Supplement of Earth Syst. Sci. Data, 11, 1603–1627, 2019
https://doi.org/10.5194/essd-11-1603-2019-supplement

Open Access Earth System Science Data

Supplement of

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

Xingdong Li et al.

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

1. Supplementary information for the developed lake data sets

Water levels in this lake data set are with respect to a certain pair of reference ellipsoid and geoid model. However, due to the altimetry water level merging scheme we used, different lakes may have different references. For each lake, a satellite altimeter with the longest observation was chosen as the baseline for merging other water level time series. Therefore, the reference ellipsoid and geoid model for each lake is the same as that of its baseline water level as listed in Table S1.

Optical water levels developed in this study were generated from regression analysis between lake shoreline positions and corresponding altimetry data. The regression relationship is robust within the regression range, or in other words, interpolation range. Optical water levels beyond the interpolation range should be used with caution. The upper and lower bounds of the interpolation range for each lake were listed in Table S1.

Table S1 Reference ellipsoid, geoid model, interpolation range for each lake

Lake Name	Reference ellipsoid	Geoid model	Lower bound of optical water level interpolation	Upper bound of optical water level interpolation
Ake Sayi Lake	WGS84	EGM08	4849	4851
Aqqikkol Lake	WGS84	EGM96	4257	4264
Ayakkum Lake	WGS84	EGM08	3879	3884
Bamco	WGS84	EGM96	4568	4570
Bangong Co	T/P	EGM96	4238.5	4240
Chibzhang Co	T/P	EGM96	4931	4934
Co Ngoin1	WGS84	EGM96	4564.8	4565.8
Cuona Lake	WGS84	EGM96	4585.6	4586.4
Dagze Co	T/P	EGM96	4468	4472
Dogai Coring	T/P	EGM96	4816	4820
Dogaicoring Qangco	WGS84	EGM08	4789	4797
Donggei Cuona Lake	WGS84	EGM08	4084.5	4086.5
Dung Co	WGS84	EGM08	4551	4553
Goren Co	WGS84	EGM96	4649	4650.4
Gozha Co	WGS84	EGM08	5082.5	5083.5
Gyaring Lake	WGS84	EGM08	4292	4294
Har Lake	WGS84	EGM08	4076	4080
Hoh Xil Lake	WGS84	EGM96	4890.5	4893
Jingyu Lake	WGS84	EGM08	4716	4723
Kusai Lake	T/P	EGM96	4474	4486

Kyebxang Co	WGS84	EGM96	4623.5	4624.5
Langa Co	T/P	EGM96	4564	4567.5
Lexiewudan Co	WGS84	EGM08	4872	4879
Lumajiangdong Co	WGS84	EGM08	4814	4818
Mapam Yumco	WGS84	EGM08	4584.6	4585.8
Margai Caka	WGS84	EGM96	4794	4800
Memar Co	WGS84	EGM08	4924	4932
Nam Co	WGS84	EGM08	4725.5	4727.5
Ngangla Ringco	WGS84	EGM08	4715.2	4716.4
Ngangze Co	T/P	EGM96	4685	4689
Ngoring Lake	T/P	EGM96	4270	4273
Paiku Co	WGS84	EGM96	4579	4580.5
Puma Yumco	WGS84	EGM96	5012	5013.5
Pung Co	WGS84	EGM08	4534	4538.5
Qinghai Lake	WGS84	EGM08	3193	3195
Rola Co	WGS84	EGM96	4818	4822
Salt Lake	WGS84	EGM96	4441	4467
Salt Water Lake	WGS84	EGM96	4902	4907
Selin Co	WGS84	EGM08	4540	4546
Tangra Yumco	WGS84	EGM08	4536	4538.5
Taro Co	WGS84	EGM96	4568	4569.5
Tu Co	WGS84	EGM96	4935	4938
Urru Co	WGS84	EGM08	4554	4555.5
Wulanwula Lake	WGS84	EGM08	4857	4861.5
Xijir Ulan Lake	WGS84	EGM08	4772	4777
Xuru Co	WGS84	EGM96	4716	4717
Yamzho Yumco	WGS84	EGM08	4436	4442
Yelusu Lake	WGS84	EGM96	4686.8	4687.8
Yibug Caka	WGS84	EGM96	4560	4561.5
Zhari Namco	T/P	EGM96	4610	4613
Zhuonai Lake	WGS84	EGM96	4743	4756
Zige Tangco	T/P	EGM96	4566	4571

