esa climate change initiative

European Space Agency

Product User Guide (PUG)

glaciers cci

Prepared by:	Glaciers_cci consortium
Contract:	4000127593/19/I-NB
Name:	Glaciers_cci+_Ph2_D4.2_PUG
Version:	0.2
Date:	22.07.2024

Contact:
Frank Paul
Department of Geography
University of Zurich
frank.paul@geo.uzh.ch

Technical Officer:
Anna Maria Trofaier
ESA Climate Office

University of Oslo

University of Zurich UZH

UCL Mullard Space Science Laboratory

GAMMA REMOTE SENSING

enveo

Document status sheet

Version	Date	Changes	Approval
0.1	26.04.2024	Initial draft	
0.2	22.07.2024	Consortium feedback included	

The work described in this report was done under ESA contract 4000127593/19/I-NB. Responsibility for the contents resides with the authors who prepared it.

Author team:
Tazio Strozzi, Andreas Wiesmann (Gamma), Thomas Nagler, Gabriele Schwaizer, Jan Wuite (ENVEO), Frank Paul (GIUZ), Andreas Kääb, Désirée Treichler (GUIO), Lin Gilbert (UCL)

Glaciers_cci+ Technical Officer at ESA:
Anna Maria Trofaier

Related documents

Acronym	Title	Document reference	Version	Date
[RD1]	CRDP	Glaciers_cci_Ph2_D3.2_CRDP	0.2	28.06.2024
[RD2]	PSD	Glaciers_cci_Ph2_D1.2_ PSDv2	0.3	07.07.2024

Table of Contents

1. Purpose and scope

This document is the Product User Guide (PUG) for the Glaciers_cci+ project. It refers to datasets which have been generated in Phase 2 of Glaciers_cci+ and which are meaningful to share. For the description of other products available in the database we refer to earlier versions of the PUG. The new products can be downloaded from cryoportal.enveo.at.

For each of the four products generated (global surge inventory, elevation changes for Greenland's peripheral glaciers, glacier facies and length changes for the Karakoram) we describe product content and format as well as known limitations and software tools to view the datasets. As a note, there is some overlap with [RD2] the product specification document (PSD) to which we will partly refer instead. Table 1.1 provides a rough dataset overview.

Table 1.1: Overview of the datasets described in this PUG.

Product	Description	Epoch	Region	Format	Size	Count
Inventory	Observed surges	2017-2022	Global	csv	KB	1 file + meta
Elevation	Altimetry time-series	2000-2023	Greenland	netCDF	100 MB	5 files + meta
Facies	Snow cover time series	2000-2015	Karakoram	netCDF	100 MB	xxx files
Length	Length changes	1965-2020	Karakoram	shape	KB	1 file + meta

2. Global inventory of glacier surges

2.1 Product content

The aim of the global inventory of glacier surges is to provide a global overview over glacier surges during a given time period. Many existing inventories of "surge-type glaciers" are limited to a specific geographic area, and they may not provide a complete overview – i.e., they only record that particular glaciers are known to have surged at some time in the past. The timing of these surges may not always be known, and surges of other (more remote) glaciers in the area may pass unknown if they didn't happen to be observed. Glacier surges and underlying processes are likely influenced by climatic factors that are currently changing rapidly, and this global inventory therefore provides a starting point to study the influence of climatic forcing on glacier surges.

Using global Sentinel-1 radar backscatter data, the locations of glaciers with surge-type activity during 2017 to 2022 are systematically mapped. Patterns of pronounced increases or decreases in the strongest backscatter between two winter seasons often indicate large changes in glacier crevassing, which are treated here as a sign of surge-type activity. Validations against velocity time series, terminus advances and crevassing found in optical satellite images confirm the robustness of this approach (Leclercq et al. 2021).

The dataset contains 116 surge-type events globally between 2017 and 2022, around 100 of which on glaciers already known as surge-type. The data reveal a pronounced spatial clustering in three regions (Fig. 2.1), (i) Karakoram, Pamirs and Western Kunlun Shan (~50 surges), (ii) Svalbard (~25) and (iii) Yukon/Alaska (~9), with only a few other scattered surges elsewhere. According to the related study by Kääb et al. (2023), this spatial clustering is significantly more pronounced than the overall global clustering of known surge-type glaciers and may indicate a climatic forcing of surge initiation.

The surge dataset is based on the interpretation of the above radar backscatter patterns by human operators. Both these human interpretations and the definition of a surge are not sharp and binary (surge vs. non-surge), but clearly indicate strong surge-type glacier-dynamic events. The dataset contains (see Section 2.2) the GLIMS and RGI identifiers of the glacier exhibiting a surge-type event, the years in which the first and last surge-type backscatter changes were observed, and a comment field, typically the name of the glacier or special circumstances. It is important to note that GLIMS and RGI IDs refer sometimes to entire glacier systems while only one branch or tributary was observed to surge. This is also noted in the comment field.

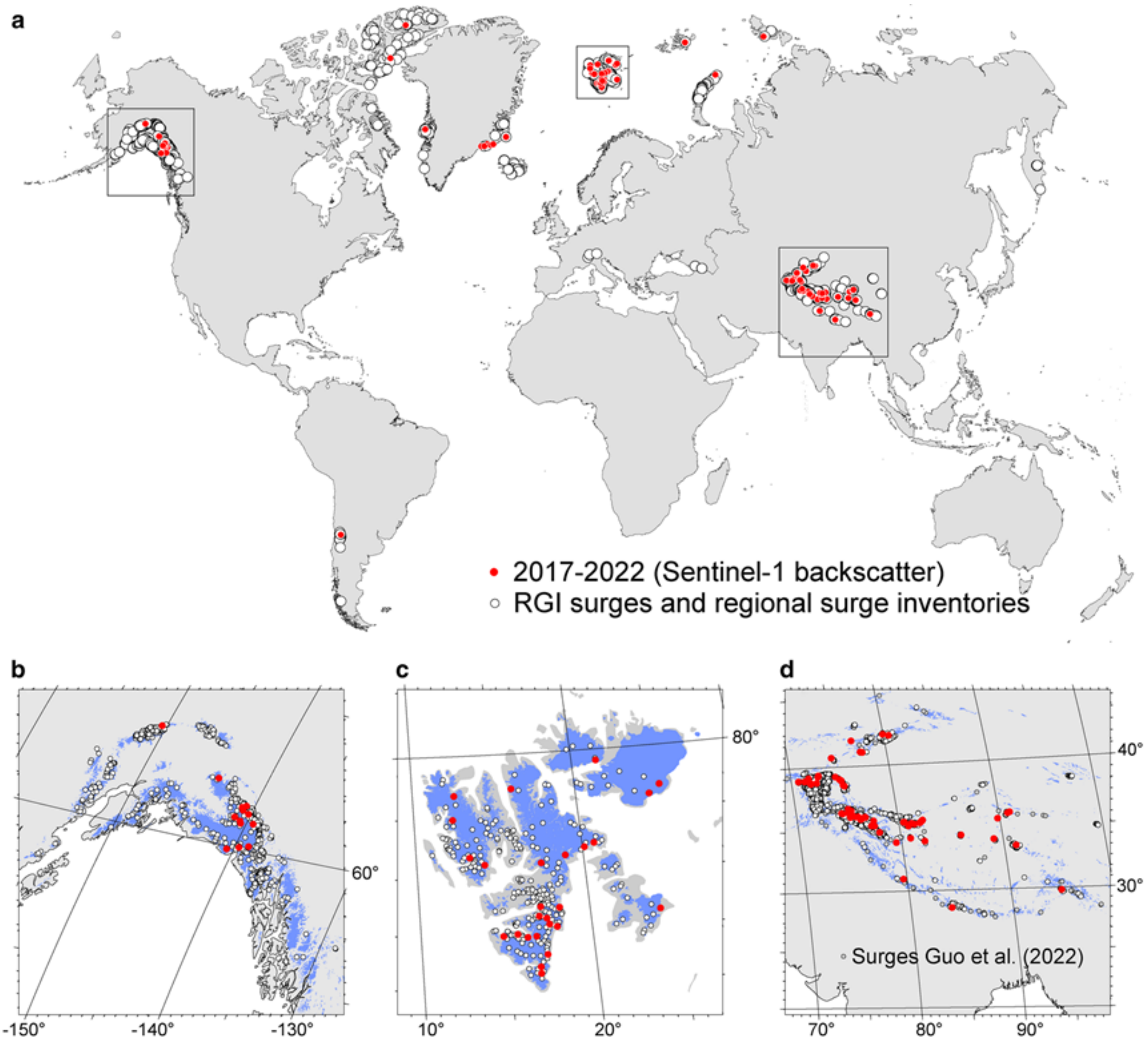

Figure 2.1: Fig. 2.1:Maps of Sentinel-1 backscatter-derived glacier surges over 2017–22. (a) Global distribution of 2017–22 surges (red dots), and glaciers with RGI v.6 surge-type flag and additional regional inventories (white circles). (b–d) Details of the main map, including RGI v.6 glacier areas. Map projection is van der Grinten as a compromise between equal-area and conformal projection.caption template

2.2 Product format

The inventory is stored as a table, where surging glaciers are identified as points (lat/lon) and with glacier IDs, directly linking it to the RGI/GLIMS glacier databases. Data columns are repeated in Table 2.1 from Table 3.2 of [RD2].

Sample rows:

```
22 ; G017096E77164N ; RGI60-07.00250 ; 17,239 ; 77,148 ; 1000 ; 2019 ; Markhambreen
23 ; G017158E77876N ; RGI60-07.00266 ; 17,114 ; 77,865 ; 2021 ; 3000 ; Vallakrabreen
24 ; G018098E77802N ; RGI60-07.00276 ; 18,213 ; 77,843 ; 2017 ; 2021 ; Arnesenbreen(two surge
phases)
```

Table 2.1: Attribute information for the glacier surge inventory.

