Comment on essd-2022-139
Anonymous Referee #1

Referee comment on "Calving fronts and where to find them: a benchmark dataset and methodology for automatic glacier calving front extraction from synthetic aperture radar imagery" by Nora Gourmelon et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-139-RC1, 2022

General Comments:

This paper uses deep learning to extract glacier termini over Greenland, Alaska, and Antarctica. The author uses two glaciers that are not included in the training set to test the generalization ability of the deep learning network. One has good results while the other has quite a large uncertainty. Overall, I think the paper is well-written and clear, and I appreciate that the author has a comprehensive review of the recent studies on a similar topic. However, I have several concerns about the manuscript, which are listed below.

Major concerns:

- Compared with previous studies, this manuscript has room to improve regarding innovation. The previous study has used deep learning to extract the glacier termini from multiple remote sensing datasets. If combining multi-sensor and long-term images brings innovations, the author needs to justify the reason (where are the challenges and how do you solve them).
- The second concern is about the data quality and quantity. The uncertainty is high for one of the two test glaciers, and the terminus results are not usable for further analysis. Also, deep learning aims to produce a large number of termini by automating terminus delineation. But the number of terminus traces derived from this study is limited. Considering the ESSD is a journal focusing on data, I think the manuscript
needs to improve in this regard.

**Specific Comments:**

**Page 2 Line 58:** It would be nice if the author could explain more about why the dataset from this study is beneficial for bridging the gap regarding the evaluation among different studies.

**Page 10 Line 212:** Please explain more about re-mapping.

**Page 10 Line 222:** What is the rationale behind identifying four zones since this study will only pick the boundary between glacier and ocean.

**Page 11 Line 246:** Morphological dilation can resolve the imbalance to some degree but will also cause uncertainty in terminus delineation. The more balanced between positive and negative pixels, the larger uncertainties the terminus will have.

**Page 13 Line 277:** I like the idea of applying Gaussian importance weighting before taking the average. It would be nice if the author could provide more details.

**Page 15 Line 298:** Is this a universal threshold? Please explain more about setting this threshold and what will happen if the threshold is too large. Usually, the threshold would be 0.5. The threshold can indeed be different as long as it is justified.

**Page 21 Line 436:** The uncertainty value might be wrong (150 m?). I couldn’t find this value in Table 7.

**Page 22, Line 461:** Maybe consider including such post-processing in this study. Using the bedrock mask can eliminate the fjord boundaries.

**Table 8, Table 9, and Figure 9:** The predicted terminus position for the Columbia glacier deviates from its true position, and the uncertainty for the Columbia glacier is too large.