Comment on essd-2021-439
Anonymous Referee #3

Referee comment on "Retrogressive thaw slumps along the Qinghai-Tibet Engineering Corridor: a comprehensive inventory and their distribution characteristics" by Zhuoxuan Xia et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-439-RC3, 2022

Major comment:

Thaw slumps are an important phenomenon of permafrost degradation and have a significant impact on engineering, ecological processes, and the carbon cycle. This paper by Xia et al. achieves mapping thaw slumps on a large scale with high precision via combining deep learning and manually inspecting. The paper is generally well organized, the objectives are clear, and the methods are also well designed, for instance, using an iterative mapping method to find more thaw slumps with limited training data and assigning a probability for each mapped uncertain thaw slump. Therefore, the results are quite robust. To date, thaw slump investigations on the QTP are still urgent, and hence I think this important dataset would potentially serve as fundamental data for understanding the impacts of thaw slumps in the warming world. I, therefore, think this paper is a nice contribution that can be published in ESSD journal after minor revisions.

Specific comments:

- P2, L34: Permafrost definition is not originally described in French (2018), please see Van Everdingen, R.O. (1998)
- P2, L30: This sentence doesn’t seem to constitute causality. Please revise the wording and grammar.
- P2, L35: … of about 1.06×10^6 km^2
- P2, L48: Please put the references behind the corresponding content respectively, rather than putting them at the end together.
- P3, L66: "cryospheric studies" rather than "cryosphere studies"
- P3, L75: There are too many "and" in this sentence
- P9, L170: Do you mean the low probability is < 100%, and the high probability is =
Tables & Figures

- **Summary of RS data:** Could you please re-organize the description RS datasets in Sec.3? What about adding one more table regarding to their summary info? For example, data coverage, used bands, spatial resolution, and purpose of each dataset.
- **Figure 1:** Could you please add the lake info here so that authors could clearly see the missing data? You could use the public land cover maps, such as the ESA CCI or TP lake inventory from TPDC.
- **Figure 6:** (f), (g): I would suggest using bar charts instead of pie charts, so that the data may be more intuitive and easier to compare. The pie charts look a little messy.

Reference