Anonymous Referee #2

General comments:

1. Currently, the authors give a wrong link to download the sharing data that is unavailable or need a registered account to download the data, it is not convenient. Actually, I find it is okay to directly download without any registration at http://www.crensed.ac.cn/portal/metadata/706ce17f-1684-4e8d-bf5e-7d517e03693c, so, why not replace with this link in the text?

It has been replaced by the new link.

2. There are many reference errors in the current text that have to be corrected carefully, I did not find all of them out, the authors have to go through. Such as, Yang et al., 2019 in the text but not listed in the reference section, P19-L24, Chaohai L, wrong surname, the same to P20-L1, L4, P21-L20, P23-L10.

We have carefully rechecked all the references in the text and reference list and corrected the errors and mistakes.

3. A technical question, how do you distinguish non-glacial lakes within 10-km of reference glaciers with defined glacial lake? For example, the 2010 Ataabad landslide dammed lake. In the alpine area of HMA, there are many landslide-dammed lakes that have no relationship with glaciology but are mainly supplied by glacier meltwater. Could you provide more detailed information about this?

The lakes related to glaciers or glaciation in the alpine cryosphere (defined within 10 km-buffer area of glacier terminals) have all been recorded as glacial lakes. Operationally, all the lakes located at the 10 km-buffer area of glacier extents have been recorded as glacial lakes in this manuscript. Some lakes which have no relation to glaciers or glaciation (i.e., non-glacial lakes) in the alpine cryosphere were possibly cataloged. We arbitrarily think the number of non-glacial lakes is relatively small and recorded all the lakes within 10 km-buffer area of glacier extents as (1) it is difficult to distinguish these non-glacial lakes from glacial lakes based on remote sensing data and (2) recoding all the lakes is conducive to evaluating the entire water source in the alpine cryosphere. Thus, the non-glacial lakes are usually not distinguished if located near the modern glaciers when glacial lake inventory is carried out in most documents.

We have tried to exclude the lake formed by man-made dam and/or long and narrow water body on river which could be river flood or wetland distributed along river. The Ataabad lake dammed by landslide in Huza river has been fed by glacier melting water since it was formed in Jan., 2010. By definition, the naturally formed Ataabad lake is located within 10 km-buffer area of glacier extent and received melting water should be recorded as glacial lake. Unfortunately, Ataabad lake was falsely catalogued in the previous inventory due to its long and narrow shape (~23km long). We have added the Ataabad lake in our updated glacial lake inventory of HMA. Furthermore, the long and narrow water bodies (usually from several kilometers long to tens) developed along river have been reexamined in the updated glacial lake inventory.

The technical comments have been discussed in the adding section of “7 comparison and limitation”.

4. About the minimum mapping unit, how do you consider selecting 0.0054 km^2 as a threshold value? Currently written as “the minimum glacial lake area recorded was set at 0.0054 km^2 (e.g., 3–4 pure lake water body pixels with approximately 12 mixed boundary pixels) because a lake area covering fewer than three pure lake water pixels could possibly have an error of >100 %”, I am confused by this writing. My understanding is that it is difficult to digitize 3-4 pixels by manual

interpretation. Otherwise, it is 6 pixels, equaling to 0.0054 km^2 using Landsat images.

I am sorry to make mistakes about the expressions. The manual delineation was required for approximately 1/2, 1/8, or 7/8 of the peripheral mixed pixels surrounding pure lake water body pixels. For example, in Fig. 1A, three pure lake water body pixels were possibly surrounded by maximum twelve peripheral mixed pixels. Theoretically, the ratio of pure lake water body pixels area to mixed peripheral mixed pixels area can be:

$$\frac{\text{Area of peripheral mixed pixels}}{\textit{Area of three pure lake water body pixels}} \times 100\% = \frac{6\times\frac{1}{2}+5\times\frac{1}{8}+1\times\frac{7}{8}}{3} \times 100\% = 150\%.$$

Similarly, in Fig. 1B, four pure lake water body pixels were also possibly surrounded by maximum twelve peripheral mixed pixels, and the ratio of pure lake water body pixels area to mixed peripheral mixed pixels area can be:

$$\frac{\text{Area of peripheral mixed pixels}}{\textit{Area of four pure lake water body pixels}} \times 100\% = \frac{8\times\frac{1}{2}+4\times\frac{1}{8}}{4} \times 100\% = 112.5\%.$$

