https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Open Access Earth System Science Data Discussions

High-resolution datasets for lake level changes in the Qinghai-Tibetan Plateau from 2002 to 2021 using multi-altimeter data

Jiaming Chen[1,3], Jingjuan Liao[1, 2*,] Yanhan Lou[1,4], Shanmu Ma [1,4], Guozhuang Shen[1,2], Lianchong Zhang[1]

[1]Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100094, China.
[2] International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China.
[3] Institute of Geodesy and Geoinformation, University of Bonn, Bonn, Germany.
[4] University of Chinese Academy of Sciences, Beijing 100049, China.

Correspondence: Jingjuan Liao (liaojj@aircas.ac.cn)

Abstract.

The Qinghai Tibet Plateau (QTP), known as the Roof of the World and the Water Tower of Asia, has the largest number of lakes in the world, and because of its high altitude and near absence of disturbances by human activity, the plateau has long been an important site for studying global climate change. Hydrological stations cannot be readily set up in this region, and *in situ* gauge data are not always publicly accessible. Satellite radar altimetry has become a very important alternative to *in situ* observations as a source of data. Estimation of the water levels of lakes via radar altimetry is often limited by temporal and spatial coverage, and, therefore, multi-altimeter data are often used to monitor lake levels. Restricted by the accuracy of waveform processing and the interval period between different altimetry missions, the accuracy and the sampling frequency of the water level series are typically low. By processing and merging data from eight different altimetry missions, the developed datasets provided the water level changes for 362 lakes (larger than 10 km^2) in the QTP from 2002 to 2021. The period for the lake level change series, which affords high accuracy, can be much longer for many lake systems. The present datasets and associated approaches are valuable for calculating the changes in lake storage, trend analyses of the lake levels, short-term monitoring of the overflow of lakes, flooding disasters on the plateau, and the relationships between changes in the lake ecosystems and changes in the water resources.

1 Introduction

As primary water reservoirs, lakes not only play an important role in the supply and adjustment of surface water but also reflect the impact of climate change and human activities on regional and global environmental change (Adrian et al., 2009; Schindler, 2009; Song et al., 2015). The water level of lakes is a key indicator for regional climate change and human disturbance. Generally, it is assumed that the changes in lake bottoms are very slight over decades, so understanding the changes in lake levels can help to evaluate the impact of climate change and human activities on regional water resources.

Observation by use of a water gauge is the traditional method to measure the changes in water levels in lakes; *in situ* gauge measurement of lakes can afford high precision but such equipment is expensive to maintain and challenging to operate in remote areas. Furthermore, the total number of monitoring stations has decreased in recent years (Frappart et al., 2006; Kleinherenbrink et al., 2014), and lake level data in many countries and regions are not freely available to the public. Alternatively, satellite altimetry technology is an effective tool that can be used to measure the dynamics of the surface elevation of the Earth and has been successful in measuring lake levels. The Qinghai-Tibetan Plateau (QTP), known as the Roof of the World and the Water Tower of Asia, has numerous and some of the largest natural lakes in the world, and because of its high altitude and the near absence of human disturbances, the plateau is an important location for studying global change. Changes in the water level in lakes are one of the important indicators for the water balance of the QTP and these are directly affected by temperature, precipitation, evaporation, glaciers, perennial snow cover, and permafrost (Zhang et al., 2012; 2013a;

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

2013b). The QTP is the source of many major rivers, and more than 1.4 billion people depend on water resources from the plateau (Pritchard, 2017). However, due to the vastness and remoteness, it is a challenge to set up *in situ* monitoring stations. There are only a few lakes (such as Qinghai Lake, Namtso, and Yamdrok Yumtso) with *in situ* gauge stations for lake level measurements (Zhang, 2018). Most lakes in the QTP lack such a measurement capability making it difficult to understand the long-term spatial and temporal characteristics regarding the evolution and dynamics of the water levels of the lakes.

Satellite altimetry has become the most important means to measure lake levels and their changes in the plateau. Numerous studies have focused on the use of satellite altimeters for measuring changes in lake levels in the QTP. For example, Gao et al. (2013) employed multi-altimeter data from Envisat, CryoSat-2, Jason-1, and Jason-2 to examine water level changes at 51 lakes between 2002 and 2012 in the OTP. Zhang et al. (2011) used Ice, Cloud, and the land Elevation Satellite (ICESat) data to determine changes in lake levels in Tibet from 2003 to 2009. Hwang et al.(2016) obtained two decades of lake level measurements at 23 lakes in the QTP from the T/P-family altimeters. Song et al. (2015) combined ICESat-1 and Cryosat-2 altimetry data to access the water level dynamics of Tibetan lakes from 2003 to 2014. Kleinherenbrink et al. (2015) and Jiang et al. (2017) used the CryoSat-2 data to measure changes in the water levels at 125 lakes and 70 lakes in the QTP, respectively. Hwang et al. (2019) constructed a lake level time series for 61 lakes on the Tibetan Plateau between 2003 and 2016 and discussed the trends of the time series. Li et al. (2019) constructed high-temporal-resolution water level datasets for 52 large lakes on the Tibetan Plateau. These studies in the QTP reveal that estimation of the levels of lake water with a given radar altimeter is often limited by temporal and spatial coverage, and, therefore, multiple altimeters are needed to obtain multiple decades of changes in the water levels of lakes. However, due to the large size of the radar altimeter footprint and contaminations from the steep lakeshore or surrounding land, the observations of lake levels via satellites are noisy, and it is difficult to obtain the distance from the altimeter to the nadir points. Therefore, waveform retracking processing may be used to remove the contamination by land signals when lake levels are retrieved from multi-altimeter data. In this study, by combining eight sets of altimeter data from Envisat, ICESat-1, CryoSat-2, Jason-1, Jason-2, Jason-3, SARAL, and Sentinel-3A, the trends of the changes in the water levels for 362 lakes (>10 km^2) in the QTP during 2002-2021 were estimated using retracking and outlier detection algorithms.

The primary objective of this study was to determine the changes in the water levels of 362 lakes in the QTP from multi-altimeters and evaluate the accuracy of the time series and the performance of the multi-altimeter data with respect to monitoring the long-term variations in the water levels of the lakes. Readers can access the dataset described in this paper at https://doi.org/10.1594/PANGAEA.939427 (Chen et al., 2021).

2 Study area and data

2.1 Study area

The QTP is in the southwest of China and covers about 27% of the total area of China (Zhang et al., 2002). There are more than 1000 lakes of >1 km^2 (Wan et al., 2016) in the QTP, most of which belong to inland drainage systems. Based on coverage by altimeter data, 364 lakes of >10 km^2 in the QTP were selected as the objects of study. Among these lakes, there were 13 lakes of > 500 km^2, 79 lakes of 100-500 km^2, 69 lakes of 50-100 km^2, and 201 lakes of 10-50 km^2. Most of these lakes are inland lakes with surface runoff, precipitation, snow and ice melting, springs, and underground runoff as their main sources of water recharge. Due to minimal impact by human activity, changes in the water levels in the lakes in the region are driven mainly by natural factors such as precipitation and temperature, which are important indicators of changes in the regional climate and the ecological environment.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

2.2 Data

2.2.1 Multi-altimeter data

Eight sets of altimeter data from Envisat, ICESat-1, CryoSat-2, Jason-1, Jason-2, Jason-3, SARAL, and Sentinel-3A were used to extract the water levels of the lakes in the QTP to obtain the lake level time series with high-space coverage. The details of the multi-altimeter data are given in Table 1. Envisat, CryroSat-2, and Sentinel-3A data provided by the European Space Agency (ESA) were available for 121, 353, and 107, lakes, respectively; Jason-1, Jason-2, and Jason-3 data provided by the Centre National d'Etudes Spatiales (CNES) were available for 48, 71 and 28 lakes, respectively, due to the relatively sparse ground tracks. ICESat-1 data provided by the National Aeronautics and Space Administration (NASA) were available for 124 lakes and afforded high spatial resolution. SARAL is a joint mission of the Indian Space Research Organization (ISRO) and CNES and is a continuation of the Envisat mission. SARAL data were available for 135 lakes in the QTP.

