

In this manuscript, the authors have described a dataset of manually digitized terminus positions for outlet glaciers of the Greenland ice sheet compiled from previously-published datasets, in order to provide a consistently-formatted training dataset for future machine learning applications. This is an excellent and timely undertaking that highlights the power of collaborative efforts.

On the whole, the manuscript does a good job describing the issues involved in combining "input" datasets from multiple authors, as well as describing the "output" dataset, and even manages to show an example application of combining data sources. Accordingly, I only have a few minor comments/suggestions to make on the manuscript. The bulk of my comments/suggestions have to do with the description of the metadata - I think a Table with a few different example entries would help clarify this for a reader.

We appreciate the constructive and positive feedback on the manuscript. We address the comments below, but also included a new figure (10) of the metadata of three glaciers to better clarify the structure. As we addressed comments, the original line numbers of the text may have changed in the final manuscript. The changed text has been noted in the responses to individual comments. Our responses are in blue below each comment.

Comments to Address:

- l. 10: is this the mean (\pm standard deviation)?
 - Yes. This was changed to "The TermPicks data set includes 39,060 individual terminus traces for 278 glaciers with a mean of 136 ± 190 and median of 93 of traces per glacier" to be more clear. The SD is higher than the mean due to the high variation of picks between certain glaciers.
- l. 52: check that months are removed from the reference dates
 - These have been checked and have removed from the text.
- l. 104: is the Howat reference here for the MODIS image?
 - The incorrect MEaSURES image was being cited here. It has been changed to "MEaSURES Greenland Ice Mapping Project (GIMP) 2000 Image Mosaic (Howat et al, 2014; Howat,2018)."
- l. 130 (Date): I found this description slightly confusing - are there 5 columns (one column for the date string, four columns for the year, month, day, and decimal date)? From the dataset, I see that it is indeed five individual columns, but the header makes it seem like there's only one column here (Date).

- Text changed to “Date Columns: The Date column represents the acquisition time for the image used to pick the terminus for that trace. There are 4 additional columns for year, month, day and decimal date” for clarity.
- l. 135 (Satellite): How is this formatted/written?
 - Added text “The names used are in listed in Table 2”. Table 2 lists the satellite names.
- l. 144 (Scene ID): here again, it would be helpful to have more information about this. The Landsat Product ID/other satellite IDs are relatively straightforward, but what about the aerial images?
 - If an author provided satellite ID information, then we do not change it - if someone is using TermPicks for machine learning, then they may need access to the original data. This assumes it is easier for them to request it with the original name. We added text “It includes information on the date and location for the original image. This may be listed as a file name the original author used and may store locally (Figure 10; Glacier 291) or a scene ID from a different satellite (e.g. Sentinel-1 product folder name)” for clarity. Figure 10 Glacier 291 shows an example of an original image name.
- l. 155 (Quality Flag): What does this entry look like for a given image? From the dataset, I see that it's comma-separated 2-digit strings (00, 01, 02, 03, 04, 05) - I'm not sure I would have gotten that from the description here.
 - Added text “If there are multiple flags, they are separated by commas (Figure 10; Glacier 278)” for clarity. Figure 10 Glacier 278 shows an example of multiple flags.
- l. 170: where do the glacier centerlines come from?
 - Text added “Centerlines are manually mapped from the MEaSURES Greenland Ice Mapping Project (GIMP) 2000 Image Mosaic (Howat et al., 2014; Howat, 2018).”
- l. 226: how many of these picks needed manual checking?
 - Only 220 traces were checked manually for this section. Text changed to “Traces with >500 m error between traces were manually checked for errors (220 traces).”
- l. 228: wouldn't it make more sense to compare the image (assuming it exists) against the different picks, rather than using the completeness of the metadata?

- The method we used to compare traces between large errors in multiple authors assumes the large error is due to mislabeling the date (i.e. the trace did not appear to be from the same front on the same date as there is a large step change in the traces). The author that included the original image likely kept detailed record of what image was used and therefore is less likely to have incorrectly listed the date. As this was a very small subset of the dataset (~0.4%) we chose not to manually check each trace.
- Figure 5: I really like this figure.
 - Thank you!
- The GEEDiT walkthrough is great - have you thought about putting it on github pages (<https://pages.github.com/>) so that it's more widely visible/available?
 - GEEDiT TermPicks has been put into a repository (<https://github.com/jmlea16/GEEDiT-TermPicks>) documenting the walkthrough and program.