General comments

The authors describe the creation of a dataset of terminus positions from multi-source satellite imagery for ~290 marine-terminating glaciers containing ~12,000 individual traces. The terminus positions have been manually delineated, and I commence the authors for the effort that goes into that. However, I fail to see the novelty of the dataset. The authors base the majority of terminus delineations on Landsat and Sentinel imagery and use existing methods such as GEEDiT, so that I find it difficult to believe that the dataset presented here has a higher accuracy than currently existing datasets. The presented dataset contains a variety of glaciers (land-, lake-, marine-terminating) so that a) the statement that the dataset contains delineations for 290 marine-terminating glaciers is incorrect, and b) the analysis of the data is skewed and needs to be revised. There are inconsistencies and a lack of detail throughout the methodology section that make it difficult to assess how the underlying satellite was processed. The validation of the data is insufficient with one of the two methods applied not being available for the majority of the time period that the dataset covers, and the other being a comparison to existing data that does not provide a measure of accuracy of the dataset presented.

Please find my more detailed comments below:

Major

1. The manuscript switches between the terms “marine-terminating outlet glaciers” and “outlet glaciers” making it unclear what the focus of this study is. When comparing the dataset presented to other existing datasets, the terminology makes a big difference as e.g. TermPicks (Goliber and Black, 2021; Goliber et al., 2022) contains only delineations for marine-terminating glaciers.

2. The validation of the terminus traces needs to be improved significantly.
    a. Currently, visual validation is carried out using PlanetScope imagery which is only available from 2014 onwards and can therefore not be applied to the majority of the dataset. The authors show only 6 examples of how their delineations match with PlanetScope imagery and present no further quantification of errors. This is insufficient as a validation method.
    b. A plausibility check is mentioned but not explained any further,
    c. The only quantifiable validation is presented by an Average Minimum Distance (AMD) which is calculated between this dataset and TermPicks and/or MEaSUREs data. The authors do not explain which delineation is used as reference line (TermPicks or MEaSUREs) and present a single value as validation. The AMD also does not provide a measure of accuracy but rather states the difference between two delineations with no indication as to which one is more accurate.

3. The authors state that panchromatically sharpened Landsat-5 imagery is used to delineate glacier fronts, which is not possible as Landsat-5 does not have a panchromatic band. In addition, there is little detail on the processing of satellite imagery outside of GEEDiT (e.g. terrain corrections, filtering). The limited detail and inconsistencies in the description of the methods make it difficult to trust the resulting data.

4. The dataset is not consistent and needs thorough revisions. In its current state, the dataset contains five different types of glaciers:
    a. Marine terminating glaciers (227)
    b. Land-terminating glaciers (21)
    c. Lake terminating glacier (5)
    d. Peripheral glaciers (22)
    e. Others including ice that I would not classify as a glacier (6), sections that are part of a larger glacier (1) or double delineations (3).

    While the majority of the data is on marine-terminating glaciers, the inclusion of other glaciers skews the data and makes the presented analysis (Figures 10, 11) redundant. It should be ensured that the dataset matches the scope of the study i.e. if the study is aimed at marine-terminating glaciers, the dataset should only contain marine-terminating glaciers; if the study is aimed at outlet glaciers, the dataset should include all lake- and land-terminating glaciers. It should be made clear whether the dataset is focused on glaciers that are directly connected to the Greenland Ice Sheet or also includes peripheral glaciers. The authors further state the dataset contains comprehensive metadata, yet it contains less than comparable datasets such as TermPicks (Goliber and Black, 2021).

5. The authors state that the presented dataset has a larger spatial coverage than existing dataset, which is misleading as it is only true when including non-marine-terminating glaciers.

6. There is no consistency in glacier names throughout the manuscript with the example glaciers being spelled wrong and differently numerous times. I suggest picking either the official name or Greenlandic name for the example glaciers and use it consistently throughout the manuscript.

