Densified multi-mission observations by developed optical water levels show marked increases in lake water storage and overflow floods on the Tibetan Plateau

Xingdong Li [1], Di Long [1], Qi Huang [1], Pengfei Han [1], Fanyu Zhao [1], Yoshihide Wada [2]

[1] State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, 100084, China
[2] Institute for Applied Systems Analysis (IIASA), Laxenburg, A-2361, Austria

Correspondence to: Di Long (dlong@tsinghua.edu.cn)

Abstract. The Tibetan Plateau (TP) known as Asia's water towers is quite sensitive to climate change, reflected by changes in hydrological state variables such as lake water storage. Given the extremely limited ground observations on the TP due to the harsh environment and complex terrain, we exploited multisource remote sensing, i.e., multiple altimetric missions and Landsat archives to create dense time series (monthly and even higher such as 10 days on average) of lake water level and storage changes across 52 large lakes (>100 km^2) on the TP during 2000–2017 (the data set is available online with a DOI: https://doi.org/10.1594/PANGAEA.898411). Field experiments were carried out in two typical lakes to validate the remotely sensed results. With Landsat archives and partial altimetry data, we developed optical water levels that cover most of TP lakes and serve as an ideal reference for merging multisource lake water levels. The optical water levels show an uncertainty of ~0.1 m that is comparable with most altimetry data and largely reduce the lack of dense altimetric observations with systematic errors well removed for most of lakes. The densified lake water levels provided critical and accurate information on the long-term and short-term monitoring of lake water level and storage changes on the TP. We found that the total storage of the 52 lakes increased by 97.3 km^3 at two stages, i.e., 6.68 km^3/yr during 2000–2012 and 2.85 km^3/yr during 2012–2017. The total overflow from Lake Kusai to Lake Haidingnuoer and Lake Salt during Nov 9–Dec 31 in 2011 was estimated to be 0.22 km^3, providing critical information on lake overflow flood monitoring and prediction as the expansion of some TP lakes becomes a serious threat to surrounding residents and infrastructure.

1 Introduction

The Tibetan Plateau (TP), providing vital water resources for more than a billion population in Asia, is a sensitive region undergoing rapid climate change (Field et al., 2014). There are more than 1,200 alpine lakes larger than 1 km^2 (Zhang et al., 2017a) on the TP where glaciers and permafrost are also widely distributed. With little disturbance by human activity in this area, lake storage changes may serve as an important indicator that reflects changes in regional hydrological processes and responses to climate change. Wang et al. (2018) showed that global endorheic basins are experiencing a decline in water

storage whereas the endorheic basin on the TP is an exception. Given the fact that TP lakes have been expanding for more than 20 years (Pekel et al., 2016), pasture, farmland, and infrastructure near the lake shore face the risk of inundation. Therefore, it is imperative to largely improve the monitoring of TP lakes.

As an important component of the hydrosphere, terrestrial water cycle, and global water balance, millions of inland water bodies such as lakes, wetlands, and human reservoirs have been investigated globally and their total storage was estimated to be $181.9 \times 10^3 \ km^3$ based on statistical models (Lehner and Döll, 2004; Messager et al., 2016; Pekel et al., 2016). Lake storage changes, which is more concerned in the regional water balance, can be derived from observed changes in lake water level and area (Crétaux et al., 2016; Crétaux et al., 2011b; Song et al., 2013; Yao et al., 2018b; Zhang et al., 2017a). Most of these two kinds of observations in related studies were obtained from satellite remote sensing due to the scarcity of in situ data across the TP where the harsh environment and complex terrain make in situ measurement costly and risky. Changes in lake water level can be monitored using satellite altimetry with a radar or laser initially designed for sea surface height or ice berg measurements. Altimeters determine the range between the nadir point and satellite by analysing the waveforms of reflected electromagnetic pulse. The waveforms of radar or laser pulse may, however, be contaminated by signal from complex terrain when applied to inland water bodies, but it is possible to remove these impacts with waveform retracking algorithms (Guo et al., 2009; Huang et al., 2018; Jiang et al., 2017). Zhang et al. (2011) mapped water level changes in 111 TP lakes for the period 2003–2009 using ICESat-1 data that have a temporal resolution of 91 days. ICESat-1 data have relatively denser ground tracks but a lower temporal resolution than most of other altimetric missions. This means that ICESat-1 covers more lakes but provide few water level observations for each lake. After ICESat-1 was decommissioned in 2010, CryoSat-2 data spanning the period 2010–2015 were adopted in related studies (Jiang et al., 2017), due to its similar dense ground tracks and competitive precision as ICESat-1. Other altimetric missions, such as TOPEX, Jason-1/2/3, ERS-1/2, and ENVISAT also have some but relatively limited applications in monitoring changes in lake water level on the TP due to sparse ground tracks. In this study, multisource altimetry data (i.e., Jason-1/2, ENVISAT, ICESat-1, and CryoSat-2) were combined if available for the study lakes, with the optical water levels developed in this study as a critical reference to densify the altimetric observations and merging data from multiple altimetric missions.

Changes in lake area can be captured by optical/SAR images from medium or high spatial resolution remote sensing data, such as Landsat TM/ETM⁺/OLI. Extraction of lake water bodies can be manually (Wan et al., 2016) or automatically (Zhang et al., 2017b) achieved. Automatic water extraction methods based on the water index and auto-thresholding are more efficient in dealing with a large amount of remote sensing images. Even so, acquisition and preprocessing of such a large amount of historical images (~10TB) covering TP lakes are still intractable for researchers with limited computation resources. With the help of cloud-based platforms, such as the Google Earth Engine (GEE) that significantly reduces data downloading and preprocessing time, tens of thousands of images may be processed online in days instead of months (Gorelick et al., 2017). In this study, more than twenty thousand Landsat TM/ETM/OLI images were processed online using GEE to extract lake water bodies based on the water index and Otsu algorithm.

There are some studies focusing on changes in lake storage on the TP over recent decades, e.g., Zhang et al. (2017a) examined changes in lake storage for ~70 lakes from the 1970s to 2015 with ICESat-1 altimetry data and Landsat archives. An individual lake area data set from the 1970s and annual area data after 1989 were used. Due to the short time span of ICESat-1, they used the hypsometric method to convert lake areas into water levels. Yao et al. (2018b) used DEMs and optical images to develop hypsometric curves for lakes on the central TP, and estimated annual changes in lake storage from 2002 to 2015 for 871 lakes. These studies have a wide spatial coverage of lakes but relatively lower temporal sampling and little altimetric information, which may limit the accuracy of trends in lake water level/storage in some cases and short-term monitoring of lake overflow flood. The LEGOS Hydroweb provides a lake data set, including multisource altimetry-based changes in lake water level and storage as well as hypsometric curves for 22 TP lakes (Crétaux et al., 2016; Crétaux et al., 2011b). The data set incorporated more altimetric information and provided data of higher temporal resolution. However, there may be a remaining bias when different sources of altimetric data were merged, due to the lack of some important reference that can be derived from optical remote sensing images to be shown in this study. We term the reference data as the 'optical water level' to be introduced in section 3.2. Here, we list recent studies and data sets (Table 1) to provide a concise summary on the progress of remote sensing-based water level and storage monitoring on TP lakes.

Table 1. Recent studies and data sets on TP lakes. H, A, V in the table denote lake water level, area, and volume, respectively.

Reference	No. of study lakes	Data type	Time span	Temporal resolution	Source data
(Song et al., 2013)	30	H, A, V, and hypsometric curve	4 records for 1970s, 1990, 2000 and 2011	Decadal	Altimetry data: ICESat-1 Optical data: Landsat TM/ETM+
(Crétaux et al., 2016)	22	H, A, V, and hypsometric curve	1995–2015 Relative bias partially removed (only for altimeters with overlapping period)	Sub-monthly for lakes with T/P, and Jason-1/2 data, and ~monthly for the rest lakes	Altimetry data: T/P, ERS-2, GFO, Envisat, Jason-1/2, SARAL, ICESat-1, and CryoSat-2 Optical data: Landsat TM/ETM+/OLI and MODIS
(Wan et al., 2016)	>1000	A	3 records for 1960s, 2005, and 2014	Decadal	Optical data: Landsat ETM+/OLI, CBERS-1 CCD, and GF-1 WFV
(Jiang et al., 2017)	70	H	2003–2015 Relative bias between ICESat and CryoSat-2 unremoved	~Monthly	Altimetry data: ICESat-1, and CryoSat-2
(Yang et al., 2017)	874	A	2009–2014	Monthly data for 24 largest lakes, and yearly data for the rest lakes	Optical data: Landsat TM/ETM+/OLI and HJ-1A/1B

(Zhang et al., 2017a)	60~70	H, A, V, and hypsometric curve	One record for 1970s, and annual data for 1989–2015	Annual	Altimetry data: ICESat-1 Optical data: Landsat TM/ETM+/OLI
(Li et al., 2017b)	167	H	2002–2012	~Monthly	Altimetry data: ICESat-1 and Envisat
(Yao et al., 2018b)	871	H, A, V, and hypsometric curve	2002–2015	Annual	Optical data: Landsat TM/ETM+/OLI and HJ-1A/1B DEM data: SRTM and ASTER
(Hwang et al., 2019)	59	H	2003–2016 Relative bias partially removed (only for lakes with Jason data/in situ data)	Sub-monthly for lakes with Jason-2 data, and ~monthly for the rest lakes	Altimetry data: Jason-2/3, SARAL, ICESat-1, and CryoSat-2 (Jason-3 data for validation)
Our study	52	H, A, V, and hypsometric curve	2000–2017 All relative biases removed	Sub-monthly for most lakes	Altimetry data: Jason-1/2/3, Envisat, ICESat-1, and CryoSat-2 Optical data: Landsat TM/ETM+/OLI

The overall objective of this study was to examine long-term and short-term changes in water level and storage across 52 lakes with surface areas larger than 150 km^2 on the TP by merging multisource altimetry and optical remote sensing images to generate a much denser (monthly or higher such as 10 days on average), more coherent lake water level/storage change data set covering the period 2000–2017 and the hypsometric curve for each study lake. Results of this study provide a comprehensive and detailed assessment of changes in lake level and storage on the TP for the recent two decades, and short-term monitoring of lake overflow flood for some lakes. This study could largely benefit more detailed investigations into lakes, lake basins, and regional climate change, because the generated data set has the highest temporal resolution during the study period. To ensure data quality, field experiments were carried out and in situ data were collected to examine the uncertainty in the derived optical water levels.

2 Study area and data

2.1 Study area

The TP can be generally divided into 12 major basins, among which the Inner/Central TP (CP) is the only endorheic basin and home to most TP lakes including ~300 TP lakes larger than 10 km^2. Therefore, it was chosen as the main study area. The endorheic basin covers an area of ~710,000 km^2 (~28% of total TP) with a mean elevation of ~4,900 m and has a semi-arid plateau climate with precipitation ranging from 100–300 mm (Yao et al., 2018b). Most lakes in the endorheic basin were expanding under the influence of climate change/variability as opposed to many other places in the TP, e.g., Lake Selin Co

exceeded Lake Nam Co in area and consequently became the largest lake in the endorheic basin between 2011–2012 and expanded by 26% over the past 40 years (Zhou et al., 2015), whereas Lake Yamzhog Yumco (outside the endorheic basin, 350 km to the southeast of Selin Co) shrunk by ~11% during 2002–2014 according to Wan et al. (2016). Located in the southeast endorheic basin, the Lake Nam Co basin covering about 10,800 km^2, with 19% of the basin lake water area and a mean lake

5 elevation of ~4,730 m was chosen as a field experiment spot. The mean annual temperature and precipitation of Lake Nam Co are 1.3 °C, and 486 mm, respectively (Li et al., 2017a). The other experiment spot was Lake Yamzhog Yumco, which has a mean lake elevation of ~4,440 m. Subjected to steep mountainous terrain, the lake has a narrow-strip shape with complex shorelines. The basin of Lake Yamzhog Yumco covers ~6,100 km^2, with mean annual temperature and precipitation of 2.8 °C and ~360 mm, respectively (Yu et al., 2011).

10 **2.2 Data**

Multisource altimetry data were used in this study as shown in Table 2. The earliest record dates back to 2002 (i.e., ENVISAT and Jason-1) and the latest record ends in 2017 (i.e., Jason-3 and CryoSat-2). Most of 52 lakes examined in this study were covered by ICESat-1, ENVISAT, and CryoSat-2 data. Jason-1/2/3 data were available only on 12 lakes due to the relatively sparse ground tracks or data quality issues. Note that Jason-2 inherited the orbit of Jason-1 after its launch in 2008, whereas

15 Jason-1 was shifted into an interleaved orbit and continued functioning until 2013, thereby increasing the spatial coverage of Jason altimetry series to some degree, e.g., Jason-1 data in Lake Qinghai, the largest lake on the TP, were only available after 2008 due to the orbit shift. ICESat-1 and CryoSat-2 data have the best spatial coverage but relatively long repeat cycles of 91 days and 369 days (Bouzinac, 2012; Zhang et al., 2011). The ENVISAT mission has a moderate set of orbital parameters to balance its spatial coverage and temporal resolution of 35 days (Benveniste et al., 2002). In order to determine if the altimetry

20 data fall into the study lakes, a lake shape data set generated by Wan et al. (2016) was used.