Table 1 Details of the multi-altimeter data used in this study

Mission	Sensor	Duration	No. of lakes	Diameter of footprint (km)
Envisat	RA-2	2002.05-2012.04	121	1.7
ICESat-1	GLAS	2003.02-2009.10	124	0.07
CryoSat-2	SIRAL	2010.07-2021.07	353	1.6 (across), 0.3 (along)
Jason-1	Poseidon-2	2002.01-2012.03	48	2.2
Jason-2	Poseidon-3	2009.12-2017.05	71	2.2
Jason-3	Poseidon-3B	2016.02-2020.12	28	1
SARAL	Altika	2013.03-2016.05	135	1
Sentinel-3A	SRAL	2016.03-2019.09	107	2 (across), 0.25 (along)

In addition, a dataset on the shapes of the lakes generated by Wan et al. (2016) was selected to determine whether the altimeter data encompassed the lakes, and a buffer of 1 km around the shape of the lake was generated to determine the change in the boundary of the lakes during the past 20 years.

2.2.2 *In situ* data

In situ data on eight lakes were used to validate reliable information on the lake level time series from the multi-altimeter data. Table 2 lists details of the *in situ* data on the eight lakes. The *in situ* data for Qinghai Lake and Ngoring Lake were from the Hydrology and Water Resources Survey Bureau in Qinghai Province and from the Yellow River Commission of the Ministry of Water Resources, respectively, and the *in situ* data on Bamco, Dagzeco, Dawaco, Namco, Pungco and Zhari Namco were from the Institute of Tibetan Plateau Research, Chinese Academy of Sciences (Lei, 2018; Wang, 2018).

Table 2 Details of the *in situ* data for eight lakes as used for validation

Lake name	Date	Coordinates	Reference	Mode
Qinghai Lake	2010.05-2019.09	100.20, 36.89	1985*	Absolution
Ngoring Lake	2010.01-2015.12	97.70, 34.90	1985	Absolution
Bamco	2013.06-2017.10	90.58, 31.27	Customize	Relative
Dagzeco	2013.06-2016.10	87.52, 31.89	Customize	Relative
Dawaco	2013.06-2016.10	84.96, 31.24	Customize	Relative
Namco	2007.04-2016.12	90.60, 30.74	Customize	Relative
Pungco	2014.05-2017.10	90.97, 31.50	Customize	Relative

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Open Access Earth System Science Data Discussions

Zhari Namco	2012.12-2017.10	85.61, 30.93	Customize	Relative

* 1985 indicates the 1985 national elevation benchmarks

3 Methods

3.1 Extraction of lake water levels

With respect to the extraction of the water level data from the satellite altimetry, there is uncertainty as to whether there is a valid footprint falling on the lakes; this problem can be addressed by comparing the geographic coordinates of the footprints with the shape of the dynamic dataset for the lake. However, it would take considerable time to extract the dynamic shape file. A static shape dataset for the Tibetan Plateau was used in this study (Wan et al., 2016); we also generated a 1 km buffer for the shape to solve the situation regarding the changes in the boundary of lakes during the past 20 years. After picking out the available footprints, the height of the lake surface height can be calculated based on using Eq. (1) for each footprint:

$$H = Alt - \left(R_{range} + \Delta R_m + \Delta R_{dry} + \Delta R_{wet} + \Delta R_{iono} + \Delta R_{tide} + \Delta R_{correction}\right) - N_{geoid} \quad (1)$$

where Alt is the satellite altitude, R_{range} is the distance between the altimeter and the lake surface, ΔR_m is the satellite centroid correction, ΔR_{dry} is the dry troposphere, ΔR_{wet} is the wet troposphere, ΔR_{iono} is the ionospheric correction, ΔR_{tide} includes the solid earth tide, the pole tide, and the ocean tide corrections, N_{geoid} is the geoid height with respect to the ellipsoid, for which the 2008 Earth Gravitational Model (EGM2008) was used in this study (Pavlis et al., 2012), and $\Delta R_{correction}$ stands for the retracking correction $\Delta R_{retrack}$ for radar altimetry and the saturation correction $\Delta R_{saturation}$ for the laser altimetry. With the exception for $\Delta R_{retrack}$, all the corrections above are included in the altimetry data product.

3.1.1 Waveform retracking

The waveforms for inland water bodies are easily contaminated or even submerged by signals from land, hence it is difficult to obtain the distance from the altimeter to the nadir points. Accordingly, the retracking correction plays an important role in removing the contamination by land signals when radar altimetry data are applied to inland water bodies (Martin et al., 1983; Lee et al., 2008). In this study, the automatic multiscale-based peak detection retracker (AMPDR) (Chen et al., 2020), which is suitable for different altimetry and different situations to derive reliable water levels, was used on the altimeter data product. The AMPDR showed excellent performance for the Jason-2/3, Sentinel-3, and Cryosat-2, but sometimes there were biases for the retracking correction caused by the hooking effect or the scatter signal of the off-nadir point for Jason-1, Envisat, and SARAL. Therefore, some modifications for AMPDR were adopted for Jason-1, Envisat, and SARAL data in this study.

To ensure that the different typologies of multi-waveforms can be dealt with, we implemented a two-step process for the modified AMPDR here. The steps of the modified retracker are illustrated in Fig. 1. The optimal retracked range was determined using several criteria:

(1) The optimal retracked levels should be within the range $H_{DEM} \pm 20$ m.
(2) The DistanceThresh in AMPDR produced the smallest difference for the median of the water levels derived from the neighboring cycles if time has continuity.
(3) The standard deviation of the water level of the current cycle was decreased if the data were not continuous; that is to say, the difference between the neighboring cycle and the current cycle was more than ten days or several months.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

