Dear Charles Amory,

Thank you for your work on our manuscript. Below you find our responses to the reviewer's questions. We implemented the corresponding technical revisions and clarifications into the manuscript. Please find the revised version together with a track change file.

Thank you and best wishes,

Erik and all co-authors

Anonymous referee #3, 16 Oct 2024

The paper by Loebel et al. presents a new dataset of calving front locations, derived from deep learning techniques, in the Antarctic Peninsula. The paper has already gone through thorough rounds of revision and is now almost ready to be published. I have only minor comments on the methodology used here:

(1) The input image size for the ANN is 512x512 pixels. Did you perform a sensitivity analysis of the final results regarding this input data size? What would be the impact of expanding the image size? Is it just more computational time, or would the model also learn more slowly (since there would be more information in a single picture)?

Our input tile size of 512 pixels by 512 pixels, together with the 30 m image resolution of Landsat-8/9 equals about 15 km by 15 km. This value was selected as the majority of Greenland outlet glacier fronts fit into this window (*Loebel et al., 2024*) while still ensuring that we don't have to downsample the input data. This value also works for AP outlet glaciers as all but one glacier (Hektoria-Green-Evans) fit in this input window.

Increasing the input tile size further would therefore not result in more ice-ocean boundary, but would simply include more land and ocean area. This would cause our reference dataset to lean more towards the land-ocean boundary and less towards the ice-ocean boundary. As the ANN model is trained to segment the entire input tile, it would be optimized more towards the land-ocean boundary than the ice-ocean boundary, which is not what we want for calving front extraction. In addition, the computational cost will increase.

In the current version of the manuscript the values of 512 pixels by 512 pixels seem a bit arbitrary. We thank the reviewer for this question and include an explanation in section 2.1 (P4 L81).

(2) The training dataset used is from the Greenland ice sheet. I believe this was mainly from Landsat 8-9. Other studies (Wood et al., 2021; Cheng et al., with Calfin) produced large-scale ice front datasets in Greenland from Landsat 4-8. I am wondering why the model was not trained with all these different types of sensors to provide longer time coverage?

Our processing system is built around maximizing the impact of the multispectral capabilities (these lead to more accurate calving front extractions (*Loebel et al. 2022*)) and higher image acquisition rate of Landsat-8/9. As the older Landsat satellites have different multispectral channels and less coverage across the spectrum, our model cannot be applied directly. It would require significant changes to reference data and retraining. In addition, a major motivation for automated calving front delineation is the higher image acquisition rate of modern satellite constellations.

However, we agree that there would be significant value in applying this or a similar approach to older satellite data, as has been done for the CALFIN product (*Cheng el al., 2021*) for Greenland. Such a more historical machine delineated calving front data record does not yet exist for the AP. We gladly include this in our manuscript's conclusions (P14 L270).

(3) Regarding the training dataset, did you provide the model with some cloudy data? This could help the model learn when not to delineate fronts.

Partially cloudy satellite images (where the calving front is not completely obscured) are included in both the training and test datasets. This is important to maximize the extraction rate for our processing system and to increase the sampling of the data product. This is already mentioned in the manuscript in section 2.1 (P5 L91). An example is also shown in Figure 3 (f).

Concerning completely clouded images (calving front not visible): Our processing system is built around classifying only between land/glacier and ocean pixels. Therefore, cloud pixels are classified as land/glacier or ocean depending on their spatial context. If there is no spatial context, pixels will still be classified as land/glacier or ocean, but with large uncertainties or even random in extreme cases. Therefore, there is no way to directly (i.e. within the model prediction) discard an extraction based on cloud cover. In our processing system, completely clouded images result in failed or incorrect calving front delineations, which are identified and discarded during post-processing.

(4) Did you compare your results in the Antarctic Peninsula with previous datasets made from MODIS imagery (https://tc.copernicus.org/articles/17/2059/2023/tc-17-2059-2023.pdf)? Additionally, there is no mention of the Green et al. dataset, which delineated calving fronts all around Antarctica using multi-sensor data (see here: https://www.nature.com/articles/s41586-022-05037-w).

The MODIS dataset (*Andreasen et al., 2023*) as well as the dataset by *Greene at al. (2022)* are concerning circum-Antarctic ice shelf calving front locations. For the AP they include the Larsen Ice Shelf but not the outlet glaciers which are the focus of our manuscript. They are comparable to the machine generated IceLines dataset (*Baumhoer et al., 2023*) which we mentioned in the introduction as a motivation for our work (as it also does not include AP outlet glaciers). Our results are therefore not directly comparable to these datasets.

To make this clearer in the manuscript, we have made some minor changes in Section 1 and expanded the description of Table 1 to better distinguish AP glacier fronts from ice shelf fronts (*Andreasen et al., 2023* and *Greene at al. (2022)* are now also mentioned).

Overall, I think the figures are in good shape, and I recommend the paper for publication after minor technical revisions.

References:

Andreasen, J. R., Hogg, A. E., and Selley, H. L.: Change in Antarctic ice shelf area from 2009 to 2019, The Cryosphere, 17, 2059–2072,https://doi.org/10.5194/tc-17-2059-2023, 2023.

Baumhoer, C. A., Dietz, A. J., Heidler, K., and Kuenzer, C.: IceLines–A new data set of Antarctic ice shelf front positions, Scientific Data, 10, 138, 2023.

Cheng, D., Hayes, W., Larour, E., Mohajerani, Y., Wood, M., Velicogna, I., and Rignot, E.: Calving Front Machine (CALFIN): Glacial Termini Dataset and Automated Deep Learning Extraction Method for Greenland, 1972–2019 , The Cryosphere, 15, https://doi.org/10.5194/tc-15-1663-2021, 2021

Loebel, E., Scheinert, M., Horwath, M., Heidler, K., Christmann, J., Phan, L. D., Humbert, A., and Zhu, X. X.: Extracting Glacier Calving Fronts by Deep Learning: The Benefit of Multispectral, Topographic, and Textural Input Features, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–12, https://doi.org/10.1109/TGRS.2022.3208454, 2022.

Loebel, E., Scheinert, M., Horwath, M., Humbert, A., Sohn, J., Heidler, K., Liebezeit, C., and Zhu, X. X.: Calving front monitoring at a subseasonal resolution: a deep learning application for Greenland glaciers, The Cryosphere, 18, 3315–3332, https://doi.org/10.5194/tc-18-3315-2024, 2024.

Greene, C. A., Gardner, A. S., Schlegel, N.-J., and Fraser, A. D.: Antarctic calving loss rivals ice-shelf thinning, Nature, 609, 948–953,350 https://doi.org/http://dx.doi.org/10.1038/s41586-022-05037-w, 2022.