

Comment on **essd-2022-224**

Anonymous Referee #2

Referee comment on "CALC-2020: a new baseline land cover map at 10 m resolution for the circumpolar Arctic" by Chong Liu et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-224-RC2>, 2022

The manuscript combined the multisource and multitemporal remote sensing imagery to develop the new circumpolar Arctic land cover product for circa 2020, containing 10 land-cover types especially for the tundra. The dataset is great and can be used to improve earth system modelling. However, I think there are several problems should be solved.

- In the second paragraph of the introduction, author stated that '**but their strength is limited by the over simplification of classification schemes**' and used the MCD12Q1 to explain the limitations of global land-cover products. In my opinion, the statement might be inaccurate, the ESA CCI_LC, Globeland30, and three global 10 m land-cover products (ESA, ESRI and dynamicworld) can capture the land-cover information with various classification systems or spatial resolutions. Therefore, I suggested that the authors further added the necessity of developing CALC-2020 when there were so many global land-cover products.
- In the third paragraph of the introduction, the reasons why authors used multi-source remote sensing data to develop the CALC-2020 map was not clear. Actually, the combination of multisource remote sensing observations has been demonstrated to improve the land-cover classification in so many studies because of importing new information. Namely, the combination of optical, SAR and terrain features was not particularly novel in this study.
- Authors used the FAST sample library and three land-cover products to derive the primary training samples and then used visual interpretation to collect lichen and moss, however, how to identify the graminoid tundra and shrub tundra samples, which was not explained in the method.
- Authors used the SOM method and further used the threshold of 75% to remove the wrong samples, why chose this threshold? Can you give the number of training samples for each land-cover class?
- In section 2.4.3, authors mentioned that "Cloud-free Sentinel-2 observations were first interpolated in each pixel at an 8-day time step using penalized cubic smoothing splines", is it mean these cloudy or snow pixels were interpolated by cubic splines method? Can you provide the details and the results of the time-series sentinel-2 observations after interpolating?
- The CALC-2020 was individually produced for each country using RFC models, how to ensure the spatial continuity between the transition areas over two neighboring

countries because of using different RFC models?

- The comparisons between CALC-2020 with three global maps were **unfair** for global land-cover products, because the CALC-2020 only focused on the circumpolar Arctic while the global land-cover maps emphasized the comprehensive performance over the globe.
- In the two comparison figures of Figure 6 and 7, I doubt the difference mainly came from the different definition for the same land-cover class especially in the Figure 7.
- I suggest to remove the statistics figure in Figure 9b, because it cannot provide useful information for the CALC-2020 map.
- Figure 10 about the feature importance is interesting and vital for developing the CALC-2020, I think the author should explain in more details why the topography had highest importance while the optical features were unimportant.
- I suggest that authors shared the validation data in the **Data availability**

Line 334, Figure 30 should be 10.