135

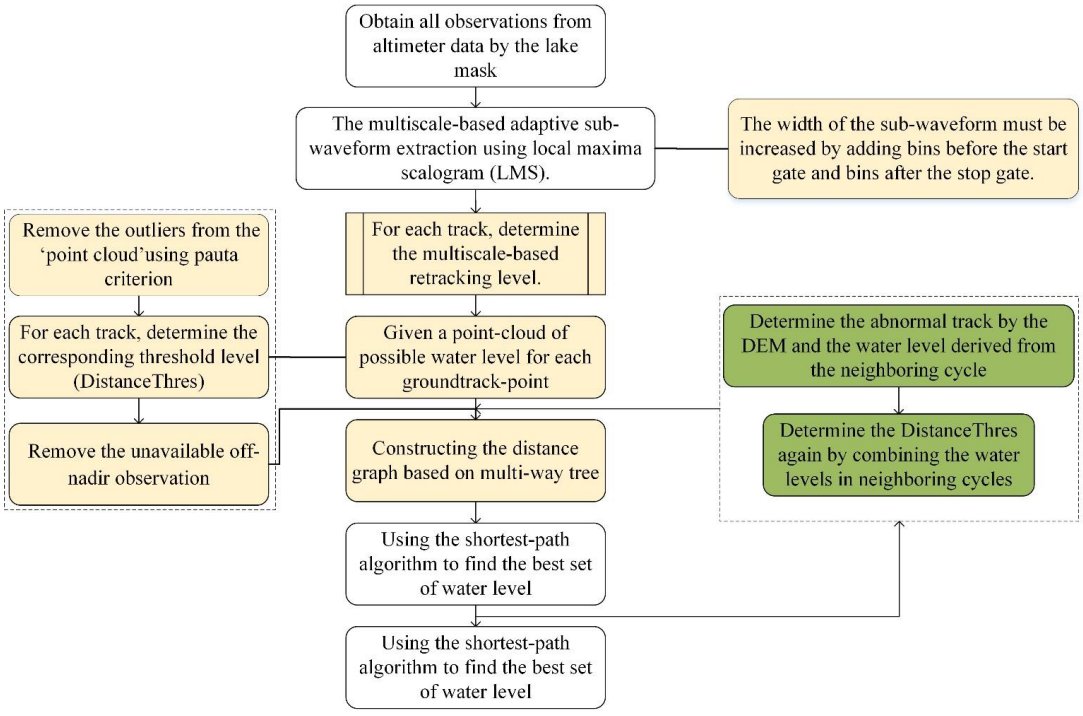


Fig. 1 Flowchart outlining the waveform retracking process. Steps with a yellow background are the preparation steps for using the shortest path algorithm. Steps with a green background are the retracking for the abnormal track by the selected DEM.

140

In the first run, the normal operation of the AMPDR was considered, and the lake level time series was calculated. Details regarding the definition and implementation of AMPDR are available elsewhere (Chen et al., 2020). Next, a second run of the retracking for the abnormal track, which was selected by the Digital Elevation Model (DEM) and the water level derived from the neighboring cycle, was implemented. However, this time the DistanceThresh in AMPDR was constructed by one of three
145 minimum second-order difference quotients of the cumulative distribution function (CDF) of the rounded water levels. In this way, it was ensured that the DistanceThresh was approached by the median of the water levels in neighboring cycles. Additionally, the retracking point from the ICE-1 algorithm was added to the construction of the "point cloud" and the CDF given that the multiscale-based adaptive threshold retracking would fail in some situations. An example of the operation of the modified two-step retracker is shown in Fig. 2.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

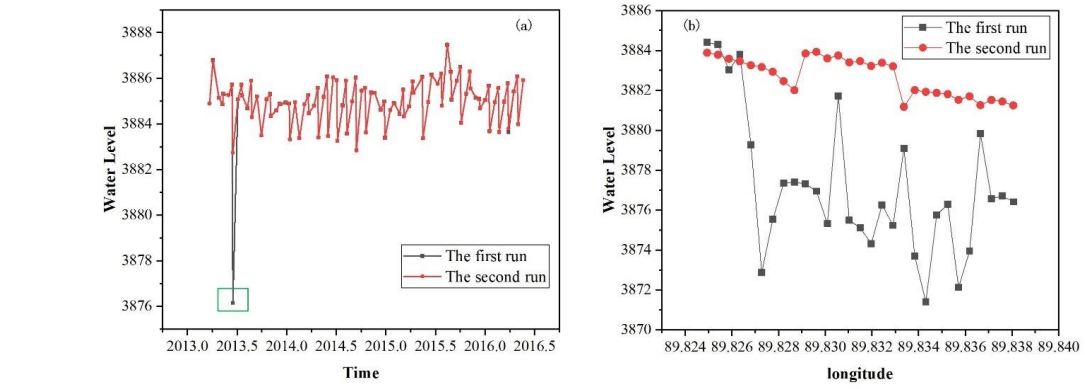


Fig. 2 An example of the operation of the modified two-step retracker. (a) shows the two water level time series for processing by the two-step retracker. (b) shows the along-track water level in the green rectangle from (a) when processing by the two-step retracker.

3.1.2 Removal of noise footprints

Due to the use of a 1 km buffer to pick out the shape of the available footprints, there would be many noise footprints caused by the reflected signals of the terrain or by the scatter signals of the off-nadir points. The noise footprints should be removed before constructing the lake level time series. Waveform classification is an effective method for identifying the noise footprints. Studies have proposed the use of various waveform classification methods and good recognition results have been achieved (Göttl et al., 2016; Lee et al., 2016; Marshall and Deng, 2016; Shen et al., 2017).

Different from the previous study whereby the waveforms are divided into multiple classes, this study only needs to divide the waveforms into noise and non-noise waveforms using a random forest (RF) classifier. The RF classifier was set up using a training set of approximately 300 waveforms over inland lakes for each altimetry, Additionally, the following features of the waveforms were selected: the pulse peakiness (Strawbridge and Laxon, 1994), the mean value of the waveform, the skewness of the waveform, the kurtosis of the waveform, the amplitude of the waveform, the width of the waveform determined by the Offset Center of Gravity (OCOG) retracker (Bamber 1994), the bin position corresponding to the center of gravity determined by the OCOG retracker, and the peakiness of the left and right pulse (Ricker et al., 2014). After discarding these noise footprints, the tracks with fewer than five observations were excluded from this study.

3.1.3 Construction of time series

Despite removing the noise footprints using waveform classification, the dataset also has outliers in the lake level time series for each cycle of a certain altimeter. Therefore, any point level in each cycle yielding a difference larger than three times the standard deviation (3σ rule) was removed. Then, the lake level time series was estimated using the R package tsHydro (https://github.com/cavios/tshydro). The core of tsHydro is a state-space model consisting of a process model and an observation model.

$$H_i^{true} = H_{i-1}^{true} + \sqrt{t_i - t_{i-1}}\sigma_{RW} z_i, \qquad z_i \sim N(0,1) \tag{2}$$

$$H_{ij}^{obs} = H_i^{true} + \sigma_{obs}\varepsilon_{ij} \tag{3}$$

The process model is used to describe the relationship between the true water level $H^{(true)}$, and the observation model is described by the observed water level $H^{(obs)}$, with an error term ε_{ij}, being used to describe the relationship between $H^{(obs)}$ and $H^{(true)}$.The model is described in detail by Nielsen et al. (2015). According to the Laplace estimation, the mean value of

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

the range was selected to represent the water level of the lake for each cycle. Meanwhile, the standard deviation of each cycle was reserved to evaluate the uncertainty of the time series.

3.2 Fusion of multi-altimeter time series

It is not uncommon that the reference plane between different altimeters should be different. Before merging the lake level from different altimeters, the reference plane should be unified as WGS84/EGM 2008. The reference system of Jason-1/2/3 is the Topex/Poseidon (T/P) ellipsoid system instead of the WGS84 system, thus it was necessary to perform an ellipsoid system transformation from T/P to WGS84 by subtracting 0.71 m from the vertical height (Bhang et al., 2007).

Due to the variations in orbits and the disparities between instruments, systematic biases existed among the lake level time series extracted from the multi-altimetry, although they were corrected to the same reference system. In most studies (Li et al., 2019; Gao et al., 2013; Huang et al., 2016), the altimeters with the longest overlap period would be merged for the first time, but there may be some special situations whereby for some lakes the lake level time series for each altimeter cannot be merged. In this study, the dynamic reference time series was used to merge the lake-level time series. We first merged the two products with the longest period for the time series and chose the altimeter-derived water level with the longer time series as the baseline. We then adjusted the time series from another altimeter by subtracting the discrepancy compared with the reference series (Lee et at., 2011; Kropáček et al., 2012) according to Eq. (4). Then, a similar process was applied to the remaining products and the merged products connecting the three altimeters. The result for the merged altimetry data when all sensors are available is shown in Fig. 3.

$$Series2_{cor}(t_i) = Series2_{ini}(t_i) + \left(\overline{Series1_{ref}} - \overline{Series2_{ini}}\right) \quad (4)$$

where $Series2_{ini}(t_i)$ is the uncorrected lake level at time t_i, $\overline{Series1_{ref}}$ is the mean value of the water level time series from the baseline, and $\overline{Series2_{ini}}$ is the mean value of the other water level time series at the same time as $\overline{Series1_{ref}}$.