We used Landsat TM/ETM$^+$/OLI Surface Reflectance data sets provided by GEE to generate information on lake shore and lake area changes. TM images covered 2000–2011, whereas ETM$^+$ images covered 2000–2017 and OLI images covered 2013–2017. Though Landsat ETM$^+$ suffered sensor failure and all the ETM$^+$ images contain gaps after 2003 (Markham et al., 2004), we managed to make use of some images with gaps in generating lake shore changes. All the images were processed online

25 using GEE API. Preprocessing such as radiometric calibration, atmospheric correction, as well as geometric correction was already performed in the production of the data sets. There were more than 20,000 images processed and a half of them were excluded from the final results due to cloud contamination or gaps.

We collected daily in situ water level measurements in Lake Yamzhog Yumco for validation purposes with a pressure-type water level sensor. The in situ water level measurements spanned a half year from May to October 2018. We also performed

30 unmanned aerial vehicle (UAV or drone) imaging over a 1 km–lake bank in Lake Yamzhog Yumco and Lake Nam Co. The UAV images were used to evaluate the accuracy of lake shore extraction from Landsat data, which is similar to Huang et al.

(2018). In addition, we performed a rigorous statistical analysis of the uncertainty in the derived optical water levels by taking the UAV-derived water area ratios as the ground truth of sub-pixel water area ratios of Landsat image pixels (see Section 4).

Table 2. Multisource altimetry data used in this study

Mission	Sensor type	Duration	Repeat cycle	Footprint interval	Footprint diameter	Lake no. with data	Data source
Jason-1	Radar	2002–2013	10 d	~300 m	2–4 km	12	Aviso
Jason-2	Radar	2008–	10 d	~300 m	2–4 km	12	Aviso
Jason-3	Radar	2016–	10 d	~300 m	2–4 km	12	Aviso
ENVISAT	Radar	2002–2010	35 d	~390 m	3.4 km	35	ESA
CryoSat-2	Interferometer	2010–	369 d (Sub cycle 30 d)	~280 m	~1.65 km (across track) ~300 m (along track)	51	ESA
ICESat-1	Laser	2003–2009	91 d	~170 m	~70 m	42	NASA

3 Methodology

5 To investigate changes in lake storage, changes in lake water level and lake area need to be derived from multisource remote sensing. First, water levels from various satellite altimeters (see Figure 1) for each lake as well as changes in lake shoreline and area from optical remote sensing images (i.e., Landsat TM/ETM⁺/OLI) were derived. Second, the systematic biases between different altimetry data were removed by either comparing the mean water level of the overlap period or comparing the two water level time series with changes in lake shoreline, depending on the length of the overlap period. The lake shoreline

10 changes, termed as the optical water levels against altimetry water levels in this study, can serve as a unique source of information reflecting water level changes as well as a data merging reference. We will show that after deriving two or three regression parameters, lake shoreline changes can well represent lake water levels with comparable accuracy as altimetry-derived water levels. Third, with information on changes in water level and water area derived from altimetry data and optical remote sensing images, the hypsometric curve that describes the relationship between the lake water level and lake water

15 storage changes can be derived. Fourth, the integral of the hypsometric curve was performed to convert water level time series into lake storage change time series. Details will be illustrated in the following sections.

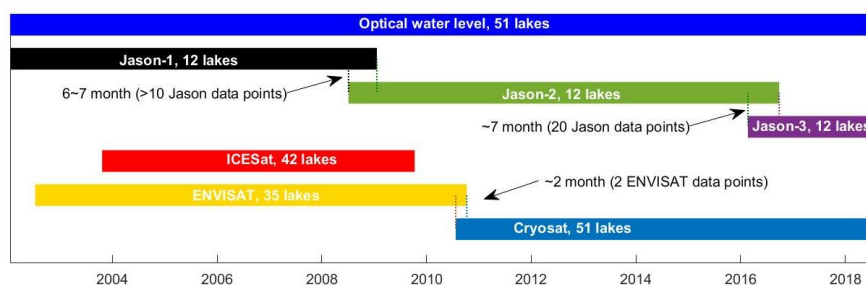



Figure 1. Spatial (the number of lakes covered) and temporal coverage and their overlap periods of multiple satellite altimetry missions used in this study, including Jason-1/2/3, ENVISAT, ICESat-1, and CryoSat-2.

3.1 Altimetry water level

5   The first step of deriving altimetry water levels is to select correct ground tracks and valid footprints falling on the lakes. Because there is a random ground track shift at ~1 km in different cycles for most altimetry missions, it is uncertain that valid lake footprints can be obtained for each cycle, even though the nominal ground track seems crossing the lake. This problem can be addressed by comparing the geographic coordinates of the footprints with a lake shape data set (Wan et al., 2016). After picking out the valid footprints, the lake surface height can be calculated for each footprint. All radar altimetry data share a

10   relationship:

$$LSH = Alt - (Range + cor)$$

Eq 1

where *LSH* represents the lake surface height with respect to the geoid. *Alt* represents the altimeter height with respect to the reference ellipsoid. *Range* represents the distance between the altimeter and lake surface. *cor* represents several range corrections including retracking correction, wet/dry troposphere correction, ionosphere correction, pole/solid tide correction, and geoid correction. The retracking range correction plays an important role in removing the contamination of land signal

15   when the altimetry data are applied to inland water bodies. In this study, Jason-1/2/3 data were retracked using a classical waveform retracking algorithm named the improved threshold method (ITR), whereas CryoSat-2 data were retracked using the narrow primary peak threshold (NTTP) method (Birkett and Beckley, 2010; Cheng et al., 2010; Jain et al., 2015). Retracking corrections were not performed for ENVISAT and ICESat data, because the ENVISAT product already provided a retracked range using the ICE-1 method and the ICESat GLAH14 product already included several corrections that are

20   sufficient for most applications including studies on inland water bodies (Zhang et al., 2011).

For each cycle of an altimeter, it is common that more than one footprint fall on a study lake, thereby providing several LSH observations on the same day. After removing outliers with the three-sigma rule, frequency distributions of the LSHs from the same cycle were calculated. The mean value of the highest bucket of the histogram was selected to represent the LSH for the cycle. Meanwhile, the frequency of the chosen bucket was reserved to evaluate the data quality for the cycle, e.g., a cycle was

25   marked as high quality if the frequency is higher than 0.8, moderate quality if it is only higher than 0.5, and poor quality if the

frequency is lower than 0.5. The LSH from each cycle constituted the lake water level time series for the study lake. LSHs that were marked as poor quality and obviously deviated from the moving average were removed from the altimetry-based lake water level time series.

It is not uncommon that systematic biases exist in different altimetry data sets due to the variation in orbit, discrepancy between
5 correction models, errors associated with sensors, and even the choice of the reference datum. After deriving lake water level time series for each altimeter, we first merged the Envisat and ICESat-1 water levels if both were available for a lake, because they have the longest overlap period (Figure 1). We chose Envisat-derived water levels as the baseline and removed the difference of the mean value of the overlap period between the two products, because Envisat data are generally denser and longer than ICESat-1 data. A similar process was applied to Jason-1/2/3, as there are two overlap periods connecting the three
10 altimeters together.

There are tradeoffs between CryoSat-2 and Jason-2/3 data in terms of spatial coverage and time span. CryoSat-2 data are available for all study lakes but they only have the overlap period with Jason-2/3 data, whereas Jason-2/3 data are only available for 12 lakes. For most lakes without Jason-2/3 data, we merged CryoSat-2 data with either ICESat-1 or Envisat using optical water levels spanning from 2000 to 2017, because there is no overlap period between these altimetry water levels. Lake
15 shoreline changes were firstly translated into optical water levels by fitting with CryoSat-2 data, functioning as extrapolation of CryoSat-2 to 1–2 years. Then, we applied the same method of merging Jason-1/2/3, to merge the extrapolated CryoSat-2 data with either Envisat or ICESat-1 data. In doing so, we were able to remove all systematic biases between multisource altimetry-derived water levels. All the water levels were with respect to EGM 96. Figure 2 provides an example of ground tracks of altimetric missions and water level time series on Lake Zhari Namco.

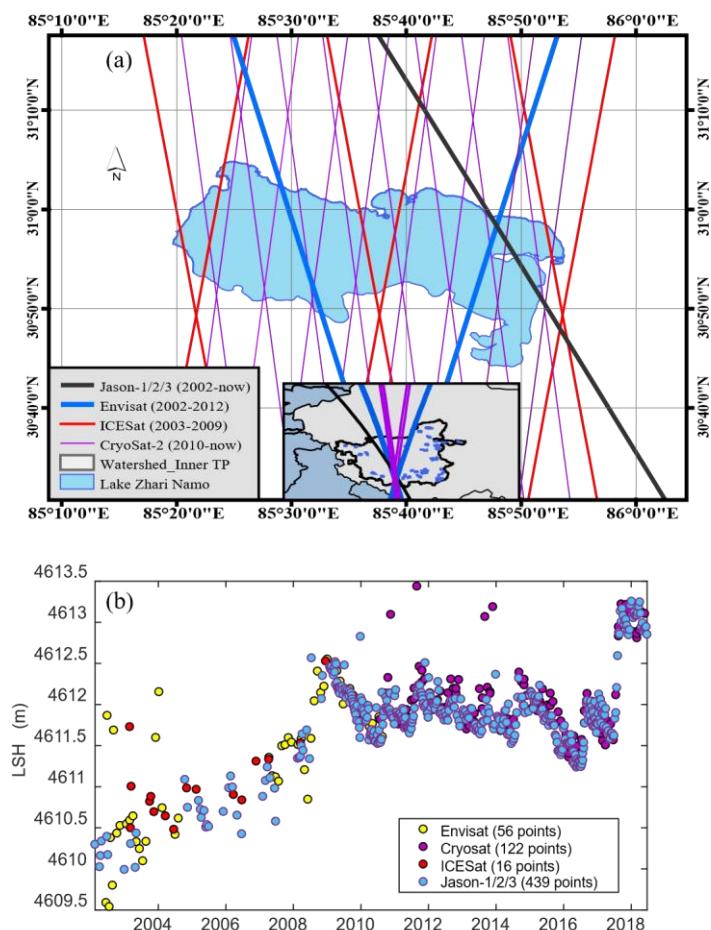



Figure 2. (a) Ground tracks of multiple altimetry missions over Lake Zhari Namco and (b) the merged altimetry water levels for Lake Zhari Namco.

3.2 Optical water level

5   For most lake basins, it is possible to find a relatively flat part of lake banks with an average slope of 1/30 or smaller, where obvious interannual or intra-annual changes in lake shorelines can be detected using high-spatial-resolution Landsat images (30 m). These locations can be found by comparing lake images from the first year and the last year of the study period if the lake shows a clear expanding/shrinking trend. Otherwise, we can compare images acquired in early summer when the LSH is at lowest with those acquired in late autumn when the lake expands to its limit. In this study, we assumed that the selected lake

10   bank was flat enough such that the relationship between changes in lake water level and shoreline can be depicted in a linear or quasi-linear (parabolic) way. The validity of this assumption can be evaluated with the coefficient of determination (R^2) for each lake as shown in Table 3. For most lakes, the goodness of fit is higher than 0.7, suggesting the generally good fitting relationship between changes in lake level and shoreline.

Though there were ~500 Landsat images obtained for the selected lake banks during the study period, many of them were largely affected by cloud or cloud shadow. In addition, the failure of Landsat 7 sensor SLC left all the Landsat 7 images with gaps after 2003 (Markham et al., 2004), making the available images even fewer. By choosing the region of interest (ROI) that is parallel to the Landsat 7 gaps, we can use most of the Landsat 7 images. However, the width of ROI must be reduced to

5 avoid shifting gaps as shown in **Figure 3 (b).** The ROI did not fill the interval of gaps, because the wider the ROI, the higher possibility of shifting gaps cross it.

Changes in lake shorelines were characterized by changes in the water surface ratio detected in the ROI. To automatically extract changes in water surface from tremendous amount of Landsat archives on GEE, the water index and Otsu threshold method were jointly used. We calculated the Normalized Difference Water Index (NDWI) and the Modified Normalized

10 Difference Water Index (MNDWI) of input images and compared their performance in different seasons. It was found that the MNDWI tends to be more sensitive to shallow water in summer, but is less effective than NDWI when the lake bank is covered by snow in the cold season as shown in **Figure 4**. Therefore, the two water indices were jointly used in this study by applying the MNDWI to images acquired during May to October and applying the NDWI to the rest months. The NDWI and MNDWI can be calculated as follows (McFeeters, 1996; Xu, 2005):

$$NDWI = \frac{B_{green} - B_{NIR}}{B_{green} + B_{NIR}}$$ **Eq 2**

$$MNDWI = \frac{B_{green} - B_{SIR}}{B_{green} + B_{SIR}}$$ **Eq 3**

15 where B_{green}, B_{NIR}, and B_{SIR} refer to reflectance of bands 2, 4, 5 for Landsat TM/ETM[+] images and bands 3, 5, 6 for Landsat OLI images.

