Automated Greenland Glacier Termini Position Time Series, Version 1

USER GUIDE

How to Cite These Data

As a condition of using these data, you must include a citation:

Catania, G. and E. Zhang. 2023. *Automated Greenland Glacier Termini Position Time Series, Version 1*. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/QUEMHTTVXMRG. [Date Accessed].

FOR QUESTIONS ABOUT THESE DATA, CONTACT NSIDC@NSIDC.ORG

FOR CURRENT INFORMATION, VISIT https://nsidc.org/data/NSIDC-0788

TABLE OF CONTENTS

1 DATA DESCRIPTION

1.1 Parameters

This data set contains shapefiles of termini traces from 294 Greenland glaciers, derived using a deep learning algorithm (AutoTerm) applied to satellite imagery. Also available are supplementary data, including temporal coverage of termini traces, time series data of termini variations, and updated land, ocean, and ice masks derived from the Greenland Ice Sheet Mapping Project (GrIMP) ice masks.

1.2 File Information

This data set is organized into 296 folders; one folder for each of the 294 identified glaciers and a single folder with the updated GrIMP masks.

1.2.1 File Format and Contents

Each of the 294 glacial data folders contains eight files. These files include a shapefile (`.shp`) and related auxiliary files for viewing geolocated glacial termini traces; these files and their functions are described in Table 1 below. Note, in order to open and view the shapefilethe four auxiliary files must be in the same folder as the shapefile and cannot be renamed.

Table 1. Termini shapefile and related auxiliary files

Extension	Description
`.shp`	Shapefile: the primary file used for viewing the termini data
`.dfb`	Database (dBASE) table: stores required feature attribute data
`.prj`	Projection file: stores coordinate system information
`.shx`	Index file: stores required feature geometry index
`.cpg`	Codepage file: stores codepage specifications

The shapefiles for each glacier include an attribute table containing the metadata for each terminus trace, including the date, glacier ID, source image satellite, and the average uncertainty of each trace, as shown in Figure 1 below.

	FID	Shape	Glacier ID	Satellite	Error	Date
1	0	Polyline	GID1	Sentinel1	24.669617	2/9/2017
2	1	Polyline	GID1	Sentinel1	24.669617	3/2/2021
3	2	Polyline	GID1	Sentinel2	19.147949	10/12/2016
4	3	Polyline	GID1	Sentinel2	19.147949	4/10/2019
5	4	Polyline	GID1	Sentinel2	19.147949	4/30/2021
6	5	Polyline	GID1	Landsat7	20.621682	10/9/1999
7	6	Polyline	GID1	Landsat8	24.53585	6/11/2018
8	7	Polyline	GID1	Landsat8	24.53585	5/15/2015
9	8	Polyline	GID1	Sentinel2	19.147949	5/19/2019
10	9	Polyline	GID1	Sentinel2	19.147949	3/14/2021

Figure 1. Example attribute table from glacier GID001 as viewed in ArcGIS Pro

Each of the individual glacial folders also includes a KML (`.kml`) file and associated image (`.png`) file and a CSV (`.csv`) file. The KML file geolocates the glacial area change graphs shown in the image file in applications such as Google Earth. Note, the `.png` file must be in the same folder as the KML and cannot be renamed.

The CSV file contains glacial terminus times series data, and includes:

- Change in glacial area
- Change in terminus length
- Decimal date
- Date formatted as Month/Day/Year

The Masks folder (masks/) contains three land, ocean, and ice masks formatted as GeoTIFF files (`.tif`) files. The files each represent different time spans, including 2015-2018, 2015-2019, and 2020.

1.2.2 Naming Convention

The individual glacier termini files conform to the following naming convention:

`NSIDC-0788_termTS_GID[NNN]_[YYYY_YYYY] _V01.0.ext`

Table 2. Naming Variable Descriptions for Individual Glacier Files

Variable	Description
NSIDC-0788	NSIDC data set ID
termTS	Terminus (term) Time Series (TS)
GID[NNN]	Glacial ID (GID) number (NNN), formatted as a 3-digit number between 001-080 and 082-295
[YYYY_YYYY]	Date range of source imagery, formatted as 4-digit start-year to 4-digit end-year
V01.0	Data set version
.ext	File type

Note: The data range listed in the file names is indicative of the temporal coverage for the entire data set and the input satellite imagery. The actual temporal coverage for each glacier may be less than that of the entire data set.

