Response to Reviewers' comments to manuscript essd-2020-57

"Annual 30-meter Dataset for Glacial Lakes in High Mountain Asia from 2008 to 2017"

Dear Editors and Reviewers:

Thank you a lot for your kind and careful reviewing. Your suggestions give us important and constructive perspective on this manuscript, and help to improve the manuscript greatly. We have fully considered all the comments of you, and have substantially revised our manuscript according to your comments. A point-by-point response to the outstanding comments raised is attached to this manuscript. The major changes are summarized as follows:

1. We have put considerable effort to update data of glacial lakes for the ten year records, and manually append their attribute information. The related statistics, figures and analysis in the article have also been modified based on the new lake inventory.

2. We have comprehensively investigated the existing works about glacial lake inventory in the Section 1. Introduction, and quantitatively analysed and compared these inventories with ours in the Section 6. Discussions, to clearly show benefits and challenges remaining in this study.

3. Detailed explanations about the mapping of some problematic ice-covered lakes have been given in the Section 3.2. The number of missed or misclassified lakes that need to be manually corrected were also described clearly.

4. A thorough and quantitative uncertainty analysis of lake area was added in the Section 4 of revised manuscript, the error bars for the lake area, and confidence intervals for the estimated trends were also added throughout the paper.

5. We have carefully modified the language deficiencies, imprecise expressions, and provided more detailed interpretations and conclusions.

The changes have been highlighted in colored text in the revised manuscript. In the following, we provide point by point responses to the outstanding comments and suggestions provided by the Anonymous Reviewers. We are indebted to you for your outstanding and constructive comments, which greatly helped us to improve the technical quality and presentation of our manuscript. Once again, thank you very much for your comments and suggestions.

Response to Comments by Reviewer #2:

Chen et al. used Google Earth Engine to map glacial lakes in High Mountain Asia (HMA) from 2008 to 2017 with Landsat imagery. Their data is given in annual time steps, which so far is the highest temporal resolution of glacial lake inventory for the HMA. Thus, this kind of dataset if with fine quality could be particularly useful for scientific researches in changes of the cryosphere of the HMA as well as for related assessments and evaluations on the hydrological responds and glacial lake outburst flood risks under a changing climate in the HMA.

We greatly appreciate the suggestions of the Reviewer for his/her accurate summary of the main contributions of our work and for the outstanding recommendations provided. Following the Reviewer's very pertinent recommendations, in the revised manuscript and the following text we have made modifications according to these questions. We thank again the Reviewer for his/her careful handling of our manuscript and for the constructive suggestions provided, which greatly helped us improve the technical quality and presentation of our manuscript.

1. *Noticed that recently there published a similar dataset produced by Wang et al. (2020), which includes two periods (1990 and 2018) of glacial inventory for the HMA and currently is also under review for the ESSD. The later one used a more traditional method and probably involved more extensive manually inspection during their investigations. When comparing these two datasets for the closest period (2017 of Cheng et al. and 2018 of Wang et al. 2020), I found there is a very large discrepancy (Table 1) in their results, although they have claimed that they used similar (not the same) area threshold (0.0081 km^2 of Chen et al. 2020 and 0.0054 km^2 of Wang et al. 2020) and distance threshold (within a 10 km from the nearest glacier terminus) for the lake mapping.*

Table 1: Differences between two datasets of HMA glacial lake inventory

Source	Year	Total Numbers	Area_Sum (km^2)	Area_Max (km^2)	Area_Mean (km^2)	Area_Min (m^2)	Notes
Wang et al. 2020	2018	28953	1955.939	6.465	0.067	5400	Altai mountains included
Wang et al. 2020	2018	22219	1672.479	6.465	0.075	8100	Altai mountains excluded and area larger than 8100 m^2
Chen et al. 2020	2017	14477	1635.939	26.598	0.113	8100	Altai mountains excluded

In general, dataset of Chen et al. has missed a quite number of glacial lakes when comparing their results with the Wang et al. 2020's. In addition, they excluded the Altay and Sayan mountains, which actually should be included for an inventory study for the HMA. Even if excludes Altay form Wang et al.'s result, there still exist remarkable discrepancy in total numbers (nearly 7800 lakes) and total area (~37 km^2) between each other (Table 1). I agree with the Reviewer #1 that the missing inventory by Chen et al. is far from the range of uncertainty, but it was indeed errors due to the lack of systematic experts' check through the results, which were greatly depended on the methods they used when applying their procedures in Google Earth Engine.

The authors should pay more attention to how to control the quality of their dataset. Although the Google Earth Engine offer the opportunity to calculate lake inventory with a higher temporal resolution, I still strongly recommend they check through their result for *each year, or maybe improve their methods to avoid errors as much as possible. Then a comparison between their results with existing datasets (such as datasets published by Zhang et al. 2015 or Wang et al. 2020) is necessary for audients judgement of the data quality. They did have do some comparison between their results with the Global Surface Water (GSW) dataset, but both these two were calculated by Google Earth Engine.*

Response:

We greatly appreciate the Reviewer's valuable comments and suggestions. This paper focus on obtaining the comprehensive knowledge of the distribution, area of glacial lakes, and also quantification of variability in their size and type at high resolution in HMA. We develop a HMA Glacial Lake Inventory (Hi-MAG) database to characterize the annual coverage of glacial lakes from 2008 to 2017 at 30 m resolution. This is the first glacial lake inventory across the HMA with annual temporal resolution, it can provide details for different types of glacial lakes and evolution patterns. Although the method of lake mapping was automatic, for quality control every lake polygon was inspected and was manually edited where needed. The related attribution were also added for each lake. This is a huge amount of work for the mapping of nearly 140,000 glacial lakes over ten time periods from 2008 to 2017.

In the last version of manuscript, the mapping work still have some deficiencies and was not fully reflected in the quality of the manuscript indeed. Therefore, we have made considerable efforts to fill these defects, which is mainly manifested in the four aspects including **supplement of glacial lake data, detailed comparison with other lake inventories, quantitative uncertainty analysis of lake area and adding confidence intervals for trends, details about the whole mapping procedure.** We have made the following improvements in our study:

1) We have been working hardly to update data of glacial lakes in the ten year, and add their attribute information. Based on the new lake inventory, we have also carefully modified the tables, figures, the related analysis and made reliable conclusions.