After calculating the water index, the grayscale image was binarized using the Otsu method. If the selected ROI comprises ~50% water and ~50% land, the performance of the method is good, as the distribution of digital numbers of the grayscale image is close to the assumption of the bimodal histogram implicit in the Otsu algorithm (Kittler and Illingworth, 1985). The

20 binarized images were further processed to provide the water surface ratio in the ROI, which represents changes in lake shoreline. The time series of changes in lake shoreline were then converted into optical lake water levels using linear regression or second-order polynomial fit with altimetry-derived water levels (**Figure 3 (c)-(d)**).

However, cloud, cloud shadow, and shifting gaps may contaminate the ROI and cause errors in the optical water levels. Therefore, the QA band of the Landsat Surface Reflectance product was used to filter the images. The data would be excluded

25 if the fraction of the cloud or cloud shadow-covered area in the ROI was higher than 5%. For every Landsat ETM[+] image acquired after 2003, the pixel number of the ROI was counted and compared with those acquired before 2003. If the loss of

pixels exceeded 2%, the ROI was considered to be affected by gap and the data were consequently excluded from the subsequent analysis.

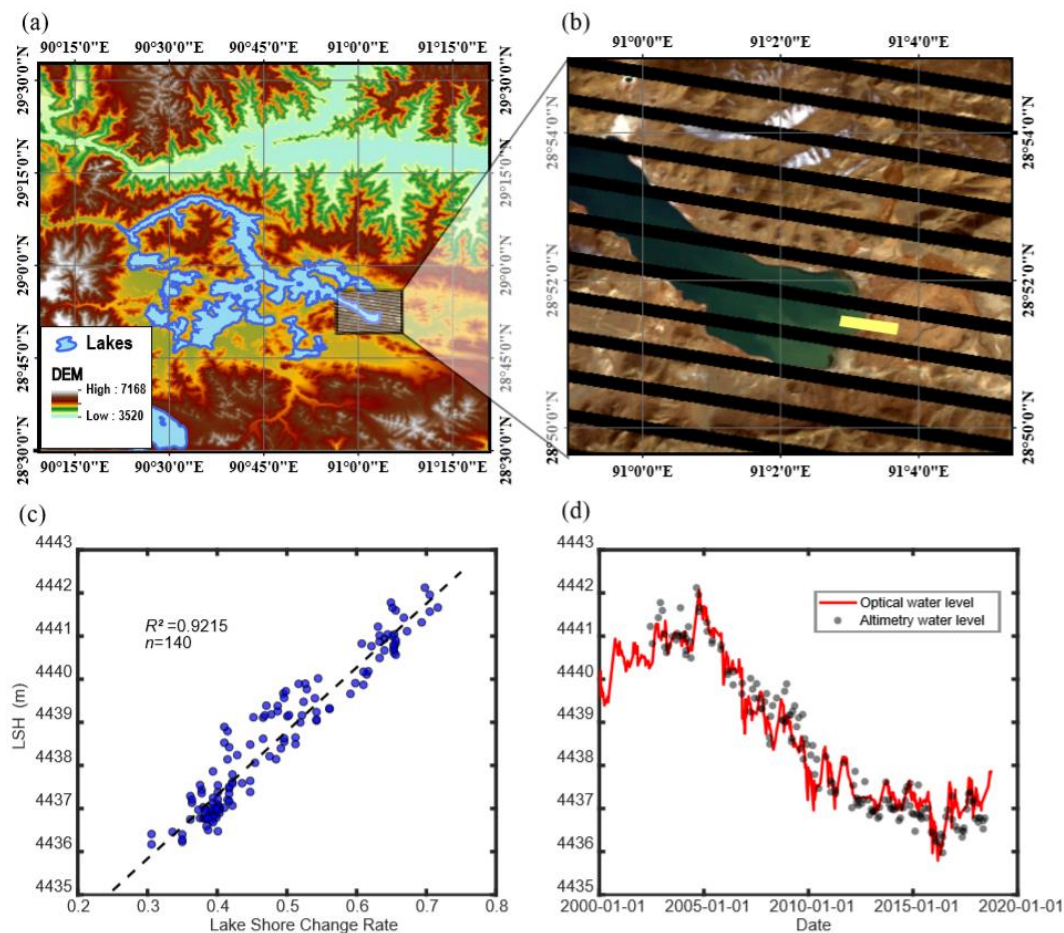



Figure 3. (a) Lake Yamzhog Yumco and its surroundings. The DEM was extracted from the STRM Global 90-m DEM product; (b)
ROI (yellow region) selected from a Landsat ETM+ image for detecting changes in lake shoreline and the gaps (black region); (c)
linear regression of the lake shoreline change and altimetry water level for Lake Yamzhog Yumco; and (d) optical water levels and
altimetry water levels for Lake Yamzhog Yumco.

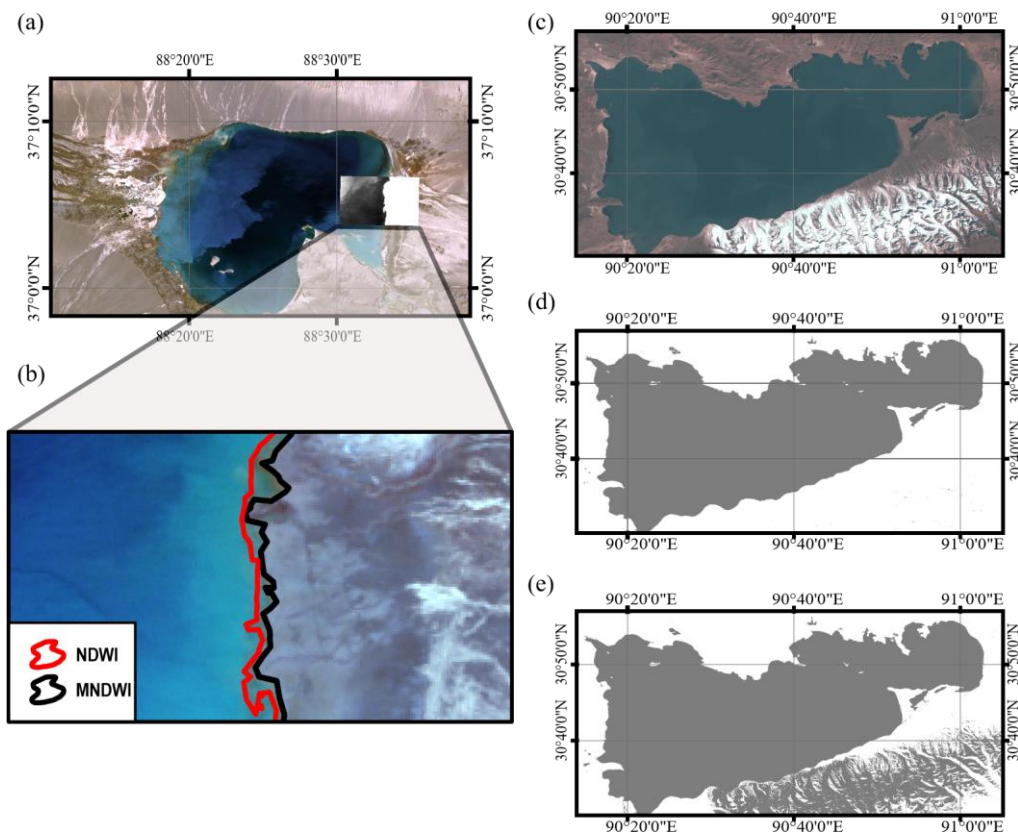



Figure 4. (a) Landsat ETM+ image of the Aqqikkol Lake acquired in summer in 2001; (b) water extraction result using the MNDWI and NDWI, showing that the MNDWI performs better in detecting shallow water; (c) Landsat OLI image of Lake Nam Co acquired in winter in 2015; (d) water extraction result of the NDWI, showing good performance in distinguishing water from snow; and (e) result of the MNDWI, showing some confusion of water and snow.

3.3 Hypsometry

We derived the hypsometric curve for each study lake by polynomial fitting of the lake area and level time series. The lake area comprises two parts: the inner invariable part and the outer variable part. As the variable water area is of more concern in this study, ROIs for extracting changes in lake area only cover the lake shoreline and its neighbouring area as shown in **Figure 5** . The inner part of the water body was calculated only once and considered as invariant, making the calculation more efficient on GEE. Meanwhile, more images are available as the area of ROI becomes smaller, because the possibility of clouds covering the ROI is reduced compared with an ROI covering the entire lake. The Landsat ETM+ images after 2003 were not included in this part of calculation as gaps negatively affected the ROI for lake area extraction. Similar to Section 3.2, we selected images with less than 5% cloud cover on an ROI to generate time series of changes in lake area, obtaining 20~30 data points on average for regression. R^2 values for each lake are listed in Table 3, indicating that most lake basins agree well with the parabolic hypsometric curve.

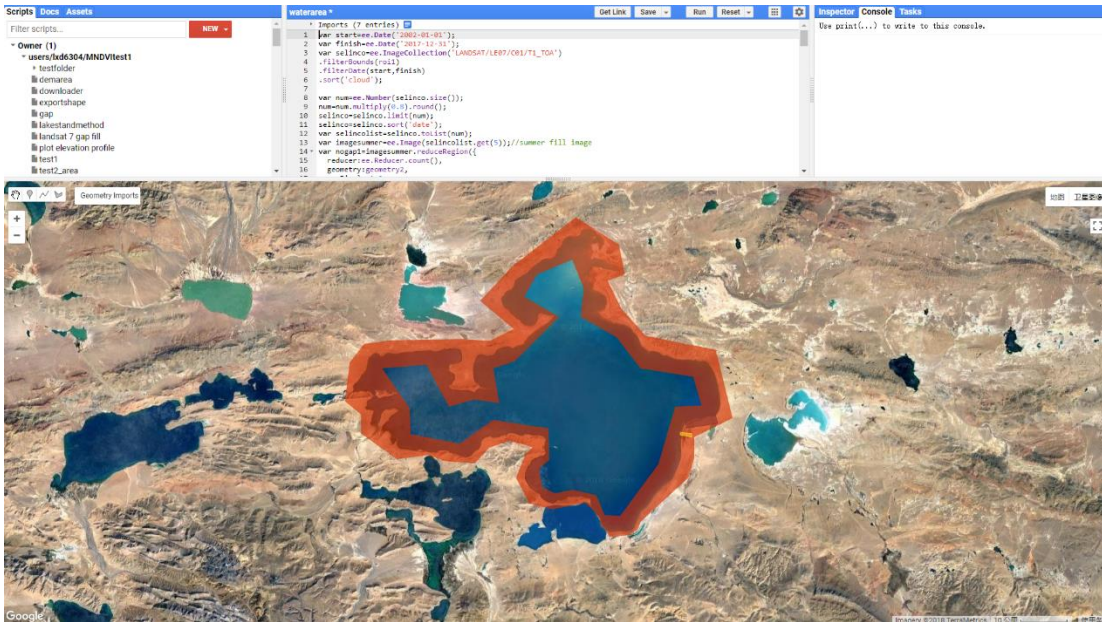



Figure 5 . Programming interface of GEE. The red polygon was the ROI for lake area change extraction of Lake Selin Co.

It can also be seen that most fitting functions have a positive parameter for the second-order term, which can be explained according to the knowledge on the hypsometric curve of catchments. This differs somewhat from the previous studies by Song et al. (2013) that shows a negative value for the second-order derivatives. After investigating a number of watersheds, researchers suggested that the general hypsometric curve for the entire catchment can be expressed as: (Strahler, 1952; Willgoose and Hancock, 1998):

$$y = \left[\frac{d-x}{x} \cdot \frac{a}{d-a}\right]^{z}$$

Eq 4

where y corresponds to the normalized elevation and x corresponds to the normalized area above the elevation. a, d, and z are fitting parameters. The hypermetric curve generated by this model always has a ′toe′ as shown in **Figure 6**, where the second-order differential of the curve is negative. This means that at the low elevation of the catchment, with decreasing (increasing) elevations, the area above (below) the elevation increases more slowly (faster). Lakes are always formed at the lowest portion of a catchment, so the hyperosmotic curve for a lake basin can be the toe part for the entire catchment. This suggests that if we use the parabolic curve to fit the lake area and water level time series, there should be a positive second-order parameter so that with increasing water levels, the lake area increases faster. The last step was to integrate the hypsometric curve to generate the volume−elevation relationship and convert the lake water levels into storage changes.

