Detailed responses for the comments

The manuscript is devoted to topical issues of determining the number and total area of thermokarst lakes (TL) in the permafrost zone of the Qinghai-Tibet Plateau (QTP). Considering the vastness of the territory and the large number of TLs of different sizes, the determination of their area from remote sensing data with a relatively high spatial resolution (10 m) in itself is a serious methodological problem. The authors provided new experimental data on the amount and area of TL within the QTP, as well as interesting information on the relationship of TL with altitude levels, types of vegetation cover, and environmental factors. At the same time, there are a number of questions about the manuscript, mainly of a methodological nature, which require clarification.

Response: Thank you very much for your detailed suggestion. We revised the manuscript according the reviewers' comments, and we believe the quality of the revised version has been greatly improved.

1. How did thermokarst lakes differ from other types of lakes in the permafrost zone?

Response: The formation of thermokarst lakes is due to the surface water accumulation following ground subsidence during permafrost degradation. There are no strict criteria to distinguish the thermokarst lakes and non-thermokarst lakes. However, it is clearly that some tectonic lakes were not thermokarst lakes, which is characterized by large areas. In addition, more than 90% of lakes along the Qinghai-Tibet Highway have an area of less than 5,000 m^2, with an average area of 5,039 m^2, and the largest thermokarst lake had an area of 4.49×10^5 m^2 (Niu et al., 2014). Thus, the area of thermokarst lakes in this study ranged from 350 m^2 to 5.0×10^5 m^2 and the tectonic lakes are excluded. Although there is a possibility that some of the lakes were formed without thermokarst genesis, this is the first dataset for the thermokarst lakes on the QTP.

These information has been added in the revised version in the Comparison and limitations section as follows:

In our study, we classified that thermokarst lakes according to the area (350 m^2 to 5.0×10^5 m^2) because the smaller ponds are likely seasonal water bodies and they are also can not be recognized from remote sensing data, while the larger lakes are maybe

tectonic lakes. However, due to there are no strict criteria to distinguish the thermokarst and non-thermokarst lakes, there is a possibility that some larger thermokarst lakes on the QTP were not included in our dataset, future studies pertained to these large lakes should paid more attention on the areas of these lakes.

2. What does 10a mean in parameter 19.5 cm / 10a, associated with the rate of increase in the active layer thickness (line 100)?

Response: This sentence has been rewritten as follows: "the active layer thickness along the Qinghai-Tibet Highway has increased at the rate of 19.5 cm/10a from 1982 to 2018".

3. It is known that the area of TL, as a rule, having an insignificant depth, significantly depends on the period of floods and precipitation that fell on the eve of the survey. Differences in the area of TL during floods and after it can reach 40%. In this regard, it is necessary to clarify - what are the dates of the space borne survey, the dates of the descent of flood waters and the dates of the passage of precipitation?

Response: Thanks for the review. The dates of the descent of flood waters and passage of precipitation is important for the area of thermokarst lakes. In this study, Sentinel-2A images on cloudless days during April 1 to October 30, 2018 were used for visual interpretation. Thus, the distribution of thermokarst lakes presents the average areas during the period, which is not affected by the extreme weather events. This has been also explained in the materials and methods section as follows: "To avoid the errors from the period floods and water passage and to obtain the high-quality data with less noise such as clouds, snow and clouds shadow, the Sentinel-2A images on cloudless days during April 1 to October 30, 2018 were used for visual interpretation."

4. Explain in comparison with what data the threshold NDWI = 0.1 was set?

Response: The use of the NDWI as 0.1 for mapping the lakes was possible in the Qinghai-Tibetan Plateau (QTP) due to its homogeneous landcover characteristics. Li and Sheng, (2012) has verified the threshold vale of 0.1, which found that the extraction

of the water body is accurate by using the mask of the potential lakes and eliminating the shadow of the mountains. What's more, through the comparison between automatic extraction of remote sensing image and visual interpretation in this study, it is also found that the threshold vale of 0.1 is best to extract the water body.

In the revised version, this has been explained as follows: "The threshold values for a large number of lake sample images were studied, and it was found that the value of 0.1 is enough to extract the area of the potential lake area (Li and Sheng, 2012). By comparing between automatic extraction of remote sensing image and visual interpretation in this study, we confirmed that this value is the best value to extract the water body."

5. How and with what data was visual interpretation carried out to ensure "lake boundary inspection with the highest quality control and consistency" (218)? From which resources and which images of bodies of water were downloaded from online for visual interpretation (220)?

