Glaciers Today

CryoTEMPO-EOLIS

Algorithms Theoretical Baseline Description Document

Issue: 1.0
Date: 20th June 2025

Prepared by:	Carolyn Michael	
Issued by:	Carolyn Michael	
Checked by:	Tristan Goss	
Approved by:	Julia Bizon	

Document Versions

Issue	Date	Reason for change
1.0	20th June 2025	First version of document

Acronyms and Abbreviations

CryoTEMPO-EOLIS	Cryosat ThEMatic PrOducts – Elevation Over Land Ice from Swath
ESA	European Space Agency
DEM	Digital Elevation Model

Table of Contents

1 Introduction

1.1 Purpose and Scope

This document serves as the Algorithm Theoretical Baseline Document (ATBD) for **Glaciers Today**, Earthwave's glacier monitoring and insight service (accessible via https://cryotempo-eolis.org/). Glaciers Today leverages data generated by ESA's **CryoTEMPO-EOLIS** project to deliver a comprehensive, regularly updated record of glacier elevation changes spanning the duration of the **CryoSat-2** mission.

The objective of this document is to provide a detailed specification of the algorithms employed within the Glaciers Today processing system. It outlines the logical structure of the processing workflows, presents mathematical formulations and descriptions of each algorithm, and specifies the necessary input and output parameters. Additionally, the document highlights the assumptions and limitations associated with each algorithm, and includes relevant citations to justify methodological choices where applicable.

2 Glaciers Today Processing Algorithms

2.1 Logical Flow

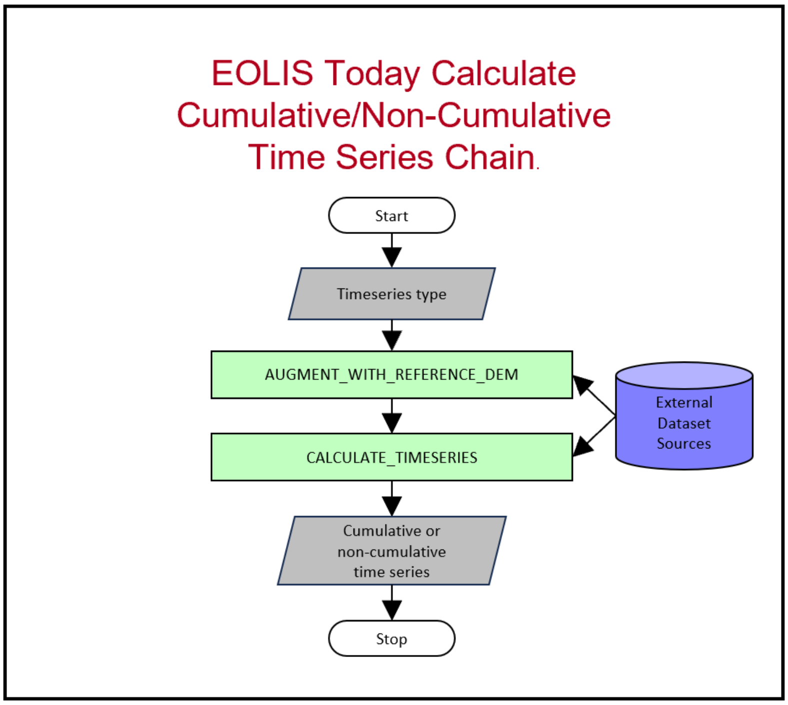

Figure 1: Logical flow of the Glaciers Today timeseries processing chain.