In this study, we produced the glacial lake dataset in High Mountain Asia in annual time steps from 2008 to 2017. The number of glacial lakes in each year is more than 12,000 totally, the ten year records have nearly 140,000 lake polygons. To improve work efficiency, obtain the accurate boundary of the glacial lakes, and meanwhile, minimize the subjective judgment errors of the operatives, we applied a systematic glacial lake detection method that combined two steps from initial glacial lake extraction and subsequently manual refinement of lake mapping results.

The initial glacial lake extraction using GEE can make sure that approximately half of all the lakes are automatically extracted. To make annual data more complete and accurate, manual inspection and refinement of individual glacial lake is necessary to supplement some missing lakes and correct the mapping errors such as mountain shadows and river segments. The glacial lake vector over the entire HMA for the years from 2008 to 2017 has been rechecked and reedited individually through dynamic cross-validation by ten trained experts. Besides, the related attribution (i.e., lake type, elevation, distance to the nearest glacier terminus, area and perimeter) were manually added for each lake. It should be noted that the type of each individual glacial lake was carefully judged according to the different formation mechanisms or growth stages. Thus the whole processing is time consuming and require a considerable amount of work. We are terrible sorry for the large discrepancy compared with dataset from

Wang et al., 2020, which may contribute to the difference in the calculated statistics, and the conclusions drawn in some sub-regions.

Therefore, following the Reviewer's very pertinent recommendations, we have carefully examined the lake data for the ten year by supplementing glacial lakes that have an area of greater than 0.04 km^2. The area of 0.04 km^2 is chosen as the threshold for re-revision of the glacial lake dataset for three reasons: **i)** glacial lakes with small sizes are more likely to be confused with surroundings due to its less effective spectral, textural and spatial information in comparison with those of relatively large glacial lakes, which results in the **large uncertainty of area** (Please see the Section 4, small glacial lakes with the area of less than 0.04km^2 have the mean area uncertainty of about 25.7 %). **ii)** small glacial lakes are highly variable in their locations, shapes and size, and also the optimal images with valid observations over the potential glacial lake area is very limited **in each year**. The image dates are not consistent across the whole HMA region because of atmospheric disturbances, also the influences from image quality, ice and shadow that obscured the lakes, this creates a great deal of uncertainty about the number and extent of small glacial lakes **for the such an annually resolved inventory**. **iii)** given the amount of work required to update the inventory and also revise our manuscript before the deadline.

Moreover, as for the definition of the geographical location of the study area, generally, the term HMA refers to a **broad high-altitude region in South and Central Asia** that covers the whole Tibetan Plateau and adjacent mountain ranges, including the Eastern Hindu Kush, Western Himalaya, Eastern Himalaya, Central Himalaya, Karakoram, Western Pamir, Pamir Alay, Northern/Western Tien Shan, Dzhungarsky Alatau, Western Kunlun Shan, Nyainqentanglha, Gangdise Mountains, Hengduan Shan, Tibetan Interior Mountains, Tanggula Shan, Eastern Tibetan Mountains, Qilian Shan, Eastern Kunlun Shan, Altun Shan, Eastern Tien Shan, Central Tien Shan, Eastern Pamir **(Yoon et al., 2019;Brun et al., 2017;Zhao et al., 2014)**. It extends from 26°N to 45°N and from 67°E to 105°E, the altitude of the plateau is about 4500 m on average.

Until now, there are still not a uniform standard about the spatial location of HMA region. When we first started working on the glacial lake mapping in the HMA region and required accurate mountain ranges division data, Tobias Bolch kindly provided us the mountain boundary shapefile in High Mountain Asia, which excluded Altay and Sayan mountains, can be downloaded from the website geo.uzh.ch/~tbolch/data/regions_hma_v03.zip. Just as the Reviewer pointed out that, the generalized concept of HMA could contain the Altay and Sayan mountains, and a comprehensive delineation of glacial lakes in these regions over the ten year periods is needed urgently in our future work.

Our team have been working hardly to update data of glacial lakes larger than 0.04 km^2 in the ten year, and add their attribute information. Correspondingly, the related statistics, figures and analysis in the article have also been modified based on the new lake inventory. We tried our best to improve the data and manuscript, and hope that the revised manuscript will meet with approval. As an annual time steps data over the HMA region from 2008 to 2017, the produced data may not be complete and perfect enough, yet our team will continuously update and share more and better glacial lake data in the future.

Finally, to provide more information about our study area, in the **Section 2.1 Study area** of the revised manuscript, we have rewritten and expanded this chapter by adding the location map of the HMA and comprehensive descriptions about the climate and topography. Please see the revised manuscript for more details.

2) We have comprehensively investigated the existing works about glacial lake inventory, and conducted deep comparisons between our Hi-MAG dataset and glacial lake inventory from Wang et al., 2020, and other studies from Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a.

To clearly show benefits and challenges remaining in this study, we firstly comprehensively investigated the existing works about glacial lake inventory.