Response: Sorry for the confusion. It has been revised as follows: "For the data process, the Sentinel-2A images are firstly processed by color synthesis, then the visual interpretation was carried out according to the information of ground feature spectrum and changes such as the hue and brightness of the images. Three common methods of visual interpretation were used in this study: (1) Direct judgment . It is used to determine the shoreline of the thermokarst lake and identify whether it is collapsing around the shore. (2) Comparison. It is used to determine thermokarst lakes by comparing with the existing materials and field monitoring thermokarst lakes. (3) Logical reasoning. It is used to determine thermokarst lakes through reasoning based on the comprehensive knowledge of geography, hydrology, soil science, and other related subjects. Although this is a time-consuming process, especially for such a large area, this method is useful to recognize the lake boundary with the highest quality control and ensured consistency.".

6. What does three months of visual interpretation mean (223)?

Response: The visual interpretation in this study is necessary but a time-consuming and labor-intensive task, and we took three months to complete that. To avoid any confusion, we deleted the three months in the revised version.

7. The data obtained show that even for large TLs, the errors are large (up to 20%) (Table 1), which indicates either the low accuracy of the method or the instability of the TL area, including large ones. For comparison with other results and errors in determining the area of TL according to remote sensing data, the authors are recommended to build a graph of the dependence of the modulus of the average and maximum error (or RMSE) on the area of water bodies for groups, for example, for groups of 400-1000 m, 1000-2000 m ... etc.

Response: Thanks for the suggestion. In the present version, the three indicators of R^2, MAE and RMSE were added for the verification of extracted data and filed monitoring data. In the results, it has been revised as follows:

We used the coefficient of determination (R^2), average absolute error (MAE) and root-mean-square error (RMSE) to present the accuracy assessment. The calculation methods (Draper and Smith, 1998) are as follows Eq. (3-5)

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \hat{y})^2} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z_{oi} - Z_{pi})^2} \quad (4)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Z_{oi} - Z_{pi}\right| \quad (5)$$

where y_i is the extracted area of Sentinel-2A thermokarst lakes, f_i is the measured value of UAV, Z_{oi} is the measured value of UAV in the i classification, Z_{pi} is the extracted value of Sentinel-2A, and N is the number of lakes in each classification.

Table 1 Accuracy verification of thermokarst lakes derived from Sentinel-2 data using unmanned aerial vehicle field monitoring

Area of thermokarst lakes (m^2)	N	R^2	RMSE	MAE
400–1000	10	0.82	149.27	108.28

1000–2000	12	0.83	285.00	202.61
2000–5000	15	0.78	599.84	493.54
5000–1000	6	0.68	1765.51	1314.39
10000–20000	5	0.68	1583.53	1369.15
20000–50000	3	0.99	1916.17	1769.44
>50000	5	0.90	30929.08	22822.81

8. The information about the errors (table 1) is in no way connected with the final result – the estimate of the total area of TLs in the QTP area. In this regard, it is not clear why table 1 is needed? At the same time, the information in table 1 can be used to assess the error in determining the total area of water bodies in the region. Any quantitative estimates are always associated with errors.

Response: In the present version, the table 1 has been revised (see response to Question 7). We also added the values of R^2, MAE and RASE to describe the accuracy of data verification.

9. What information about the number and total area of thermokarst lakes in the region of the entire QTP was previously obtained by other authors, for example, topographic maps of different scales? If there are such data (maps), then it is necessary to show what are the advantages of the data obtained by the authors in comparison with the known data (maps).

Response: The previous studies mainly focused on the thermokarst lakes at catchment scale on the middle of QTP, such as the area and changes of thermokarst lakes. The methods were quite different, and the results were obtained from different time. Therefore, it is difficult to compare with our findings. These references (Niu, F., Lin, Z., Hua, L., Lu, J., 2011. Characteristics of thermokarst lakes and their influence on permafrost in Qinghai–Tibet Plateau. Geomorphology 132, 233.

Niu, F., Luo, J., Lin, Z., Liu, M., Yin, G., 2014. Morphological Characteristics of Thermokarst Lakes along the Qinghai-Tibet Engineering Corridor. Arctic Antarctic & Alpine Research 46, 963-974.) were cited in the study.

10. In fig. 6 some data do not match the description in the text (268, 269, 270). So in

the text for an altitude of 4750-5000 m, the number of TLs is 59, 314 and the area they occupy is 874.24 km2. On the graph for this height, the number of TLs is more than 70,000, and the area is about 1800 km2.