Attribute	Format	Content
No	Integer	Running number
GLIMS_ID	Text	GLIMS ID of glacier
RGI_ID	Text	RGI6.0 ID of Glacier
Lon	Float	Longitude (degree WGS 1984)
Lat	Float	Latitude (degree WGS 1984)
Surge_start	Integer	approx. year of first visibility of surge as backscatter anomaly 1000: surge start before 2017
Surge_end	Integer	approx. year of last visibility of surge as backscatter anomaly 3000: surge end after 2022
Name_or_comment	Text	Glacier name and/or other comments

2.3 Known limitations

Surge activity was mapped by the operators from visual inspection. We cannot be sure to have mapped every such event that happened over 2017–22, but a large number of cross-checks with literature, own data (velocity time series, optical satellite data, frontal advance, elevation changes), and discussion with the operator team makes us confident that we have identified a large percentage of events that qualify as surges with significant backscatter changes. The novel method to map surge-type events globally and at annual scale from repeat satellite radar images appears quite robust and well suited for this purpose (Leclercq et al., 2021).

2.4 Software tools and data access

The data are in text/table format that can be read by any text software, or spatial software using the provided coordinates. GLIMS/RGI IDs ensure compatibility with existing glacier inventories that provide further information about the surging glaciers.

3. Elevation changes of Greenlands' peripheral glaciers

3.1 Product content

The aim of the five products is to investigate the usage of radar altimetry to produce long-term time-series of elevation change over the challenging terrain of Greenland's peripheral glaciers. Following on from products made over the East Arctic islands in phase 1 of the project, new sensors, input products and methods are used.

In each product elevation change time-series and the surface elevation change rates derived from them are provided, where possible, for individual glaciers, for grid cells and for along-track points. The datasets share similar formats, but differ in detail due to the different satellite missions involved. Change rates are given over the longest possible period and also for a set of five year long moving windows that start at midnight beginning the first full year of the time-series, and step by one year at a time.

The Greenland peripheral glaciers form a narrow rim around the Greenland ice sheet. In order to use a grid layout consistent across the whole of Greenland, but without a large majority of empty cells, only cells containing data are stored, and their central locations in the EPSG 3413 projection are given. Individual glaciers are defined by the outline polygons given by the Randolph Glacier Inventory v7.0 (RGI 7.0 Consortium, 2023). A brief description of each product is given below.

3.1.1 Sentinel 3 radar altimetry

The input data to this product was the recent S3 Land Thematic product (Jettou, 2022), which reprocessed mission data from both Sentinel-3 satellites (A and B) using an extended radar window in order to capture full radar echoes over fast-changing terrain. The two satellites, each follow a 27-day repeating orbit, offset so that the tracks of one are centred between the tracks of the other.

In the product, each satellite is treated separately. Time-series are created at along-track points spaced 300 m apart. These are then aggregated into 500 m by 500 m cells to match the CryoSat-2 gridding, and also aggregated over each individual glacier. Surface elevation change rates are derived for the full mission period for each satellite, and for Sentinel 3A in two five year long overlapping window periods. Sentinel 3B has not been operating for long enough to provide a five year window.

3.1.2 CryoSat-2 point swath radar altimetry

The input data to this product was the recent point swath product (Jakob and Gourmelen, 2023), where each radar echo is used to provide multiple elevation data points for a strip across-track rather than a single elevation value at the point of closest approach. This approach dramatically increased the amount of data available. CryoSat-2 follows an orbit that repeats every 369 days, with a near-repeat every 31 days

In the product, since CryoSat-2's orbit is not well suited to the along-track methodology, time-series are only produced for 500 m by 500 m cells, and also aggregated over each individual glacier. Surface elevation change rates are derived for the full mission and for eight five year long overlapping windows.

3.1.3 CryoSat-2 gridded swath radar altimetry

The input data to this product was the gridded swath product (Jakob and Gourmelen, 2023), which aggregates the point swath product (see section 3.1.2 above) onto a grid at 2km resolution. This grid aligns with the 500 m grid used for the point swath product, so that each cell contains sixteen of the point swath product cells. The grids are provided monthly, and stacked before time-series are produced for 2km by 2km cells and individual glaciers. As for the point swath product, surface elevation change rates are derived for the full mission and for eight five year long overlapping windows.

3.1.4 Sentinel 3 and Cryosat-2 joint radar altimetry

This product was made by combining and cross-calibrating the time-series from the Sentinel 3 and CryoSat-2 point swath products (see sections 3.1.1 and 3.1.2 above) in each cell and individual glacier where at least two satellites were represented. Surface elevation change rates are derived from the combined time-series, for the full length of the time-series (dependent on the missions involved) and as many five year long overlapping windows as possible.

3.1.5 ICESat-2 laser altimetry

The input data to this product were two related products from the Arctic Land Ice Height dataset (Smith et al., 2023), the gridded digital elevation model (DEM) from product ATL14 and the 3-monthly elevation changes to it from product ATL15. The elevation changes are provided at several grid resolutions, so the highest, 1km, is used. This grid aligns with the 500 .m grid used for the point swath product, so that each cell contains four of the point swath product cells. The changes are stacked to provide time-series, and change rates derived where possible. The full mission time-series do not yet cover a five year period, so a four year window is used to cover the same seasons as the five year windows (i.e. starting and finishing on the midnight beginning a year).

3.2 Product format

Since there is a lot of commonality between products, the variable names in each product are similar. Complete lists of variable names, dimensions, data types and units for each product are given in the PSD [RD2]. Each variable name is constructed from a set of elements, described below, and listed in table 3.2.1.

3.2.1 Time-series elements

Each product contains time-series of height or height change for glaciers, for grid cells, and in the case of Sentinel 3 for discrete along-track locations called ‘nodes’. For any one mission, every time-series is sampled at the same set of timestamps.

ts_dh is a time-series of height changes
ts_h is a time-series of heights (CryoSat-2 gridded swath product only)

These variable names will be followed by _glac, _cell or _node, to indicate the area the time-series relates to, eg ts_dh_glac is an array of time-series for individual glaciers.

ts_t gives the timestamps for all time-series

Since there are two Sentinel 3 satellites, in the Sentinel 3 product they are distinguished by either _a or _b at the end of the variable name, eg ts_dh_glac_a is the array of time-series for individual glaciers from Sentinel 3A data, and its timestamps will be given by ts_t_a.

Finally, every time-series has associated uncertainties, which are indicated by _uncert at the end of the variable name, eg ts_dh_glac_a_uncert contains the uncertainties associated with ts_dh_glac_a.

3.2.2 Surface elevation change rate elements

Surface elevation change rates are derived from each time-series, and for a set of 5 year long windows within the time-series.

mission_sec is the surface elevation change rate for full time-series
windowed_sec is the surface elevation change rate for windows within a time-series

As for the time-series, these variable names will be followed by _glac, _cell or _node. In the case of Sentinel 3 they will be followed by _a or _b. Again, their associated uncertainties will be indicated by _uncert. Eg the full mission surface elevation change rates per glacier from Sentinel 3A will be given by mission_sec_glac_a, and their uncertainties by mission_sec_glac_a_uncert.

3.2.3 Ancillary information

If a time-series represents a node or cell location, then its location in EPSG 3413 in m, will be given by arrays x_node and y_node or x_cell and y_cell, where the latter define the cell centre. If it represents a glacier, then the RGI v7 ID of the glacier will be given by array glac_id.

Where cell-based data is given, the cell resolution in x and y respectively is given by cell_resolution. Where windowed SEC is given, the arrays window_start and window_end give the start and end times of each window.

In the joint product, it is necessary to distinguish between missions. Time-series contain data from two or three missions, biased to a common level, but not combined. The set of timestamps used by each mission are always the same, so the ts_mission_flag is provided to distinguish which timestamp came from which mission. SEC values likewise may contain data from two or three missions, and the combined_mission_flag distinguishes what combination of missions was used. The values for both mission flags are given in table 3.2.1.

3.3 Product format

Since there is a lot of commonality between products, the variable names in each product are similar. Complete lists of variable names, dimensions, data types and units for each product are given in the PSD [RD2]. Each variable name is constructed from a set of elements, described below, and listed in Table 3.1. A set of use cases is provided in the Appendix.

3.3.1 Time-series elements

Each product contains time-series of height or height change for glaciers, for grid cells, and in the case of Sentinel 3 for discrete along-track locations called ‘nodes’. For any one mission, every time-series is sampled at the same set of timestamps.

ts_dh is a time-series of height changes
ts_h is a time-series of heights (CryoSat-2 gridded swath product only)

These variable names will be followed by _glac, _cell or _node, to indicate the area the time-series relates to, eg ts_dh_glac is an array of time-series for individual glaciers.

ts_t gives the timestamps for all time-series

Since there are two Sentinel 3 satellites, in the Sentinel 3 product they are distinguished by either _a or _b at the end of the variable name, eg ts_dh_glac_a is the array of time-series for individual glaciers from Sentinel 3A data, and its timestamps will be given by ts_t_a.

Finally, every time-series has associated uncertainties, which are indicated by _uncert at the end of the variable name, eg ts_dh_glac_a_uncert contains the uncertainties associated with ts_dh_glac_a.

3.3.2 Surface elevation change rate elements

Surface elevation change rates are derived from each time-series, and for a set of five-year long windows within the time-series.

mission_sec is the surface elevation change rate for full time-series
windowed_sec is the surface elevation change rate for windows within a time-series

As for the time-series, these variable names will be followed by _glac, _cell or _node. In the case of Sentinel 3 they will be followed by _a or _b. Again, their associated uncertainties will be indicated by _uncert. Eg the full mission surface elevation change rates per glacier from Sentinel 3A will be given by mission_sec_glac_a, and their uncertainties by mission_sec_glac_a_uncert.

3.3.3 Ancillary information

If a time-series represents a node or cell location, then its location in EPSG 3413 in m, will be given by arrays x_node and y_node or x_cell and y_cell, where the latter define the cell centre. If it represents a glacier, then the RGI v7 ID of the glacier will be given by array glac_id. Where cell-based data is given, the cell resolution in x and y respectively is given by cell_resolution. Where windowed SEC is given, the arrays window_start and window_end give the start and end times of each window.

In the joint product, it is necessary to distinguish between missions. Time-series contain data from two or three missions, biased to a common level, but not combined. The set of timestamps used by each mission are always the same, so the ts_mission_flag is provided to distinguish which timestamp came from which mission. SEC values likewise may contain data from two or three missions, and the combined_mission_flag distinguishes what combination of missions was used. The values for both mission flags are given in Table 3.1.

Table 3.1: Variable name elements and descriptions.