Many previously published researches have devoted to the glacial lakes mapping with remotely sensed data over the different regions of HMA. Some works mainly focus on the investigation of the development of relatively large glacial lakes. Rounce et al. identified 131 glacial lakes in Nepal in 2015 that are greater than 0.1 km^2 (Rounce et al., 2017). Li et al. compiled an inventory of glacial lakes (≥0.01 km^2) with a spatial resolution of 30 m in the Karakoram mountains (Li et al., 2020). Aggarwal et al. shared a new dataset of glacial and high-altitude lakes that have an area > 0.01 km^2 for Sikkim, Eastern Himalaya from 1972–2015 (Aggarwal et al., 2017). Ukita et al. constructed a glacial lake inventory of Bhutan in the Himalaya from the period 2006-2010 based on high-resolution PRISM and AVNIR-2 data from ALOS. Considering small lakes present less of a GLOF risk. They set 0.01 km^2 as the minimum lake size (Ukita et al., 2011). Ashraf et al. used Landsat-7 ETM+ images for the 2000-2001 period to delineate glacial lakes greater than 0.02 km^2 in the Hindukush-Karakoram-Himalaya (HKH) Region of Pakistan (Ashraf et al., 2012). Because small glacial lakes experience highly variable in their shape, location, and occurrence, and were clearly sensitive to the warming climate and glacier wastage, a growing number of scholars have paied attention to the abundance of small glacial lakes. Salerno et al. provided a complete mapping of glacial lakes (including lake size less than 0.001 km^2) and debris-covered glaciers with 10-m spatial resolution in the Mount Everest region in 2008 (Salerno et al., 2012). Wang et al. utilized Landsat TM/ETM+ images for the years 1990, 2000 and 2010 to map glacial lakes with area more than 0.002 km^2 in Tien Shan Mountains (Wang et al., 2013). Luo et al. examined glacial lake changes (lake area >0.0036 km^2) for the entire western Nyainqentanglha range for the five periods between 1976 and 2018 using multi-temporal Landsat images (Luo et al., 2020). International Centre for Integrated Mountain Development (ICIMOD) provided comprehensive information about the glacial lakes (greater than or equal to 0.003 km^2) of five major river basins of the Hindu Kush Himalaya (HKH) using Landsat images for the year 2005 (Sudan et al., 2018). Nie et al. mapped the distribution of glacial lakes across the entire Himalaya in the year of 2015 using a total of 348 Landsat images at 30 m resolution. They set the minimum mapping unit to 0.0081 km^2 (Nie et al., 2017). Zhang et al. presented a database of glacial lakes larger than 0.003 km^2 in the Third Pole for the years ~1990, 2000, and 2010 (Zhang et al., 2015). All these researches greatly help to fill the data gap of glacial lakes information in the HMA region. At the global scale, Pekel et al. used millions of Landsat satellite images to record global surface water over the past 32 years at 30-metre resolution (Pekel et al., 2016), many large and visible glacial lakes were also included. More recently, Shugar et al. mapped glacial lakes with areas >0.05 km^2 around the world using 254,795 satellite images from 1990 to 2018 (Shugar et al., 2020). Wang et al. developed a glacial lake inventory (with size of larger than 0.0054 km^2) across the High Mountain Asia at two time periods (1990 and 2018) using manual mapping on 30 m Landsat images (Wang et al., 2020). They firstly introduce glacial lake inventory at such a large-scale, the data shared will be served as a baseline for the further studies related to water resource assessment or glacier hazard risk.

For more rigorous cross-validation and assessment of the data quality, in the **Section 6.1** of revised manuscript, we have conducted a deep analysis and comparison between our Hi-MAG dataset and glacial lake inventory from Wang et al., 2020, who mapped glacial lake larger than 0.0054 km^2 in a much larger High Mountain Asia region in two time periods (1990 and 2018), and made their data public recently. The comparative discussions with other studies (e.g. Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a) were also used as a general reference.

Firstly, we compared our dataset with that of Wang et al., 2020 for the closest period (2017 from our Hi-MAG database and 2018 from Wang et al. 2020) over the spatial extent of our HMA region. It should be noted that in the last version of database, some huge glacial lakes with an area of larger than 6.5km^2 have also been mapped in the Hi-MAG dataset, this is because these lakes all located in the range of 10 km from the nearest glacier terminus according to the definition of glacial lake in HMA. However, in the dataset produced by Wang et al., 2020, these huge lakes were removed. For the sake of a reliable comparative analysis between different studies under the equal conditions, glacial lakes with an area of larger than 6.5km^2 were deleted in our new lake inventory. Besides, **serval hundreds of glacial lakes in the dataset from Wang et al. are located outside of 10km buffer to the glacier terminus, which were also excluded for the comparison**.

The differences in the total number and area between these two dataset are 6206 and 223.97 km^2, respectively. **We also found that 2077 glacial lakes with a total area of 178.77km^2 in our Hi-MAG dataset were not detected by Wang et al.** As for the lakes we have not mapped, 96.1% of the glacial lakes are smaller than 0.04km^2, which means that 3.9 % (323 glacial lakes) of the difference in the total number is composed by the glacial lakes larger than 0.04 km^2. The main reasons for these missed 323 glacial lakes in our dataset are because the interference of some bad observations (cloud or snow), glacial lake dried up or outburst, or located in the middle of the river, which are summarized and explained in the following:

i) The limited or lack of high-quality images in a whole year. Many glacial lakes were always covered by cloud or snow, and no effective observation data are available, as shown in Figure I.

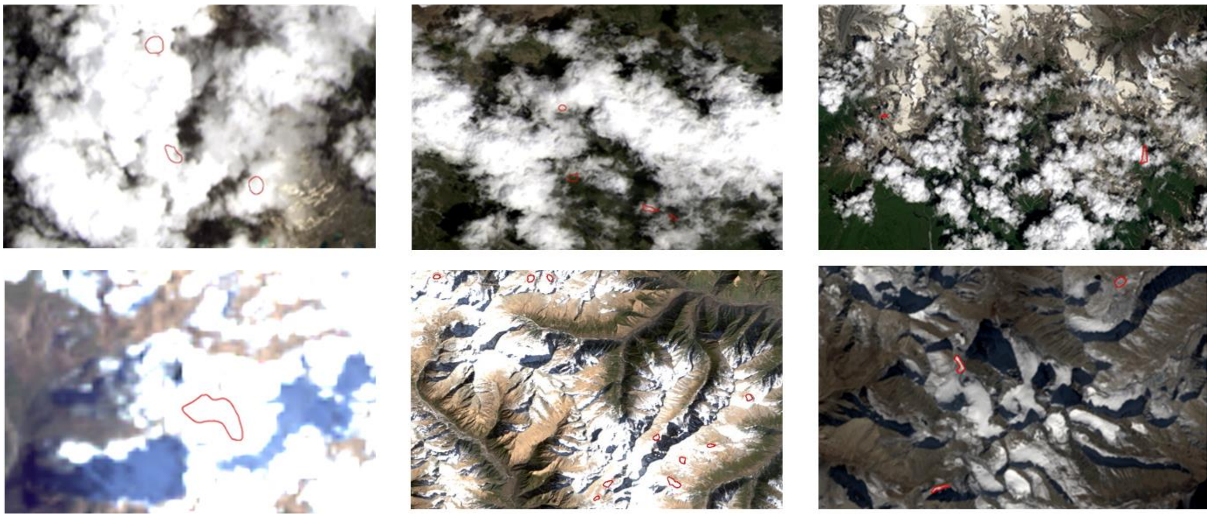


Figure I. Some glacial lakes that are always covered by cloud (first row) or snow (second row) in the true color composites (Bands:7, 4, 3) of the Landsat 8 OLI images for the year of 2017. The red contours refer to the 2018 glacial lake outlines digitized by Wang et al.

ii) Many glacial lakes dried up or outburst, and thus vanished in the image of 2017 (Figure II).