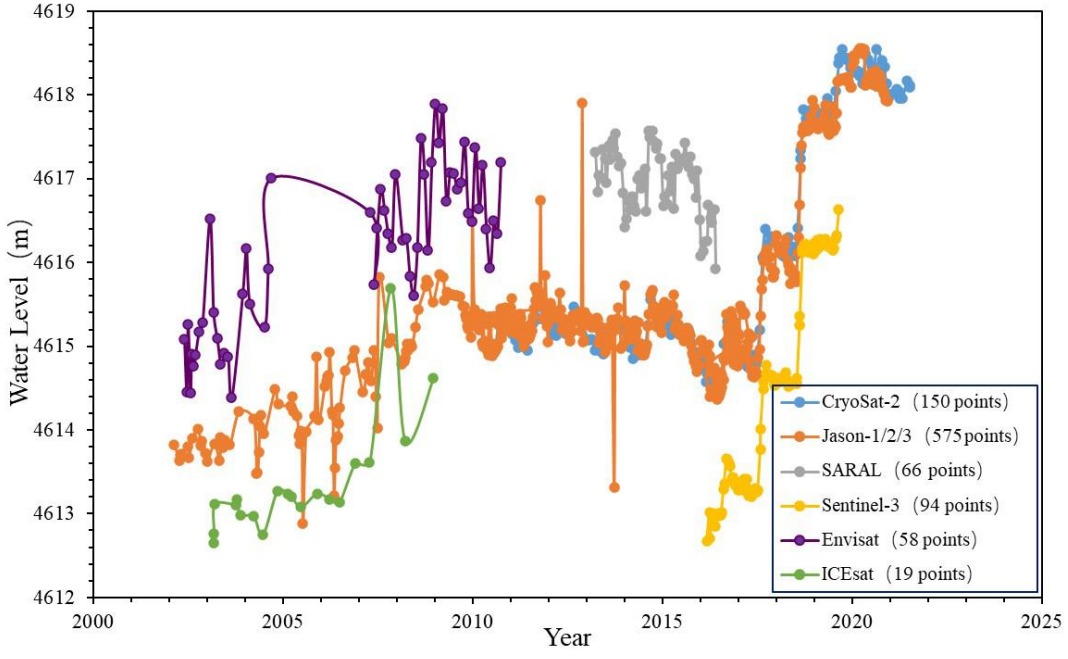


Fig. 3 The process of merging multi-altimetry in Zhari Namco which contains six altimeters.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Nevertheless, not all the lake-level time series can be merged successfully following the steps outlined above. For instance, Cuona Lake, Xiasa'er Co, and Bei Hulsan Lake cannot be merged successfully because only ICESat and Cryosat-2 were available on these lakes before 2013, while there is no overlap period between ICESat and Cryosat-2. In this study, 18 lakes were found to have similar problems.

A combined linear-periodic-residual model (Liao et al., 2014) was used to simulate and forecast the lake-level time series in the no-overlap period to merge the two altimeters with no overlap period. Numerous studies (Medina et al., 2008; Irvine et al., 1992; Kropáček et al., 2012; Lee et al., 2011) have indicated that the changes in the lake-level exhibited a clear linear trend and inter-periodic fluctuations at some scales such as 10 or 20 years in line with Eq. (5).

$$x_i = a + bt + \sum_{i=1}^{p}\left(\alpha_i cos\frac{2\pi}{T_i}t + \beta_i sin\frac{2\pi}{T_i}t\right) + \varepsilon_t \quad (5)$$

where a and b are the linear components of the lake-level time series, T_i indicates the ith periodic component, and ε_t is the remaining random component after removal of the linear and periodic components.

A result for the merged altimetry data of Cuona Lake is presented in Fig. 4. First, singular spectrum analysis (SSA) algorithms are used to reduce the noise of the lake-level time series and to extract the effective fluctuating signal. Second, we decomposed the fluctuating signal into a linear component, a periodic component, and the remaining residuals using a simple linear fitting, wavelet analysis; then simple regression analysis, trigonometric function fitting, and the autoregressive-moving-average (ARMA) model were used to fit each component, respectively. Finally, we combined the modeling data of each component and obtain the simulated water level. The diagram for fusion processing is shown in Fig. 5.

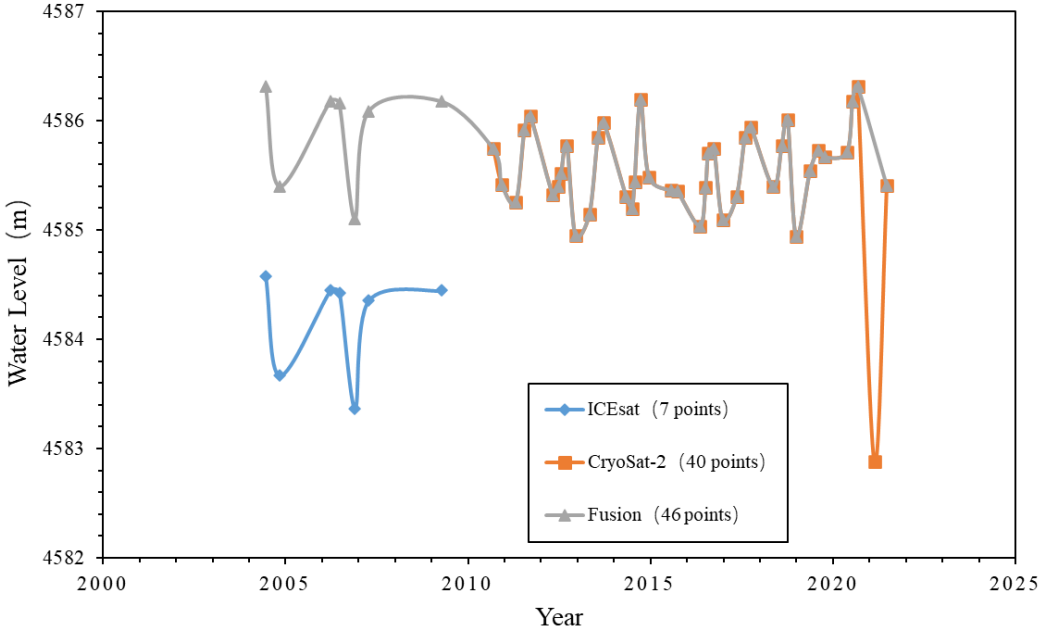


Fig. 4 The process of merging multi-altimetry in Cuona Lake.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

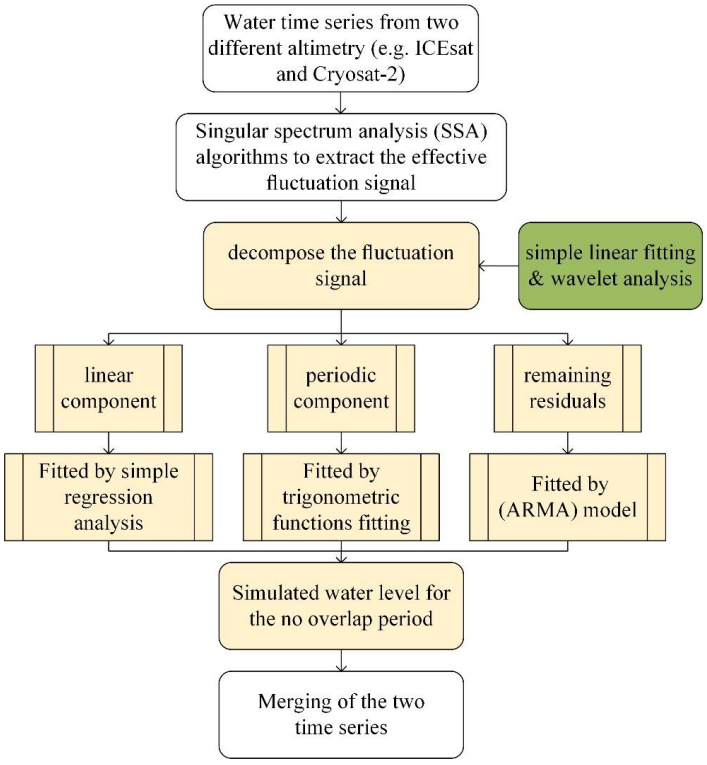


Fig. 5 Flowchart of fusion processing for the water level time series from different altimeters. Steps with a yellow background indicate preparation for merging the time series.