The mask files conform to the following naming convention:

```
NSIDC-0788_GRIMPMASK_15M_[YYYY-YYYY] _V01.0.ext
```

Table 3. Naming Variable Descriptions for Mask Files

Variable	Description
NSIDC-0788	NSIDC data set ID
GRIMPMASK	GrIMP mask
15M	Image resolution (15 meters)
[YYYY-YYYY]	Dates of mask data, formatted as 4-digit year; one file only contains one year of data and is formatted as YYYY
V01.0	Data set version
.tif	File type (GeoTIFF)

1.3 Spatial Information

1.3.1 Coverage

Greenland

Northernmost Latitude: 85 N

Southernmost Latitude: 57 N

Westernmost Latitude: 75 W

Easternmost Latitude: 8 W

1.3.2 Resolution

Varies

1.3.3 Geolocation

The following tables provide information for geolocating this data set

Table 4. Geolocation Details

Geographic coordinate system	WGS 84
EPSG code	4326
PROJ4 string	+proj=longlat +datum=WGS84 +no_defs +type=crs
Reference	https://epsg.io/4326

1.4 Temporal Information

1.4.1 Coverage

Temporal coverage for all possible input source imagery is 1 January 1972 to 1 January 2022; temporal coverage for individual files may vary.

1.4.2 Resolution

Varies

2 DATA ACQUISITION AND PROCESSING

2.1 Background

This data set contains glacial termini positional data and time series data for the Greenland ice sheet outlet glaciers, derived using an automated machine learning model, AutoTerm. The model functions as a pipeline, imputing publicly availably satellite imagery from Google Earth Engine (GEE) and outputting shapefiles of glacial termini positions for each image. The glacial termini position data was used to create updated versions of the available GrIMP (Greenland Ice Mapping Project) the ice, ocean, and land masks.

A reference CSV file identifying each glacier for which termini data is available can be found here.

2.2 Acquisition

The satellite imagery used as input data was accessed automatically using the Google Earth Engine (GEE) Python application programming interface (API) and associated tools. Bounding boxes which geographically constrain glacial termini variations were manually delineated and used as selection criteria for the satellite imagery sourced from GEE. In total, ~430,00 satellite images of varying resolution from five different publicly accessible satellite datasets were sourced from GEE. Satellite dataset details are available in Table 5 below, and additional information about imagery sources and the usage of GEE can be found in Zhang et al., 2023.

Table 5. Satellite imagery details

Satellite	Resolution	Time Range
Landsat 5	30 meters	1972-2013
Landsat 7	15 meters	1999-2013
Landsat 8	15 meters	2013-present
Sentinel-2	10 meters	2015-present
Sentinel-1	10 meters	2015-present

2.3 Processing

2.3.1 Deep Learning Model

AutoTerm was built by training a deep learning image segregation algorithm (DeepLabv3+; Chen et al., 2018) to identify glacial termini traces using manually delineated Greenland glacial termini from the TermPicks data set (Goliber et al., 2022). This required converting the TermPicks data from polylines to labeled polygons to be used as model test and training data. Additional training images were manually selected to represent specific conditions that were identified as being more difficult for the model to interpret accurately, including: images with clouds, shadows, ice mélange, or snow cover; blurry image boundaries caused by coarse resolution; images where tabular icebergs are adjacent to a glacier terminus; and images where the ice mélange and the glacier have similar textures. Following model training, the test data was used to calculate model test error. Additional details regarding the training process for the deep learning model can be found in Zhang et al., 2023.

2.3.2 Termini Data Set

Prior to input in the AutoTerm model, satellite imagery was first cropped, then histogram normalized to reduce differences in image texture, resolution, and pixel value. The images were also size normalized to adjust the apparent physical size of each glacier to be more equal when viewed by the deep learning model. The satellite images were then input into the model to produce the termini data set. A georeferencing adjustment was necessary for termini data produced from Sentinel-1 satellite imagery, due to known location errors (Small and Shubert, 2019). Full code for the AutoTerm model is available for download. The termini data is available as shapefiles, which show termini traces over time for each glacier. The data was also used to produce time series of termini variations, available as a CSV and KML file for each glacier.