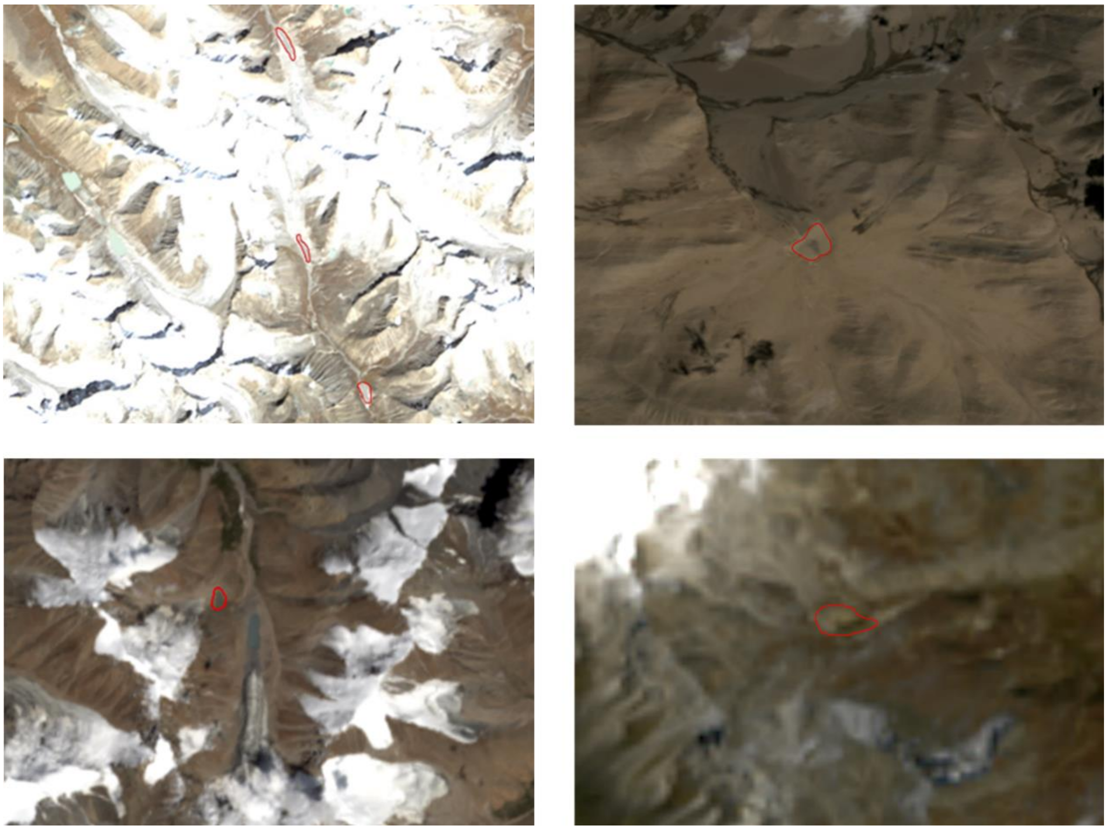


Figure II. Some glacial lakes dried up or outburst in 2017. Background images are the true color composites (Bands:7, 4, 3) of the Landsat 8 OLI images for the year of 2017. The red contours refer to the 2018 glacial lake outlines digitized by Wang et al.

iii) Lakes or ponds located in the middle of the river, which were not judged to be glacial lakes in our Hi-MAG database (Figure III).

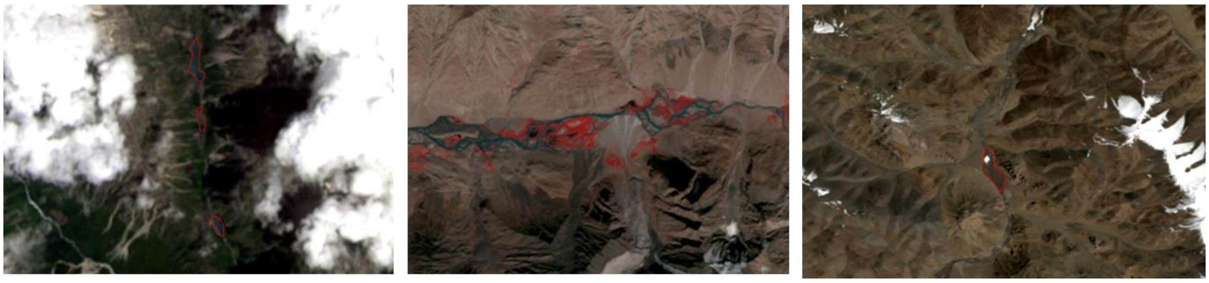


Figure III. Lakes or ponds located in the middle of the river, and were not judged to be glacial lakes in our inventory. Background images are the true color composites (Bands:7, 4, 3) of the Landsat 8 OLI images for the year of 2017. The red contours refer to the 2018 glacial lake outlines digitized by Wang et al.

To test the spatial correlation of glacial lakes distribution in two datasets, we compared the statistics in glacial lake number and area aggregated on a 0.1°×0.1° grid for the HMA regions. The results for the total glacial lake area, areas for glacial lakes larger than $0.04km^2$, and number for glacial lakes larger than $0.04km^2$ are depicted in Figure IV. A clear and strong correlation can be observed for all the statistics between our Hi-MAG dataset and glacial lake data by Wang et al. Most of the points being distributed around the 1:1 line, which shows that there is great consistency in the results.