4 Validation of data quality

4.1 Validation and accuracy of lake level time series

Due to the lack of *in situ* data for the water levels of lakes in the QTP, only *in situ* data for eight lakes were collected to validate the accuracy of the lake level time series. Table 3 gives the statistical results for a comparison between the lake level time series and the in-situ data for the eight lakes. The results show that the accuracy for all eight lakes was less than 0.35 m, and the average accuracy was 0.213 m. Dawaco had the lowest root-mean-square errors (RMSEs) (0.149 m), and Ngoring Lake had the highest RMSEs (0.335 m), indicating that the results of this study are reliable and the accuracy of the time series can reach the decimeter level with respect to the monitoring inland lakes. At the same time, except for Dawaco, the lake levels obtained in this study agreed well with those from the *in situ* gauges, showing a good correlation (the correlation coefficients >0.60). Furthermore, it can be seen from the comparison between the satellite-derived lake levels and the *in situ* water levels for the eight lakes that the satellite-derived lake level series followed the gauged data quite well, especially for Qinghai Lake, Bamco, and Pungco (correlation coefficients >0.90).

Table 3 Comparison between the lake levels in this study and the *in situ* water levels

Lake	Correlation coefficient	RMSE (m)	Number of validation points
Qinghai_Lake	0.977	0.190	570
Ngoring_Lake	0.635	0.335	284
Bamco	0.930	0.181	19

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Dagzeco	0.744	0.199	156
Dawaco	0.209	0.149	7
Namco	0.738	0.179	60
Pungco	0.924	0.222	29
Zhari Namco	0.762	0.251	314

240

4.2 Cross-validation with similar products

We compared our product with the lake level data set provided by the LEGOS Hydroweb. There were 11 lakes that were featured in both studies conducted from 2002 to 2020. The annual trends for the lake levels for 2002-2020 are compared in Fig. 6, and the results indicate that the two products are consistent with each other, and the R^2 of the linear fit for the two
245 products is 0.83.

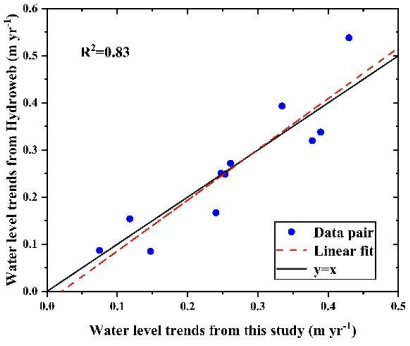


Fig. 6 Cross-validation of the lake levels in the QTP derived from the present study with those provided by the LEGOS Hydroweb database.

4.3 Description of the data set

250 The lake-level change time series for 362 lakes (196 lakes for the time series from 2002 to 2021 and 168 lakes for time the series from 2010 to 2021) are available on the datasets. The water level time series for each lake are archived as 362 entities based on the names of the lakes, with a table describing all the information about each lake. The first part of each file describes the basic information of the lake-level time series, such as the geographic information, the start date of the time series, the end date of the time series, and the number of data points. Next is the main part for each file: the first row stands for the time, the
255 second row records the water level, the third row is the uncertainty of the water level, and the final row stands for the source of the data. It should be noted that the uncertainty of the water level time series was calculated using the standard deviation for the processing in constructing the time series with the "R" package.

5 Applications

5.1 Spatio-temporal analysis of changes in lake levels in the QTP

260 Based on the changes in the water levels of the lakes, the spatial patterns for the trends in the lake levels during 2002-2021 are shown in Fig. 7. Overall, the lake levels in the QTP show a clear rising trend, and the overall average annual rate of change is 0.175 m/a; further, the number of lakes with rising water levels accounts for 78% of all lakes. The total area of lakes with rising water levels (35213 km^2) is much larger than the total area of lakes with falling water levels (6364km^2), indicating that

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

the water storage of lakes on the QTP is growing. From the distribution of the annual average rate of change of lake levels (Fig. 8), among the monitored lakes between 2002 and 2021, there are more lakes with rising water levels than those with falling water levels. Among the lakes with an average annual rate of change greater than 0.20 m/a, the number of lakes with an increasing trend in the water levels is much higher (281 lakes) than the number of lakes with a decreasing trend (81 lakes).

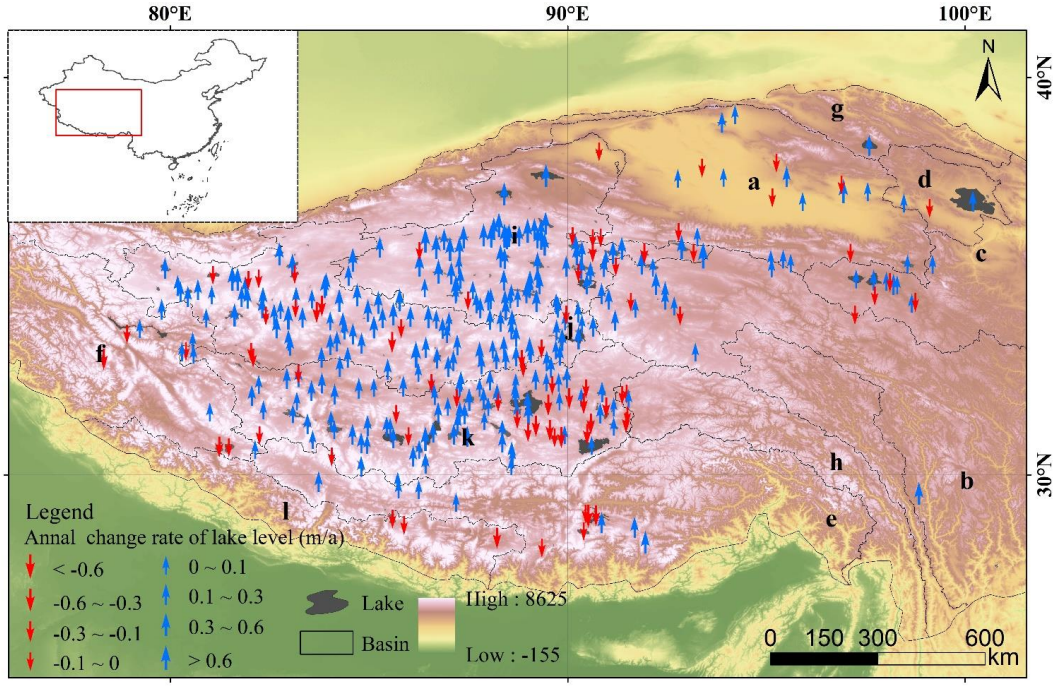


Fig. 7 Spatial distribution of trends in the changes in the water levels of lakes on the QTP during 2002-2021. The black line shows the boundary of the basin of the QTP (refered to Wan et al., 2016). The lowercase letters indicate different basins. The DEM of the base map is from the Global Multi-resolution Terrain Elevation Data 2010（GMTED2010）(GMTED: https://topotools.cr.usgs.gov/gtmed_viewer/)