Attribute	Content
ts_	Time-series
_dh	Surface height changes, in m
_h	Surface heights, in m
_t	Times, in years
mission_sec_	Surface elevation change rate over the whole time-series, in m/y
windowed_sec_	Surface elevation change rate over a time window within the time-series, in m/y
_glac	Relates to a glacier as a whole
_cell	Relates to a grid cell
_node	Relates to an along-track location
_a	Data from Sentinel 3A
_b	Data from Sentinel 3B
_uncert	Uncertainty, in the same units as the variable of the same name
x_	X co-ordinate of central location of time-series, in m in EPSG 3413
y_	Y co-ordinate of central location of time-series, in m in EPSG 3413
cell_resolution	Cell size in x and y, in m
window_start	Start time of window, for windowed_sec, in y
window_end	End time of window, for windowed_sec, in y
ts_mission_flag	Joint product only - array of flags giving the mission at each datapoint in the time-series, 1 = CryoSat-2, 2 = Sentinel 3A, 3 = Sentinel 3B
combined_mission_flag	Joint product only - array of flags indicating the missions used in the surface elevation change rate derivation, 1 = CryoSat-2 and Sentinel 3 A and B, 2 = CryoSat-2 and Sentinel 3A, 3 = CryoSat-2 and Sentinel 3B, 4 = Sentinel 3 A and B

3.4 Known limitations

3.4.1 Coverage

Sentinel 3 cannot view the northern periphery of Greenland, as its northern orbital limit is 81.35N. Both CryoSat-2 and ICESat-2 can view the whole of Greenland, but CryoSat-2's gridded swath product at time of writing only covered glaciers near to the ice sheet. Figure 3.3.1 below shows the Sentinel-3A reference orbits, and the peripheral glacier map overlaid on an example month of data from the gridded swath product.

It is also challenging to recover elevations from strongly sloping terrain, so the input datasets contain many data gaps. The altimetry products were made in part to address whether sufficient data could be accumulated to be of use in elevation change determination in this region.

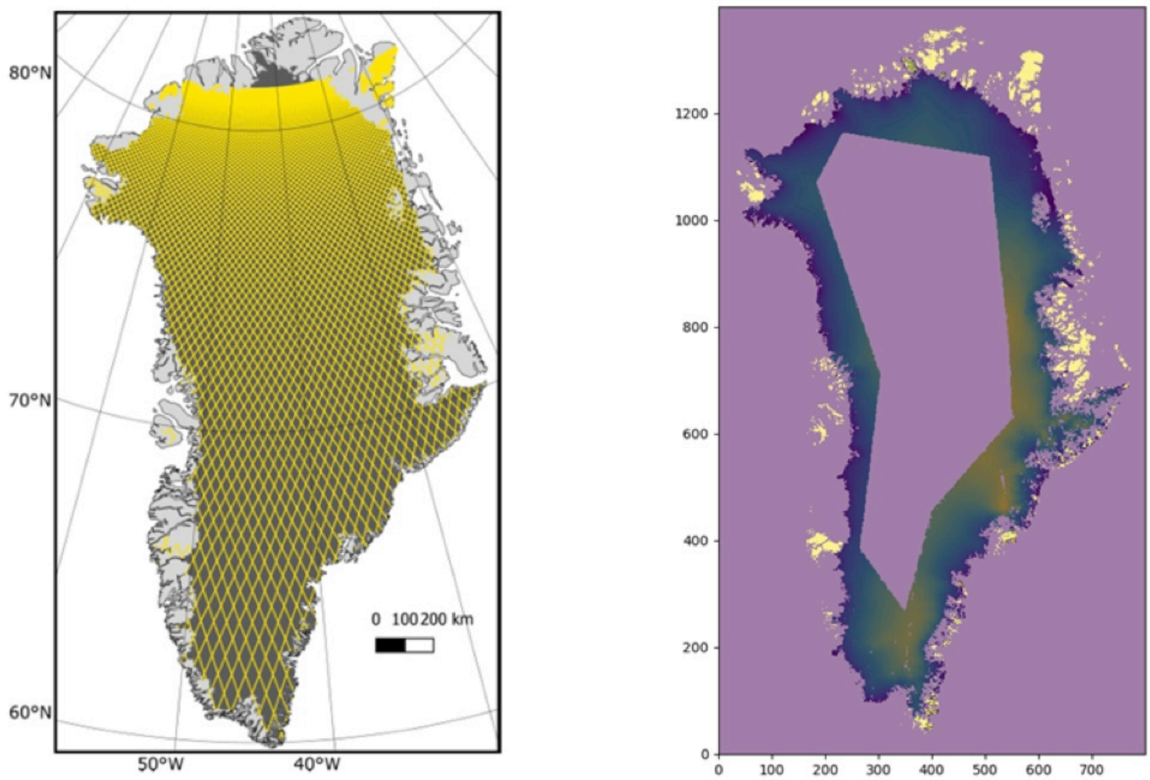

Fig. 3.1: Sentinel-3A reference tracks over Greenland (left). CryoSat-2 gridded swath product elevations (right), pale yellow where there is no data over a peripheral glacier and purple where there is no data over other surfaces - land, ice sheet, ocean.

3.4.2 Uncertainties

Some uncertainty values in the products are extremely high. This is due to high uncertainties in some of the input datasets, which were not filtered out in order to retain as much input and output data as possible. The user is welcome to apply their own filtering. Notes on the input elevation uncertainties are given below:
CryoSat-2 gridded swath: The highest uncertainty is 7 m.
CryoSat-2 point swath: The highest uncertainty is 20 m.
Sentinel 3: The highest uncertainty, in high slope regions, is 7.53 m.
ICESat-2 ATL14 (the DEM): The vast majority of uncertainties are below 7 m, but there are outliers up to 94 km.

3.4.3 Unrealistic change in input elevations, CryoSat-2 gridded swath

In a few cases there are unrealistic changes in elevations between months in the gridded swath product. One is shown in Fig 3.2, where the change is much larger than the uncertainties of the data points.

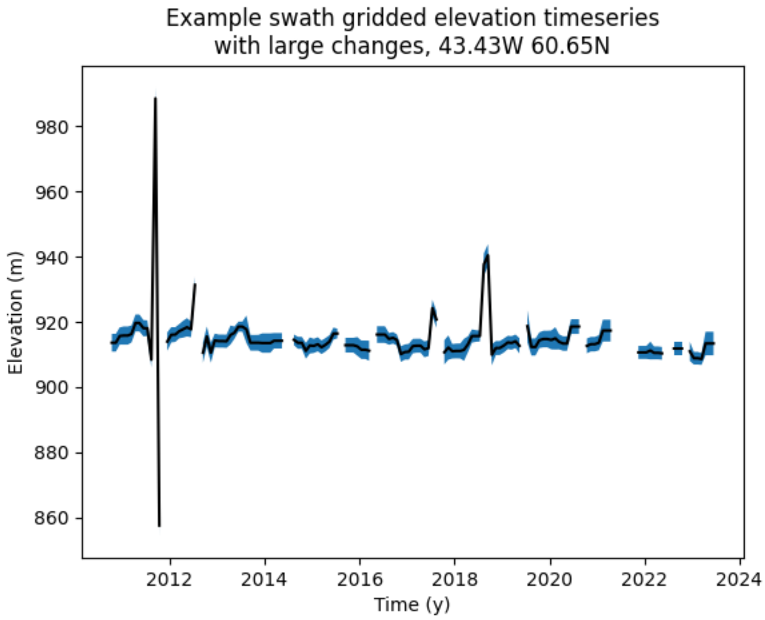

Fig. 3.2: An elevation time-series from the CryoSat-2 gridded swath product, showing an unrealistic change in elevation.

3.4.4 Sentinel-3 excess data loss at low elevations due to slope model

Elevations in the Sentinel-3 land ice thematic products are located to the point of closest approach using a slope model derived from the Alfred Wegener Institut (AWI) DEM from 2014 (Helm et. al., 2014). This DEM is much smoother in Greenlands peripheral regions than more modern DEMs, such as the Arctic DEM used for this project. Input elevations were filtered out if they were more than 100 m different to the Arctic DEM. This has led to excess loss of data at lower elevations. For example, the AWI and Arctic DEMs for the ice cap RGI 19164 are shown in Fig. 3.3. In a test dataset of 27 cycles of Sentinel-3A data, almost 60% of input elevations marked as good quality over this ice cap were rejected due to the discrepancy.

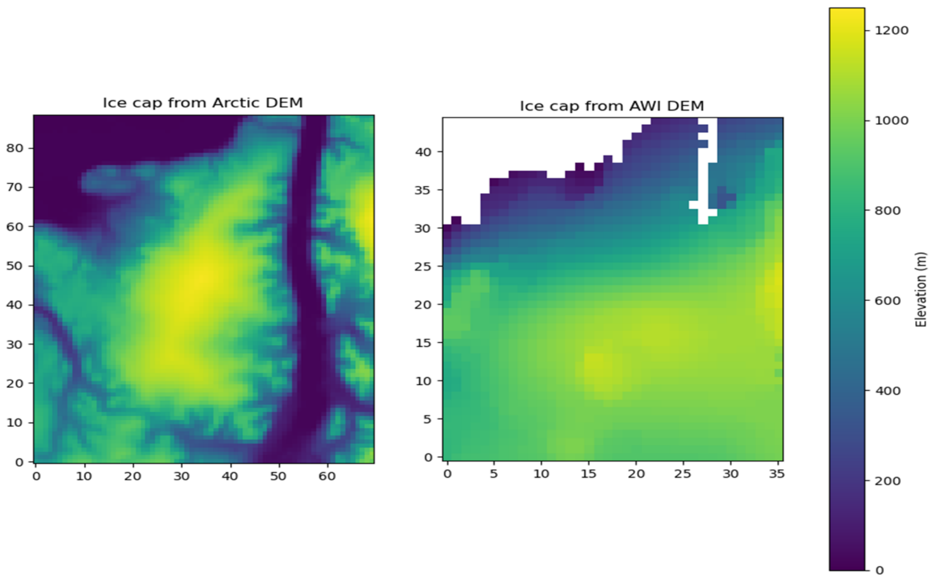

Fig. 3.3: Ice cap 19164 elevations, in Arctic DEM (500 m resolution, left) and AWI DEM (1 km resolution, right).

3.4.5 Mismatch between ICESat-2 and Arctic DEMs

In the Greenland peripheral glacier regions there are some small areas of mismatch between the ICESat-2 DEM from the ATL14 product, and the Arctic DEM. Similar mismatches were found against the Copernicus DEM in the example region shown below. This should not affect the output products, which rely on the ATL15 differences to the ATL14 DEM, but is noted here as an example of the difficulties faced by different instruments in these regions.