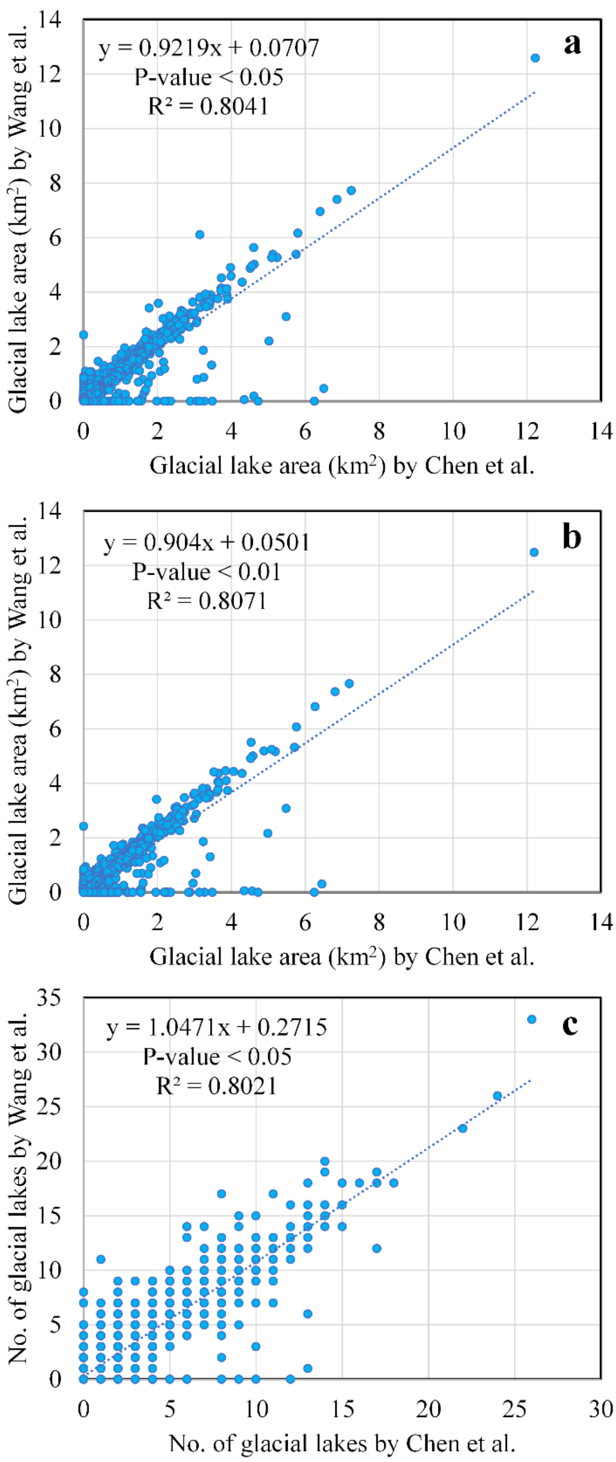


Figure IV. Comparison the results of (a) total glacial lake area, (b) areas for glacial lakes larger than 0.04km^2, and (c) number for glacial lakes larger than 0.04km^2 summed over a 0.1°×0.1° grid between Hi-MAG database and inventory by Wang et al., 2020.

In order to quantitatively and systematically evaluate the accuracy of our product, we implemented a stratified random sampling (Song et al., 2017;Stehman and SV, 2012), where the glacial lakes were divided into four strata. The sample sizes are the spatial resolution (30 m) of the data, and the strata are designed as: C0W0. both the results are non glacial lakes;

C0W1. non glacial lake for Chen’s and glacial lake for Wang’s; C1W0. glacial lake for Chen’s and non glacial lake for Wang’s; C1W1. Both the results are glacial lakes.

A total of 4,000 points were randomly selected, as shown in Figure V. The sample number for C1W1 and C1W0 are 1300 and 700, respectively, which is almost the same ratio between the total areas for the two strata (1450.50 km^2 vs 732.77 km^2). Because of the approximate total area with C1W0, we also randomly selected 700 samples from stratum C0W1. The rest 1300 samples are from C0W0.

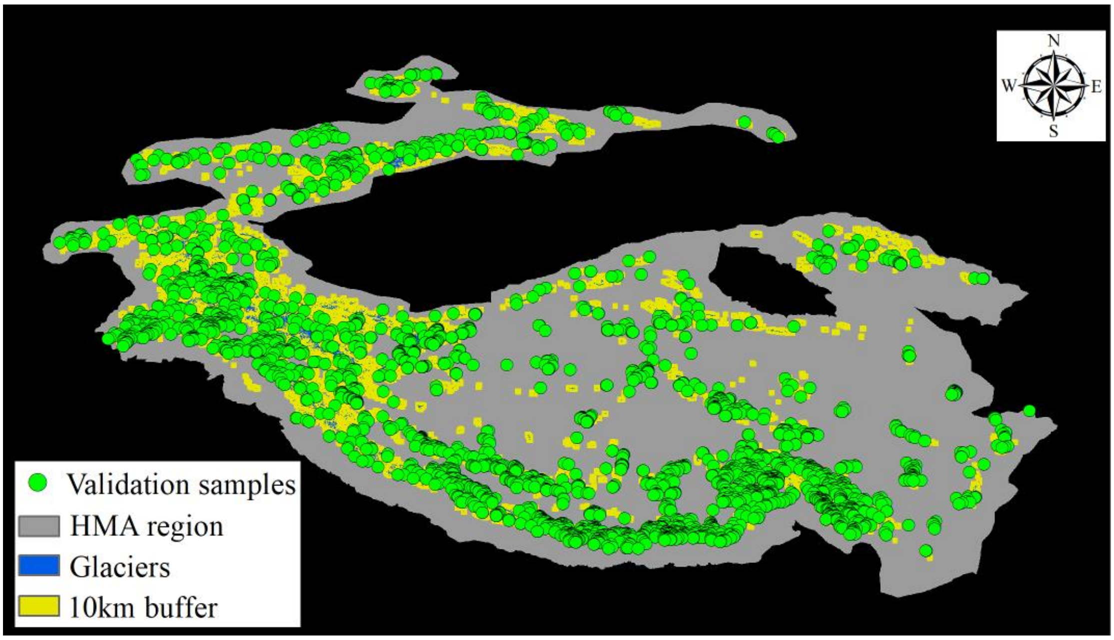


Figure V. Distribution of validation samples selected using stratified random sampling. Blue polygons are glacier outlines taken from the Randolph Glacier Inventory (RGI v5.0). Yellow polygons refer to buffer area within 10 km of glacier terminals.

Every validation sample was visually examined using Landsat imagery and higher-resolution imagery in Google Earth. Sample pixels were interpreted by a regional glacial lake mapping expert, and ambiguous samples were cross-validated by a second observer. If a sample is difficult to interpret, it was marked as ambiguous sample and excluded for the accuracy assessment. The sample number estimates were produced for each of the four strata (Table I), and these strata totals were then summed to obtain the total accuracy.

For the 1300 pixel samples that were considered to be non glacial lake by both datasets, after the pixel by pixel verification, 1215 were indeed non-glacial lakes, while 37 were the missed glacial lakes. In contrary, 1260 out of 1300 pixels belongs to the class of glacial lake, and 25 pixels were misclassified as glacial lake by the two inventories. 307 error pixels were found in the results from Wang et al., constitute about half of the total validation number. For the glacial lakes identified only by our inventory, 678 out of 700 were corrected classified. Our results yielded high overall classification accuracy (88%), user’s accuracy (97%), and producer’s accuracy (82%) for glacial lake classification using Landsat data, which further confirm the improved quality of the Hi-MAG database.