Then when area error within one standard deviation (1σ) (Hanshaw and Bookhagen, 2014) was considered (i.e., the error adjusted coefficient of 0.6872 was used), the error (1σ) of lake area with three pure lake water body pixels and four lake water body pixels can be expressed as:

$$\mathrm{E} = \frac{\text{Area of peripheral mixed pixels } \times 0.6872}{\textit{Area of three pure lake water body pixels}} \times 100\ \% = 1.5 \times 0.6872 \times 100\% = 103.1\%.$$

And

$$\mathrm{E} = \frac{\text{Area of peripheral mixed pixels } \times 0.6872}{\textit{Area of four pure lake water body pixels}} \times 100\ \% = 1.125 \times 0.6872 \times 100\% = 77.3\%.$$

So, in the Landsat images, 3–4 pure lake water pixels could possibly be surrounded by approximately 12 mixed boundary pixels with the total area equivalent to about 3 peripheral mixed pixels, i.e., 6 pixels area (equals to 0.0054 km^2) or more in total lake water body. When the threshold of 3–4 pure lake water pixels is set as minimum recorded lake water body, the uncertainty could theoretically amount to about 100%, and a lake area covering fewer than three pure lake water pixels could possibly have a relative error of >100 %.

We have further rewritten the arduous expressions (P12–L16–30) and Figure 4. has also been reedited. (P13–L1–8)

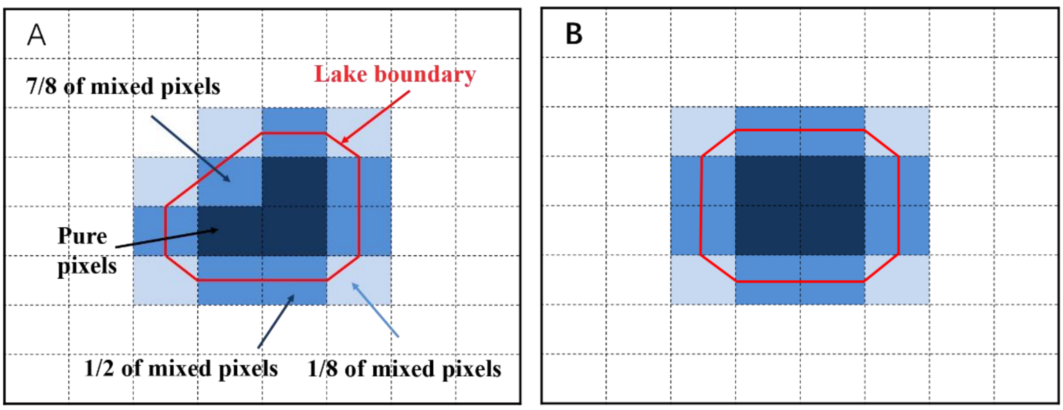

Fig. 1 Sketch maps showing the 3–4 pure lake water body pixels with approximately 12 mixed boundary pixels

Reference:

Hanshaw, M. N., and Bookhagen, B.: Glacial areas, lake areas, and snow lines from 1975 to 2012: status of the Cordillera Vilcanota, including the Quelccaya Ice Cap, northern central Andes, Peru, The Cryosphere, 8, 359-376, doi:10.5194/tc-8-359-2014, 2014.

5.I suggest writing a further revision plan in the end of this draft to point out the shortage of current glacial lake data. Actually, the data have been published in a data sharing platform, and some errors exist inevitably in terms of two times manually vectorization for the same lake, for example, induced by wrongly digitizing, maybe operated by different operatives. Do you have any plan to update the data? Once the data updated where to share? In what kind of ways?