Figures 3.4 and 3.5 below show the ICESat-2 and Arctic DEMs for a 100 km by 100 km region of northern Greenland, and the differences between them when the Arctic DEM is interpolated to the ICESat-2 data locations. The ICESat-2 DEM is supplied for the ice areas only, all data is shown. Differences range from -750 m to +1 km.

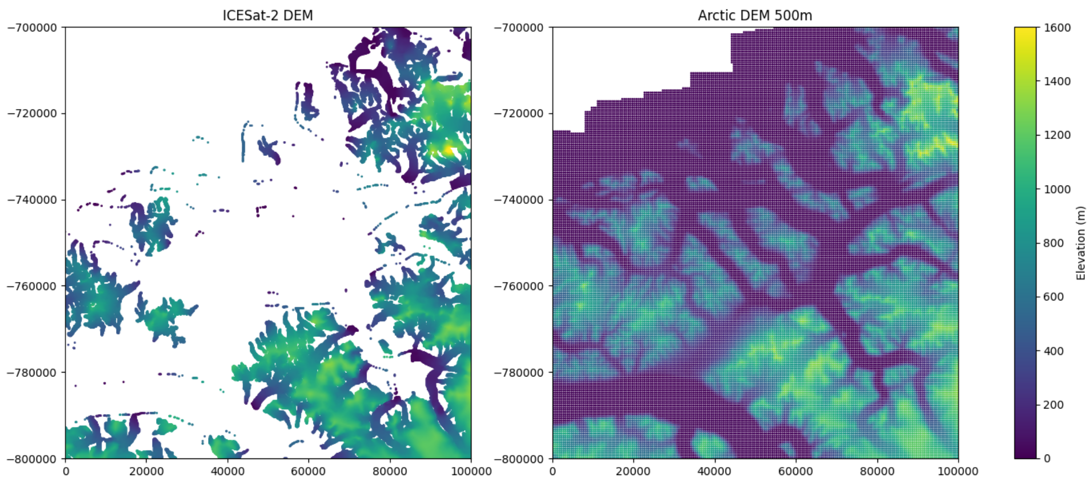

Fig. 3.4: ICESat-2 DEM (100 m resolution, left) and Arctic DEM (500 m resolution, right) over example region of northern Greenland periphery.

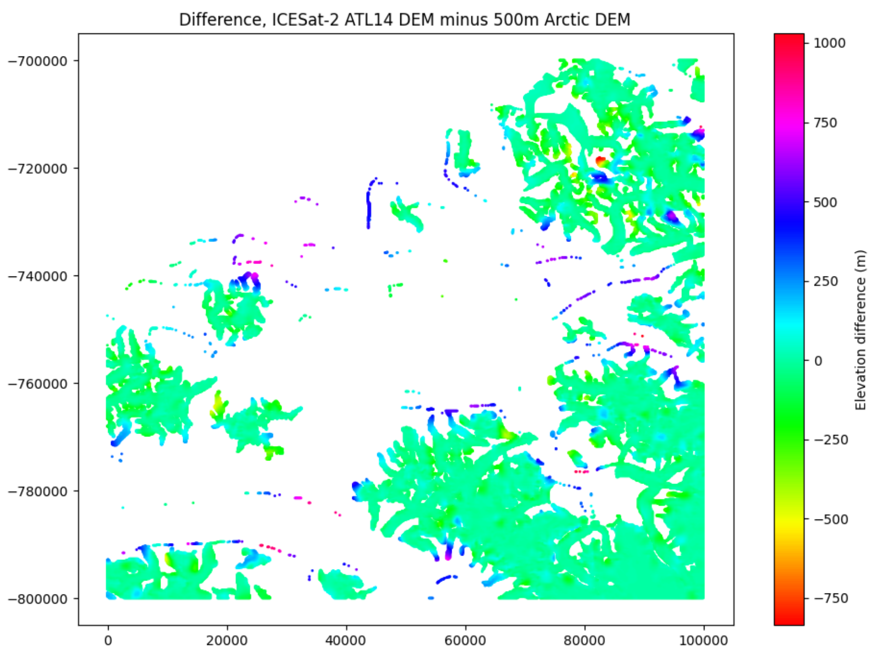

Fig. 3.5: Differences of ICESat-2 and the ArcticDEM interpolated to the ICESat-2 locations.

3.5 Software tools and data access

The data files are in netCDF format, a machine-independent format supported by the Unidata Program Center, and described here; https://doi.org/10.5065/D6H70CW6.

A list of software packages, both commercial, and free and open-source, for manipulating and displaying netCDF files, is maintained at:
https://www.unidata.ucar.edu/software/netcdf/software.html

All input datasets are freely available from:
Sentinel 3 land thematic datasets: https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-3-altimetry/products-and-algorithms/baseline-collection-005-processing-status
CryoSat-2 swath products: https://cryotempo-eolis.org
ICESat-2 land ice height datasets: https://nsidc.org/data/atl14/versions/3
Randolph Glacier Inventory v7.0: https://www.glims.org/RGI/
EPSG 3413 map projection: https://epsg.io/3413

4. Snow and glacier facies Karakoram

4.1 Product content

The surface classification of glaciers in the Karakoram provides information on the the snow cover on glaciers, the snow free clean ice areas on glaciers, and the snow free debris covered glacier areas at particular dates during the melting period per year, depending on the availability of cloud-free observations from Sentinel-2 and Landsat satellite data. Clouds over glaciers and non-glacierized areas are masked.

The product provides the classification of the glacier surface per pixel. Glacier outlines of the Randolph Glacier Inventory version 7.0 (RGI 7.0) are used as input, which were derived primarily from Landsat TM and Landsat ETM+ data of the year 2002 (RGI 7.0 Consortium, 2023). On glacierized areas, the following glacier surface classes are identified: snow cover extent, firn / old snow / bright ice, clean glacier ice and debris cover.

Clouds are masked over glacierized areas. Pixels outside the glacierized areas are set to no data. For each classified pixel, the associated uncertainty estimation is provided. Three categories of uncertainties are classified for each, illuminated and shaded pixels. Clouds and no data values in the uncertainty layer are each the same as in the product classification. Products are generated per Sentinel-2 track or Landsat path, respectively.

4.2 Product format

The glacier surface classification (GLS) product is provided as a two-layer raster file in netCDF format meeting the CCI Data Standards v2.3 (2021). The first layer provides the GLS classification, the second layer contains the associated uncertainty estimation per pixel. Additionally, a preview per product is provided as PNG.

Table 4.1: Format of the glacier surface classification product.

Attribute	Content
File format	netCDF, PNG (preview)
File name	When following the specifications for the Enveo database the filename for a single year (2018) is: 'glaciers_cci_facies_rgi14_S2_2018.nc' When following the filename convention of the CCI Data standards '<Indicative Date><Indicative Time>-ESACCI-L3U_GLACIERS-<Data Type>-<Product String>-<Additional Segregator>-fv<File version>.nc' the filename is '20220721053649-ESACCI-L3U_GLACIERS-GLS-MSI_S2B-005-fv1.0.nc'
Map projection	geographic coordinates, WGS84
Spatial extent	Karakoram
Pixel spacing	0.0002° x 0.0002°
Data type	8-bit unsigned integer
Product coding	0 = unclassified 1 = snow 2 = firn / old snow / bright ice 3 = glacier ice (clean) 4 = debris cover 5 = cloud 255 = no data

Uncertainty coding	0 = unclassified 1 = low uncertainty for illuminated pixels 2 = medium uncertainty for illuminated pixels 3 = high uncertainty for illuminated pixels 5 = clouds 11 = low uncertainty for shaded pixels 12 = medium uncertainty for shaded pixels 13 = high uncertainty for shaded pixels 255 = no data

4.3 Known limitations

Clouds are masked based on existing methods or products, which have been primarily developed for non-glacierized and snow free areas. Cloud shadows are not included in these methods and products. Misclassified clouds can obscure the glacier facies classification. Cloud shadows which are not matching the areas in cast shadow caused by the topography remain uncorrected and can introduce errors or major uncertainties in the classification. The shadow detection might miss some shadowed areas. Such pixels are thus handled as fully illuminated pixels and not corrected for any shadow effects, which can result in misclassified glacier facies classes.

Gaps in time series can occur if the satellite acquisitions do not meet the selection criteria. Each processing step and the classification itself introduces uncertainties to the final glacier facies classification. Misclassifications can be caused by any of these processing steps. The full error estimation of all processing steps is not yet included, as the assessment with a full error propagation approach would require further R&D and detailed analyses.

Changes in the total glacier areas in the Karakoram in the recent decades are not taken into account, but the glacier outlines of the year 2002 are used as static map for the full time series. The user should be aware that some changes in the debris cover might actually be the result of changing glacier areas, which is not taken into account in the product generation.

4.4 Software tools and data access

The products can be used with any software package that can handle the netCDF file format.

5. Glacier length changes Karakoram

5.1 Product content

The purpose of the glacier length change product was to reveal length changes of glaciers that are not surging and compare these to surge-type glaciers. As the latter might change their length also for non-climatic reasons, only length changes of not surging glaciers might have a relation to climatic forcing. From various investigations of geodetic glacier mass balance it is known that also the not surging glaciers have roughly a zero net mass balance over the past decades. Accordingly, these glaciers should only show minimal glacier front variations, i.e. length changes. The dataset has thus been created to answer a science question rather than feeding into a database. However, we think it can be shared and have thus included it in the CRDP [RD1] of the Glaciers_cci project.

The point shape file is listing in its attribute table the length for each of the investigated glaciers and the available time periods. Only those being fully covered by the 1965 Corona scene mosaic have entries for 1965, surge-type glaciers are flagged. For glaciers >1 km^2 length values are available for the years 1990, 2000, 2010 and 2020 usually over a 2 year time window (Fig. 5.1). Cumulative length changes can be obtained by subtracting the length from year t+1 from year t (e.g. the 1990 length from the 1965 length) and summing the values up, either with the original length or 0 as a starting point.

