Review of the manuscript
"High-resolution dataset of thermokarst lakes on the Qinghai-Tibetan Plateau"
Xu Chen, Cuicui Mu, Lin Jia, Zhilong Li, Chengyan Fan, Mei Mu, Xiaoqing Peng, Xiaodong Wu

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.

Questions and comments

1. How did thermokarst lakes differ from other types of lakes in the permafrost zone?
2. What does 10a mean in parameter 19.5 cm / 10a, associated with the rate of increase in the active layer thickness (line 100)?
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?
4. Explain in comparison with what data the threshold NDWI = 0.1 was set?
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)?
6. What does three months of visual interpretation mean (223)?
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.
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.
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).
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 km^2. On the graph for this height, the number of TLs is more than 70,000, and the area is about 1800 km^2.
11. The map in Fig. 9 does not coincide in contours with the map in Fig. 4.

The manuscript can be published after serious revision.