(a Qaidam; b Yangtze; c Yellow; d Qinghai Lake; e Brahmaputra; f. Indus; g Hexi Corridor; h Salween; i Northern Inner Plateau; j Central Inner Plateau; k Southern Inner Plateau; l Ganges)

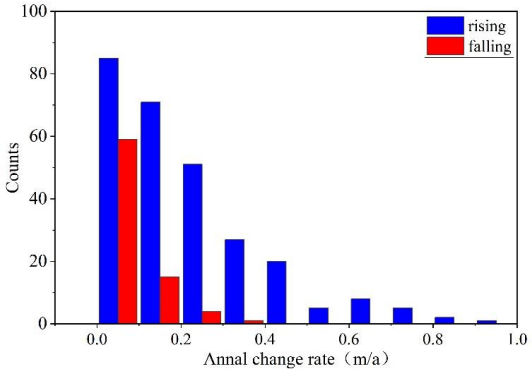


Fig. 8 Histogram of trends in the lake level changes on the QTP during 2002-2021.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Analysis of the trends in the changes in the water levels based on the lake areas shows that there is a clear rising trend in the water level of lakes on the QTP, the most significant trends in the case of rising water levels being for larger-size lakes (>500 km^2) and also for smaller size (<50 km^2) lakes, and intermediate size lakes showing significant rising trend (Table 4).

Table 4 The trends for changes in the lake water levels in the QTP during 2002-2021

Lake area/km^2	No. of lakes	Annual rate of change ($m \cdot a^{-1}$)	No. of lakes with rising water levels	Mean rate of rise ($m \cdot a^{-1}$)	No. of lakes with decreasing water levels	Mean rate of decrease ($m \cdot a^{-1}$)
>500	13	0.167	12	0.198	1	-0.201
[200,500]	31	0.186	24	0.262	7	-0.075
[100,200)	48	0.232	36	0.333	12	-0.069
[50,100)	69	0.157	50	0.243	19	-0.069
[10,50)	201	0.142	159	0.195	42	-0.058

To better understand the spatial distribution pattern of the changes in the water levels of the lakes, the trends for the changes in the water levels of the lakes in each basin of the QTP were analyzed (Table 5). Overall, during the period 2002-2021, the water levels of the lakes in all basins increased significantly, except for the Brahmaputra basin. The area of lakes with rising water levels was larger than that for lakes with decreasing water levels. Amongst all the basins, the lakes with a decreasing water level were distributed mainly in the Brahmaputra, Ganges, and Salween basins (Fig. 9).

Table 5 The trends in the changes in the water levels of the lakes in the different basins of the QTP during 2002-2021

Basin	No. of lakes	No. of lakes with rising water levels	Annual rate of rise (m/a)	Area of lakes with rising water levels (km^2)	No. of lakes with decreasing water levels	Annual rate of fall (m/a)	Area of lakes with decreasing water levels (km^2)
Qaidam	22	13	0.115	986	9	-0.027	685
Yangtze	15	12	0.158	590	3	-0.008	97
Yellow	11	7	0.069	1285	4	-0.019	82
Qinghai Lake	3	2	0.124	4391	1	-0.005	115
Brahmaputra	13	7	0.163	224	6	-0.114	1048
Indus	8	4	0.062	762	4	-0.077	869
Northern Inner Plateau	80	73	0.378	7285	7	-0.079	506
Central Inner Plateau	105	86	0.236	5681	19	-0.050	1013
Southern Inner Plateau	98	75	0.137	13617	23	-0.094	1466

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Salween	3	1	0.003	17	2	-0.008	212
Ganges	3	0	/	/	3	-0.101	335
Hexi Corridor	1	1	0.189	609	0	/	/

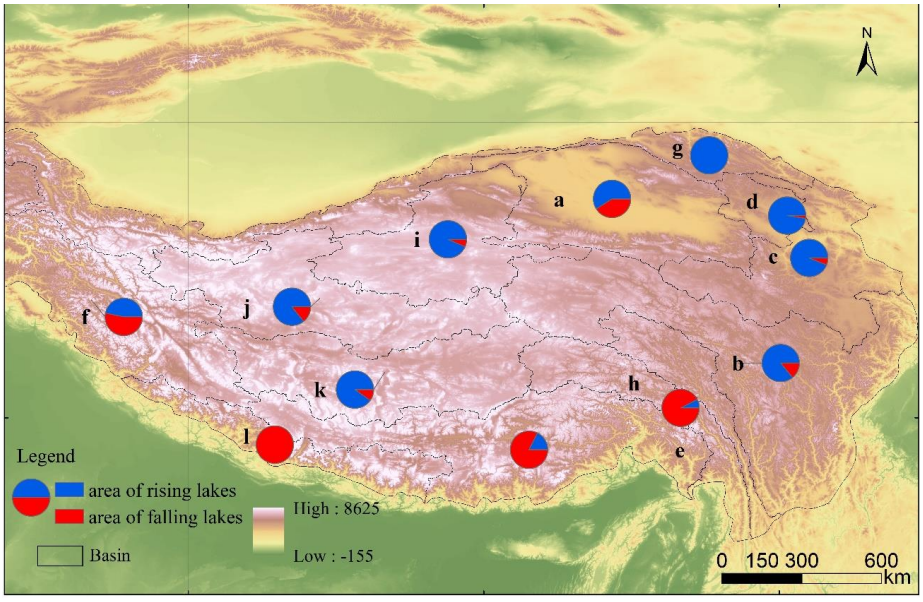
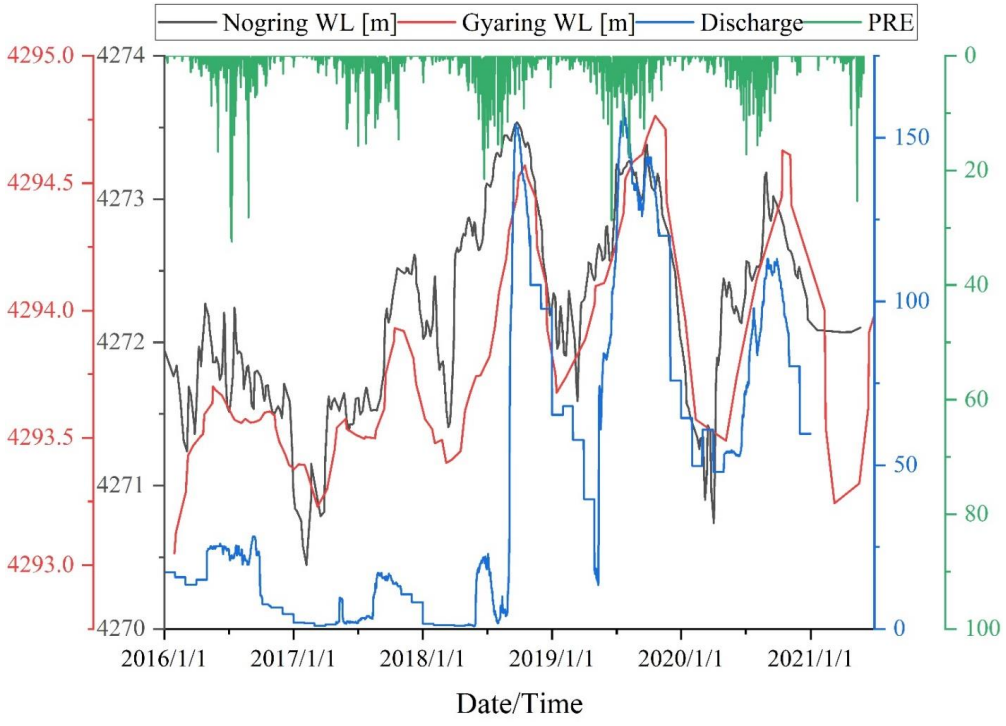


Fig. 9 Relative proportions of the trends in the changing levels of water in the lakes in each basin. The boundry of each basin is referred to Wan et al.(2016). The DEM of the base map is from the Global Multi-resolution Terrain Elevation Data 2010（GMTED2010）(GMTED: https://topotools.cr.usgs.gov/gtmed_viewer/) (the lowercase letters indicate the different lake basins studied as in Fig. 7).