MINOR

Abstract

Line 19-20: Maybe rephrase. It is not quite clear for what application it would be insufficient detail for. Also, did you test the performance and transferability of current datasets in automated calving front detection? If not, I would leave this out.

Line 21: I fail to see how your dataset can be high accuracy (or higher than other datasets) given that you use Landsat and Sentinel imagery for the majority of the delineations. Similarly, when comparing your data to TermPicks I cannot see a higher spatial extent than other datasets. Consider omitting this sentence or rephrasing.

Line 23: Why do you say approximately 12,000 traces for approximately 290 glaciers. You should state the accurate numbers, which are easily determined.

Line 28: It is unclear what the offsets of 40-100m relate to in this context. Consider rephrasing.

Introduction

Line 37-40: I would combine these two sentences e.g. Mass loss from the Greenland Ice Sheet contributes significantly to global sea level rise (Sheperd et al., 2018; Frederikse et al., 2020), with nearly half of this dynamic ice loss being attributed to frontal ablation (Enderlin et al., 2014; Mouginot et al., 2019).

Line 44: Consider adding references for Fahrner et al., (2025) and KC et al., (2025) as these publications address frontal ablation over large spatio-temporal scales.

Line 47: Consider adding Fahrner et al., (2021) to the reference list as you do refer to it later on.

Line 66: "…marine-terminating glaciers show highly non-linear behaviour "– While this is true for individual glaciers, many studies have shown that their behaviour on a regional scale is linear (e.g. Cowton et al., 2018; Fahrner et al., 2021). Consider rephrasing.

Figure 1: Please use the full names of the glaciers (e.g. Petermann Glacier, Jakobshavn Isbræ), check the spelling (e.g. Isstrøm) and be consistent (Gletscher vs Glacier). Panels should be labelled from A) overview map to G) Kangilinguata Sermia. Also "Sermia" should be spelled with a captial "S". This glacier is also not marine-terminating as is clearly visible in the image which contradicts previous statements that the dataset is looking at marine-terminating glaciers.

Data and Methodology

Line 108: GEEDiT is the acronym for the “Google Earth Engine Digitisation Tool” not the Glacier Extraction and Evaluation Dataset Tool as stated (Lea, 2018).

Line 113 - 114: How does the use of these auxiliary images impact delineation uncertainties - e.g. how can you be certain that the calving front has not changed within ±15 days?

Line 120: Be consistent with the naming of the satellites e.g. most satellites are introduced by the written-out name followed by the abbreviation in parenthesis except for ERS1/2.

Table 2: Landsat-5 does not have a panchromatic band so it cannot be resampled to a 15 m resolution.

Line 133: “…quality control and metadata creation” replace “and” with “,”

Line 136 – 137: Repetition from line 125-126. Remove

Line 137: “ MODIS OD09GQ?” - Is that specific MODIS imagery and therefore different from the other MODIS imagery you are using? Be consistent with the naming of data - if this is the MODIS data you use throughout, stick with the actual name. Also, auxiliary dataset for what? MODIS data?

Line 138: Is the NDWI only calculated for MODIS?

Line 141-142: I struggle to see how an image with a resolution of 250m provides a reliable source for detecting a calving front even as a reference. Maybe provide more detail here.

Line 142-143: “ENVISAT ASAR (Level-1B) and ERS-1/2 SAR (Level-1.5) scenes were downloaded from the ESA archives” can be removed as you already state prior that the data was downloaded from ESA.

Line 143-144: Be more specific regarding the calibration and terrain correction, e.g., what DEM was used for the terrain correction?

Line 145: Same as above, what DEM was used and what filter applied.

Line 146: Why was the data not reprojected to EPSG:3413 Polar Stereographic? This is the common format for datasets that cover Greenland.

Line 148-149: These sentences contradict each other as you first say that delineations were done on native-resolution imagery and then say that you used pan-sharpened images.

