

Earth Syst. Sci. Data Discuss., referee comment RC2
<https://doi.org/10.5194/essd-2021-257-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.



Comment on **essd-2021-257**

Anonymous Referee #2

Referee comment on "An inventory of supraglacial lakes and channels across the West Antarctic Ice Sheet" by Diarmuid Corr et al., Earth Syst. Sci. Data Discuss.,
<https://doi.org/10.5194/essd-2021-257-RC2>, 2021

Review of the manuscript "An inventory of supraglacial lakes and channels across the West Antarctic Ice Sheet" submitted to the ESSD special issue on "Benchmark datasets and machine learning algorithms for Earth system science data"

The manuscript presents an approach to map supraglacial lakes and channels for the year 2017 across the West Antarctic Ice Sheet on the basis of Sentinel-2 and Landsat-8 data and spectral thresholding. It complements the study of Stokes et al 2019 that used a similar approach to map supraglacial lakes around the margin of the East Antarctic Ice Sheet. The resulting inventory of mapped lakes and channels is suggested as benchmark datasets for Earth system science data.

Although the manuscript is generally well written and well presented, I have some major concerns, especially regarding the fit to the SI.

I don't think that the resulting inventory can be regarded as a benchmark dataset for machine learning. First, the inventory is the result of a modelling procedure (by spectral thresholding) and hence does not present reference data with "ground truth quality", which however would be required for a good benchmark dataset to be useful for algorithm and model comparison. Second, the dataset is limited to the labeled locations of lakes and channels but no predictor data are given so that the data cannot be used to test different machine learning approaches without extensive further data acquisition.

Further, I have concerns about the thresholding approach used to delineate water pixels. The authors applied a series of thresholds (see Fig. 2 and 3) to the spectral channels but it is unclear to me how these thresholds were derived. Is that based on try and error, and if so, how was the error assessed (visual? statistically by comparing to the manually digitized reference data?). I also wonder why the authors did not use a supervised classification algorithm but decided for a series of manually selected thresholds instead.

The validation is not complete. If I get it correctly, the accuracy/sensitivity/specificity values that are given in the manuscript refer to the water delineation. What is missing is the validation for the classification into channels and lakes.