2. Derivation of the probability density function in the uncertainty analysis

To evaluate the uncertainty of the automatic water extraction method, a random variable X was created to represent the difference between the observed/classified water area X_1 and real water area X_0.

$$X = X_1 - X_0 \quad \text{Eq S1}$$

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:

$$F(X) = \begin{cases} P(X = -X_0|X_0) \cdot P(X_0), & X < 0 \\ P(X_0 = 0, X_1 = 0) + P(X_0 = 1, X_1 = 1), & X = 0 \\ P(X = 1 - X_0|X_0) \cdot P(X_0), & X > 0 \end{cases} \quad \text{Eq S1}$$

$$P(X = -X_0|X_0) = P(X_1 = 0|X_0) = 1 - f(X_0) = 1 - f(-X) \quad \text{Eq S2}$$

$$P(X = 1 - X_0|X_0) = P(X_1 = 1|X_0) = f(X_0) = f(1 - X) \quad \text{Eq S3}$$

where $f(x)$ is referred to as the classifier function defined in Eq. (4) in the manuscript:

$$f(x) = x^n \quad \text{Eq 4}$$

$f(x)$ describes the probability of a Landsat shoreline pixel being classified as water, given that the water area ratio of that pixel is x. n is a parameter controlling the shape of the curve. In our study, n equals 1.43 based on the maximum likelihood analysis of 4128 Landsat shoreline pixels.

Because X_0 is uniformly distributed between [0,1], $P(X_0) = 1$. Combining Eqs. (S2)–(S4) with the explanations on the specific case $X = 0$ results in the following:

$$F(X) = \begin{cases} 1 - f(-X), & X < 0 \\ 1 + 1 - f(0), & X = 0 \\ f(1 - X), & X > 0 \end{cases} = \begin{cases} 1 - f(-X), X < 0 \\ 2 - f(0), & X = 0 \\ f(1 - X), & X > 0 \end{cases} \quad \text{Eq S5}$$

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.

3. Extrapolation issues

Extrapolation in this study means the extrapolation of the linear relationship developed from the regression analysis between altimetry water levels and lake shoreline positions. For instance, if the altimetry water levels used for the regression analysis have a range of 4500–4502 m, then the generated optical water levels beyond this range are regarded as

extrapolated values. On the other hand, if an optical water level H_1 acquired in 2003 is within 4500–4502 m, though the altimetry water levels used for such a regression were from 2010 to 2017, H_1 is still regarded as an interpolated value because it is within the elevation range of the linear regression. The interpolation ranges for each lake were listed in Table S1.

There are two periods when extrapolation of the regression relationship may occur. The first period is 2000–2002, when little altimetry information is available. Another period is the time gap between two altimetry missions, e.g., the time gap between ICESat and CryoSat-2. It is shown in the following paragraphs that if the water level and the lake slope both intensively changed during these two periods and corresponding elevations, the extrapolation may result in uncertainties in the optical water levels. In addition, if such an extrapolation error occurred during the time gaps between two altimetric missions, it would impact the altimetry data merging process because optical water levels were used as the reference.

Note that we have performed two regressions to generate the optical water levels. For the first regression, we only used one altimetry data product and optical images-derived lake shoreline positions. After merging the altimetry water levels, we performed the second regression using the merged altimetry water levels and the optical water levels temporally close to the altimetry water levels throughout the entire study period. Here we show that the extrapolation problem during the time gap of altimetric missions is evitable in nature with the second regression.

As shown in Figure S1 (Figure S1 and S2 are both conceptualized examples, optical water levels are fitted with the second altimetry product), when seasonal signal is dominated in the time series, there is no need for extrapolation. The red line in the optical water levels (which serves as the merging reference to altimetry data 1) are within the range of the linear regression. The merging between the two altimetry water levels can subsequently be achieved by removing the difference (symmetrical bias) of the mean water levels between altimetry data 1 and altimetry data 2 during the reference period (the red solid line) from altimetry data 1 (typically ICESat data).