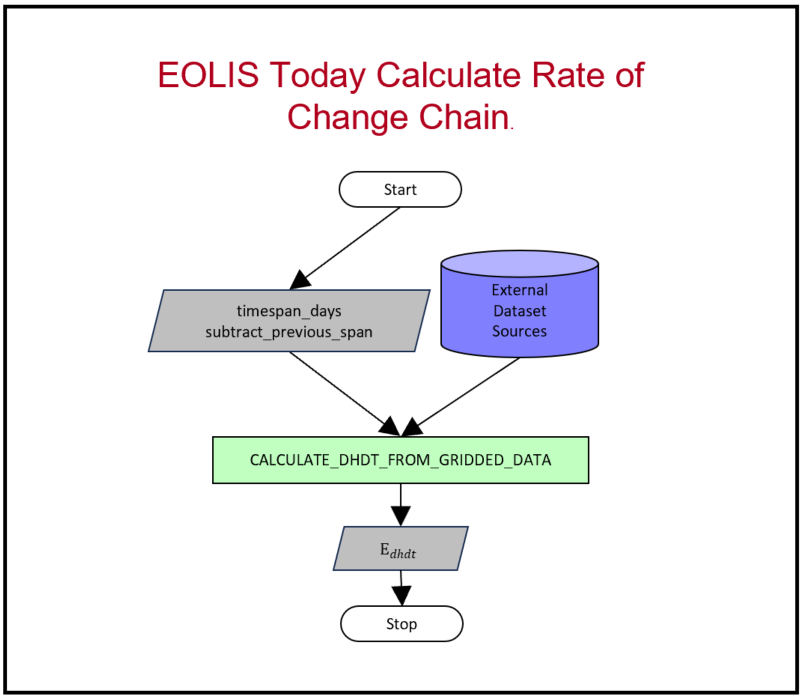

Figure 2: Logical flow of the Glaciers Today rate of change processing chain.

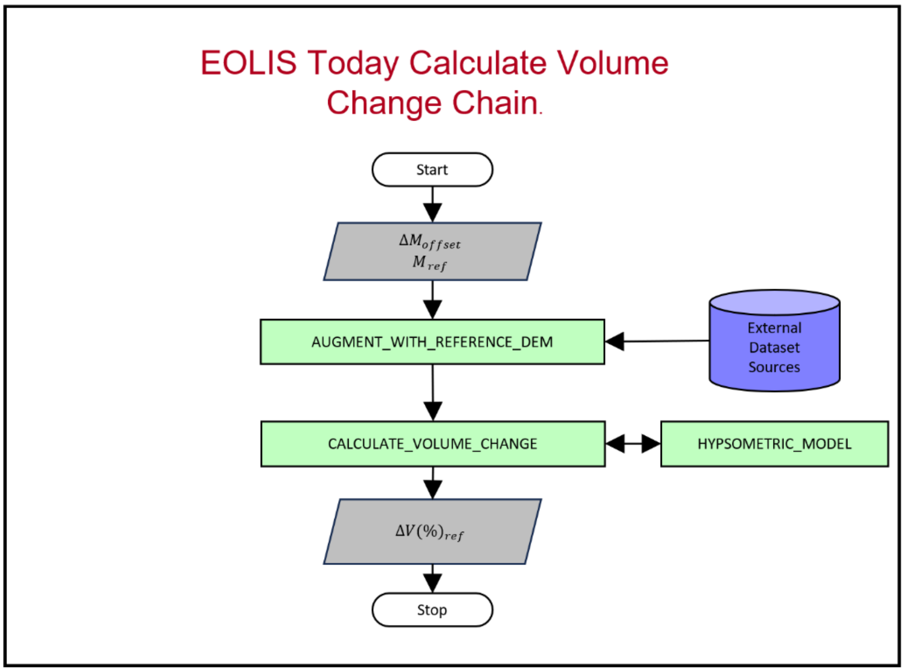

Figure 3: Logical flow of the Glaciers Today volume change processing chain.

2.2 Algorithm Descriptions

In this section, we describe each of the algorithms that make up the processing chain.

2.2.1 AUGMENT_WITH_REFERENCE_DEM – Purpose of algorithm

Statement of Function

Calculate a reference DEM from a gridded elevation product time series. Evaluate cumulative elevation change between the gridded elevation product at each time step and the reference DEM.

Limitations and Assumptions

The reference DEM, derived from the first N timesteps of elevation gridded data, is computed using a basic mean calculation. No adjustments are made for outliers or missing data, meaning any such anomalies will propagate through the subsequent timeseries analysis.