Several limitations deserve proper consideration when using the glacial lake inventory data. First, a degree of uncertainty has resulted from using Landsat image data that covered different periods, i.e., both interannually and seasonally. Although images acquired in summer or autumn (June–November) have been set as optimal choices, the selected images covered most seasons of the year, e.g., the images selected in June–November accounted for only 72.3 and 88.8 % of the total number in 2018 and 1990, respectively. Interannually, images were selected from a span of 10 years (1986–1995) and 4 years (2016–2019) to obtain sufficient high-quality images of the HMA area. Second, this study has recorded all lakes located within the 10 km buffer area of glacier extent as glacial lakes. Therefore, certain lakes that have no relation to glaciers or glaciation (i.e., non-glacial lakes) in the alpine cryosphere were potentially catalogued in error because of the difficulty in distinguishing non-glacial lakes from glacial lakes based on remote sensing data. Third, we have identified water bodies related to glaciers or glaciation in the alpine cryosphere as glacial lakes. However, in many cases, it was difficult to determine whether such bodies should be recorded as glacial lakes, e.g., cases of long narrow water bodies on rivers and cases where the number of pure water body pixels was small. Thus, some errors and inconsistences were inevitable because the lake boundary vectorization and inspection were performed by different operatives. In the future, this glacial lake inventory will be updated and shared on the National Special Environment and Function of Observation and Research Stations Shared Service Platform (China), and further water source evaluation and hazards assessment would be carried out in our next research schedule.

The main limitations of current glacial lake inventory data have been added in line (P20–L5–21).

6. I also suggest authors to polish the language once again, and some sentences are arduous to follow.

The language of the revised manuscript has been polished by a professional language retouching company.

Specific comments:

P1-L26, update the link

It has been updated.

P2-L25, recognizing→revealing

It has been modified.

P3-L6, Yang et al., 2019, not listing in the reference; L18, a Landsat imagery series?

The reference of Yang et al., 2019 has been added; “a Landsat imagery series” has been modified as “Landsat images”

P4-L5, only Antarctic? What about Arctic? L8 “the primary source of both lake basin formation”, I think no relationship.

L5: It has been modified as “The HMA area has the largest surviving glaciers of any region other than the polar regions”

L8: It has been modified as “which was the primary source of water supply for the development of glacial lakes”

P5-L13, suggest revise as circa 1990 and circa 2018.

It has been revised.

P6-L19, 20, what scale do you keep while an operator did computer screen vectorization of mixed pixels?

The viewing scale of 1： 10000 has been kept and it has been modified as “e.g., viewing scale of 1:10,000 on a computer screen vectorization of mixed pixels”.

P7-L12, -0.1 of NDWI to extract lake extent? Generally, this value is greater than 0.

Liu et al. (2016) and Du et al. (2014) suggested that it might be preferable to set the optimized threshold of $NDWI_{GREEN/NIR}$ of Landsat OLI images to −0.05. By considering the edge effects according to the mixed pixels, this study has initially selected a lower optimal threshold (approx. −0.1) for specific images to obtain the maximum water body (Fig. 2). Then, higher thresholds have been tested for visual water extraction before selecting the suitable threshold (varied in the range of −0.10 to 0.20).

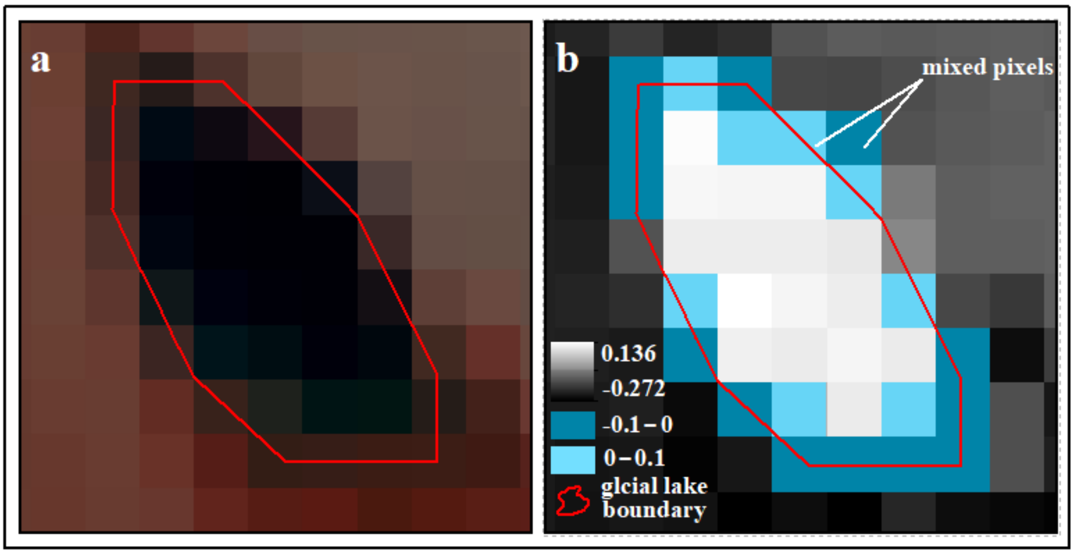

Fig. 2 An example showing the optimal thresholds for $NDWI_{GREEN/NIR}$. (a) original OLI image taken in 2018; (b) the $NDWI_{GREEN/NIR}$ image

Reference:

Du, Z., Li, W., Zhou, D., Tian, L., Ling, F., Wang, H., Gui, Y., and Sun, B.: Analysis of Landsat-8 OLI imagery for land surface water mapping, Remote Sens. Lett., 5, 672-681, doi:10.1080/2150704x.2014.960606, 2014.