Table I. Statistical results of stratified random sampling.

Strata	Total pixel number	Total area (km^2)	Sample number	Sample No. of non glacial lake	Sample No. of glacial lake	No. of ambiguous sample
C0W0	2,022,448,650	1,820,203.78	1300	1215	37	48
C0W1	925,449	832.90	700	307	362	31
C1W0	814,196	732.77	700	21	678	1
C1W1	1,611,668	1,450.50	1300	25	1260	15

Based on the new version of glacial lake inventory, we also made comparisons of the statistical results between our Hi-MAG database and other studies from Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a. It should be noted that these inventory data have not been released publicly yet, except for GSW dataset produced by Pekel et al. Pekel et al. used millions of Landsat satellite images to map global surface water, many transient melt ponds and streams on the debris-covered glaciers were also identified here as permanent water surfaces, meanwhile, large quantity of small glacial lakes were missing. For the sake of a reliable comparative analysis, glacial lakes in the GSW were further extracted using the range of glacier buffer (10 km). We therefore conducted a rough comparisons with statistical results from Nie et al., 2017;Zhang et al., 2015, and qualitative comparison of the lake extent between GSW dataset and our Hi-MAG database summed by mountain range in 2015.

Hi-MAG lake number was 7268 higher and area was 644.26 km^2 higher than the estimation for the Tibetan Plateau (Zhang et al., 2015). The largest discrepancy is in the Gangdise, Himalaya and Nyainqentanglha Mountains in 2010. Across the Himalaya, we found 476.09 km^2 of glacial lakes, 4.57% more than previous estimates in 2015 (Nie et al., 2017). In addition, we qualitatively compared the lake extent between publicly available high-resolution Global Surface Water (GSW) dataset and our Hi-MAG database summed by mountain range in 2015. GSW data can be accessed at https://global-surface-water.appspot.com/download. For the sake of a reliable comparative analysis, lake polygons in the Hi-MAG dataset were converted into the grid format, and glacial lakes in the GSW were further extracted using the range of glacier buffer (10 km). Hi-MAG detected more glacial lakes in the Himalaya, Eastern Hindu Kush, and Tien Shan, and fewer in Eastern Pamir and Western Kunlun Shan. Fig. A5 illustrates the differences between our Hi-MAG glacial lake results and GSW-derived lake area for the whole HMA region.

The glacial lake area observed in our lake dataset in the Eastern Pamir and Western Kunlun Mountains does not conform to the mapped surface water in the GSW for these sub-regions. While there are numerous glacial lakes from an open water perspective, actually part of them are river segments. Additionally, the Himalaya, Eastern Hindu Kush, and some other Tien Shan host thousands of glacial lakes that are not readily observable in the GSW product. Large discrepancies in mountainous glacial lake estimates preclude a significant consistency between GSW and our Hi-MAG lake data over the HMA region. The region with the highest consistency between GSW and Hi-MAG product is interior Tibet. There is little agreement for Tien Shan, where the weather is rainy and snowy in the region above 3000 m, and large amounts of ancient glacial deposits have been accumulated. Here glacial lakes are featured by small size, due to the influence of source glaciers and lake beds, as well as the water depth and sediment inflow, glacial lakes appear heterogeneous reflectance in the image. Errors could exist in datasets produced by automated classification, but we also did a detailed manual editing, so we were not relying exclusively on automatic classification. Karakoram regions seem to have fewer glacial lakes in our estimate, owing to the overestimation of surface water on debris covered glaciers in the GSW dataset.

Finally, we have added the related content in the **Section 1. Introduction**, to give a comprehensive introduction about the existing glacial lake inventories. In the **Section 6.1** of revised manuscript, we have conducted a deep analysis and comparison between our Hi-MAG dataset and glacial lake inventory from Wang et al., 2020 and other studies from Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a. Please see the revised manuscript for more details.

3) A thorough and quantitative uncertainty analysis of lake area was added in the Section 4 of revised manuscript, the error bars for the lake area, and confidence intervals for the estimated trends were also added throughout the paper.

A thorough and quantitative uncertainty analysis of lake area was added in the Section 4 of revised manuscript. In order to highlight differences in the tails of area uncertainties in each year, we have added a new figure (Fig. 4a) that plotted the distributions of uncertainties from 2008 to 2017. Fig. 4b was also added for the better visualization of the relationship of area uncertainties against areas of all the glacial lakes in HMA (taking the results in 2017 as an example).

Assuming an uncertainty of 1 pixel for the detected glacial lake boundaries, we calculated the systematic errors for the whole HMA region (Fig. 4). For the year between 2008 and 2017, the area uncertainty of each glacial lake generally ranged from 0.30% to 50%, with the mean value falling around the 17%, and standard deviation around 11% (Fig. 4a). The maximum and mean value of area uncertainty for the glacial lakes in 2010 are the lowest, while for the year of 2016, the corresponding statistics are highest, this can be attributed to the different factors. The maximum of area uncertainty of glacial lakes is related with the shape and size of a certain lake (as can be seen from equation (1)), but its mean value is equal to the sum of the area uncertainties of each glacial lake divided by the total number, which depend on the total number of glacial lakes in a year, and also the shape and area of each lake. Besides, a close relationship can be found between the area uncertainties and the sizes of the glacial lakes (Fig. 4b). Most of the large glacial lakes (area $\geq 0.04 km^2$) have the mean area uncertainty of about 7%. This systematic errors were more significant for the small-sized glacial lakes. We measured glacial lake down to $0.0081\ km^2$ (nine pixels in Landsat imagery), where systematic errors calculated by equation (1) were ~50%.