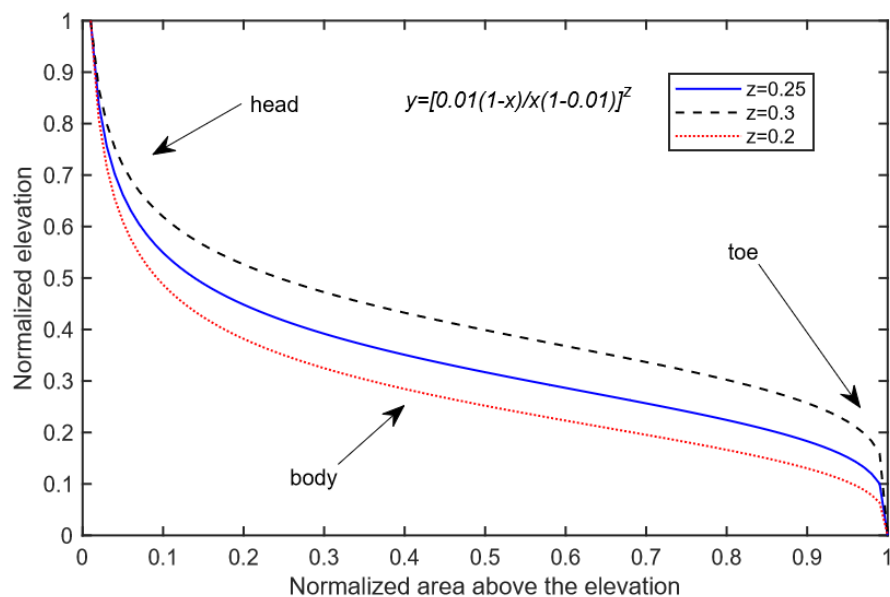



Figure 6. Strahler's hypsometric model for catchment with _a_=0.01 and _d_=1.

Table 3. Information on regression analysis of study lakes.

Lake Name	Lake Area km²	No. of optical water level	R^2 of optical water level	R^2 of hypsometric curve	Hypsometric function
Ake Sayi Lake	260.74	113	0.8951	0.9556	S=0.45dh²+11.26dh+163.97, dh=H-4846
Aqqikkol Lake	538.21	354	0.9717	0.9353	S=2.36dh²+0.21dh+370.29, dh=H-4252
Ayakkum Lake	987.23	183	0.9651	0.9695	S=0.16dh²+65.72dh+658.28, dh=H-3878
Bamco	255.81	209	0.901	0.9287	S=0.28dh²+2.84dh+206.46, dh=H-4560.5
Bangong Co	661.64	232	0.5164	0.7991	S=1.43dh²+15.67dh+619.28, dh=H-4238
Chibzhang Co	541.96	49	0.8766	0.9792	S=0.69dh²+3.36dh+475.79, dh=H-4930
Co Ngoin1	268.37	174	0.6637	0.8803	S=3.67dh²+-1.33dh+263.1, dh=H-4564.5
Cuona Lake	192.15	254	0.7607	0.8876	S=1.77dh²+3.6dh+184.79, dh=H-4585.5
Dagze Co	310.8	192	0.8334	0.8862	S=0.08dh²+6.14dh+230.51, dh=H-4460
Dogai Coring	492.39	257	0.8624	0.9048	S=3.2dh²+5.66dh+427.17, dh=H-4816
Dogaicoring Qangco	403.18	162	0.9202	0.9218	S=0.53dh²+3.93dh+279.6, dh=H-4786
Donggei Cuona Lake	247.83	561	0.8776	0.925	S=0.54dh²+7.22dh+222.19, dh=H-4084
Dung Co	139.4	145	0.9218	0.8652	S=0.07dh²+2.3dh+137.06, dh=H-4547
Goren Co	477.95	191	0.6166	0.9096	S=2.91dh²+-0.03dh+468.33, dh=H-4648.5
Gozha Co	246.91	96	0.4297	0.5564	S=1.57dh²+-0.06dh+254.43, dh=H-5082
Gyaring Lake	535.84	253	0.6217	0.3451	S=1.99dh²+2.8dh+517.18, dh=H-4292
Har Lake	621.52	370	0.8652	0.9893	S=1.1dh²+1.52dh+582.34, dh=H-4075
Hoh Xil Lake	351.3	132	0.9038	0.9355	S=1dh²+5.29dh+300.5, dh=H-4887.1

Jingyu Lake	339.69	224	0.8978	0.989	$S=0.37dh^2+4.77dh+238.43$, dh=H-4710
Kusai Lake	328.8	295	0.9787	0.8987	$S=0.52dh^2+5.04dh+254.67$, dh=H-4473
Kyebxang Co	187.32	233	0.75	0.8753	$S=0.16dh^2+5.4dh+150.9$, dh=H-4619
Langa Co	256.03	167	0.859	0.888	$S=-0.19dh^2+4dh+249.28$, dh=H-4564
Lexiewudan Co	273.63	286	0.9216	0.9496	$S=0.13dh^2+4.63dh+219.65$, dh=H-4868
Lumajiangdong Co	386.71	220	0.9135	0.9708	$S=0.62dh^2+2.09dh+353.95$, dh=H-4812
Mapam Yumco	412.69	163	0.7096	0.9973	$S=1.18dh^2+5.16dh+399.68$, dh=H-4584
Margai Caka	137.7	247	0.9399	0.9955	$S=0.03dh^2+5.14dh+112.12$, dh=H-4791
Memar Co	167.3	193	0.911	0.8626	$S=0.27dh^2+3.17dh+134.69$, dh=H-4923
Nam Co	2028.5	187	0.9064	0.8749	$S=2.43dh^2+5.55dh+1970.1$, dh=H-4724.5
Ngangla Ringco	492.86	88	0.4652	0.9498	$S=3.87dh^2+3.86dh+490.69$, dh=H-4715
Ngangze Co	471.21	245	0.9538	0.9332	$S=0.2dh^2+7.03dh+391.21$, dh=H-4680
Ngoring Lake	656.83	86	0.844	0.8613	$S=4.69dh^2+-5.04dh+613.66$, dh=H-4270
Paiku Co	272.85	231	0.8341	0.9079	$S=0.91dh^2+2.64dh+264.89$, dh=H-4578.5
Puma Yumco	290.98	250	0.6871	0.5629	$S=0.48dh^2+0.8dh+286.34$, dh=H-5011
Pung Co	176.93	187	0.8017	0.9841	$S=0.03dh^2+3.75dh+151.66$, dh=H-4526
Qinghai Lake	4495.33	323	0.9011	0.8181	$S=3.45dh^2+155.03dh+4084.73$, dh=H-3193
Rola Co	169.83	347	0.7842	0.9403	$S=-0.88dh^2+14.87dh+115.59$, dh=H-4816
Salt Lake	144.3	206	0.9344	0.9858	$S=0.16dh^2+-0.69dh+37.42$, dh=H-4430
Salt Water Lake	212.47	347	0.9086	0.9494	$S=-0.82dh^2+17.31dh+133.71$, dh=H-4901
Selin Co	2300.49	179	0.9777	0.945	$S=1.05dh^2+45.86dh+1754.31$, dh=H-4536.4
Tangra Yumco	862.94	100	0.9155	0.8072	$S=0.94dh^2+-0.28dh+862.94$, dh=H-4536
Taro Co	485.15	268	0.8903	0.9576	$S=0.18dh^2+4.97dh+477.32$, dh=H-4567.3
Tu Co	448.64	257	0.9276	0.9875	$S=0.02dh^2+4.91dh+396.59$, dh=H-4926
Urru Co	356.35	260	0.71	0.8994	$S=1.35dh^2+2.67dh+345.34$, dh=H-4553
Wulanwula Lake	652.08	225	0.9679	0.9285	$S=2.05dh^2+16.49dh+513.15$, dh=H-4856
Xijir Ulan Lake	463.36	316	0.9736	0.9691	$S=0.93dh^2+13.3dh+366.35$, dh=H-4770.8
Xuru Co	209.87	144	0.5984	0.5527	$S=0.12dh^2+0.22dh+206.53$, dh=H-4714
Yamzho Yumco	549.61	398	0.9215	0.9364	$S=0.51dh^2+9.63dh+531.79$, dh=H-4436
Yelusu Lake	203.4	486	0.7014	0.8352	$S=14.84dh^2+-5.77dh+185.15$, dh=H-4686.5
Yibug Caka	178.22	118	0.9206	0.9615	$S=-1.25dh^2+15.79dh+147.03$, dh=H-4558.5
Zhari Namco	1000.18	143	0.9177	0.8388	$S=2.66dh^2+10.07dh+962.57$, dh=H-4610
Zhuonai Lake	160.1	260	0.9528	0.973	$S=0dh^2+10.06dh+124.29$, dh=H-4742
Zige Tangco	238.67	171	0.9008	0.976	$S=0.06dh^2+4.62dh+212.71$, dh=H-4565

4 Validation of data quality

4.1 Field experiment

Most Tibetan lakes are located in remote and inaccessible regions, resulting in the scarcity of ground-based in situ measurements that are, however, vital for data quality assessment. We made some in situ measurements in two study lakes to

5   validate the data quality of optical water levels developed in this study, instead of the satellite altimetry data whose quality on lakes or rivers has been widely known. Many studies used in situ water levels to calculate certain statistical metrics, e.g., root mean squired error (RMSE). However, results provided by different studies vary and could be associated with the cross-section width of the study water body in the ground track panel (Nielsen et al., 2017). This means that these results may not be comparable due to their unique applications. In addition, it is not rigorous to use in situ data of only one lake to represent the

10   overall situation of study lakes in the uncertainty assessment for altimetry water levels. Instead, we used the standard deviation of valid footprints acquired in the same cycle as an estimate of uncertainty in altimetry-derived water levels. In contrast, the applicable condition of optical water levels is not so variable as that of altimetry data. Derivation of optical water levels requires relative flat bank as well as some altimetric information, which were available in all study lakes. Since these selected bank slopes were similarly small (~1/30), it is possible to use a few typical lakes to represent all study lakes. Therefore, we carried

15   out a field experiment in Lake Yamzhog Yumco and Lake Nam Co to validate the derived optical water levels.

There were two main goals in our experiment: (1) collecting daily in situ water level data in a certain TP lake to validate the corresponding optical water levels statistically; (2) imaging a certain length of the lake shore with UAV to test the performance of lake shore extraction so as to provide a theoretical uncertainty analysis of optical water levels. On Lake Yamzhog Yumco, we installed a pressure type water level sensor (type H5110−DY, manufactured by Shenzhen Hongdian technologies Co., Ltd.),

20   which measures water pressure and temperature of the installation depth and converts them into water depths with a relative accuracy of ~0.1%. The device was carried onto the lake and put ~20 m below the water surface and 0.5 m above the lake bottom, which suggests an absolute error of ~2 cm. We chose a typical landscape (i.e., mild slope, little vegetation, and gravel lake beach) to perform UAV imaging in both Lake Yamzhog Yumco and Lake Nam Co in mid-May, 2018. The UAV was operating at a height of 200 m and imaging the ground at a constant rate in visible bands with a wide-angle lens. With a GPS

25   tracker onboard, these UAV images were well georeferenced. Then image mosaic was performed using the Piz4D mapper and converted into a digital orthophoto map (DOM), which is a similar process described by Huang et al. (2018). The spatial resolution of UAV data reached ~5 cm as we compared a real ground object with its size in images.

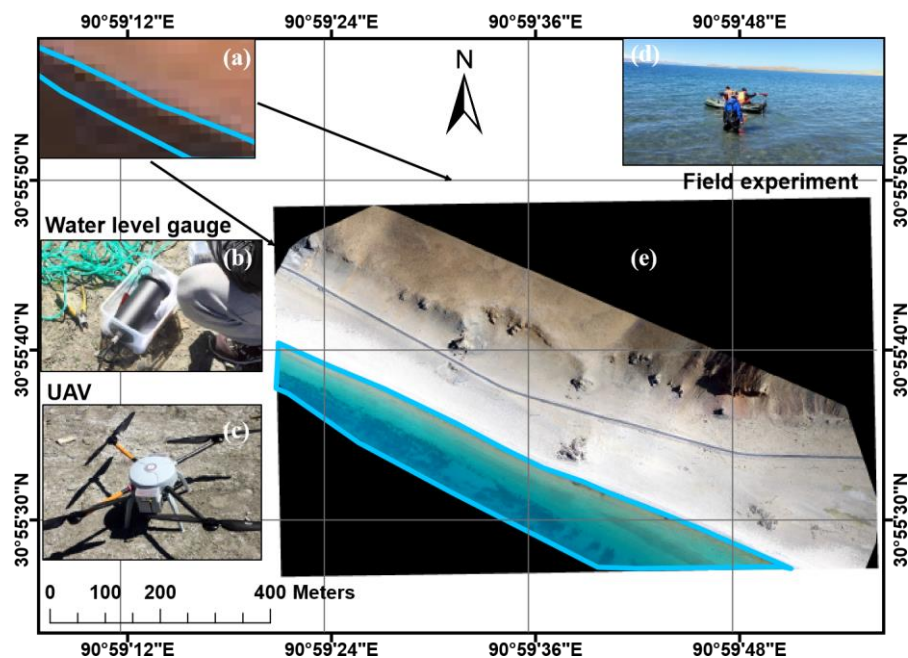



Figure 7. Field experiments in two study lakes: (a) Landsat 8 image of the experiment spot; (b) pressure-type water level sensor; (c) unmanned aerial vehicle; (d) installation of the water level sensor; and (e) UAV image of a portion of the bank of Lake Nam Co.

4.2 Uncertainty analysis of optical water levels

5   Based on the in situ water level measurements made by the pressure-type water level sensor, we evaluated the accuracy of optical water levels statistically. There were 16 optical water level records available for the comparison against the in situ measurements. The RMSE of the water level anomaly was 0.11 m. The linear fit had a slope close to 1 and an R^2 of 0.89, suggesting the good consistency of the in situ water level measurements and the derived optical water levels (Figure 8 (b)). It should be noted that the optical water levels used for validation here were translated from lake shore changes using parameters

10   derived from fitting with CryoSat-2 data, i.e., there is no in situ information involved in generating the optical water levels shown in Figure 8.