5.2 Exploring the responses of the lake levels to river regulation

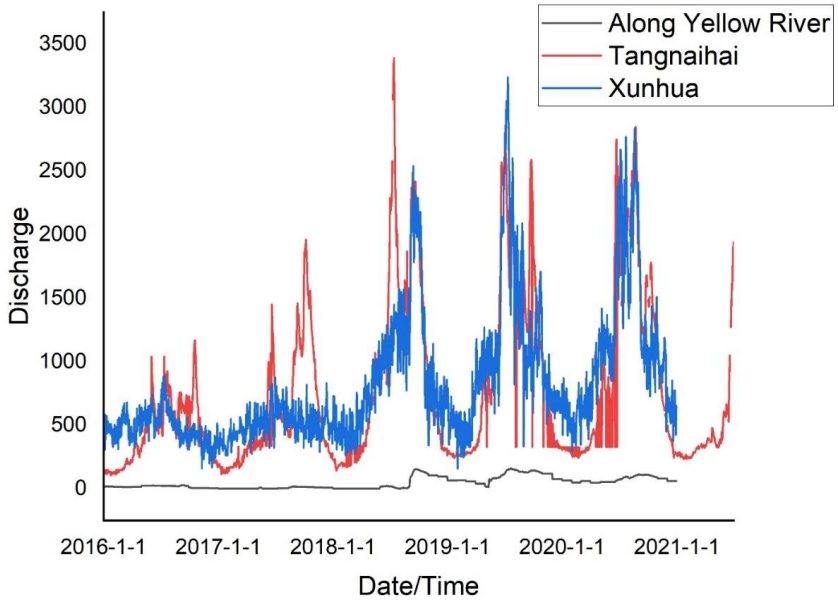
Aided by the availability of the high-space-coverage lake level time series, it is possible to explore the responses of the water levels in the lakes to river regulation to provide support for integrated management of the lake water resources. Here, the streamflow of the rivers and lakes in the source area of the Yellow River are taken as an example to analyze the relationships between changes in the streamflow of the source area of the Yellow River and the Nogring Lake and Gyaring Lake. From inspection of Fig. 10, the discharge at the source of the Yellow River is directly affected by the regulation and storage of these two lakes, which results in changes to the annual distribution of the discharge at the source; thus, the discharge along the Yellow River is more uniform than that in the lower reaches of Tangnaihai and Xunhua (Fig. 11). However, the correlation between precipitation and the changes in the discharge at the source of the Yellow River is very poor, and there is a certain time lag. The changes in water levels of the lakes are basically synchronized with the changes in the streamflow at the source of the Yellow River (Fig. 10), indicating that the streamflow is interannual regulated.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022



310

Fig. 10 Responses of water levels in the lakes to regulation of streamflow in the river (Nogring WL represents the water level of Nogring Lake; Gyaring WL represents the water level of Gyaring Lake; PRE represents precipitation).



315 Fig. 11 Changes in the discharge along the Yellow River and the lower reaches of Tangnaihai and Xunhua.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

6 Data availability

The derived water levels in the lakes of the QTP are archived and available at https://doi.org/10.1594/PANGAEA.939427 (Chen et al., 2021).

7 Conclusion

In this study, high-resolution datasets for changes in the water levels for 362 lakes on the QTP during 2002-2021 were developed based on multi-altimeter data from Envisat, ICESat-1, CryoSat-2, Jason-1, Jason-2, Jason-3, SARAL, and Sentinel-3A. A modified waveform retracker and a noise-footprint removal method were used to extract the water levels, and the lake level time series were then estimated using the R package tsHydro. The dynamic reference time series was then used to merge the lake-level time series from the multi-altimeter data. It was found that the merged water levels based on the altimetry increased the overall sampling frequency regardless of the lake size. The water levels derived from the altimeter data were validated with *in situ* data, and the accuracy of the time series for monitoring lakes reached the decimeter level. Based on comparison with the LEGOS Hydroweb dataset, the new product was found to be consistent with the LEGOS Hydroweb product, and the R^2 of the linear fit of the two products was 0.83, indicating that the datasets were reliable. In addition, the spatio-temporal changes in the water levels of the lakes on the QTP during 2002-2021 were explored. Overall, the measured lake levels on the QTP were indicative of a rising trend with an overall average annual rate of change of 0.175 m/a; moreover, the number of lakes with rising water levels accounted for 78% of the total examined. The lakes with the most significant rises in the water levels were those of large size (>500 km^2) and small size (<50 km^2), and the intermediate size lakes showed the significant rising trend in the water levels The water levels of lakes in all basins have been increasing significantly over the period 2002 to 2021 except for the Brahmaputra basin. The lakes with decreasing water levels were distributed mainly in Brahmaputra, Ganges, and Salween basins. Further applications of the lake level dataset of the QTP are anticipated. For example, the dataset may be used to analyze the responses of the lake levels to river regulation to provide support for managing lake water resources.

Author contributions

Liao J and Chen J designed the research plan. Chen J developed the approaches and the dataset. Liao J, Lou Y and Ma S contributed to the analysis of the results. Shen G and Zhang L contributed to the data processing. Liao J and Chen J wrote the manuscript.

Competing interests. The authors declare that there are no conflicts of interest.

Acknowledgements

We thank the European Space Agency and Centre National d'Etudes Spatiales for providing the altimeter data and the Bureau of Hydrology and Water Resources of Qinghai Province, the Yellow River Commission of the Ministry of Water Resources, and the Institute of Tibetan Plateau Research, Chinese Academy of Sciences for providing *in situ* gauge measurements of the water levels.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Financial support

This work was supported by the National Natural Science Foundation of China (Grant 41871256).

Review statement

References

Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., Livingstone, D. M., Sommaruga, R., Straile, D., Donk, E. V., Weyhenmeyer, G. A., and Winder, M.： Lakes as sentinels of climate change, Limnol. Oceanogr., 54(6): 2283-2297, 2009.

Bamber, J. L.: Ice sheet altimeter processing scheme. Int. J. Remote Sens.,15(4):925–938, 1994.

Bhang, K. J., Schwartz, F. W., and Braun, A.: Verification of the Vertical Error in C-Band SRTM DEM Using ICESat and Landsat-7, Otter Tail County, MN, IEEE Trans. Geosci. Remote Sens., 45: 36-44, 2007.

Chen, J., and Liao, J.: Monitoring lake level changes in China using multi-altimeter data (2016–2019), J. Hydrol., 590: 125544, 2020.