Algorithm Definition

The algorithm computes a reference DEM by averaging elevations from the first N timestamps of an input elevation dataset for each pixel. Once the reference DEM is calculated, the elevation change between the reference DEM and the input elevation dataset for each timestamp is calculated.

Input Parameters

Parameter Description	Symbol	Units	Source (default value)
Elevation gridded dataset at regular time intervals.	E_t	Metres	Auxiliary data source
Number of earliest unique timestamps to be used as input to calculate the reference DEM.	n_times	Integer	Input, default is 6

Output Parameters

Parameter Description	Symbol	Units
Computed reference DEM elevations and elevation change between E_t and the reference DEM for each pixel coordinate.	$E_{AUGMENTED}$	Metres

Mathematical Statement

E_t is an elevation gridded data product at regular time intervals:

$$\mathrm{E}_t = [\mathrm{E}_{t_0}, \mathrm{E}_{t_1}, \mathrm{E}_{t_2} \dots \mathrm{E}_{t_i}]$$
$$dates = [t_0, t_1, t_2 \dots t_i]$$

A reference DEM (E_{REF}) is calculated by taking the mean elevation value per grid coordinate (x, y) of the first six elevation gridded products ($\mathrm{E}_{t_0}, \mathrm{E}_{t_1} \dots \mathrm{E}_{t_5}$).

$$\mathrm{E}_{REF}(x, y) = \frac{1}{6}\sum_{i=0}^{5} \mathrm{E}_{t_i}(x, y)$$

A timeseries of the cumulative difference between the reference DEM and the elevations at each step for each pixel is calculated:

$$\Delta\mathrm{E} = \mathrm{E}_t(x, y, t) - \mathrm{E}_{REF}(x, y)$$

The elevation differences, ΔE, along with the calculated reference DEM, $\mathrm{E}_{REF}(x, y)$, are thus referred to as $\mathrm{E}_{AUGMENTED}$.

Implementation Notes

None.

References

None.

2.2.2 CALCULATE_TIMESERIES – Purpose of algorithm

Statement of Function

Create a cumulative time series of glacier elevation changes in a specified region. The resultant timeseries is presented on the Glaciers Today website to provide insights into glacier changes observed by CryoSat-2.

Limitations and Assumptions

The uncertainties of the time series represent the uncertainty on the mean value accounting for spatial coverage, however, they do not account for the uncertainty of the input data itself.

Algorithm Definition

1. **Cumulative Time Series Calculation**: The algorithm computes a cumulative time series of elevation changes by averaging the supplied gridded elevation change products at each timestep. The resultant time series is smoothed using an exponentially weighted averaging technique.
2. **Percentage Coverage**: The percentage spatial coverage per timestep is evaluated as a fraction of the total expected number of pixels in the region of interest.
3. **Uncertainty Estimation**: Uncertainties on the cumulative time series are defined as the standard error of the mean per timestamp, scaled by spatial coverage.
4. **Non-Cumulative Time Series Handling**: If the specified time series type is non-cumulative, the cumulative time series of elevation changes is differenced between subsequent timesteps to obtain non-cumulative values and the cumulative time series uncertainties are propagated from the contributing timesteps.

Input Parameters

Parameter Description	Symbol	Units	Source (default value)
Elevation gridded products for regular time intervals, augmented with a reference DEM as described in section 2.2.1.	$E_{AUGMENTED}$	Metres	Chain
The total number of glaciated pixels over the area of interest covered by $E_{AUGMENTED}$.	$C_{glacier}$	Integer	Auxiliary data source

Type of time series to calculate: cumulative or non-cumulative.	timeseries_type	String	Input
The half-life used for smoothing the time series. It represents the time (in days) over which the exponentially weighted average assigns decreasing weights to older observations.	λ	Days	Input, default is 30

Output Parameters

Parameter Description	Symbol
Elevation change, associated uncertainties, and percentage glacial coverage time series.	time_series_data