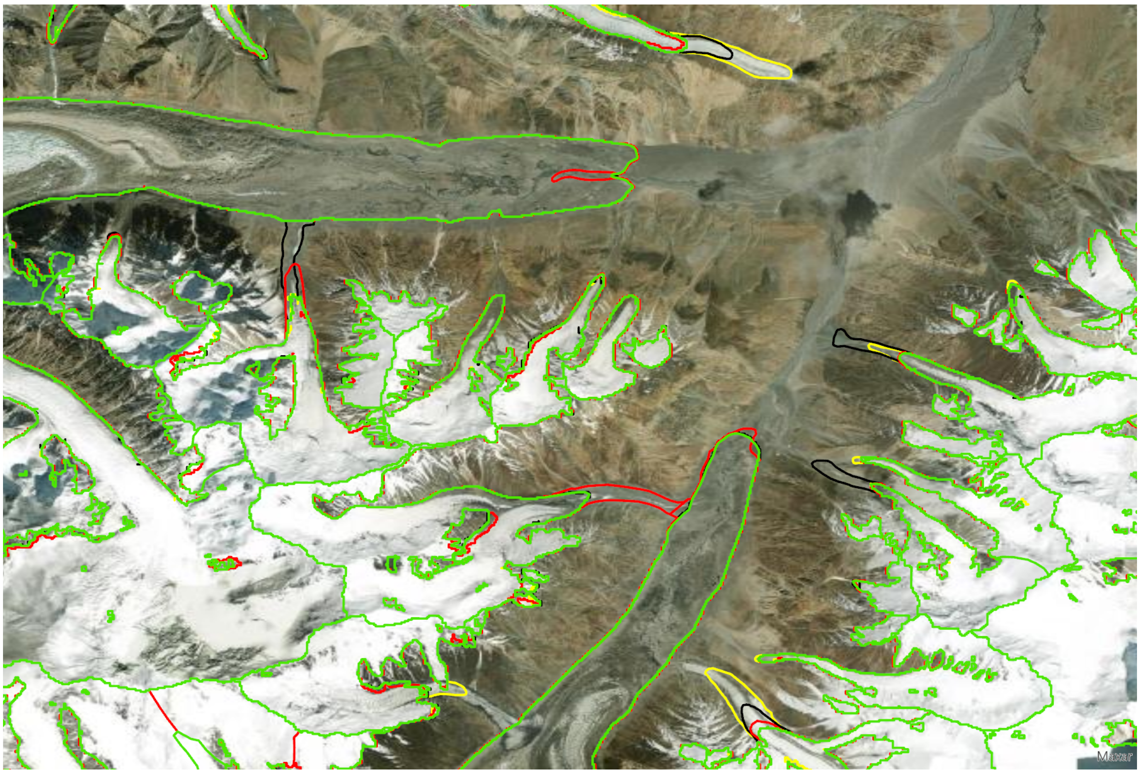

Fig. 5.1: Glacier outlines from 1990 (green), 2000 (red), 2010 (black) and 2020 (yellow) for the region around Skamri Glacier (upper left) in the central Karakoram. Glaciers that surged over this period show very colourful temini. The ESRI Basemap in the background is from 2023 and shows further advanced glaciers (yellow tongues at the lower centre).

5.2 Product format

The product is provided as a zipped point shape file along with a metadata sheet providing further details about the processing. Details of the included variables are listed in the PSD [RD2] and are shown as an example in Fig. 5.2.

	RGI_ID	Long	Lat	Length_65	Length_90	Length_00	Length_10	Length_20	Surge	Select
1238	RGI60-14.11074	75.30210000000	36.56130000000	NULL	2750	2768	2768	2768	0	0
1239	RGI60-14.11167	74.47590000000	36.43200000000	NULL	2924	2968	2976	2968	0	0
1240	RGI60-14.11173	73.92990000000	36.42310000000	NULL	881	881	881	881	0	0
1241	RGI60-14.11178	74.08380000000	36.43730000000	NULL	342	342	342	342	0	0
1242	RGI60-14.11179	74.20280000000	36.41970000000	NULL	2082	2096	1804	1533	0	0
1243	RGI60-14.11184	74.20210000000	36.43780000000	NULL	3835	3769	3448	3835	0	0
1244	RGI60-14.11184	74.20830000000	36.44910000000	NULL	3211	3255	3034	3255	0	0
1245	RGI60-14.11184	74.20310000000	36.44570000000	NULL	189	0	189	189	0	0
1246	RGI60-14.11212	73.99400000000	36.26130000000	NULL	1916	1955	1955	1751	0	0
1247	RGI60-14.11213	73.98550000000	36.24630000000	NULL	2322	2346	2346	2200	0	0
1248	RGI60-14.11327	76.19780000000	36.07910000000	NULL	2764	2764	2764	2764	0	0
1249	RGI60-14.11350	76.90310000000	35.24930000000	NULL	3186	3221	3221	3221	0	0
1250	NULL	74.49390000000	36.88550000000	3483	2641	2387	3211	2209	0	1
1251	RGI60-14.00005	75.72890000000	35.86290000000	85433	85433	85490	85490	85490	1	1
1252	RGI60-14.00005	75.67370000000	36.22030000000	51071	51066	51071	51071	50785	1	1
1253	RGI60-14.00005	75.50680000000	36.19900000000	50468	50444	50468	50468	49148	1	1
1254	RGI60-14.00032	74.88010000000	36.14120000000	37285	37285	37363	37363	37152	1	1
1255	RGI60-14.00032	75.11430000000	35.95050000000	54118	54092	54118	54118	54033	1	1
1256	RGI60-14.00036	74.44050000000	36.13650000000	9787	9978	9978	9978	9978	0	1
1257	RGI60-14.00042	74.02060000000	36.33850000000	9991	9852	8506	9863	9230	0	1

Fig. 5.2: Example of the information stored in the attribute table of the dataset.

5.3 Known limitations

Most glaciers in the Corona scene mosaic from 1965 are suffering from adverse snow conditions, i.e. they are covered down to their lowest points by snow. Identification of the terminus was thus only impossible for the largest glaciers with a terminus reaching down to low elevations (see Fig. 5.3). The dataset of glacier outlines from 1990 to 2020 by Xie et al. (2023) has been used as is, despite some interpretation errors. We acknowledge that length changes are often minimal, i.e. well within the geolocation and mapping uncertainties. Several glaciers do thus have a zero length change that might be in reality any value between ±30 m.

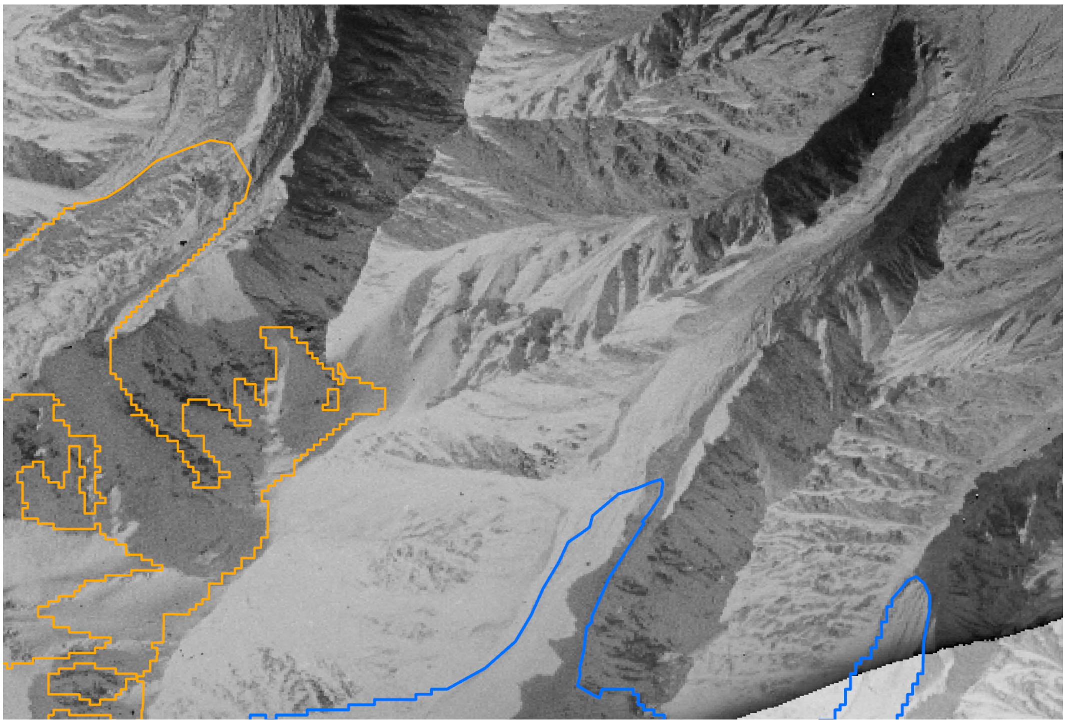

Fig. 5.3: Difficult glacier identification in the 1965 Corona mosaic. Terminus identification has some uncertainty for the (blue) glacier in the middle, the orange glacier is of surge type.