Chen, J., Liao, J., Deng, W., Shen, G., Zhang, L., Wang, C. High-space-coverage lake level change data sets on the Tibetan Plateau from 2002 to 2021 using multiple altimeter data. PANGAEA, 2021, https://doi.org/10.1594/PANGAEA.939427

Frappart, F., Calmant, S., Cauhopé, M., Seyler, F., and Cazenave, A.: Preliminary results of Envisat RA-2-derived water levels validation over the Amazon basin, Remote Sens. Environ., 100: 252-264, 2006.

Gao, L., Liao, J., and Shen, G.: Monitoring lake-level changes in the Qinghai-Tibetan Plateau using radar altimeter data (2002-2012), J. Applied Remote Sens., 7, 073470, 2013.

Göttl, F., Dettmering, D., Müller, F., and Christian, S.: Lake level estimation based on Cryosat-2 SAR altimetry and multi-looked waveform classification, Remote Sens., 8(11): 885, 2016.

Hwang, C., Cheng, Y., Han, J., Kao, R., Huang, C., Wei, S., and Wang, H.: Multi-decadal monitoring of lake level changes in the Qinghai-Tibet Plateau by the TOPEX/Poseidon-Family altimeters: Climate implication, Remote Sens., 8, 446, 2016.

Hwang, C., Cheng, Y., Yang, W., Zhang, G., Huang, Y., Shen, W., and Pan, Y.: Lake level changes in the Tibetan Plateau from Cryosat-2, SARAL, ICESat, and Jason-2 altimeters, Terr. Atmos. Ocean. Sci. 30, 1–18, 2019.

Irvine, K. N., and Eberhardt, A. J.: Multiplicative, seasonal ARIMA models for Lake Erie and Lake Ontario water levels. Water Resour. Bull., 28(2): 385–396, 1992.

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, J. Hydrol., 544: 109-124, 2017.

Kleinherenbrink, M., Ditmar, P. G., and Lindenbergh, R. C.: Retracking Cryosat data in the SARIn mode and robust lake level extraction, Remote Sens. Environ., 152: 38-50, 2014.

Kleinherenbrink, M., Lindenbergh, R. C., and Ditmar, P. G.: Monitoring of lake level changes on the Tibetan Plateau and Tian Shan by retracking Cryosat SARIn waveforms, J. Hydrol., 521: 119-131, 2015.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Kropáček, J., Braun, A., Kang, S., Feng, C., Ye, Q., and Hochschild, V.: Analysis of lake level changes in Nam Co in central Tibet utilizing synergistic satellite altimetry and optical imagery. Int. J. Appl. Earth Obs. Geoinf. 17, 3–11, 2012.

Lee, H., Shum, C.K., Kuo, C.Y., Yi, Y., and Braun, A.: Application of TOPEX altimetry for solid earth deformation studies, Terr. Atmos. Ocean. Sci. 19: 37–46, 2008.

Lee, H., Shum, C.K., Tseng, K.H., Guo, J.Y., and Kuo, C.Y.: Present-day lake level variation from envisat altimetry over the northeastern Qinghai-Tibetan Plateau: links with precipitation and temperature, Terr. Atmos. Ocean. Sci. 22: 169–175, 2011.

Lee, S., Im, J., Kim, J., Kim, M., Shin, M.. Kim, H.C., and Quackenbush, L.J.: Arctic sea ice thickness estimation from CryoSat-2 satellite data using machine learning-based lead detection, Remote Sens., 8(9): 698, 2016.

Lei, Y.: The water level observation of lakes on the Tibetan Plateau (2010-2017), National Tibetan Plateau Data Center, 2018. DOI: 10.11888/Hydrology.tpe.249464.db.

Li, X., Long, D., Huang, Q., Han, P., Zhao, F., and Wada, Y.: High-temporal-resolution water level and storage change data sets for lakes on the Tibetan Plateau during 2000–2017 using multiple altimetric missions and Landsat-derived lake shoreline positions, Earth Syst. Sci. Data, 11: 1603–1627, 2019.

Liao, J., Gao, L., and Wang, X., Numerical simulation and forecasting of water level for Qinghai Lake using multi-altimeter data between 2002 and 2012, IEEE J. Selected Topics Appl. Earth Obs. Remote Sens., 7(7): 609-622, 2014.

Marshall, A., and Deng, X.L.: Image analysis for altimetry waveform selection over heterogeneous inland waters, IEEE Geosci. Remote Sens. Lett., 13(8): 1198-1202, 2016.

Martin, T.V., Zwally, H.J., Brenner, A.C., and Bindschadler, R. A.: Analysis and retracking of continental ice sheet radar altimeter waveforms, J. Geophys. Res. [Atmos.] 88: 1608–1616, 1983.

Medina, C. E., Gomez-Enri, J., Alonso, J.J., and Villares, P.: Water level fluctuations derived from ENVISAT Radar Altimeter (RA-2) and in-situ measurements in a subtropical waterbody: Lake Izabal (Guatemala), Remote Sens. Environ. 112, 3604–3617, 2008.

Nielsen, K., Stenseng, L., Andersen, O. B., Villadsen, H., and Knudsen, P.: Validation of CryoSat-2 SAR mode based lake levels, Remote Sens. Environ., 171: 162-170, 2015.

Pavlis, N.K., Holmes, S.A., Kenyon, S.C., and Factor, J.K.: The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), J. Geophys. Res. Solid Earth 117: 1–38, 2012.

Pritchard, H.D.: Asia's glaciers are a regionally important buffer against drought, Nature, 545(7653):169-174, 2017.

Schindler, D. W.: Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes, Limnol. Oceanogr., 54(6_part_2): 2349-2358, 2009.

Shen, X., Zhang, J., Zhang, X., Meng, J., and Ke, C.: Sea ice classification using Cryosat-2 altimeter data by optimal classifier feature assembly, IEEE Geosci. Remote Sens. Lett., 14(11): 1948-1952, 2017.

Song, C., Ye, Q., and Cheng,, X.: Shifts in water-level variation of Namco in the central Tibetan Plateau from ICESat and Cryosat-2 altimetry and station observations, Sci. Chin., 60(14): 1287-1297, 2015.

Song, C., Ye, Q., Sheng, Y., and Gong, T.: Combined ICESat and CryoSat-2 Altimetry for accessing water level dynamics of Tibetan lakes over 2003–2014, Water, 7, 4685–4700, 2015.

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, Sci. Data, 3(3): 160039, 2016.

https://doi.org/10.5194/essd-2022-313
Preprint. Discussion started: 18 October 2022

Wang, J.: The lake level observation data of Lake Namco from the Integrated Observation and Research Station of Multisphere in Namco (2007-2016), Monitoring & Big Data Center for Three Poles, 2018.

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 Sens. Environ. 115, 1733–1742, 2011.

Zhang, G., Xie, H., Yao, T., and Kang S.: Water balance estimates of ten greatest lakes in China using ICESat and Landsat data, Chin. Sci. Bull. 58, 3815–3829, 2013a.

Zhang, G., Xie, H., Yao, T., Liang, T., and Kang, S.: Snow cover dynamics of four lake basins over Tibetan Plateau using time series MODIS data (2001–2010), Water Resour. Res., 48(10),W10529, 2012.

Zhang, G., Yao, T., Xie, H., Kang, S., and Lei, Y.: Increased mass over the Tibetan Plateau: from lakes or glaciers?, Geophys. Res. Lett., 40(10): 2125-2130, 2013b.

Zhang, G.: Changes in lakes on the Tibetan Plateau observed from satellite data and their responses to climate variations, Progr. Geogr., 37(2): 214-223, 2018.

Zhang, Y., Li, B., and Zheng, D.: A discussion on the boundary and area of the Tibetan Plateau in China, Geogr. Res., 21(1): 1-8, 2002.