Line 150-151: You say in Line 113-114 that this was only done for Landsat-7, but now it reads like it was done for other satellites too?

Line 152: How did you manually adjust the delineations? Did you use a different image or the same image? Be more specific.

Line 153: Please explain what the plausibility check entails as this is the only mention of it.

Figure 2: The figure needs to be redone as it is currently confusing. Under preprocessing you mention everything that has been done to the satellite imagery, but you don't specify what was done to which imagery. It is somewhat clear that the imagery on the left corresponds with the boxes, but I think it could be made clearer.

I don't think the "Sets parameter.." is necessary unless you specify which parameters.

I am still confused by the editing and modification of the delineations.

The QA process is not clearly explained at all and needs additional details.

Also, again check the glacier names for consistency, spelling etc.

Line 162: You mention uncertainties in delineation of 25m, however Brough et al., (2019) and (Fahrner et al., (2021) contradict this statement as they performed error checks for Landsat 4 (30 m) and Landsat 8 (15 m) and came to the conclusion that the error is much less than the pixel resolution.

Line 165-166: Planet imagery is only available from 2014, so there is no data for over half of the time period that your dataset covers. I fail to see how this can be a quality check, especially since it is also only a visual validation. I suggest the authors should find a way to quantify the uncertainties.

Line 175: “Its fine spatial detail and temporal coverage make it a valuable benchmark for evaluating glacier front delineations.” That might be correct, but it is only available from 2014 onwards so that it is not useful for longer time series such as your dataset.

Line 177: Glacier names are incomplete and spelled wrong.

Line 183: “…capturing key features with sub-pixel accuracy…” if I understand this correctly you say that your delineations capture sub-pixel features. If so, I think that might just be coincidence as I don’t think you’d be able to accurately delineate features within a pixel. Otherwise, please rephrase.

Line 188-189: I don’t think the fact that this glacier is included in this dataset but not in others warrants the claim that this dataset has a greater spatial coverage. Especially since some of the other datasets do not contain land-terminating glaciers such as Kangilinguata Sermia. Consider rephrasing.

Line 201: Double space between “Kangilinguata” and “glacier”.

Line 201: “…which is not included in any of the existing datasets,…” see previous comment on Line 188-189.

Line 207: Rephrase. I think you're trying to say that your dataset outperforms automated approaches but not all manually delineated datasets.

Line 227: Make sure the citation is correct – You are citing Goliber et al., (2022) which is the publication, but you should be citing the dataset Goliber and Black, 2021)

Line 226 – 228: I struggle to see how a close alignment to existing datasets is showing that your dataset is novel or more accurate. Maybe rephrase?

Line 236: It is unclear here if TermPicks of MEaSUREs is used as reference line. I assume that the AMD was calculated for bother and only the maximum is mentioned? Please be more specific.

Line 279-280: Replace “high positional consistency with” with “corresponds well with”

Line 280-281: “… depending on glacier geometry, image quality, and surface contrast.” I would leave this out as you do not provide any measure on how any of these affect your accuracy.

Line 281: This is the first time semi-automated methods are mentioned. Unless it is introduced prior with a specific reference, I would leave it out.

Line 286 – 287: From the figures in the manuscript, I struggle to see the increased spatial and temporal coverage compared to TermPicks. If anything, the dataset presented here seems to have less temporal coverage than existing datasets (e.g. Figure 7).

Line 287 – 288: Given that the dataset contains a multitude of glaciers that are land- and lake-terminating and has no consistent temporal resolution, I would remove this sentence. Currently the dataset contains 227 marine-terminating glaciers, and the temporal resolution varies from annually to monthly.

Line 289 -290: It states here that there is “high-precision agreement with PlanetScope imagery” however earlier it is stated that this was only accessed visually. There is no proof in the manuscript that would support the claim of “positional discrepancies are consistently below 10m”. I suggest the authors rephrase or provide a detailed description of how this was determined.