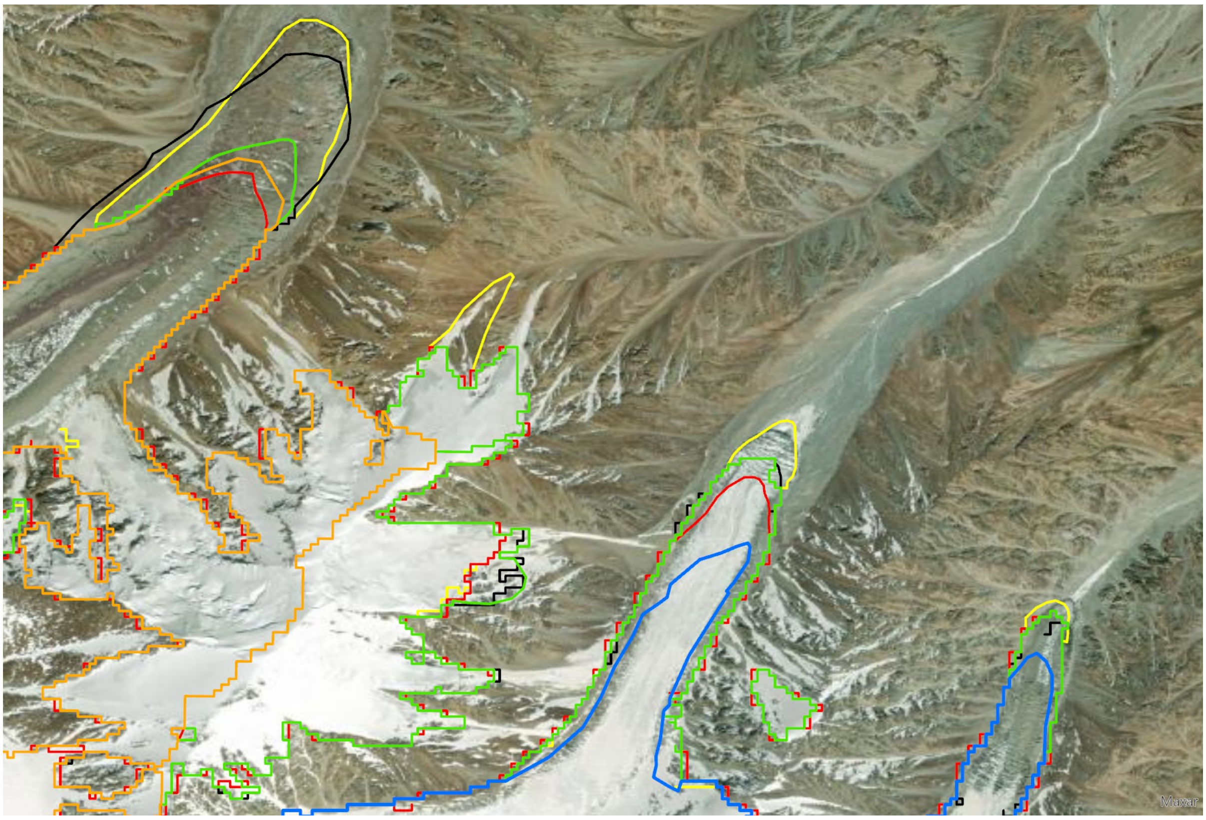

Fig. 5.4: The three glaciers from Fig. 5.3 have advanced from 1965 (shown by orange and blue outlines) to 2020. Whereas the orange glacier to the left has indeed recently surged, it is difficult to say for the two blue glaciers if this is really a surge (probably not for climatic reasons) or an advance (for climatic reasons). The strong advance of the glacier near the middle (yellow tongue) is likely due to a new interpretation rather than due to a real surge,

5.4 Software tools and data access

The shape file format is open and can be import by a wide range of free (e.g. QGIS) and commercial (e.g. ESRI ArcGIS) software packages. The glacier-specific information is stored in a .dbf file that can also be read by most software packages (e.g. OpenOffice, Excel) and converted to other formats such as .xls or csv.

References

Bhambri, R., Hewitt, K., Kawishwar, P. and Pratap, B. (2017). Surge-type and surge-modified glaciers in the Karakoram. Scientific Reports 7(1), 15391; doi.org/10.1038/s41598-017-15473-8

Farnsworth, W.R., Ingolfsson, O., Retelle, M. and Schomacker, A. (2016). Over 400 previously undocumented Svalbard surge-type glaciers identified. Geomorphology 264, 52–60; doi.org/ 10.1016/j.geomorph.2016.03.025

Goerlich, F., Bolch, T. and Paul, F. (2020). More dynamic than expected: an updated survey of surging glaciers in the Pamir. Earth System Science Data 12(4), 3161–3176; doi.org/ 10.5194/essd-12-3161-2020

Guan, W.J. and 7 others (2022). Updated surge-type glacier inventory in the West Kunlun mountains, Tibetan Plateau, and implications for glacier change. Journal of Geophysical Research-Earth Surface 127(1), e2021JF006369; doi.org/10.1029/2021JF006369

Guillet, G. and 6 others (2022). A regionally resolved inventory of High Mountain Asia surge-type glaciers, derived from a multi-factor remote sensing approach. Cryosphere 16(2), 603–623; doi.org/10.5194/tc-16-603-2022

Guo, L. and 5 others (2022). A new inventory of High Mountain Asia surge-type glaciers derived from multiple elevation datasets since the 1970s. Earth System Science Data, 15, 2841-2861; doi.org/10.5194/essd-15-2841-2023

Hugonnet, R. and 10 others (2021). Accelerated global glacier mass loss in the early twenty-first century. Nature 592(7856), 726–731; doi.org/10.1038/s41586-021-03436-z

Jiskoot, H., Murray, T. and Luckman, A. (2003). Surge potential and drainage-basin characteristics in East Greenland. Annals of Glaciology 36, 142–148; doi.org/ 10.3189/172756403781816220

Kääb, A, Bazilova, V., Leclercq, P.W., Mannerfelt, E.S. and Strozzi, T. (2023). Global clustering of recent glacier surges from radar backscatter data, 2017–2022. Journal of Glaciology, 69 (277), 1515-1523; doi.org/10.1017/jog.2023.35

Leclercq, P.W., Kääb, A. and Altena, B. (2021). Brief communication: detection of glacier surge activity using cloud computing of Sentinel-1 radar data. The Cryosphere 15(10), 4901–4907; doi.org/10.5194/tc-15-4901-2021

Sevestre, H., and D.I. Benn (2015). Climatic and geometric controls on the global distribution of surge-type glaciers: implications for a unifying model of surging, Journal of Glaciology, 61(226), 646-662; doi.org10.3189/2015JoG14J136

Xie, F., Liu, S., Gao, Y., Zhu, Y., Bolch, T., Kääb, A., Duan, S., Miao, W., Kang, J., Zhang, Y., Pan, X., Qin, C., Wu, K., Qi, M., Zhang, X., Yi, Y., Han, F., Yao, X., Liu, Q., Wang, X., Jiang, Z., Shangguan, D., Zhang, Y., Grünwald, R., Adnan, M., Karki, J., and Saifullah, M. (2022). Interdecadal glacier inventories in the Karakoram since the 1990s, Earth Syst. Sci. Data, 15, 847–867; doi.org/10.5194/essd-15-847-202

Acronyms

AWI	Alfred Wegener Institut
CCI	Climate Change Initiative
CRDP	Climate Research Data Package
DEM	Digital Elevation Model
ESRI	Environmental Systems Research Institute
ETM+	Enhanced Thematic Mapper plus
EPSG	European Petroleum Survey group Geodesy
ESRI	Environmental Systems Research Institute
ETM+	Enhanced Thematic Mapper plus
GLIMS	Global Land Ice Measurement from Space
GLS	Glacier Surface
ICESat	Ice, Cloud, and land Elevation Satellite
QGIS	Quantum GIS
RA	Radar Altimeter
RADAR	Radio Detection and Ranging
R&D	Research and Development
RGI	Randolph Glacier Inventory
SEC	Surface Elevation Change
TM	Thematic Mapper

Appendix

Use cases for the elevation change products

A set of examples of python code to read in and plot various parts of the dataset are given in the following. The output plot is given below each example. Conversion from the polar stereographic projection to latitude and longitude is demonstrated, and further comments are given in the code. For glacier outlines, the RGI Greenland peripheral glaciers shape file is used - it can be downloaded for free from the site linked in Section 3.4.

Use case 1: Time-series for a glacier, colour coded

```
#### Plot a joint time-series for a glacier, colour coded.
# In this case use glacier 17550, which is part of Flade Isblink, and covers 52 km2.

from netCDF4 import Dataset      # NetCDF file reader
import matplotlib.pyplot as plt  # Plotter
import numpy as np                      # Maths

# Open the netCDF4 joint product file

ds=Dataset('ESACCI-GLACIERS-L3-SEC-GREENPERI_JOINT-2010_2023-v001.nc')

# Read in the list of glaciers IDs, the time-series array and its timestamps

glac_id=ds['glac_id'][:]
ts_dh_glac=ds['ts_dh_glac'][:]
ts_t=ds['ts_t'][:]

# Read in the units of the time-series and timestamps from the variable attributes

ts_dh_glac_units=ds['ts_dh_glac'].getncattr('units')
ts_t_units=ds['ts_t'].getncattr('units')

# Read in the mission flags and get the flag values and their meanings from
# the variable attributes

ts_mission_flag=ds['ts_mission_flag'][:]
ts_mission_flag_values=ds['ts_mission_flag'].getncattr('flag_values')
ts_mission_flag_meanings=ds['ts_mission_flag'].getncattr('flag_meanings')

# The flag meanings are given in a space-separated string, so convert them to
# an iterable list.

ts_mission_flag_meanings=ts_mission_flag_meanings.split()

# Select the desired glacier by finding its index in the array of glacier IDs.
# This will be the primary index into the time-series array. Select its time-series.

this_glac=np.where(glac_id==17550)             # Returns a tuple
this_ts_dh_glac=ts_dh_glac[this_glac[0][0],:]  # [0][0] extracts first entry of tuple

# Iterate through the missions, plotting and labelling each separately
```

```
for i,v in enumerate(ts_mission_flag_values):
    r=np.where(ts_mission_flag == v)
    plt.plot(ts_t[r], this_ts_dh_glac[r], '.', label=ts_mission_flag_meanings[i])
plt.legend()                        # Adds a box with the flag meanings
plt.xlabel(ts_t_units)              # Labels x axis
plt.ylabel(ts_dh_glac_units)        # Labels y axis
plt.title('Joint product dh time-series, Glacier 17550')       # Set title
plt.show()                  # Display plot
```