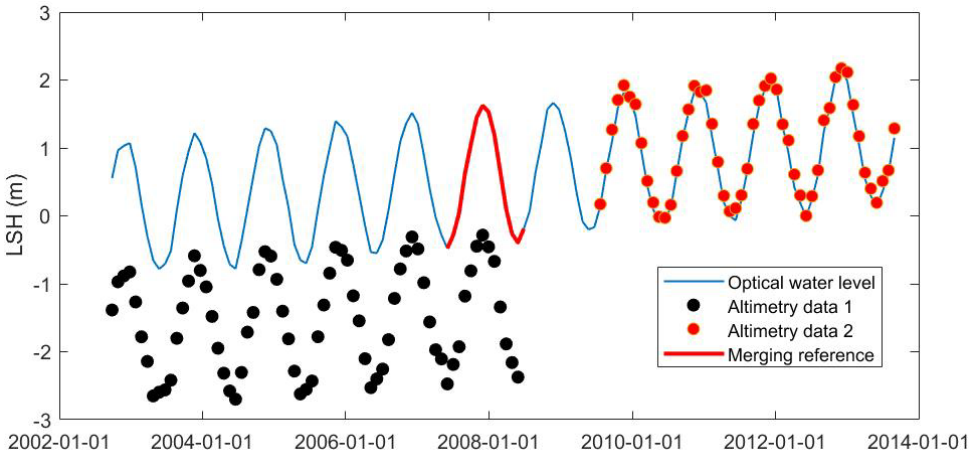

Figure S1 Conceptual plot of a seasonal signal dominated lake water levels time series: black dots denote the first altimetry product, red dots denote the second altimetry product, blue line represents optical water levels generated by fitting with the second altimetry product, red line represents the optical water levels used as reference for altimetry data merging.

When a multiyear trend is dominated in the time series as shown in Figure S2, the merging reference is out of the range of the regression relationship, and then extrapolation may occur. Both situations are common in our study. The first situation comprises 60% of all study lakes, and extrapolation can take place in ~40% lakes. The two altimetry datasets in the extrapolation case can still be merged using the similar procedure and optical water levels shown in the interpolation case above.

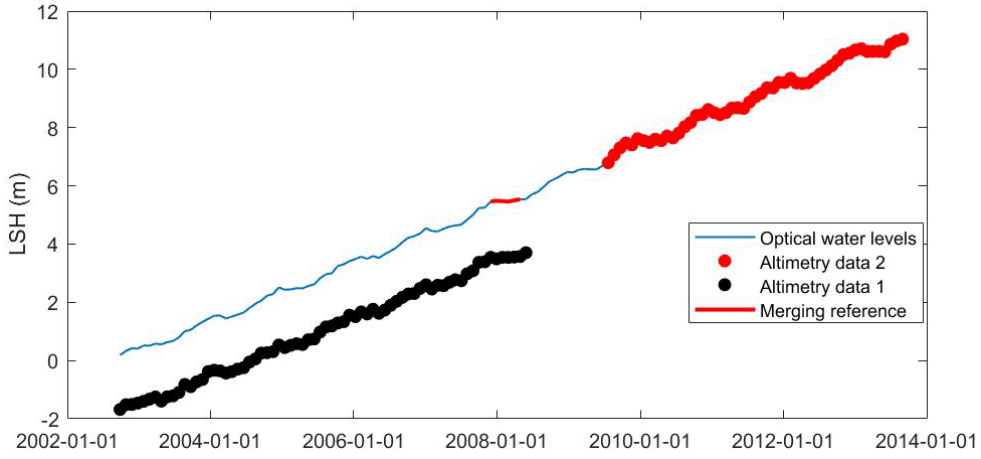

Figure S2 Conceptual plot of a multiyear signal dominated lake water levels time series.

In the merging process, extrapolation becomes a problem only if the lake bank slope experiences an abrupt change at the exact elevation where both altimetry products fail to cover, as illustrated in Figure S3. Such a situation may happen, but the possibility is

relatively low. If it happens, the extrapolation will result in a remaining systematic bias in the merged altimetry water levels and consequently jeopardizing the accuracy of the optical water levels.