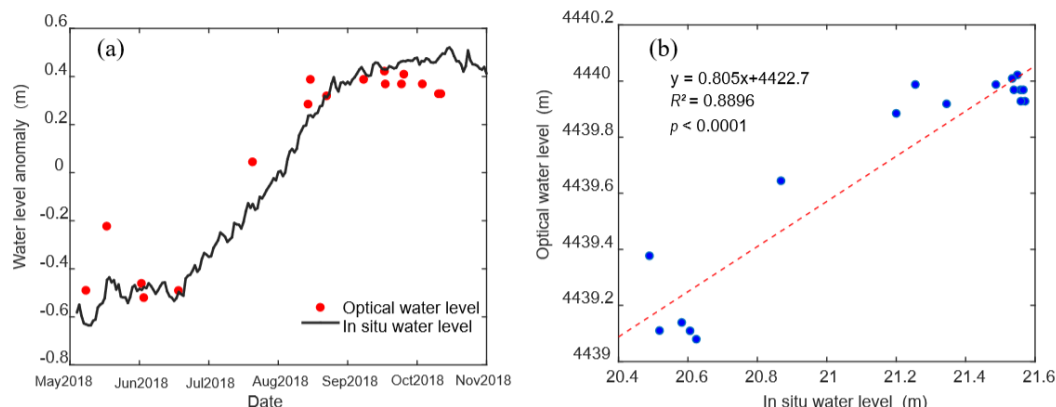



Figure 8. (a) In situ water level anomaly versus optical water level anomaly in Lake Yamzhog Yumco; and (b) linear regression of the optical water levels and concurrent in situ water levels.

To be more convincing, we performed a more theoretical uncertainty analysis of the optical water levels by looking at the original optical data and generation processes with the help of UAV images. First, we took UAV images as the ground truth to determine the accurate position and shape of the lake shore line. Second, we performed water classification based on the concurrent Landsat OLI image that contained the UAV scanned area, with the combined water index method and Otsu algorithm to derive the binarized image. Landsat image pixels where the lake shoreline determined using the UAV images crosses were delineated manually and marked as shoreline pixels as shown in Figure 9 (a). Then the water area in each shoreline pixel was calculated.

Given that these shoreline pixels were classified as either water or land, a relationship between the water area ratio of the shoreline pixel and the probability of the pixel being classified as water can be derived. This relationship generally describes the function of the water classification method by telling how likely a pixel will be determined as water, given the water area ratio of the pixel. Based on the observations of a total of 64 UAV-scanned shoreline pixels, a logistic type curve with two parameters that determine the position and shape of the curve was chosen to represent the water classifier as Eq 5 shows:

$$f(x) = \frac{1}{1 + e^{-a(x-b)}}$$

Eq 5

where x represents the water area ratio in the shoreline pixel, $f(x)$ represents the probability of the shoreline pixel being classified as water pixel, a and b are parameters that determine the shape and position of the curve, respectively. The parameters were determined using the maximum likelihood method. Results are shown in Figure 9 (b).

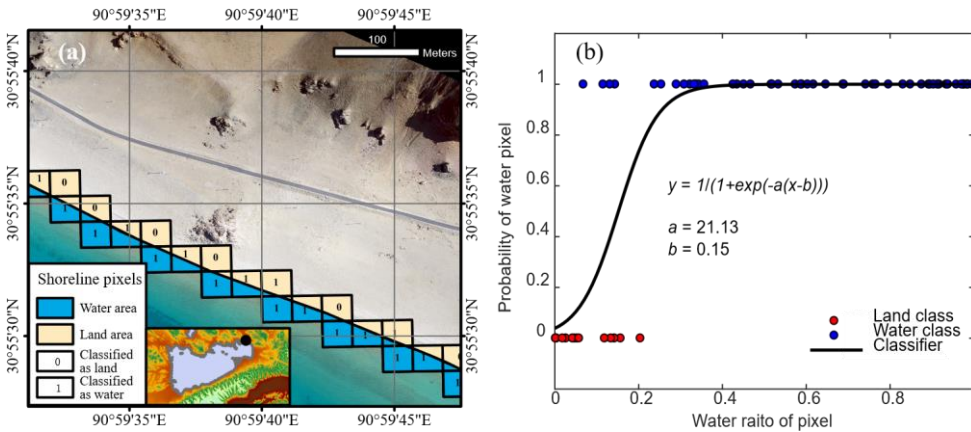



Figure 9. (a) Landsat OLI shoreline pixels on Lake Nam Co (the background is the UAV image), pixels marked with 1 were classified as water, and pixels marked with 0 were classified as land. The black dot in the insert figure represents the field experiment spot. (b) The chosen classifier, with two parameters a=21.13 and b=0.15 determined from the maximum likelihood method. The y axis represents the probability that a certain pixel is classified as water, and the x axis is the water area ratio in that pixel. Red dots represent pixels classified as land that have relatively lower water area ratios. Blue dots represent pixels classified as water that have relatively higher water area ratios.

It should be noted that even the water ratio in a shoreline pixel is zero, there is a small probability that this pixel may be mistaken as a water pixel, because the surface reflectance information may be contaminated by adjacent water pixels. Therefore, the classifier has a small value (~0.02) when the water ratio is zero. As can be seen from Figure 9 (b), when the water ratio is larger than 0.3, the probability of the pixel classified as water is close to 1. This suggests that there may be a higher probability of the occurrence of water pixels that is associated with a systematic bias of the lake shoreline detection. Note that the systematic bias can be removed when linearly fitting the lake shore changes and altimetry-derived water levels as long as the bias is stable. Therefore, uncertainty in optical water levels developed in this study arises from the variation in this systematic bias.

To describe the variation in the systematic bias, a new random variable X was introduced to represent the bias between the classified water area and the real water area in a shoreline pixel. Given the shape and position of the shoreline, the real water area in each shoreline pixel is a complex function of the relative position between the pixel and the shoreline. To simplify the derivation, we assumed that the water area ratio in a shoreline pixel is uniformly distributed on [0,1], meaning that the probability of any value between 0 and 1 is equal. If we use X_0 to represent the true water area ratio in the shoreline pixel and X_1 to represent the classified results based on the water ratio, the random variable X can be expressed as:

$$X = X_1 - X_0 \qquad\qquad\qquad \text{Eq 6}$$

where X_1 can take on 0 or 1 (i.e., the classified results only tell us whether a pixel is water pixel or not), so X can only take on either $-X_0$ or $1-X_0$. Because the range of X_0 is [0,1], it is obvious that the range of X is [-1,1]. A derivation of $F(X)$, i.e., the probability density function (PDF) of X is given as follows:

$$F(X) = \begin{cases} P(X = -X_0|X_0) \cdot P(X_0), & X < 0 \\ P(X_0 = 0, X_1 = 0) + P(X_0 = 1, X_1 = 1), & X = 0 \\ P(X = 1 - X_0|X_0) \cdot P(X_0), & X > 0 \end{cases} \qquad \text{Eq 7}$$

$$P(X = -X_0|X_0) = P(X_1 = 0|X_0) = 1 - f(X_0) = 1 - f(-X) \qquad \text{Eq 8}$$

$$P(X = 1 - X_0|X_0) = P(X_1 = 1|X_0) = f(X_0) = f(1 - X) \qquad \text{Eq 9}$$

where *f(x)* is referred to as the classifier function defined in Eq **5.**

Because X_0 is uniformly distributed between [0,1], $P(X_0) = 1$. Note that $P(X_0 = 1, X_1 = 1) = 1$, this means that if X_0 is 1, the probability of a pixel classified as water is 1. However, $P(X_0 = 0, X_1 = 0)$ is not necessarily 1 (i.e., 1-*f*(0)), because the pixel with $X_0 = 0$ still has a low probability, i.e., *f*(0) of being classified as water based on the illustration above. Combining Eqs.(7)–(9) with the explanations on the specific case $X = 0$ results in the following:

$$F(X) = \begin{cases} 1 - f(-X), & X < 0 \\ 1 + 1 - f(0), & X = 0 \\ f(1 - X), & X > 0 \end{cases} = \begin{cases} 1 - f(-X), & X < 0 \\ 2 - f(0), & X = 0 \\ f(1 - X), & X > 0 \end{cases} \qquad \text{Eq 10}$$

It is obvious that $F(X)$ is not a continuous function, but it can be integrated and the integral of $F(X)$ on [-1,1] equals 1, meaning that it satisfies the basic property of PDF.

Overall, $F(X)$ describes how the systematic bias between the classified water ratio and real water ratio in shoreline pixels is distributed as shown in Figure 10. If there are N shoreline pixels in an ROI, we can take them as N independent observations of X and calculate the mean value $\bar{X}$. This value $\bar{X}$ can represent an average shift of the detected lake shoreline from the real lake shoreline in the unit of one-pixel width (30 m). As we mentioned above, the systematic bias can be removed in the regression between lake shore changes from optical remote sensing and the corresponding water levels from satellite altimetry. As such, it is the variation of the bias that determines the accuracy of the optical water levels. Therefore, we can calculate the standard variation of $\bar{X}$ to represent the uncertainty in lake shoreline changes. Note that there is a simple relationship between $\sigma_{\bar{x}}$ and σ_x:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}} \qquad \text{Eq 11}$$

One only needs to calculate σ_x:

$$\bar{X} = \int_{-1}^{1} F(X) \cdot X dX = 0.346 \qquad \text{Eq 12}$$

$$\sigma_x = \sqrt{\int_{-1}^{1} F(X) \cdot (X - \bar{X})^2 dX} = 0.2999 \approx 0.300$$

Eq 13

Eq 13 in combinaions with Eq 5, Eq 10, and Eq 12 was resolved numarically, resuling in ~0.3 pixel width. Substituting σ_x in Eq 11 with Eq 13 gives:

$$\sigma_{\bar{x}} = \frac{0.3}{\sqrt{N}}$$

Eq 14

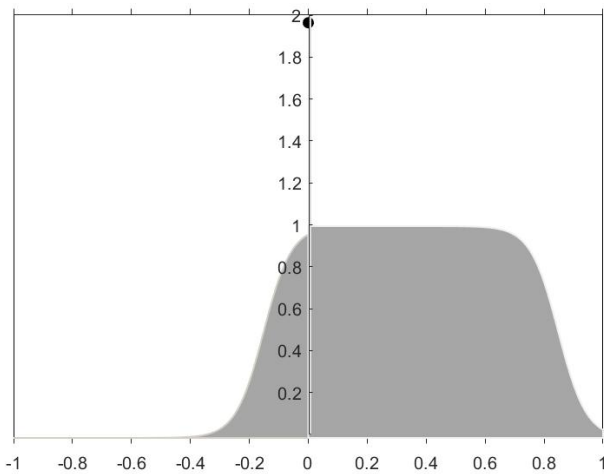



Figure 10. Probability density function of the systematic bias X between the classified water ratio (X_1) and real water ratio (X_0) in a shoreline pixel.

If the slope of the shoreline is known, e.g., $tan\theta$, the uncertainty of the optical water level can be expressed as:

$$\sigma_{ho} = \sigma_{\bar{x}} \cdot d \cdot tan\theta = \frac{0.3 \times 30 \times tan\theta}{\sqrt{N}}$$

Eq 15

where σ_{ho} is the uncertainty of optical water levels and d is the spatial resolution of the satellite image (30 m). In this study, a typical width of ROI for deriving optical water levels is ~10-pixel width, meaning that N is ~10. In addition, lake shores used for generating optical water levels here generally have a relatively mild slope of ~1/30 or even smaller, which can be roughly estimated from the maximun shoreline change and altimetry water level change within a year. Here if we use 1/30 as the slope $tan\theta$, the uncertainty of the optical water levels can untimately be estimated to be ~0.1 m, which is very close to the RMSE (~0.09 m) based on the comparision of the opitcal water levels and in situ water level measurements mentioned earlier.

Overall, the uncertainty quantification of the optical water levels developed in this study indicates clearly that the accuracy of optical water levels depends on the width of an ROI, e.g., the number of pixels/observations, slope of the lake shore, and the

effectiveness of the water classification method. One of the advantages of the optical water level is that an ROI does not necessarily cover a large area of lake shores, which maximizes the potential of optical remote sensing images to increase the spatial coverage and temporal resolution of lake water level estimates that may not, however, be achievable by using satellite altimetry alone. Optical remote sensing images provide important complementary information on altimetry-derived water
5   levels and would subsequently facilitate lake water storage estimation.