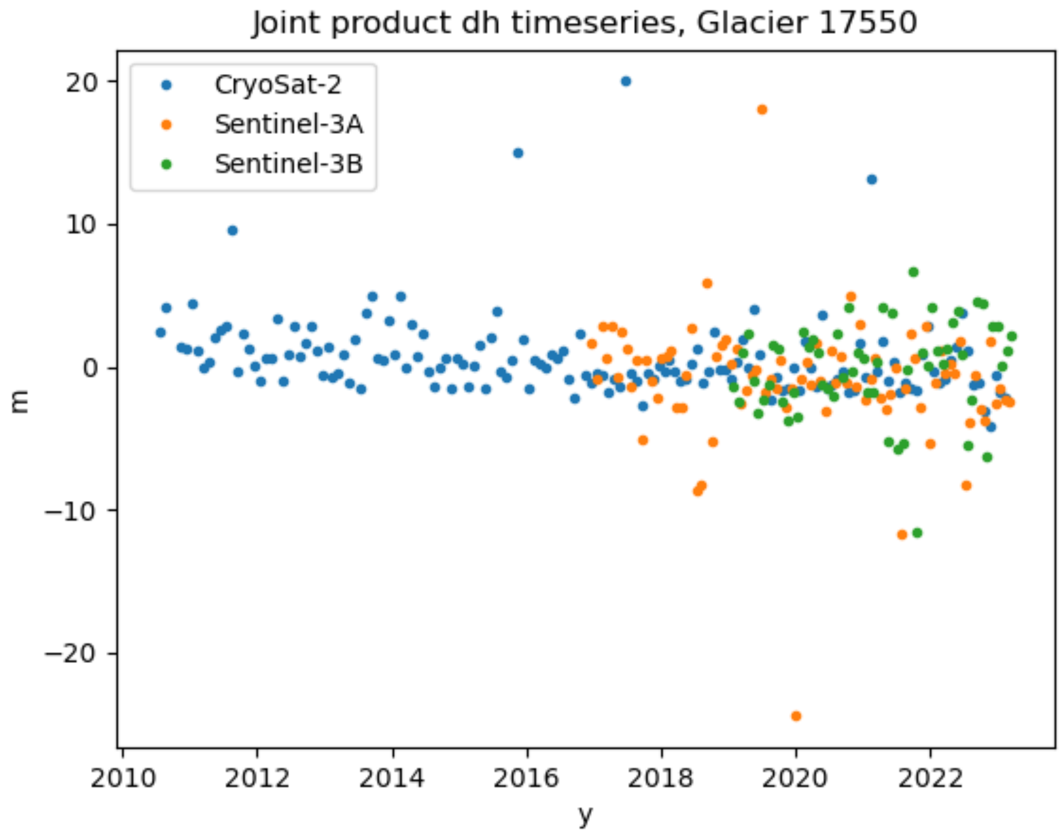

Fig. A.1: Output from joint product time-series example code.

5.4.1 Use case 2: Time-series with uncertainties

```
#### Plot a time-series with uncertainties, selecting one of the Sentinels and
# a cell, and find its latitude and longitude

from netCDF4 import Dataset     # NetCDF file reader
import matplotlib.pyplot as plt  # Plotter
import numpy as np                          # Maths
from pyproj import Proj          # Map projections

# Open the netCDF4 Sentinel 3 product file.

ds=Dataset('ESACCI-GLACIERS-L3-SEC-GREENPERI_S3-2016_2023-v001.nc')

# Read in the cell centre locations for Sentinel 3A

x_cell_a=ds['x_cell_a'][:]
y_cell_a=ds['y_cell_a'][:]

# Convert to latitude and longitude

proj=Proj('EPSG:3413')

lon_cell_a,lat_cell_a=proj(x_cell_a,y_cell_a,inverse=True)

# Read in the time-series and uncertainty arrays and their timestamps

ts_dh_cell_a=ds['ts_dh_cell_a'][:]
ts_dh_cell_a_uncert=ds['ts_dh_cell_a_uncert'][:]
ts_t_a=ds['ts_t_a'][:]

# Read in the units of the time-series and timestamps from the variable attributes

ts_dh_cell_a_units=ds['ts_dh_cell_a'].getncattr('units')
ts_t_a_units=ds['ts_t_a'].getncattr('units')

# Plot as an example the cell at index 2444, which is inside the glacier above,
# and put its location in the plot title. Only plot unmasked points, otherwise any
# data gaps will show in the plot. Include the uncertainty limits.

this_index=2444
this_lon=("{:.2f}".format(lon_cell_a[this_index]))  # Get as a string to 2 decimal places
this_lat=("{:.2f}".format(lat_cell_a[this_index]))

this_ts_dh_cell_a=ts_dh_cell_a[this_index,:]         # Extract the time-series for the cell
this_ts_dh_cell_a_uncert=ts_dh_cell_a_uncert[this_index,:]

r=(this_ts_dh_cell_a.mask==False)    # Find all unmasked points,

plt.plot(ts_t_a[r], this_ts_dh_cell_a[r], color='black')  # Main time-series, bridging data gaps
plt.fill_between(ts_t_a[r], this_ts_dh_cell_a[r]-this_ts_dh_cell_a_uncert[r],
this_ts_dh_cell_a[r]+this_ts_dh_cell_a_uncert[r])
plt.xlabel(ts_t_units)            # Labels x axis
plt.ylabel(ts_dh_glac_units)      # Labels y axis
```

```
plt.title('S3A dh cell time-series, longitude '+this_lon+' latitude '+this_lat)  # Set title
plt.show()
```

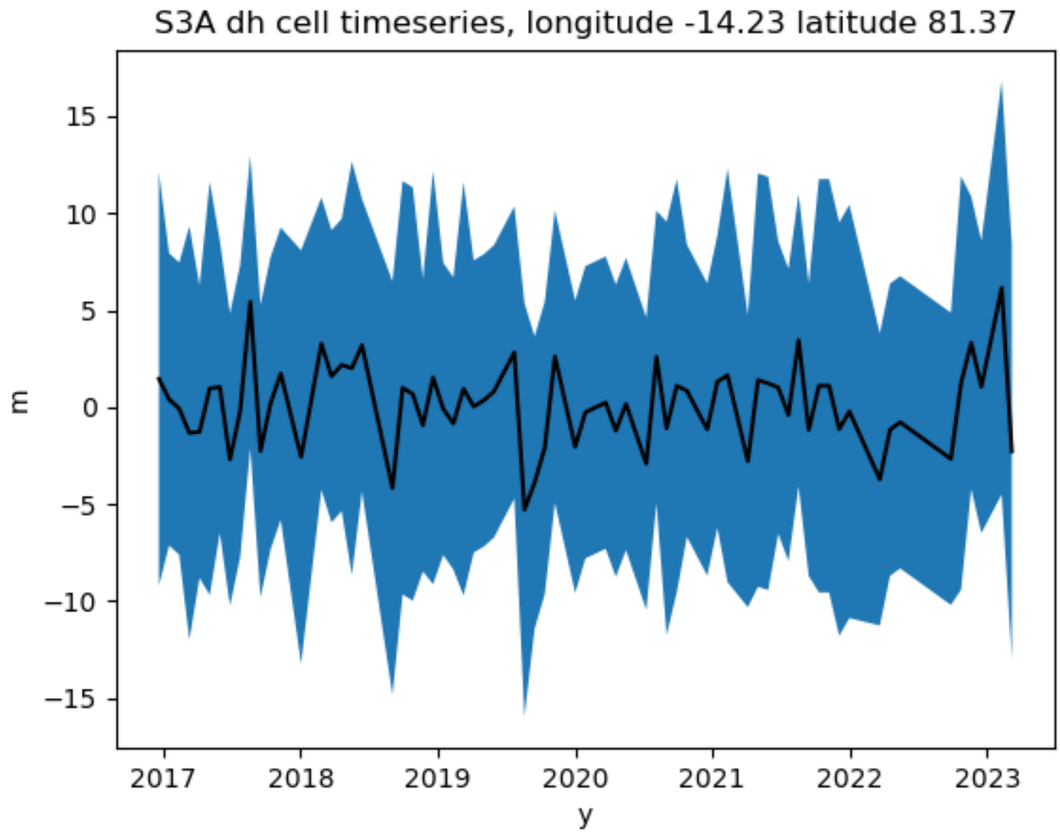

Fig. A.2: Output from Sentinel 3 product time-series example code

5.4.2 Use case 3: CryoSat-2 full mission SEC (Flade Isblink)

```
#### Plot the CryoSat-2 full mission SEC per glacier for Flade Isblink, as a map

from netCDF4 import Dataset     # NetCDF file reader
import matplotlib               # Plotter settings
import matplotlib.pyplot as plt # Plotter
import numpy as np              # Maths
from pyproj import Proj         # Map projections
import shapefile                # RGI reader
import copy                     # Copying utility

# Open the RGI 7 glacier outlines file

rgi_file='RGI2000-v7.0-G-05_greenland_periphery.shp'

sf=shapefile.Reader(rgi_file)
shaperecs=sf.shapeRecords()

# Find the approximate central location of each glacier. The records are ordered by
# glacier ID, with ID 1 at index 0, ID 2 at index 1, and so on. To make it easier
# to follow, make arrays where the index matches the glacier ID except for index 0,
# which is NaN filled.

num_shapes=len(shaperecs)

cen_lon=np.full(num_shapes+1, np.nan)
cen_lat=np.full(num_shapes+1, np.nan)

for i in range(num_shapes):
   cen_lon[i+1]=shaperecs[i].record.cenlon
   cen_lat[i+1]=shaperecs[i].record.cenlat

# Use the map projection to convert longitude and latitude to x and y locations.

proj=Proj('EPSG:3413')

cen_x, cen_y=proj(cen_lon, cen_lat)

# Open the netCDF4 CryoSat-2 point swath product file.

ds=Dataset('ESACCI-GLACIERS-L3-SEC-GREENPERI_CS2_POINT_SWATH-2010_2023-
v001.nc')

# Read in the list of glacier IDs and the full mission SEC for the glaciers

glac_id=ds['glac_id'][:]
mission_sec_glac=ds['mission_sec_glac'][:]

# Read in the units of SEC from the variable attributes

mission_sec_glac_units=ds['mission_sec_glac'].getncattr('units')
```

```
# Find the central locations of the glaciers for which a SEC has been found

cen_x_glac=cen_x[glac_id.astype(int)]
cen_y_glac=cen_y[glac_id.astype(int)]

# Find which of these are within Flade Isblink, and get their IDs and SEC values
# Set up the known limits of Flade Isblink in this map projection, in m

fi_min_x=400000
fi_max_x=510000
fi_min_y= -900000
fi_max_y= -750000

glac_in_flade_isblink=np.where((cen_x_glac >= fi_min_x) & (cen_x_glac <= fi_max_x) &
(cen_y_glac >= fi_min_y) & (cen_y_glac <= fi_max_y))

glac_id_in_flade_isblink=glac_id[glac_in_flade_isblink].astype(int)
sec_in_flade_isblink=mission_sec_glac[glac_in_flade_isblink]

# Set up the plotting colours to red to blue, limited to the range -1 to +1 m/y

cmap=matplotlib.cm.get_cmap('bwr_r')
norm=matplotlib.colors.Normalize(vmin=-1, vmax=1)

# Set up a plot, and loop through the chosen glaciers. Select the correct glacier outline
# (remembering that the index of a glacier in shaperecs is one less than its ID number)
# and fill with colour selected from the colourmap.

fig,ax=plt.subplots(figsize=[9,9])

for i,id in enumerate(glac_id_in_flade_isblink):

    this_shape_lonlat=np.array(shaperecs[id-1].shape.points)    # Get outline
    this_shape_xy=np.array(proj(this_shape_lonlat[:,0],this_shape_lonlat[:,1])) # Convert to
                                                                                  # EPSG 3413

    this_parts=copy.deepcopy(shaperecs[id-1].shape.parts) # Copy to allow alteration
    num_parts=len(this_parts)
    this_parts.append(this_shape_xy.shape[1])      # Add last point to close outline

    this_color=cmap(norm(sec_in_flade_isblink[i]))   # Norm shifts the value to the right point
                                                      # on the colormap given the limits selected above

    im=ax.fill(this_shape_xy[0,this_parts[0]:this_parts[1]], this_shape_xy[1,this_parts[0]:this_parts[1]],
color=this_color)
    ax.plot(this_shape_xy[0,this_parts[0]:this_parts[1]], this_shape_xy[1,this_parts[0]:this_parts[1]],
color='black')  # Outline

ax.set_xlim([fi_min_x,fi_max_x])      # Trim the plot to Flade Isblink limits
ax.set_ylim([fi_min_y,fi_max_y])
ax.set_title('Flade Isblink CryoSat-2 mission SEC per glacier')  # Add title
```

fig.colorbar(matplotlib.cm.ScalarMappable(norm=matplotlib.colors.Normalize(-1, 1), cmap='bwr_r'), ax=ax, orientation='vertical', label='SEC '+mission_sec_glac_units, extend='both') # Separate colour-bar

plt.show()

