left panel, and too low around 2005 vs. too high from ca. 2013 in the middle right panel. Can you explain this?

Response:

Thanks for this insightful comment. Though there seems to be an offset at around 2005, the actual deviation between the optical and altimetry water levels here (ICESat data) is about 0.2~0.3 m, which is within the uncertainty range of altimetry measurements for inland water bodies. Instead of a time-dependent offset, we think it is more like a random error, which can be caused by the loss of valid altimeter footprints of that cycle, e.g., a random shift of ground tracks resulting in a smaller cross section and fewer footprints on the lake. It is also suggested that optical water levels may be more robust and less noisy than altimetry data. This is the same for the middle right panel. It should be noted that in the middle right panel, Envisat/RA-2 was used, which has a larger uncertainty than ICESat. Therefore, it is not surprising that altimetry dots seem to be more randomly distributed.

11) P28, figure 17: What does the blue shaded area show? What data are you showing

in these time series? Is the right panel a zoom-in of the left panel?

Response:

The blue shade shows the period when an outburst happens. The data we show in Figure 17 are lake water storage changes for relevant lakes during the outburst event. Their locations are shown in Figure 18 (b) and (c). And yes, the right panel is a magnified plot of the blue shade in the left panel.

12) P28, L17: "Team, 2017": check author name

Response:

Yes, we have checked the reference. It has been cited dozens of times in other journal papers according to the Google Scholar.

[CITATION] **Planet Application Program Interface**: In Space for Life on Earth
P Team - San Francisco, CA, 2017
Cited by 69 Related articles

13) P29, figure 18: acquisition dates of the images in b) and c)?

Response:

Figure 18 (b) was acquired in December, 2010. Figure 18 (c) was acquired in December, 2013. The outburst took place at the end of 2011. These are images from the Google Earth (i.e., the image source is Landsat but experienced merging processes, e.g., merging of images acquired from the same month) and we do not know the exact acquisition date.

14) P29ff: The entire overflow analysis (lots of new methods introduced) seems to be a study on its own and somewhat out of place in the applications (results) section of this paper.

Response:

Thanks for this comment. We have shortened this section and moved some of the analyses into the supplementary file. But we would like to keep this part, because some information (e.g., height and width of the outlet) of the overflow lake, Lake Kusai, is critical to downstream residents and emergency administrations, given that there are reports showing high overflow/outburst risks of Lake Salt in the near future.

Conclusions

1) A short summary of your methods should be provided, in particular the novelty of using shoreline positions from optical data to interpolate between available lake level measurements.

Response:

Yes, this has been added.

2) P31, L7: rephrase the sentence to avoid brackets.

Response:

Done.

3) P31, L10f: Unclear what you mean. From the comparison you provide currently, I am not yet convinced that your dataset is more correct than the Hydroweb data.

Response:

Thank you for this comment. We have put more detailed explanations (most of them are already discussed in this response letter) in the second paragraph of section 5.2 and hopefully this would convince the reviewers and readers. Based on the overall comparison shown in previous Figure 14 (now Figure 11), our product is generally consistent with Hydroweb data, and has a higher temporal resolution.

But there are indeed some discrepancies between the two products over some lakes during some time windows as what we illustrated earlier. Hydroweb is a decent global dataset whereas our dataset is more a regionally based product. It is not uncommon in the remote sensing community that a regionally based dataset may have some advantages than a global dataset in some aspects due to the improvement of the algorithm for the data generation and use of more detailed (a priori) information derived from optical images to densify the spaceborne altimetry water levels with systematic errors being well removed. The developed method we present has potential to improve lake water level and storage changes in different regions globally at large.

4) P31, L18: "rigorous uncertainty analysis": As mentioned above, I am not convinced about the theoretical uncertainty exercise you provide.

Response:

Thanks for this comment. We have redone the uncertainty analysis with more high-resolution images and corresponding Landsat images. We have also provided co-registration accuracy and considered different seasons and locations as the reviewer suggested. This part should now be convincing the reviewer and general readers.