5 Applications

5.1 Spatiotemporal analysis of changes in Tibetan Lakes

Based on the lake water storage changes we derived, spatial patterns of lake storage trends from 2000 to 2017 were shown in Figure 11. In the endorheic basin of the TP, similar to some reported results (Yao et al., 2018b; Zhang et al., 2017a), most
10   lakes have been expanding rapidly, e.g., Lake Selin Co (89.00 E, 31.80 N) gained ~19 km^3 of water during the study period, Lake Kusai (92.90 E, 35.70 N) experienced an abrupt expansion due to flood and gained ~2.2 km^3 of water in 2011, as reported in related work (Xiaojun et al., 2012). In contrast, some lakes in the southern part of the TP experienced shrinkage, e.g., Lake Yamzhog Yumco (90.70 E, 28.93 N) gained a total of ~0.8 km^3 water during 2000–2004 but has been shrinking during the remaining 13 years (2005–2017) at a rate of ~ -0.2 km^3/yr. In contrast to Lake Yamzhog Yumco, Lake Qinghai (100.00 E,
15   36.90 N) lost ~1.9 km^3 water during 2000–2004 but gained ~8.1 km^3 of water during 2005–2017. Similar patterns can be detected in adjacent lakes of Lake Qinghai, e.g., Lake Donggei Cuona (98.55 E, 35.28 N) and Lake Ngoring (97.70 E, 34.90 N).

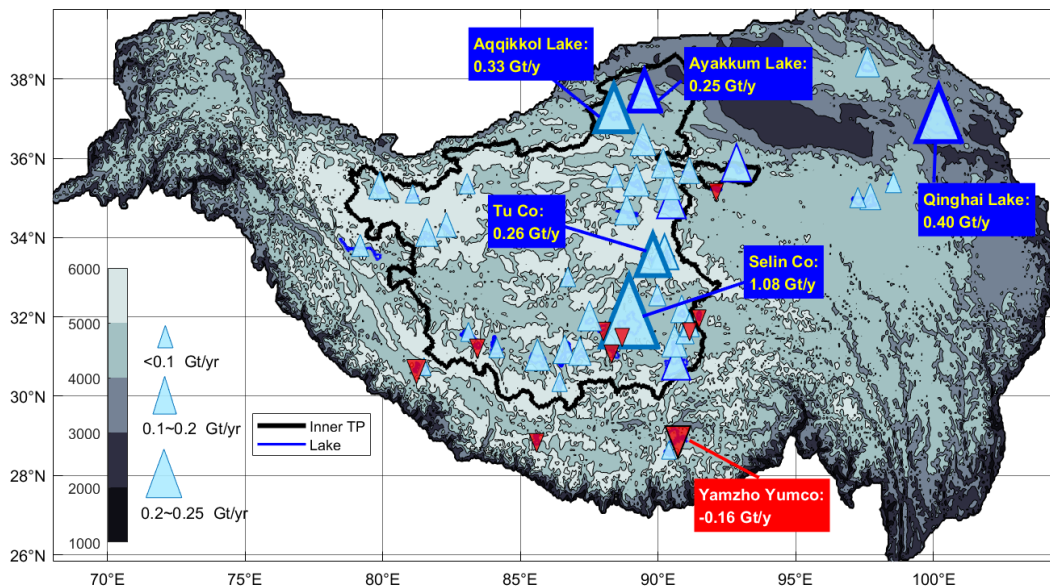



Figure 11. Spatial distribution of trends in lake storage on the TP during 2000–2017. The back polygon shows the boundary of the endorheic basin of the TP including 39 study lakes. The other 13 study lakes are located outside the endorheic basin.

However, spatial proximity cannot fully explain the intricate trend distribution in the Selin Co basin, where large lakes such as Lake Selin Co were expanding whereas smaller adjacent lakes showed an opposite decreasing trend, e.g., Lake Urru (88.00 E, 31.70 N), Lake Co Ngoin (88.77 E, 31.60 N), and Lake Goren Co (88.37 E, 31.10 N). In fact, we found that the decreasing trends in some small lakes like Lake Goren Co were not detected in Yao et al. (2018b), which is likely due to the lower sampling frequency as shown in Figure 12. The three shrinking lakes are located in the upstream region and feed Lake Selin Co through two small rivers. One of the rivers links Lakes Goren Co, Urru, and Selin Co, whereas the other links Lakes Co Ngoin and Selin Co.

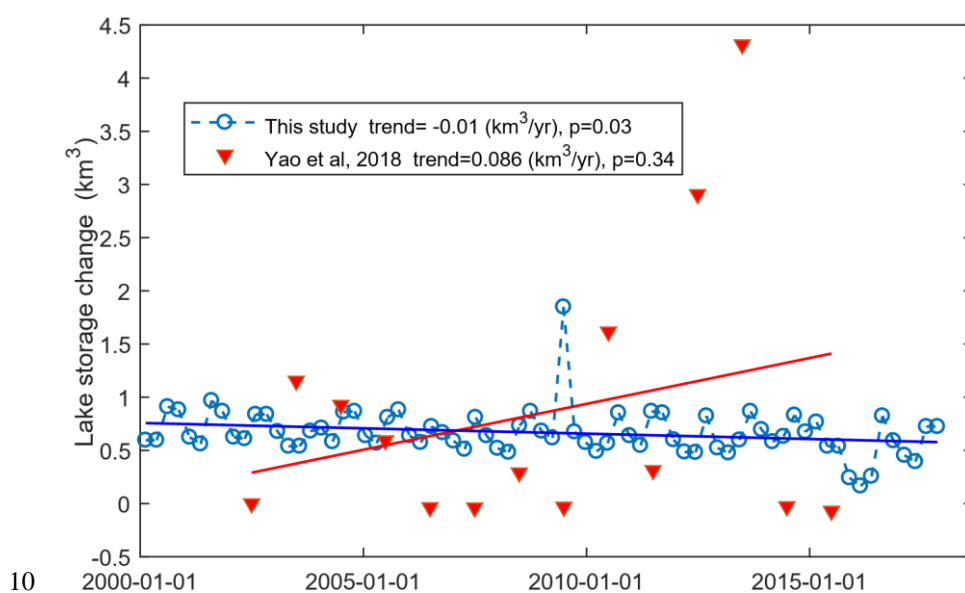



Figure 12. Discrepancy of lake storage trends in Lake Goren Co between Yao et al. (2018a) and our study.

A possible explanation of the disparity of changes in lake water storage in the Selin Co basin could be the principle of minimum potential energy. If we simplify the basin with the tank model and take the upstream small lake as a tank with a leaking hole, the storage of the small lake is mainly controlled by the height of the leaking hole. Given that surface water of the small lake increased, most of the increased water would flow into the large lake (a lower tank), and the outflow discharge of the small lake at higher elevations would increase accordingly. The height of the leaking hole would decline (erosion) so as to increase the overflow capacity, which eventually results in the decrease in small lake storage. Another possible situation is that the height of the leaking hole remains the same and the water surface height of the small lake increases, but this situation is not consistent with the minimum potential energy principle, as more water potential energy is stored in the small lake. This phenomenon shows that river-lake interactions may cause complex patterns of the regional surface water distribution. Therefore, decreases in small lake water storage and increases in water storage of Lake Selin Co in the basin detected by our

study seem reasonable. Increases in small lake water storage in this basin reported in some published studies may be associated with the sparse sampling of lake water levels.

We averaged the total lake water storage change in each season to generate the time series shown in Figure 13 **(a)**. The overall storage change in the 52 study lakes is 98.3±2.1 km^3. The total lake water storage was increasing rapidly during the first 12
5 years but became relatively stable since 2012.

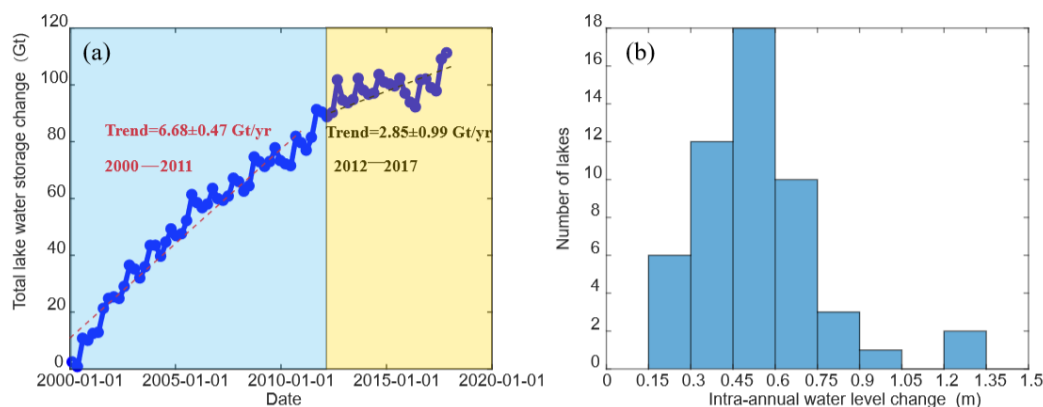



Figure 13. (a) Total storage changes in the study lakes (52) on the TP, which can be generally divided into two stages: (1) a rapidly increasing stage (2000–2011) with a higher increasing rate of 6.68 km^3/yr and (2) a mildly increasing stage (2012–2017) with an increasing rate of 2.85 km^3/yr. (b) Histogram of changes in lake water levels of the study lakes on the TP.

10 Intra-annual variation in the TP lakes can also be investigated using the densified time series generated by this study. We removed the linear trend (sometimes there were multiple linear trends for a lake in different periods, which were removed in a stepwise fashion) and calculated the mean monthly water level anomaly for each lake across the study period. Then the intra-annual water level change was represented by the difference between the maximum and minimum values of the monthly water level anomaly. The histogram of the intra-annual water level change in Figure 13 (b) shows that most of the TP lakes have
15 water level variations ranging from 0.3–0.75 m in a year on average. Similar work was performed by Lei et al. (2017) but only a small number of lakes were investigated in their study.

5.2 Quality assessment of similar data products

We made a comparison with a widely used lake water level/storage data set provided by the LEOGS Hydroweb, indicating that our product may perform better in terms of the consistency as well as sampling frequency. Both advantages are important
20 in improving our understanding of responses of lakes to climate change. There are 21 same lakes in both our study and LEOGS Hydroweb. Annual trends in water level and lake storage during 2003–2015 were compared (Figure 14). Overall, the two products are consistent in terms of R^2 of the linear fit.

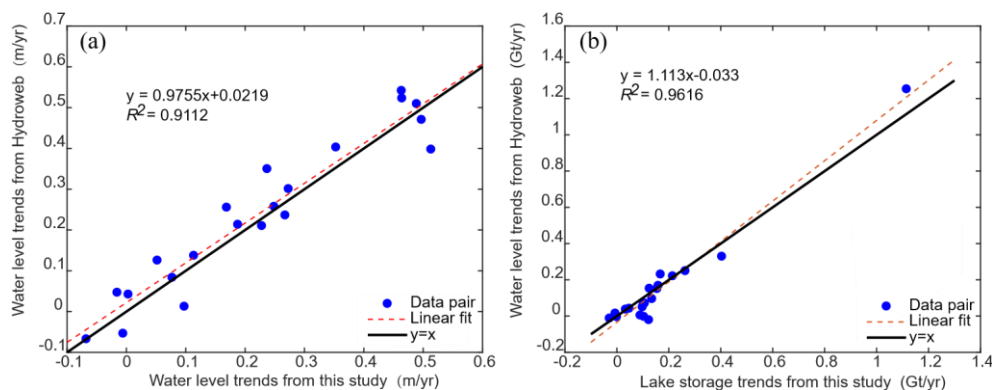



Figure 14. Cross validation of the TP lake level and storage changes derived from our study with those provided by the LEGOS Hydroweb database (Crétaux et al., 2011a): (a) trends in lake water levels from 2003 to 2015 and (b) trends in lake water storage from 2003 to 2015.

5 However, some obvious discrepancies between the two data sources still exist, e.g., water levels of Lake Taro Co. Both Hydroweb data and our estimation used ICESat-1 and CryoSat-2 data. The difference lies in the fact that our CryoSat-2 product was more updated with a longer time span but Hydroweb used an additional altimetry satellite SARAL. Since both products were performed some kind of removal of the systematic bias, it is possible that we chose different baselines that resulted in the overall shift as shown in Figure 15 **(a)**.

10 The black curve shows the optical water level we derived, which is a critical reference when connecting two different altimetry data time series without an overlap period. The optical water level shows that the last two samples of ICESat-1 data should not be lower than the first few samples of the CryoSat-2/SARAL data (see the dashed boxes). However, it is apparent that Hydroweb data display a reverse relationship. Though the optical water levels were derived by linearly fitting lake shore changes with altimetry data, the relative water levels during different periods should not be affected by the fitting parameters.

15 This is the main reason for us to use optical water levels as reference data. Therefore, Hydroweb data may underestimate the decreasing trends in the water levels of Lake Taro Co since 2009. This problem may also exist in some similar studies when multisource altimetry data without overlap periods were used.