Liu, Z., Yao, Z., and Wang, R.: Assessing methods of identifying open water bodies using Landsat 8 OLI imagery, Environ. Earth Sci., 75, doi:10.1007/s12665-016-5686-2, 2016.

P9-L3, Weicai et al.,?

It has been revised as "Wang et al., 2014".

P10-L3, 10 km from modern glacier terminals? Or glacier extent?

The "glacier extent" may be more exact and the relative expressions has been revised.

P10-L12-15, given a reasonable classification of lakes, why did not you take this?

The classification schema by Yao et al. (2018) is difficult to carry out based on available remote sensing data for so large area as HMA. For it is a little difficult to distinguish glacial lake type in terms of material properties, topographic features, and phase of lake formation using remote sensing imagery.

P10-L19-22, it is not clear what your point is? "because of the lack of sufficient amounts of remote sensing data with appropriate resolution."

It has been revised as "because of the lack of remote sensing data with satisfied spatial resolution."

P10-L23, two types: glacier-fed lakes and non-glacier-fed lakes? The significance becomes very limited by too simple classification system. Maybe more types, such as pro-connected lake and supraglacial lake, be cataloged. But being cautious, once you modified the data, meaning that you have to update the sharing data on the platform online.

The glacier-fed lakes were further subdivided into three sub-classes: supraglacial lakes (lakes developed on glacier surface), ice-contacted lakes (lakes contacting the glacier terminal or margin), and ice-uncontacted lakes (lakes not contacting the glacier but fed directly by glacial meltwater). We have updated attribute items in the datasets and resubmitted them on the platform online (http://www.crensed.ac.cn/portal/metadata/706ce17f-1684-4e8d-bf5e-7d517e03693c).

It has been added in P10–L24-27.

P13-L14,15, why did not you record the date of used images? Only recorded the month and year?

The date of used images has been input in the attribute item of Lake time phase.

P14-L21, Narrate the accuracy of Trimble GeoXH6000 for a better understanding about the validation.

The accuracy of Trimble GeoXH6000 is in decimeter and has been added in line of P15–L18, L19.

P16, Figure 6, suggest adding a scale bar for each subset.

It has been added.

P16-L13 The HMA glacial lakes are located within the elevation range of 1600–6300m. while, P17-L10, L11, maximum distribution elevation of 6078 m in 1990 rising to 6247 m in 2018. Maybe use the relatively accurate value of elevation.

The accurate elevation value of 1357–6247m has been used to replace the elevation range of 1600–6300m in P17–L13.

P17, Back up to previous error, in Figure 7, the maximum X axis value is 6000 m, so you miss your lakes with maximum elevation.

I am sorry to miss the maximum X axis value and it has been updated to 6300m.

P18, L1, L2, How to prove that no observable trends were discovered in Karakoram and Western Kun Lun, Western Himalaya?

I am sorry that we presented an arduous expression. The sentence has been revised as "The expansion rate varied markedly and no observable trend in the rate of increase or decrease with elevation were discovered in Karakoram and Western Kun Lun, Western Himalaya, and Inner Tibet".

P18, suggest adding a section about the shortage and updating plan for this data, putting before Data availability

A section about the comparison, shortage and updating plan for this data has been added.

P18, replacing the existing link

It has been replaced.

P18, rewrite the sentences in L23-26, it is unclear. "Lake area expanded most in the higher elevation bands during 1990–2018. The data set has been developed as basic data for cryosphere hydrology research; however, it is expected that it could support practical utilization and management of water resources and assessment of glacier-related hazards in the HMA region"

It has been rewritten as "The data set is expected to provide basic data to support the cryosphere hydrology research, water resources utilization and management, and assessment of glacier-related hazards in the HMA region."