5) P31, L25ff: These insights about extrapolating using the derived statistical relationship are very important, but currently not quantified, mentioned or discussed anywhere else in the paper.

Response:

Thanks for this constructive comment. We will put the discussion of extrapolation we made in this response letter (specific comment 8 in the method section) into the supplementary file. Clearly, our discussion mainly focuses on the period during which various altimetry data sources are merged, but does not include the period before 2002 when little altimetry information is available and DEM is too course (for instance, SRTM DEM has a 1 m vertical resolution with more than 10 m vertical uncertainty according to Mukherjee et al. (2013)) to provide a detailed description on the lake shore micro topography. Therefore, we do not have much information and materials to discuss about the extrapolation before 2002.

We have informed readers in the manuscript that this is a possible issue but it may only exist in the first 2–3 years of the dataset for lakes with strong signal from multiyear trends as opposed to seasonal variations. After all, compared with the 18-year study period, the impact of extrapolation of the optical water levels during 2000–2002 would be quite limited.

Reference

Frappart, F., Seyler, F., Martinez, J.-M., León, J. G., and Cazenave, A.: Floodplain water storage in the Negro River basin estimated from microwave remote sensing of inundation area and water levels, Remote Sensing of Environment, 99, 387-399, 2005.

Huang, Q., Long, D., Du, M., Zeng, C., Qiao, G., Li, X., Hou, A., and Hong, Y.: Discharge estimation in high-mountain regions with improved methods using multisource remote sensing: A case study of the Upper Brahmaputra River, Remote Sensing of Environment, 219, 115-134, 2018.

Hwang, C., Cheng, Y.-S., Yang, W.-H., Zhang, G., Huang, Y.-R., Shen, W.-B., and Pan, Y.: Lake level changes in the Tibetan Plateau from Cryosat-2, SARAL, ICESat, and Jason-2 altimeters, Terrestrial, Atmospheric, and Oceanic sciences journal, 30, 1-18, 2019.

Jiang, L., Nielsen, K., Andersen, O. B., and Bauer-Gottwein, P.: Monitoring recent lake level variations on the Tibetan Plateau using CryoSat-2 SARIn mode data, Journal of Hydrology, 544, 109-124, 2017.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O.: Estimating the volume and age of water stored in global lakes using a geo-statistical approach, Nature communications, 7, 13603, 2016.

Mukherjee, S., Joshi, P. K., Mukherjee, S., Ghosh, A., Garg, R., and Mukhopadhyay, A.: Evaluation of vertical accuracy of open source Digital Elevation Model (DEM), International Journal of Applied Earth Observation and Geoinformation, 21, 205-217, 2013.

Song, C., Huang, B., and Ke, L.: Modeling and analysis of lake water storage changes on the Tibetan Plateau using multi-mission satellite data, Remote Sensing of Environment, 135, 25-35, 2013.

Wang, J., Song, C., Reager, J. T., Yao, F., Famiglietti, J. S., Sheng, Y., MacDonald, G. M., Brun, F., Schmied, H. M., and Marston, R. A.: Recent global decline in endorheic basin water storages, Nature geoscience, 11, 926, 2018.

Yao, F., Wang, J., Yang, K., Wang, C., Walter, B. A., and Crétaux, J.-F.: Lake storage variation on the endorheic Tibetan Plateau and its attribution to climate change since the new millennium, Environmental Research Letters, 13, 064011, 2018.

Zhang, G., Yao, T., Chen, W., Zheng, G., Shum, C., Yang, K., Piao, S., Sheng, Y., Yi, S., and Li, J.: Regional differences of lake evolution across China during 1960s–2015 and its natural and anthropogenic causes, Remote sensing of environment, 221, 386-404, 2019.

Zhou, J., Wang, L., Zhang, Y., Guo, Y., Li, X., and Liu, W.: Exploring the water storage changes in the largest lake (Selin Co) over the Tibetan Plateau during 2003–2012 from a basin-wide hydrological modeling, Water Resources Research, 51, 8060-8086, 2015.