Mathematical Statement

Time Series Calculation

$\mathrm{E}_{AUGMENTED}$ contains cumulative elevation change measurements per pixel over time. Firstly, these elevation change measurements are grouped per timestep and a time series is calculated via simple averaging such that the cumulative elevation change at timestamp t_i is:

$$y(t_i) = \frac{1}{N}\sum_{j}^{N} \mathrm{E}_{AUGMENTED\,j}(t_i)$$

The cumulative timeseries is smoothed using an exponentially weighted average. Weights are defined as $w_i = (1 - \alpha)^i$ where alpha is the specific decay in terms of half-life (λ):

$$\alpha = 1 - e^{-\ln 2/\lambda}$$

here we by default define λ to be equal to the temporal resolution of the timeseries - for simplification $\lambda = 30\ days$. The weighted average of the timeseries is evaluated as:

$$y(t_i) = (y(t_i) + y(t_{i-1})(1 - \alpha)\ldots y(t_0)(1 - \alpha)^i)\,/\,(1 + (1 - \alpha)\ldots + (1 - \alpha)^i)$$

If timeseries_type is ‘cumulative’ then the timeseries is shifted to start at 0:

$$y(t_i) = y(t_i) - y(t_0)$$

If timeseries_type is ‘non-cumulative’ then the timeseries is converted to derivatives:

$$\Delta y(t_i) = y(t_i) - y(t_{i-1})$$

Percentage Coverage

The total count of non-NaN values for each timestep in $\mathrm{E}_{AUGMENTED}$ is calculated as:

$$C_{t_i} = \sum_{j=1}^{n} \mathbb{I}_{finite}\left(\mathrm{E}_{AUGMENTED\,j}(t_i)\right)$$

$$\mathbb{I}_{finite}\left(\mathrm{E}_{AUGMENTED\,j}(t_i)\right) = \begin{cases} 1, & if\ \mathrm{E}_{AUGMENTED\,j}(t_i)\ is\ finite \\ 0, & otherwise \end{cases}$$

Where:

- C_{t_i} is the count of finite values in the i-th timestep
- n is the total number of grid points
- $\mathrm{E}_{AUGMENTED\,j}(t_i)$ represents the j-th element of the $\mathrm{E}_{AUGMENTED}$ at time t_i

Percentage glacier coverage per timestep of $\mathrm{E}_{AUGMENTED}$ is defined as:

$$P_{t_i} = \left(\frac{C_{t_i}}{C_{glacier}}\right) * 100\ [\%]$$

This implicitly assumes that the processing mask used to produce $\mathrm{E}_{AUGMENTED}$ does not contain any non-glaciated regions, where glaciated regions are defined as per the above source.

Time Series Uncertainties

Uncertainties on the cumulative timeseries are defined as the standard error on the mean per timestamp scaled by spatial coverage. Implicitly, then, the algorithm describes the uncertainty in the elevation change of the full glaciated region, *not* the portion of the region covered by the gridded elevation dataset:

$$\sigma_y(t_i) = SEM(y(t_i)) * \left(\frac{100}{P_{t_i}}\right)^{0.5}$$

Uncertainties on the derivative timeseries are propagated from the cumulative timeseries uncertainties:

$$\sigma_{\Delta y}(t_i) = (\sigma_y(t_i)^2 + \sigma_y(t_{i-1})^2)^{0.5}$$

Implementation Notes

Auxiliary data is retrieved from Specklia.

References

None.

2.2.3 CALCULATE_DHDT_FROM_GRIDDED_DATA – Purpose of algorithm

Statement of Function

Calculate the rate of elevation change from gridded elevation measurements over time for plotting on the Glaciers Today website and estimating glacier volume change over time.

Limitations and Assumptions

All years are assumed to have 365.25 days.

Algorithm Definition

The algorithm generates a rate of elevation change dataset by applying a least squares linear regression to elevations at each grid coordinate through time over a specified timespan. If requested, then two rate of change calculations are performed over subsequent timespans and the anomalies between the two are evaluated.