Data product and usage notes:

Line 301: The name of the individual files “Calving_front_positions_for_Greenland_outlet_glaciers_2002_2021 — New_NeighborGID115” is unnecessarily long. I would recommend to either just use the

Glacier ID or use the glacier name. In the current format, the key information (GlacierID) is almost hidden and I have to use half my screen in QGIS to see it.

Line 308: The metadata should contain at least the official glacier name as specified in TermPicks or Bjork et al., (2015).

Line 313: The “season” field can be left out as it is quite clear from the “Date” field what season the delineation was taken in.

Line 318: Check the glacier names for spelling and consistency.

Line 344: Write these values as values±STD km.

Line 350 – 351: If you only model one glacier it's not really the dataset that you are using but a subset. I would leave this sentence out as it doesn’t really add anything.

Conclusions

Line 376: It should be mentioned here that the validation against PlanetScope imagery was visual only, and that the AMD was only determined for TermPicks and/or (?) MEaSUREs. Currently it reads like the validation of the dataset was extensive, which unfortunately, it was not.

References

Brough, S., Carr, J.R., Ross, N. and Lea, J.M., (2019) Exceptional retreat of Kangerlussuaq Glacier, east Greenland, between 2016 and 2018. *Frontiers in Earth Science*, 7123.

Cowton, T.R., Sole, A.J., Nienow, P.W., Slater, D.A. and Christoffersen, P., (2018) Linear response of east Greenland’s tidewater glaciers to ocean/atmosphere warming. *Proceedings of the National Academy of Sciences*, 11531, pp.7907–7912.

Fahrner, D., Lea, J.M., Brough, S., Mair, D.W.F. and Abermann, J., (2021) Linear response of the Greenland ice sheet’s tidewater glacier terminus positions to climate. *Journal of Glaciology*, [online] 67262, pp.193–203. Available at: https://www.cambridge.org/core/article/linear-response-of-the-greenland-ice-sheets-tidewater-glacier-terminus-positions-to-climate/6B3723E3A0E94012A1DB9D3E49246AF2.

Fahrner, D., Slater, D.A., KC, A., Cenedese, C., Sutherland, D.A., Enderlin, E., de Jong, M.F., Kjeldsen, K.K., Wood, M., Nienow, P., Nowicki, S. and Wagner, T.J.W., (2025) A Frontal Ablation Dataset for 49 Tidewater Glaciers in Greenland. *Nature Scientific Data*.

Goliber, S. and Black, T., (2021) *TermPicks: A century of Greenland glacier terminus data for use inmachine learning applications (Version 1)*. Available at: https://zenodo.org/records/5117931.

Goliber, S., Black, T., Catania, G., Lea, J.M., Olsen, H., Cheng, D., Bevan, S., Bjørk, A., Bunce, C., Brough, S., Carr, J.R., Cowton, T., Gardner, A., Fahrner, D., Hill, E., Joughin, I., Korsgaard, N.J., Luckman, A., Moon, T., Murray, T., Sole, A., Wood, M. and Zhang, E., (2022) TermPicks: a century of Greenland glacier terminus data for use in scientific and machine learning applications. *Cryosphere*, [online] 168, pp.3215–3233. Available at: https://zenodo.org/record/5117931 [Accessed 27 May 2022].

KC, A., Enderlin, E.M., Fahrner, D., Moon, T. and Carroll, D., (2025) Seasonality in terminus ablation rates for the glaciers in Greenland (Kalaallit Nunaat). *The Cryosphere*, [online] 198, pp.3089–3106. Available at: https://tc.copernicus.org/articles/19/3089/2025/.

Lea, J.M., (2018) The Google Earth Engine Digitisation Tool (GEEDiT) and the Margin change Quantification Tool (MaQiT)--simple tools for the rapid mapping and quantification of changing Earth surface margins. *Earth Surface Dynamics*, 63, pp.551–561.