Response: Sorry for the wrong description because there were many statistical charts during our analysis. These chart have been revised as follows:

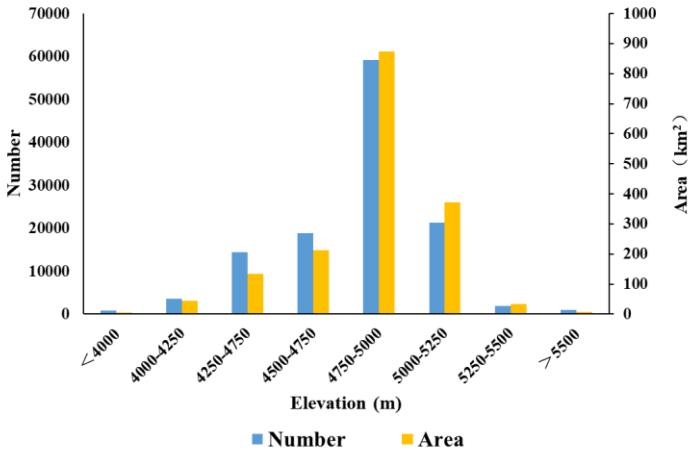

Figure 6: Number and area of thermokarst lakes at different altitudes on the QTP

11. The map in Fig. 9 does not coincide in contours with the map in Fig. 4.

Response: We recreated this figure using the same contours with Fig. 4.

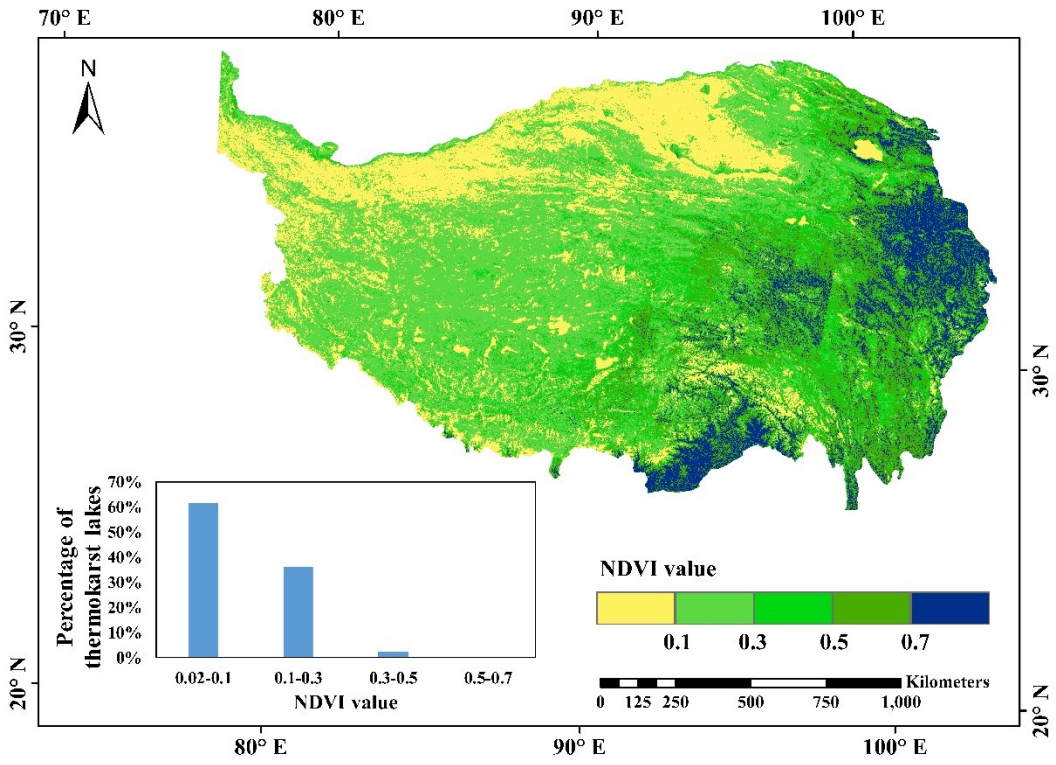

Figure 9: Distribution of thermokarst lakes under different NDVI values

The manuscript can be published after serious revision.

Response: We appreciate your constructive comments. We believe the quality of this manuscript has been highly improved. We are happy to address additional concerns.