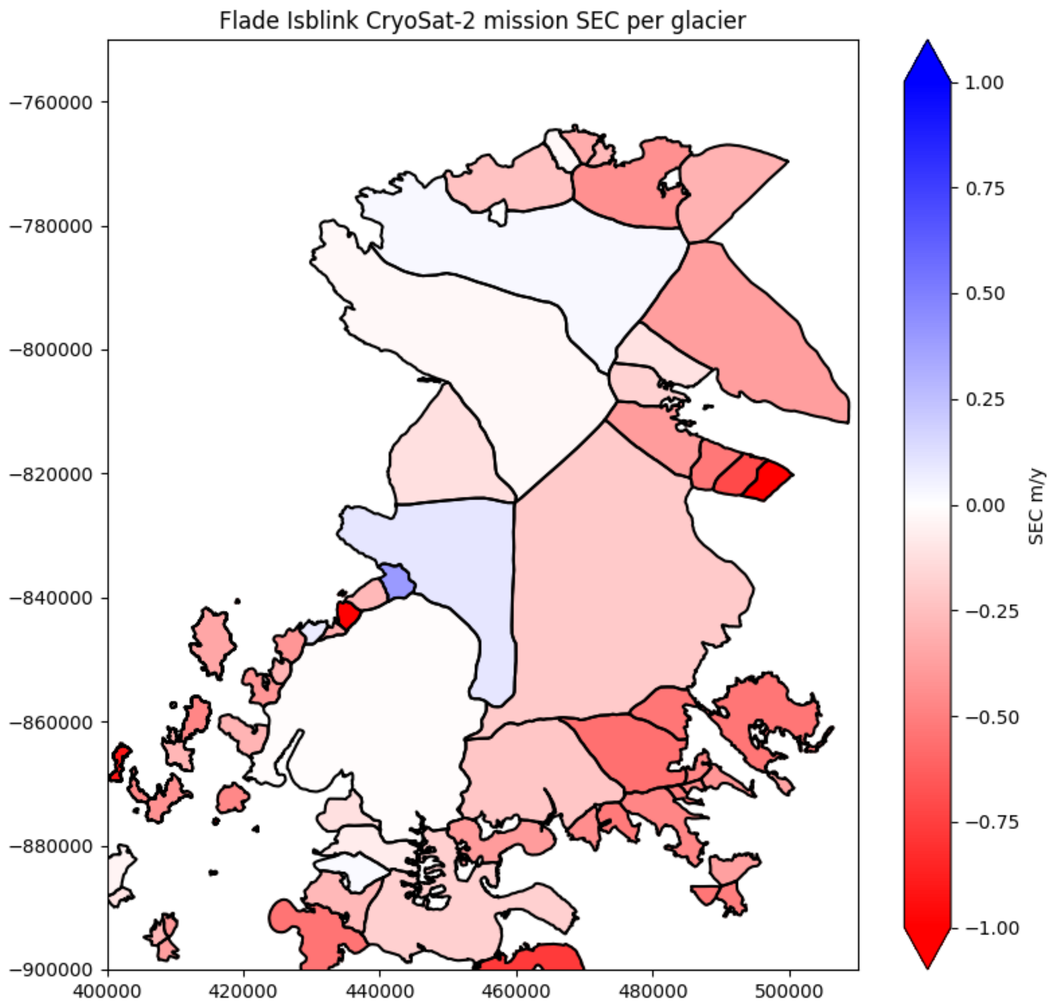

Fig. A.3: Output from CryoSat-2 product glacier SEC example code.

5.4.3 Use case 4: Plot all SEC windows, per cell

```
#### Plot all SEC windows, per cell, for a small area of Flade Isblink from CryoSat2

from netCDF4 import Dataset          # NetCDF file reader
import matplotlib.pyplot as plt      # Plotter
import numpy as np                   # Maths

# Open the netCDF4 CryoSat-2 point swath product file.

ds=Dataset('ESACCI-GLACIERS-L3-SEC-GREENPERI_CS2_POINT_SWATH-
2010_2023-v001.nc')

# Read in the per cell windowed SEC, cell centres and window time ranges

windowed_sec_cell=ds['windowed_sec_cell'][:]
x_cell=ds['x_cell'][:]
y_cell=ds['y_cell'][:]
window_start=ds['window_start'][:]
window_end=ds['window_end'][:]

# Read in the units of SEC from the variable attributes

windowed_sec_cell_units=ds['windowed_sec_cell'].getncattr('units')

# Select cells within the small region - in this case a region above Nord Station where ice loss
# is increasing. Set up the limits of the region, in this map projection, in m.

fi_min_x=433000
fi_max_x=441000
fi_min_y= -836000
fi_max_y= -824000

cell_in_region=np.where((x_cell > fi_min_x) & (x_cell < fi_max_x) & (y_cell > fi_min_y) &
(y_cell < fi_max_y))

# Set up a plot with one panel per window. Loop through the panels and plot, adding titles
# that give the window time ranges

fig,axes=plt.subplots(nrows=2, ncols=4, figsize=[16,8])
ax=axes.ravel()   # Flatten axes array so they can be looped through easily

for i in range(len(window_start)):
   im=ax[i].scatter(x_cell[cell_in_region],                    y_cell[cell_in_region],
c=windowed_sec_cell[cell_in_region,i], cmap='bwr_r', vmin=-1, vmax=1)
   ax[i].set_title('CryoSat-2 SEC per cell\n'+str(window_start[i])+' to '+str(window_end[i]))
   ax[i].set_xlim([fi_min_x,fi_max_x])     # Trim the plot to Flade Isblink region limits
   ax[i].set_ylim([fi_min_y,fi_max_y])
```

```
    ax[i].tick_params(axis='x', labelrotation=20)  # Avoid tick labels overlapping

fig.subplots_adjust(left=0.05, right=0.95, top=0.9, bottom=0.05, hspace=0.3, wspace=0.6)
                                                                  # Add space between panels
fig.colorbar(im, ax=axes, orientation='vertical', label='SEC '+windowed_sec_cell_units, extend='both')  # Separate colourbar

plt.show()
```

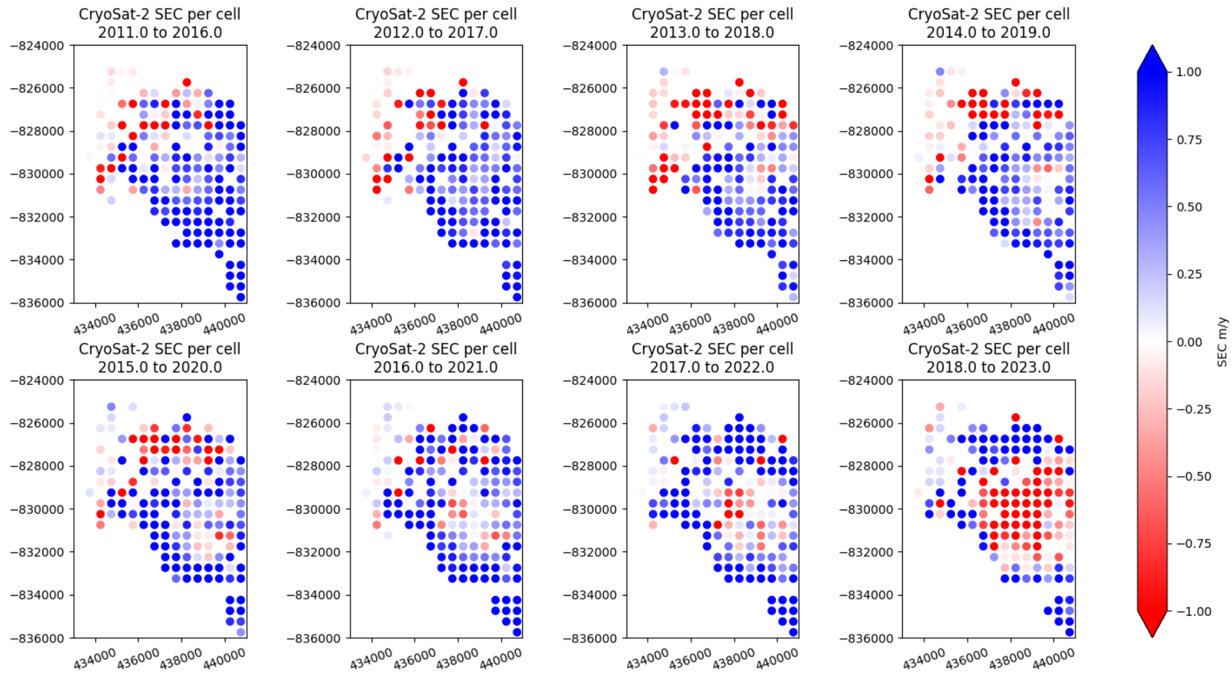

Fig. A.4: Output from CryoSat-2 product windowed cell SEC example code