2.3.3 Ice and Ocean Masks

From the newly derived glacial termini traces, updates were made to existing ice and ocean masks (MEaSUREs Greenland Ice Mapping Project (GIMP) Land Ice and Ocean Classification Mask, Version 1) to reflect annual changes in glacial termini positions in 2018, 2019, and 2020. This was accomplished by converting the original GrIMP ocean mask to a vector shapefile, cropping the shapefile using the newly delineated minimum glacial termini positions for each year, then reconverting the cropped shapefile into a rasterized ocean mask. An updated ice mask is created from the residual between the original GrIMP bedrock mask and the new ocean mask.

2.4 Quality, Errors, and Limitations

Multiple means of mitigating model error were employed when training the AutoTerm pipeline, including: calculating test error, training the model with specific image types that were deemed more difficult for the model to interpret, and using validation data to decrease model overfitting. A complete discussion of model error and termini data quality and can be found in Zhang et al., 2023.

3 VERSION HISTORY

Table 6. Version History Summary

Version	Release Date	Description of Changes
1	February 2024	Initial release

4 RELATED DATA SETS

MEaSUREs Greenland Ice Mapping Project (GIMP) Land Ice and Ocean Classification Mask, Version 1

5 RELATED WEBSITES

Greenland Ice Sheet Mapping Project (GrIMP)

Global Land Ice Measurements from Space (GLIMS)

6 REFERENCES

Baumhoer, C. A., Dietz, A. J., Kneisel, C., and Kuenzer, C. 2019, Automated Extraction of Antarctic Glacier and Ice Shelf Fronts from Sentinel-1 Imagery Using Deep Learning, Remote Sens., 11, 2529, https://doi.org/10.3390/rs11212529.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, in: Proceedings of the European Conference on Computer Vision (ECCV), 8–14 September 2018, Munich, Germany, 801–818, https://doi.org/10.48550/arXiv.1802.02611.

Cheng, D., Hayes, W., Larour, E., Mohajerani, Y., Wood, M., Velicogna, I., and Rignot, E. 2021. Calving Front Machine (CALFIN): glacial termini dataset and automated deep learning extraction method for Greenland, 1972–2019, The Cryosphere, 15, 1663–1675, https://doi.org/10.5194/tc-15-1663-2021.

Gal, Y. and Ghahramani, Z. 2016. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, Proceedings of the 33rd International Conference on Machine Learning, 19–24 June 2016, New York, United States, 1050–1059, https://doi.org/10.48550/arXiv.1506.02142

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, The Cryosphere, 16, 3215–3233, https://doi.org/10.5194/tc-16-3215-2022.

Mohajerani, Y., Wood, M. H., Velicogna, I., and Rignot, E. J. 2019. Detection of Glacier Calving Margins with Convolutional Neural Networks: A Case Study, Remote Sensing, 11, 74, https://doi.org/10.3390/rs11010074.

Small, D. and Schubert, A. 2019. Guide to sentinel-1 geocoding, Remote Sensing Lab. Univ. Zurich (RSL), Zürich, Switzerland, Tech. Rep. UZHS1-GC-AD..

Zhang, E., Liu, L., and Huang, L. 2019. Automatically delineating the calving front of Jakobshavn Isbræ from multitemporal TerraSAR-X images: a deep learning approach, The Cryosphere, 13, 1729–1741, https://doi.org/10.5194/tc-13-1729-2019.

Zhang, E., Catania, G., and Trugman, D. T. 2023. AutoTerm: an automated pipeline for glacier terminus extraction using machine learning and a “big data” repository of Greenland glacier termini, The Cryosphere, 17, 3485–3503, https://doi.org/10.5194/tc-17-3485-2023.

7 DOCUMENT INFORMATION

7.1 Publication Date

February 2024

7.2 Date Last Updated

February 2024