Input Parameters

Parameter Description	Symbol	Units	Source (default value)
Elevation gridded dataset at regular time intervals.	E_t	Metres	Auxiliary data source
Timespan over which to calculate the rate of elevation change. Rate of change will be calculated from the maximum timestamp in $E_{MONTHLY}$ (t_{max}) to t_{max} - timespan_days.	timespan_days	Integer	Input

Output Parameters

Parameter Description	Symbol
Rate of surface elevation change, or rate of change anomalies, per year per grid coordinate and y-intercepts of regression.	E_{dhdt}

Mathematical Statement

First, E_t is filtered to keep timestamps between t_{max}, the maximum timestamp in E_t, and $t_{max} - timespan\_days$.

$$\mathrm{E}_{FILTERED} = \{\mathrm{E}_i \in \mathrm{E}_t \mid t_{max} - timespan\_days \leq t_i \leq t_{max}\}$$

Subsequently all elements of $\mathrm{E}_{FILTERED}$ are concatenated to create one table of gridded elevation measurements through time, E. All elevation measurements for all timestamps for each grid coordinate (E(x, y)) are extracted and a least squares regression is fit.

Here, the relationship between elevation and time is modelled as:

$$E(x, y, t) \approx \beta_1 * t + \beta_0$$

where:

- β_1 = rate of elevation change over time
- β_0 = elevation at t_0

The goal of least squares is to find values of β_0 and β_1 that minimise the difference between the observations, $E(x, y, t)$, and the model predictions, $\beta_1 * t + \beta_0$. In matrix form this can be expressed as:

$$\mathbf{E}(x, y, t) \approx \mathbf{T}\boldsymbol{\beta}$$

Where:

- $\mathbf{E}(x, y, t) = \begin{bmatrix} E(x, y, t_0) \\ \vdots \\ E(x, y, t_i) \end{bmatrix}$, is the vector of elevation measurements for the grid point $E(x, y)$ over time.
- $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, a vector of the unknown slope and y-intercept parameters.
- $\mathbf{T} = \begin{bmatrix} 1 & t_0 \\ \vdots & \vdots \\ 1 & t_i \end{bmatrix}$, is the design matrix, with a column of 1s for the intercept and time values in the second column.

The least squares solution is given by:

$$\boldsymbol{\beta} = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{E}(x, y, t)$$

Time coordinates are stored in seconds thus β_1 is in units of meters per second. This is converted to meters per year using a simple conversion:

$$\beta_1[m/y] = \beta_1[m/s] * 60 * 60 * 24 * 365.25$$

A table is constructed containing all unique pixel coordinates in $\mathrm{E}_{MONTHLY}$ and their fit slope and y-intercept values β_1, β_0 respectively, denoted as E_{dhdt}.

Implementation Notes

Auxiliary data is retrieved from Specklia.

References

None.

2.2.4 HYPSOMETRIC_MODEL – Purpose of algorithm

Statement of Function

Determine the best fitting model to describe the relationship between elevation and elevation change. This model can be used to interpolate missing data to predict elevation change values given an underlying reference elevation.

Limitations and Assumptions

It is assumed that the relationship between elevation and elevation change can be described using a polynomial model.

Algorithm Definition

1. **Input Data Preparation**: A reference DEM elevation value is merged with an elevation change gridded dataset to determine the corresponding elevation for each pixel.
2. **Binning Process**: Bins are defined based on the reference DEM. For each bin, the average elevation change value from the gridded dataset is calculated using the associated elevation values of the pixels.
3. **Polynomial Fitting**: The central elevation and mean elevation change values for each bin are used to fit polynomials of varying orders. Each polynomial is evaluated using the chi-squared score.
4. **Model Selection**: The polynomial with the minimum chi-squared score is selected as the best fitting model. This best fitting polynomial is then output from the algorithm.