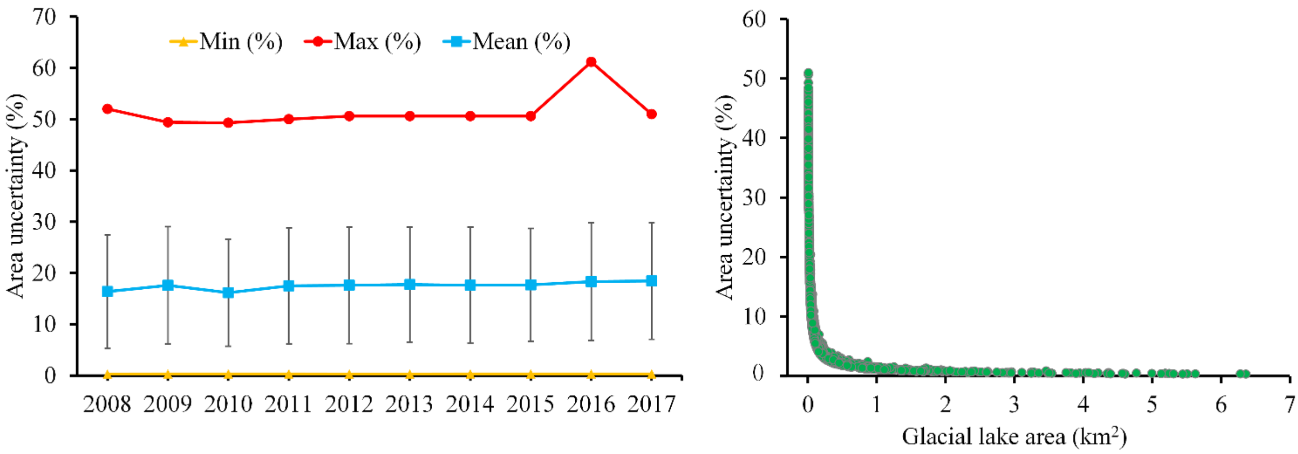
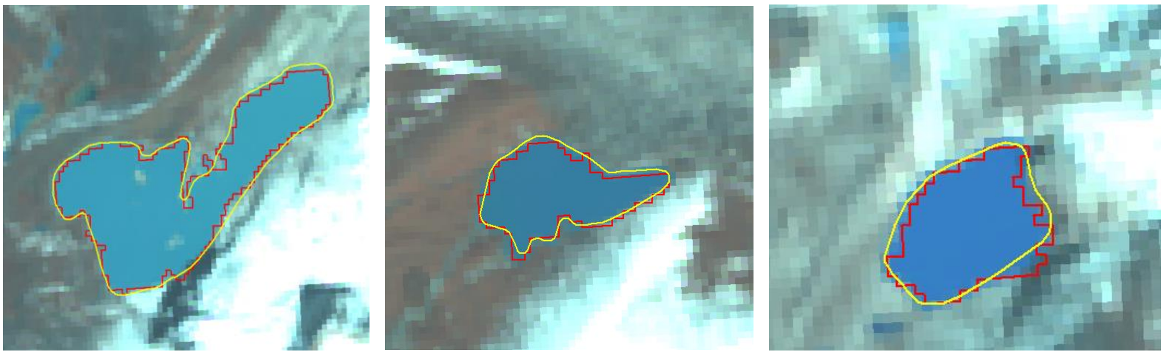


Fig. 4. (a) Statistics of area uncertainty (%) of glacial lakes for the years from 2008 to 2017. (b) Relationship between area uncertainties and areas of all the glacial lakes in HMA in 2017.

Besides, given the mapping uncertainty of glacial lakes and thus for the further improvement of reliability of calculated statistics, the error bars for the lake area, and confidence intervals for the estimated trends, and measure of uncertainty for the annual changes were also added throughout the paper.

4) Detailed descriptions about the key procedures for glacial lake mapping were provided in the Section 3.2, such as advantages of automated mapping, and the amount of work required for the manual editing.

For the development of HMA Glacial Lake Inventory (Hi-MAG) database, we applied a systematic glacial lake detection method that combined two steps from initial glacial lake extraction and subsequently manual refinement of lake mapping results. As for the automatic classification using GEE, there are four main procedures including (i) the clipped Landsat images by the extent of the glacier buffers and assembled into a time-series dataset; (ii) identified and masked some poor quality observations such as cloud, cloud shadow, topographic shadow; (iii) calculated MNDWI; and (iv) extracted the potential glacial lake areas by applying adaptive MNDWI threshold. Based on the automated processing, nearly 60% glacial lakes in each year can be correctly classified, of the other lakes that were not properly classified, 30% were missed and 10% were misclassified (mainly mountain shadows and river segments). For such a large-scale area that characterized by various and complex climatic, geological and terrain conditions, this classification method is simple but effective, the results are also reasonable since it provides very low commission errors. As shown in Figure VI, glacial lake outlines extracted using the automated classification method in our study fit the real boundary of the glacial lake very well, while manually delineated glacial lake outlines are largely influenced by people's subjective experience and manual operation, resulting in overestimation for most part of a glacial lake region, and underestimation for a few areas. Moreover, results from manual digitization show poor performance for the delineation of the some glacial lakes with complex curved shapes.



Manually delineated lake outlines from Wang et al. 2020
Automatically extracted glacial lake outlines

Figure VI. Examples of extraction of glacial lake outlines using the automated classification method in our study and manual digitization from Wang et al. 2020.

The estimated 30% errors in the initially classified map were from the missing small glacial lakes. Glacial lakes with small sizes are more likely to be confused with surroundings due to its less effective spectral, textural and spatial information in comparison with those of relatively large glacial lakes. Because the automated method is mainly based on the spectral features for the glacial lake mapping over the large mountainous area, the spectral information provided by the small glacial lakes will be incomprehensive and insufficient for the accurate detection of this kind of glacial lakes under various and complex environmental conditions. Therefore, visual interpretation and manual editing is still an effective way to ensure the high accuracy of lake inventories. In the updated version of dataset, we have carefully examined the lake data for the ten year by supplementing glacial lakes that have an area of greater than 0.04 km^2, and add their attribute information. We tried our best to improve the quality of glacial lake dataset, and make sure it is greatly optimized.

Finally, in the **Section 3.2 Adaptive glacial lake mapping method** of revised manuscript, we have added some text to give detailed explanations about the whole mapping procedures. The amount of work required for the manual editing were also described clearly. Please see the revised manuscript for more details.

2. *The current attribute table of this inventory is too sample, that it even did not give an ID for each lake. Glacial lakes should be indexed with unique ID that could be used to connect with RGI or GLIMS glacier inventory dataset. In addition, the abbreviations used in the dataset (PGL, UCL and SGL) were totally not mentioned in the main text of the paper, we don't know what they meant.*

Response:

We are terrible sorry for the incomplete attribute information of our inventory. We have carefully modified the attribute table of the new version of inventory by adding the unique ID for each glacial lake (attribute item GL_ID), for example, GL075720E40943N that formed by 'GL'+ 'longitude of centroid' + 'latitude of centroid', and retain three decimal places. Besides, the abbreviations of glacial lake type including PGL, UCL, SGL and IGL were all replaced by their full name, i.e. proglacial lake, unconnected glacial lake, supraglacial lake and ice-marginal lake.