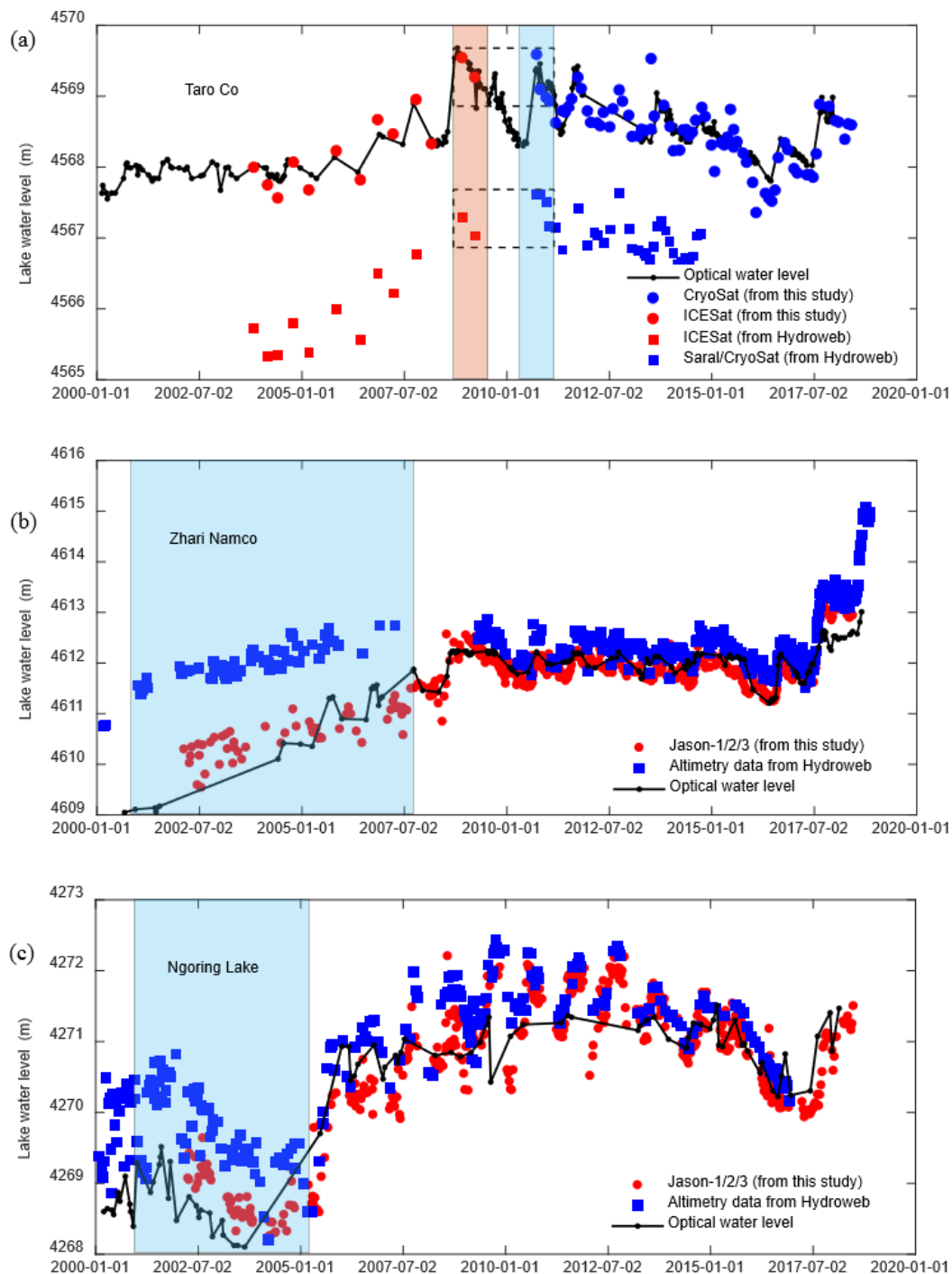



Figure 15. Similarities and differences between water level time series from the LEGOS Hydroweb database and this study (Crétaux et al., 2011a). (a) Lake Taro Co (84.12 E, 31.14 N); (b) Lake Zhari Namco (85.61 E, 30.93 N); and (c) Ngoring Lake (97.70 E, 34.90 N).

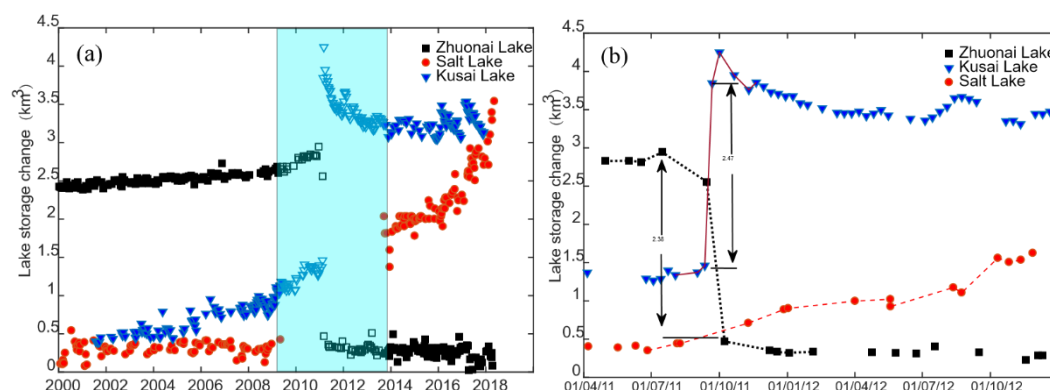
As shown in Figure 16, optical data can be less noisy than altimetry data in certain lakes and significantly improved the continuity of lake monitoring. In addition, a more apparent seasonality in lake change can also be seen from the densified lake level time series. These advantages would largely benefit a better understanding of responses of TP lakes to climate change and facilitate hydrological modelling of lake basins, regional water balance analysis, and even hydrodynamic analysis of lake water bodies.

Figure 16. Lake water level (left *y* axis) estimates from our approach for six TP lakes. Black lines represent optical data and red dots represent Altimetry data.

5.3 Lake overflow flood monitoring

As mentioned earlier in Section 5.1, Lake Kusai experienced an abrupt expansion in 2011, resulting from dike-break of an upstream lake (Hwang et al., 2019; Liu et al., 2016; Xiaojun et al., 2012), named Lake Zhuonai (91.93 E, 35.54 N). The

outburst of Lake Zhuonai occurred on Sep 14 (Liu et al., 2016) and ~2.4 km³ of water leaked into the Kusai River (as shown in Figure 17 (b)), the main inflow of Lake Kusai. The water level of Lake Kusai increased by up to ~8 m within 20 days (from Sep 11 to Oct 1 in 2011) based on Jason-2 data, and then started to drop as water overflowed from the southeast corner into Lake Haidingnuoer (93.16 E, 35.55 N) and Lake Salt (93.40 E, 35.52 N). Lake Salt, the lowest part of the basin close to the

5   basin boundary, has gained ~3.2 km³ of water since 2011 and became a critical threat to the surrounding residents and railway ~10 km southeast to the boundary. Note that there are few satellite altimetry data available for Lake Salt except several CryoSat-2 observations, where optical water levels can provide a near real-time monitoring of changes in lake water level and storage that are crucial to flood early warning and risk management.

10   **Figure 17. (a) Lake storage changes in Lake Zhuonai, Lake Kusai, and Lake Salt corresponding to the outburst event in Sep 2011 and (b) storage changes in relevant lakes during the outburst event.**

Aided by the densified water level series, it was possible to estimate the height of the outlet of Lake Kusai, an important parameter for overflow estimation. The overflow of Lake Kusai can help predict the water level rise in Lake Salt and even serve as an indicator of flood forecast, as Jason-3 data with a 10-day revisit cycle are now available on Lake Kusai. Several

15   pairs of concurrent Landsat OLI images and lake water levels were compared to provide a range of possible outlet heights, which are likely to be 4483.9 m to 4484.1 m, as shown in Figure 18 (a). Then we measured the mean width of the outlet from high resolution optical images provided by Planet Explorer (Team, 2017), which is relatively stable in Dec at 31.5±2.3 m in recent years.

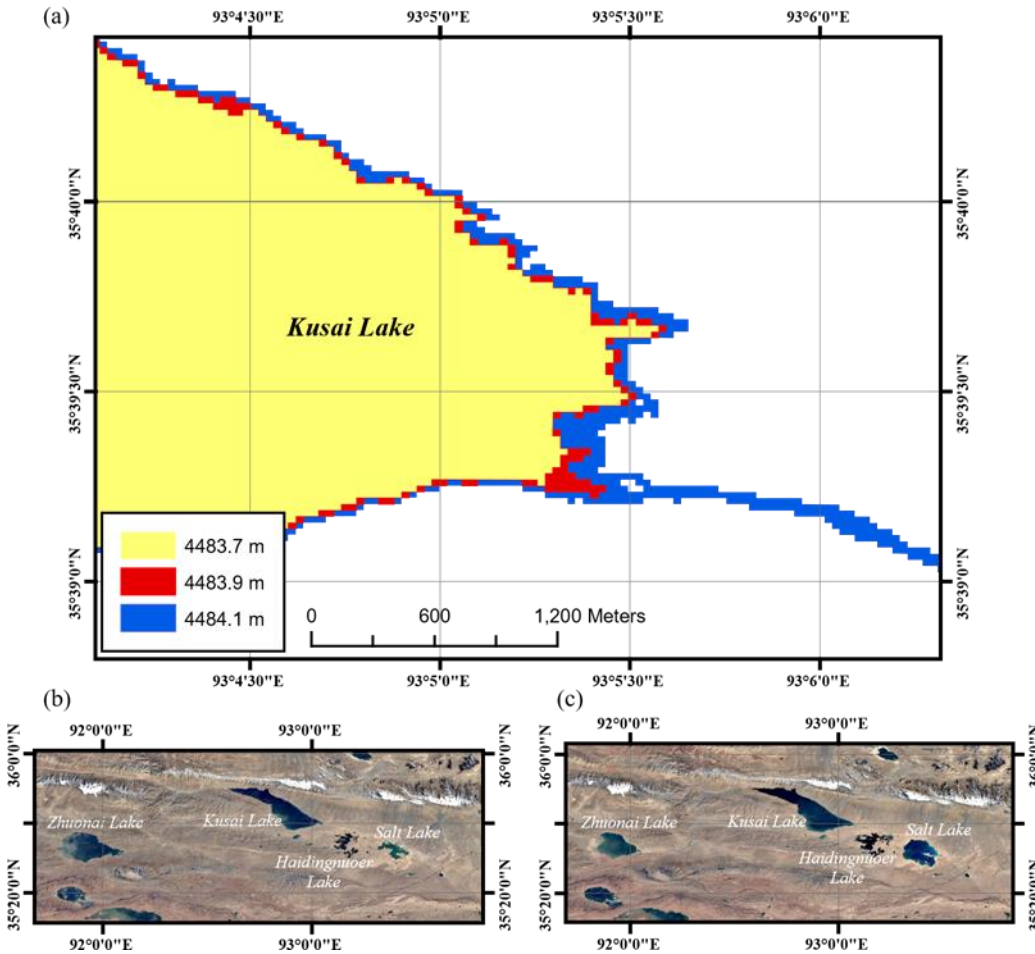



Figure 18. (a) Height variations in the outlet of Lake Kusai, the overflow would occur when the water level increases from 4483.9 m to 4484.1 m; (b) Landsat images before the outburst of Lake Zhuonai; and (c) Landsat images after the outburst event.

Given lake water levels, the outlet height and width, estimation of overflow can be made using the broad crest weir formula:

$$Q = C \cdot b \cdot H^{1.5} \sqrt{2g} \qquad\qquad\qquad \text{Eq 16}$$

5 where C is a parameter mainly reflecting geometric characteristics of the weir that mainly varies from 0.3–0.4, b is the width of the weir, H is the water head with respect to the top of the weir, and g is the acceleration of gravity.

It is difficult to obtain the exact value of C without performing field investigations. Nevertheless, the range of C can be narrowed down by investigating the lake storage change process of Lake Kusai. As shown in Figure 19 (a), from Oct 1 to Nov 9 in 2011, the water level of Lake Kusai decreased rapidly by ~1.2 m. Given that the rainy season has ended (Liu et al., 2016),

10 the water level of Lake Zhuonai became stabilized, providing minimum inflow to Lake Kusai. Meanwhile, the magnitude of total evaporation during the period would not exceed 0.1~0.2 m as the mean annual potential evaporation of the region is around 1000 mm (Zhang et al., 2007) and stage 1 (shown in Figure 19) only lasted for 40 days. In addition, the evaporation

loss could be partially compensated by inflow. Overflow should be the driving factor of the lake storage balance of Lake Kusai during the period. Therefore, we used the following function to reproduce changes in water level and storage in Lake Kusai during Oct 1 to Nov 9 in 2011 (stage 1 shown in Figure 19):

$$-\frac{dV(H)}{dt} = Q = C \cdot b \cdot H^{1.5}\sqrt{2g}$$ Eq 17

where V is referred to as the lake storage change, which is a function of water head H.

5 Equation 17 can be theoretically solved to describe the relationship between H and t if $V(H)$ is simple, e.g., a cubic curve. Otherwise, it can be solved with a numeric algorithm, such as the finite difference method. In a word, by solving the equation with different values of parameter C and optimizing the mean absolute error between the simulated result and the remotely sensed observations (shown in Figure 19 (a)), we suggested that C equals to 0.30 in this case, which is a reasonable value in hydrodynamic calculations.