Input Parameters

Parameter Description	Symbol	Units	Source (default value)
Elevation gridded products at regular time intervals, augmented with a reference DEM as described in section 2.2.1.	$\mathrm{E}_{AUGMENTED}$	Metres	Chain
Full coverage external reference DEM at the same or greater spatial resolution as $\mathrm{E}_{AUGMENTED}$.	DEM_{ext}	Metres	Auxiliary data source
Width of the elevation bands used to perform hypsometric averaging for filling missing data.	B_{width}	Metres	Input, default=50

Interval of the elevation bands used to perform hypsometric averaging for filling missing data.	$B_{interval}$	Metres	Input, default=100
The minimum number of points in a bin, used to perform hypsometric averaging for filling missing data, for it to be input into the polynomial model.	$B_{\min\_count}$	Integer	Input, default=100

Output Parameters

Parameter Description	Symbol
Best fit polynomial model, describing the relationship between elevation change and elevation, of order d.	P_d

Mathematical Statement

A hypsometric model is built with elevation as the independent variable and elevation difference, $\mathrm{E}_{AUGMENTED}$, as the dependent variable. The elevation bands are retrieved from a full coverage external reference DEM (DEM_{ext}) at the same or greater spatial resolution than $\mathrm{E}_{AUGMENTED}$.

The reference DEM elevation is retrieved for each elevation difference grid point ($\Delta E(x,y)$) and the data binned into elevation bands of width B_{width} and interval $B_{interval}$:

$$B_i = \{(x,y) \in \mathbb{R}^2 \mid DEM_{ext,i} \leq DEM_{ext}(x,y) \leq DEM_{ext,i+1}\}$$

where $DEM_{ext,i}$ and $DEM_{ext,i+1}$ are the lower and upper bands of the i-th elevation band.

For each band the mean elevation change ΔE, $\mu(\Delta E)_{B_i}$, and the mean elevation, $\mu(DEM_{ext})_{B_i}$, is calculated. Bins with fewer than $B_{\min\_count}$ points are removed.

Polynomials of order 1-3 are fit to the binned data where the independent variable is $\mu(DEM_{ext})_{B_i}$ and the dependant variable is $\mu(\Delta E)_{B_i}$. The fit polynomials are compared to the observations, $\mu(\Delta E)_{B_i}$, via a chi-squared statistic:

$$\chi_d^2 = \sum_{i=1}^{n} \frac{(\mu(\Delta E)_{B_i} - P_{d,i})^2}{P_{d,i}}$$

Where P is the evaluated polynomial fit to the observations with order d. The polynomial with the minimum χ_d^2 value is chosen and any missing grid points are assigned a ΔE by inputting the $DEM_{ext}(x,y)$ into the polynomial model.

Implementation Notes

Auxiliary data is retrieved from Specklia.

References

None.

2.2.5 CALCULATE_VOLUME_CHANGE – Purpose of algorithm

Statement of Function

Calculate the total ice volume loss for a region for presentation on the Glaciers Today website.

Limitations and Assumptions

It is assumed that elevation changes in the region are correlated with altitude and that the relationship can be represented using a polynomial function. The ice density is treated as a constant value throughout the analysis.

Algorithm Definition

1. **Calculate Elevation Change**: Compute the elevation change across the entire elevation gridded product timeseries supplied using the algorithm described in section 2.2.1.
2. **Interpolation Using Linear Regression & Hypsometric Averaging**: Fill pixels where gridded elevation data is available for some timesteps but not others using linear regression (see section 2.2.3) and then fill any remaining pixels that have no coverage at any timestep using hypsometric averaging (see section 2.2.4).
3. **Calculate Glacier Level Mass Change**: Assign all pixels to one or more intersecting glaciers. Calculate the mean elevation change per glacier and then convert to mass change using glacier area and ice density.
4. **Calculate Mass Change for Region**: Sum mass changes for all glaciers within the region of interest to determine the mass change for the full region.
5. **Calculate Volume Change Percentage**: Augment the mass change for the region with the mass change between some reference date and the start date of the elevation gridded dataset. Calculate the percentage change between the augmented mass change and the mass of the region at that reference date.