3. *For a dataset paper, it should avoid including any further analysis on the data (for example, the inter - annual variability of lake area presented in the section 5.2), especially when their results exist large uncertainty or errors.*

Response:

We greatly appreciate the suggestion of the Reviewer. According to the response to your Comment 1, the dataset was updated by supplementing glacial lakes that have an area of greater than 0.04 km^2 for the ten year. As for the estimation of area error, most of the large glacial lakes (area $\geqslant$ 0.04km^2) have the mean relative error of about 7%. We also measured glacial lake down to 0.0081 km^2 (nine pixels in Landsat imagery), where relative errors calculated were ~50%. It can be found that this systematic error was more significant for the small-sized glacial lakes. Given the large uncertainty in the area and number of small glacial lakes, we should avoid including any further analysis on the data, especially for small glacial lakes that have large uncertainty. Therefore, we have carefully modified the whole manuscript to make sure the statistics and related analyses all focus on large-sized and all sizes of glacial lakes over the whole HMA region or different mountain ranges to discuss spatial and temporal variability for the 10-year record, and meantime, to avoid analysing and explaining the trends and number for small-sized glacial lakes. Fig. 5 shows the inter-annual variations in the number of small glacial lakes for different HMA mountain regions, we have removed the Section 5.2 in the revised manuscript.

Aggarwal, S., Rai, S. C., Thakur, P. K., and Emmer, A.: Inventory and recently increasing GLOF susceptibility of glacial lakes in Sikkim, Eastern Himalaya, Geomorphology, 295, 39-54, https://doi.org/10.1016/j.geomorph.2017.06.014, 2017.

Ashraf, A., Naz, R., and Roohi, R.: Glacial lake outburst flood hazards in Hindukush, Karakoram and Himalayan Ranges of Pakistan: implications and risk analysis, Geomatics, Natural Hazards and Risk, 3, 113-132, 10.1080/19475705.2011.615344, 2012.

Brun, F., Berthier, E., Wagnon, P., Kääb, A., and Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016, Nature Geoscience, 10, 668-673, 2017.

Li, D., Shangguan, D., and Naveed, A. M.: Glacial Lake Inventory Derived from Landsat 8 OLI in 2016–2018 in China–Pakistan Economic Corridor, International Journal of Geo-Information, 9, 294, 2020.

Luo, W., Zhang, G., Chen, W., and Xu, F.: Response of glacial lakes to glacier and climate changes in the western Nyainqentanglha range, Science of The Total Environment, 735, 139607, https://doi.org/10.1016/j.scitotenv.2020.139607, 2020.

Nie, Y., Sheng, Y., Liu, Q., Liu, L., Liu, S., Zhang, Y., and Song, C.: A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015, Remote Sensing of Environment, 189, 1-13, 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, 418-422, 2016.

Rounce, D., Watson, C. S., and Mckinney, D.: Identification of Hazard and Risk for Glacial Lakes in the Nepal Himalaya Using Satellite Imagery from 2000–2015, Remote Sensing, 9, 654, 2017.

Salerno, F., Thakuri, S., D'Agata, C., Smiraglia, C., Manfredi, E. C., Viviano, G., and Tartari, G.: Glacial lake distribution in the Mount Everest region: Uncertainty of measurement and conditions of formation, Global & Planetary Change, 92-93, 30-39, 2012.

Shugar, D. H., Burr, A., Haritashya, U. K., Kargel, J. S., Watson, C. S., Kennedy, M. C., Bevington, A. R., Betts, R. A., Harrison, S., and Strattman, K.: Rapid worldwide growth of glacial lakes since 1990, Nature Climate Change, 10.1038/s41558-020-0855-4, 2020.

Song, X. P., Potapov, P. V., Krylov, A., King, L. A., Bella, C. M. D., Hudson, A., Khan, A., Adusei, B., Stehman, S. V., and Hansen, M. C.: National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey, Remote Sensing of Environment, 190, 383-395, 2017.

Stehman, and SV: Impact of sample size allocation when using stratified random sampling to estimate accuracy and area of land-cover change, Remote Sens Lett, 2012.

Sudan, B. M., Pradeep, K. M., Wu, L., Gao, X., Finu, S., Rajendra, B. S., Narendra, R. K., Samjwal, R. B., Sharad, J., Sonika, S., and Prasant, B.: The status of glacial lakes in the Hindu Kush Himalaya-ICIMOD Research Report 2018/1, ICIMOD Research Report 2018/1, 2018.

Ukita, J., Narama, C., Tadono, T., Yamanokuchi, T., Tomiyama, N., Kawamoto, S., Abe, C., Uda, T., Yabuki, H., and Fujita, K.: Glacial lake inventory of Bhutan using ALOS data: methods and preliminary results, Annals of Glaciology, 52, 65-71, 2011.

Wang, X., Ding, Y., Liu, S., Jiang, L., Wu, K., Jiang, Z., and Guo, W.: Changes of glacial lakes and implications in Tian Shan, central Asia, based on remote sensing data from 1990 to 2010, Environmental Research Letters, 8, 575-591, 2013.

Wang, X., Guo, X., Yang, C., Liu, Q., Wei, J., Zhang, Y., Liu, S., Zhang, Y., Jiang, Z., and Tang, Z.: Glacial lake inventory of High Mountain Asia (1990-2018) derived from Landsat images, Earth System Science Data, 1-23, doi: https://doi.org/10.5194/essd-2019-212, 2020.

Yoon, Y., Kumar, S. V., Forman, B. A., Zaitchik, B., and F., K., Yonghwan, Qian, Yun, Rupper, Summer, Maggioni, Viviana, Houser, Paul, Kirschbaum, Dalia, Richey, Alexandra, Arendt, Anthony, Mocko, David, Jacob, Jossy, Bhanja, Soumendra, and Mukherjee, Abhijit.: Evaluating the Uncertainty of Terrestrial Water Budget Components Over High Mountain Asia, Frontiers in Earth Science, 7, 1-21, 2019.

Zhang, G., Yao, T., Xie, H., Wang, W., and Yang, W.: An inventory of glacial lakes in the Third Pole region and their changes in response to global warming, Global & Planetary Change, 131, 148-157, 2015.

Zhao, L., Ding, R., and Moore, J. C.: Glacier volume and area change by 2050 in high mountain Asia, Global and Planetary Change, 122, 197-207, 2014.