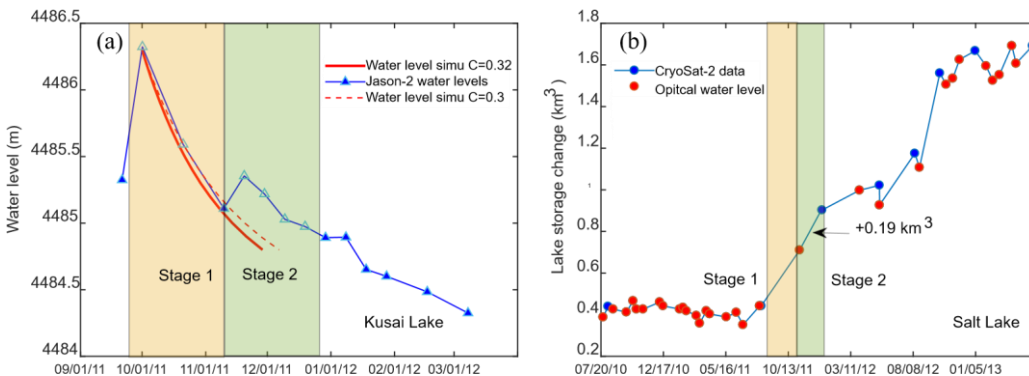



Figure 19. (a) Changes in the water level of Lake Kusai after receiving the outburst flood from Lake Zhuonai. Water level in stage 1 was simulated using Eq 17, which provided a referencing range of parameter C. Water level in stage 2 provided water level input for Eq 17 to calculate total outflow, which was compared with the concurrent water gain of Lake Salt downstream; and (b) changes in water storage of Lake Salt derived from remote sensing using our developed method. There was 0.19 km³ of water gained in stage 15 2, which was comparable to the outflow estimate of Lake Kusai (0.22 km³).

In stage 2 (Nov 9–Dec 31 in 2011) shown in Figure 19 (a), a temporary increase in water level occurred in Lake Kusai, implying that the overflow is not the only driving factor in stage 2. Nevertheless, stage 2 can provide water level input for Eq 17; thus the total outflow can be simulated using parameter C determined in stage 1. Since Lake Salt downstream mainly relied on the replenishment of Lake Kusai during that period, with little precipitation input and negligible glacier melt water in winter,

20 the outflow of Lake Kusai can be comparable with the water gain in Lake Salt derived from remote sensing, though there was a small amount of evaporation loss. This relationship can provide a straightforward validation of our developed method. However, it was not available in stage 1, because the outflow of Lake Kusai first replenished Lake Haidingnuoer until the later began overflowing. The total outflow from Lake Kusai in stage 2 was calculated to be 0.21–0.22 km³, whereas the water gain in Lake Salt was 0.19 km³. This result showed that our densified lake water level time series from multiple optical and altimetric

missions and the developed method are valuable in monitoring and predicting the outflow flood risk that is crucial for the safety of downstream residents and infrastructure.

6 Data availability

The derived TP lake water levels, hypsometric curves, and water storage changes are archived and available at
5 https://doi.org/10.1594/PANGAEA.898411. (Li et al., 2019)

7 Conclusion

In this study, we generated a dense (monthly and ever higher such as 10 days on average) continuous 18-year data set of changes in lake water level and storage for 52 large lakes on the TP by combining multisource optical and altimetric information. The optical data served as a unique reference covering the entire study period, enabling a more consistent
10 conjunction of multisource altimetry time series. By comparison with a widely used LEGOS Hydroweb data set, we showed that without such a reference data set, there may be a remaining bias in the combined altimetry water levels. Our study has considerably improved the temporal resolution of the monitoring of TP lake water level and storage changes. For most lakes examined in the published studies, to our best knowledge, the estimates from our study provided the densest observations that can better reveal the interannual and intra-annual variability and trends in lake water level and storage, even in some relatively
15 small TP lakes whose annual trends may, however, be incorrectly estimated by sparse sampling of lake water levels. The densified data set can also facilitate the monitoring of some rapidly expanding lakes with overflow risks and provide important information on flood prediction and early warning.

We evaluated the uncertainty in the optical water levels by field experiments and rigorous uncertainty analysis. Both methods are consistent that the magnitude of the uncertainty is around 0.1 m, which suggests that optical water levels are often more
20 efficient and less noisy than altimetry data when the altimeter footprints on the lake surface are insufficient, especially for small lakes. Based on our estimates, 52 large TP lakes accounting for ~60% of the total TP lake area have gained 98.3±2.1 km^3 of water during the past 18 years. Lakes in the endorheic basin on the TP were mostly expanding. Water loss was more likely to be found among the southern TP lakes. In the Selin Co basin, a more complicated spatial pattern of lake storage changes was detected, as small lakes were slowly losing water whereas the large lake was gaining water, which we speculated
25 to be caused by lake-river interactions that need further investigation. Note that the quality of the optical water levels before 2002 may not be as good as those obtained after 2002, because no altimetry data before 2002 were used in this study. Extrapolation of the lake shore change-water level relationship may not be stable if the water level during 2000–2001 was much lower or higher than those from 2002–2017.

8 Author contribution

LD and LX designed the research. LX, LD, HQ, and ZF developed the approaches and the data set. LX, HQ, HP, and LD carried out the field experiment. LX, LD, and YW contributed to the analysis of results and writing of the paper.

9 Competing interests

5   The authors declare that they have no conflict of interests.

10 Acknowledgements

This study was supported by the National Nature Science Foundation of China (Grant Nos. 91547210 and 51722903). Mr. Mingda Du's assistance in the field experiments, discussion on the use of satellite altimetry data for lake monitoring with Dr. Gang Qiao from Tongji University, and efforts in improving and archiving the data set made by Dr. Daniela Ransby from
10   PANGAEA data publisher are acknowledged here.

References

Benveniste, J., Baker, S., Bombaci, O., Zeli, C., Venditti, P., Zanife, O., Soussi, B., Dumont, J., Stum, J., and Milagro-Perezin, M. P.: Envisat RA-2/MWR Product Handbook, Frascati, Italy: Eur. Space Agency, 2002. 2002.

Birkett, C. M. and Beckley, B.: Investigating the performance of the Jason-2/OSTM radar altimeter over lakes and reservoirs,
15   Marine Geodesy, 33, 204-238, 2010.

Bouzinac,   C.:   CryoSat   product   handbook,   ESA,   UCL,   URL   https://earth.   esa. int/documents/10174/125272/CryoSat_Product_Handbook, 2012. 2012.

Cheng, K.-C., Kuo, C.-Y., Tseng, H.-Z., Yi, Y., and Shum, C.: Lake surface height calibration of Jason-1 and Jason-2 over the Great Lakes, Marine Geodesy, 33, 186-203, 2010.

20   Crétaux, J.-F., Abarca-del-Río, R., Berge-Nguyen, M., Arsen, A., Drolon, V., Clos, G., and Maisongrande, P.: Lake volume monitoring from space, Surveys in Geophysics, 37, 269-305, 2016.

Crétaux, J.-F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.-C., Nino, F., Del Rio, R. A., and Cazenave, A.: SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data, Advances in space research, 47, 1497-1507, 2011a.

25   Crétaux, J., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Nino, G., Del Rio, R., Cazenave, A., and Maisongrande, P.: Hydrolare/Hydroweb: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data, Advances in Space Research, 47, 1497-1507, 2011b.

Field, C. B., Barros, V. R., Dokken, D., Mach, K., Mastrandrea, M., Bilir, T., Chatterjee, M., Ebi, K., Estrada, Y., and Genova, R.: IPCC, 2014: Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution
30   of working group II to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment, 202, 18-27, 2017.

Guo, J., Chang, X., Gao, Y., Sun, J., and Hwang, C.: Lake level variations monitored with satellite altimetry waveform
35   retracking, IEEE journal of selected topics in applied earth observations and remote sensing, 2, 80-86, 2009.

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.

Jain, M., Andersen, O. B., Dall, J., and Stenseng, L.: Sea surface height determination in the Arctic using Cryosat-2 SAR data from primary peak empirical retrackers, Advances in Space Research, 55, 40-50, 2015.

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.

Kittler, J. and Illingworth, J.: On threshold selection using clustering criteria, IEEE transactions on systems, man, and cybernetics, 1985. 652-655, 1985.

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, Journal of Hydrology, 296, 1-22, 2004.

Lei, Y., Yao, T., Yang, K., Sheng, Y., Kleinherenbrink, M., Yi, S., Bird, B. W., Zhang, X., Zhu, L., and Zhang, G.: Lake seasonality across the Tibetan Plateau and their varying relationship with regional mass changes and local hydrology, Geophysical Research Letters, 44, 892-900, 2017.

Li, B., Zhang, J., Yu, Z., Liang, Z., Chen, L., and Acharya, K.: Climate change driven water budget dynamics of a Tibetan inland lake, Global and Planetary Change, 150, 70-80, 2017a.

Li, H., Qiao, G., Wu, Y., Cao, Y., and Mi, H.: Water Level Monitoring on Tibetan Lakes Based on Icesat and Envisat Data Series, The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 1529, 2017b.

Li, X., Long, D., Huang, Q., Han, P., Zhao, F., and Wada, Y.: A high temporal resolution lake data set from multisource altimetric missions and Landsat archives of water level and storage changes on the Tibetan Plateau during 2000-2017, https://doi.org/10.1594/PANGAEA.898411. PANGAEA, 2019.

Liu, B., Lin, L. I., Yue, D. U., Liang, T., Duan, S., Hou, F., and Ren, J.: Causes of the outburst of Zonag Lake in Hoh Xil,Tibetan Plateau,and its impact on surrounding environment, Journal of Glaciology & Geocryology, 2016. 2016.

Markham, B. L., Storey, J. C., Williams, D. L., and Irons, J. R.: Landsat sensor performance: history and current status, IEEE Transactions on Geoscience and Remote Sensing, 42, 2691-2694, 2004.

McFeeters, S. K.: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, International journal of remote sensing, 17, 1425-1432, 1996.

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.

Nielsen, K., Stenseng, L., Andersen, O. B., and Knudsen, P.: The performance and potentials of the CryoSat-2 SAR and SARIn modes for lake level estimation, Water, 9, 374, 2017.

Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418, 2016.

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.

Strahler, A. N.: Hypsometric (area-altitude) analysis of erosional topography, Geological Society of America Bulletin, 63, 1117-1142, 1952.

Team, P.: Planet Application Program Interface: In Space for Life on Earth, San Francisco, CA, 2017. 2017.

Wan, W., Long, D., Hong, Y., Ma, Y., Yuan, Y., Xiao, P., Duan, H., Han, Z., and Gu, X.: A lake data set for the Tibetan Plateau from the 1960s, 2005, and 2014, Scientific data, 3, 160039, 2016.

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.

Willgoose, G. and Hancock, G.: Revisiting the hypsometric curve as an indicator of form and process in transport-limited catchment, Earth Surface Processes and Landforms: The Journal of the British Geomorphological Group, 23, 611-623, 1998.

Xiaojun, Y., Shiyin, L., and Meiping, S.: Changes of Kusai Lake in Hoh Xil region and causes of its water overflowing, Acta Geographica Sinica, 67, 689-698, 2012.

Xu, H.: A study on information extraction of water body with the modified normalized difference water index (MNDWI), JOURNAL OF REMOTE SENSING-BEIJING-, 9, 595, 2005.

Yang, K., Yao, F., Wang, J., Luo, J., Shen, Z., Wang, C., and Song, C.: Recent dynamics of alpine lakes on the endorheic Changtang Plateau from multi-mission satellite data, Journal of Hydrology, 552, 633-645, 2017.

Yao, F., Wang, J., Yang, K., Wang, C., Walter, B. A., and Crétaux, J.-F.: High resolution data set of annual lake areas and water storage across the Inner Tibet, 2002-2015. In: Supplement to: Yao, F et al. (2018): Lake storage variation on the endorheic Tibetan Plateau and its attribution to climate change since the new millennium. Environmental Research Letters, 13(6), 064011, https://doi.org/10.1088/1748-9326/aab5d3, PANGAEA, 2018a.

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, 2018b.

Yu, S., Liu, J., Xu, J., and Wang, H.: Evaporation and energy balance estimates over a large inland lake in the Tibet-Himalaya, Environmental Earth Sciences, 64, 1169-1176, 2011.

Zhang, G., Xie, H., Kang, S., Yi, D., and Ackley, S. F.: Monitoring lake level changes on the Tibetan Plateau using ICESat altimetry data (2003–2009), Remote Sensing of Environment, 115, 1733-1742, 2011.

Zhang, G., Yao, T., Shum, C., Yi, S., Yang, K., Xie, H., Feng, W., Bolch, T., Wang, L., and Behrangi, A.: Lake volume and groundwater storage variations in Tibetan Plateau's endorheic basin, Geophysical Research Letters, 44, 5550-5560, 2017a.

Zhang, G., Zheng, G., Gao, Y., Xiang, Y., Lei, Y., and Li, J.: Automated water classification in the Tibetan plateau using Chinese GF-1 WFV data, Photogrammetric Engineering & Remote Sensing, 83, 509-519, 2017b.

Zhang, Y., Liu, C., Tang, Y., and Yang, Y.: Trends in pan evaporation and reference and actual evapotranspiration across the Tibetan Plateau. J Geophys Res 112:D12110, Journal of Geophysical Research Atmospheres, 112, -, 2007.

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.