Input Parameters

Parameter Description	Symbol	Units	Source (default value)
Elevation gridded products at regular time intervals, augmented with a reference DEM as described in section 2.2.1.	$E_{AUGMENTED}$	Metres	Chain
Glacier extent definition for region. Each geometry defines an individual glacier.	glacier_extent		Auxiliary data source
This many gigatonnes will be added to the mass loss	ΔM_{offset}	Gigatonnes	Input

calculated from $E_{AUGMENTED}$. This value represents the mass change between some reference date and the start date of $E_{AUGMENTED}$ and allows the calculation of percentage volume change since the reference date.			
The glacier mass at the reference date for all glaciers in the region.	M_{ref}	Gigatonnes	Input
Estimated density of ice.	ρ_{ice}	kg/m^3	Input, default=850

Output Parameters

Parameter Description	Symbol	Units
Percentage ice volume change for the region since the reference date.	$\Delta V(\%)_{ref}$	Percentage

Mathematical Statement

The elevation change between the most recent timestep and the reference DEM is calculated:

$$\Delta E = E_{t_i} - E_{REF}$$

where E_{REF} is supplied in $E_{AUGMENTED}$ and is calculated as described in Section 2.2.1. ΔE provides an estimation of the elevation change over the full CryoSat-2 mission duration.

To fill data gaps (i.e. time-pixel combinations where data in $E_{AUGMENTED}$ does not exist, but where data does exist for that pixel in other timestamps) a least squares regression is fit to each grid point over all timestamps (see section 2.2.3) and the ΔE value determined by:

$$\Delta E = (\beta_1 * t_i + \beta_0) - (\beta_1 * t_{ref} + \beta_0)$$

where β_1 and β_0 are the slope and gradient parameters determined through the least squares regression for a particular grid point, t_i is the timestamp of the most recent observations and t_{ref} is the mean timestamp of E_{REF}.

Outliers in ΔE are identified as being points more than three standard deviations from the mean and are removed:

$$\mu - 3\sigma \leq \Delta E \leq \mu + 3\sigma$$

where μ is the mean of the dataset, and σ is the standard deviation.

Any additional missing grid points, defined as pixels in the E$_{AUGMENTED}$ coordinate reference system that intersect glaciers as defined within glacier_extent, but where there is no data in E$_{AUGMENTED}$ for any timestamp, are filled using hypsometric averaging as outlined in Section 2.2.4.

Each pixel in ΔE is defined as a square polygon with x coordinate boundaries [x–res/2, x+res/2] and y coordinate boundaries [y–res/2, y+res/2] where res is the spatial resolution of E$_{AUGMENTED}$. For each geometry in glacier_extent, defining a single glacier extent, all pixels in ΔE that intersect it are aggregated together to calculate the mean mass change over the full time span of E$_{AUGMENTED}$ for that particular glacier:

$$\Delta M_g = A_g * \rho_{ice} * \left(\frac{1}{n}\sum_{i=1}^{n} \Delta E_{g,i}\right) * 10^{-12}$$

where A_g is the glacier area in units m^2, $\Delta E_{g,i}$ is the i-th elevation change measurement for a particular glacier, ρ_{ice} is the density estimate for ice, and 10^{-12} is a conversion factor to convert the mass change units from kilograms to gigatonnes.

The total mass change over the full time span of E$_{AUGMENTED}$ for the full region is calculated by:

$$\Delta M_{CS2} = \sum_{g=1}^{G} \Delta M_g$$

The mass change is then presented as the percentage volume (equivalent to mass) change since a chosen reference date:

$$\Delta V(\%)_{ref} = \left(\frac{\Delta M_{offset} + \Delta M_{CS2}}{M_{ref}}\right) * 100$$

Implementation Notes

The mass offset, ΔM_{offset}, and reference mass, M_{ref}, are taken from the Glacier Mass Balance Intercomparison Exercise (GlaMBIE).
Auxiliary data is retrieved from Specklia.

References

GlaMBIE: https://doi.org/10.1038/s41586-024-08545-z