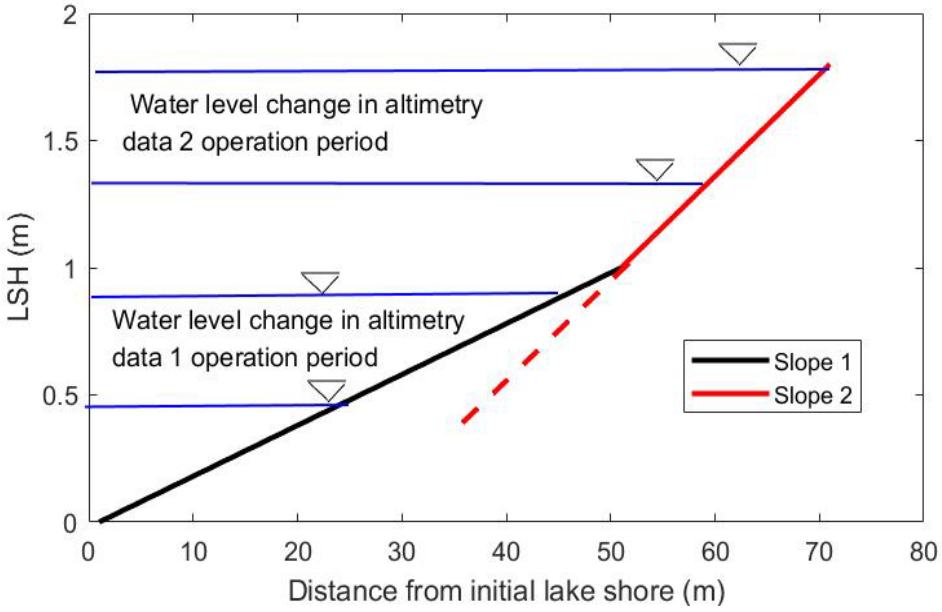

Figure S3 Conceptual plot of a two-slope lake shore topography, the slope change occurred at an elevation where both altimetry products failed to cover.

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

4. Determination of parameter C in lake overflow flood modeling

In section 6.3 of the manuscript, the broad crest formula shown in Eq. (13) was used to calculate the lake overflow discharge of Lake Kusai. *H*, representing water head over outlet in Eq. (13), can be determined from merged water levels of Lake Kusai, if the height of the lake outlet is known. Therefore, several Landsat-based lake outlet water extent and same-period water levels were compared to determine the outlet height (see Figure 18 in the manuscript). Parameter *b* represents the width of the outlet, which can be measured directly from optical images. However, parameter C is empirical and needs to be estimated or calibrated.

$$Q = C \cdot b \cdot H^{1.5}\sqrt{2g} \quad \text{Eq 13}$$

C is a parameter that mainly reflects geometric characteristics of the weir, mainly varying from 0.3–0.4. 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 ± 0.2 m. Given that the rainy season has ended (Liu et al., 2016), 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):

$$-\frac{dV(H)}{dt} = Q = C \cdot b \cdot H^{1.5}\sqrt{2g} \quad \text{Eq S6}$$

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

Eq. (S6) 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 0.30 in this case, which is a reasonable value in hydrodynamic calculations. Then we validated our method with Stage 2. Results showed that the simulated lake over flow from Lake Kusai in Stage 2 (0.21 km^3) is close to the gained water in Lake Salt during the same period (0.19 km^3), suggesting that the method and the parameters were effective in providing an initial estimation of upstream overflow flood from Lake Kusai. Given that few ground observations are available, such an initial estimation is critical for the downstream flood risk control.

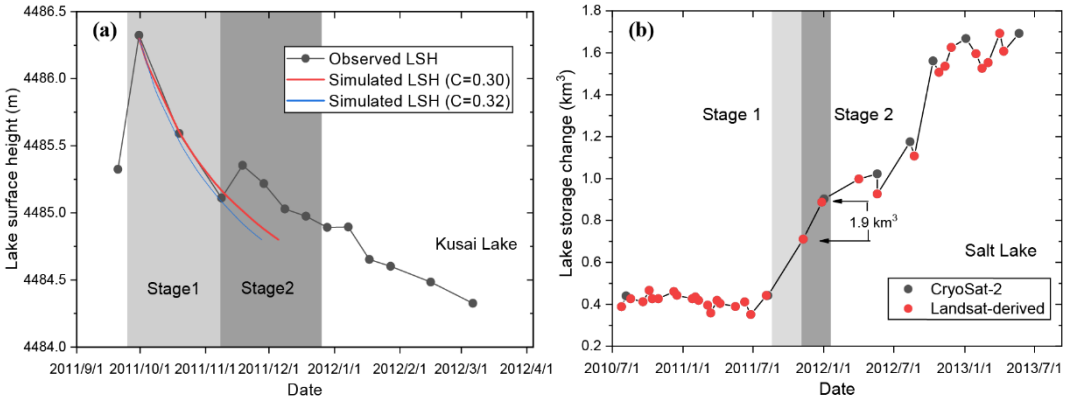

Figure 19. (a) Changes in the water level of Lake Kusai after receiving the outburst flood from Lake Zhuonai. Water level in stage 1 was simulated using Eq. (S6), which provided a referencing range of parameter *C*. Water level in stage 2 provided water level input to calculate total outflow, which was compared with the same period 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 km^3 of water gained in stage 2, which was comparable to the outflow estimate of Lake Kusai (0.22 km^3).

Reference

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.

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.