ECV Uncertainty Budget (EUB) Version 2

glaciers cci

Prepared by:	Glaciers_cci consortium
Contract:	4000127593/19/I-NB
Name:	Glaciers_cci+D2.2-2_EUBv2
Version:	1.1
Date:	11.03.2022

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]

UNIVERSITY OF LEEDS

GAMMA REMOTE SENSING

enveo

Document status sheet

Version	Date	Changes	Approval
0.1	10.08. 2020	Initial draft	
0.2	15.09. 2020	Final draft version compiled	A. Kääb
0.3	10.10. 2020	Comments from TO implemented	
1.0	07.02. 2022	Updated version v2	
1.1	11.03. 2022	Comments from TO on v2 implemented	F. Paul

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:
Andreas Kääb, Livia Piermattei, Désirée Treichler (GUIO), Thomas Nagler, Jan Wuite, Ludivine Libert (ENVEO), Tazio Strozzi, Andreas Wiesmann (Gamma), Frank Paul, Philipp Rastner (GIUZ), Lin Gilbert, Andrew Shepherd (SEEL)

Glaciers_cci Technical Officer at ESA:
Anna Maria Trofaier

Related documents

Acronym	Title	Document reference	Date
[RD1]	UCR CCI	Uncertainty Characterisation Reportv2	11.11.2016
[RD2]	ATBD CCI+	Algorithm Theoretical Basis Document	10.10.2020
[RD3]	PVIR CCI+	Product Validation and Intercomparison Report	21.12.2021

Table of Contents

1. Purpose of this document

This document provides a description of the ECV uncertainty budget for the new algorithms developed in Glaciers_cci+ (see [RD1] for the algorithms of Glaciers_cci Phase 2) and documented in its ATBD [RD2]. This document is thus also structured along the Glaciers_cci+ products, i.e. glacier area, elevation change and velocity. According to the Statement of Work (SoW), the Glaciers_cci+ EUB shall present major sources of error and estimate uncertainties according to each step of the retrieval process. Moreover, potential sources of error from ancillary data and external processes shall be investigated. Finally, sources of errors that are difficult or impractical to quantify shall be investigated.

We thus cover for each product and algorithm of the ATBD [RD2] the following aspects:

- Major error sources
- Uncertainties per processing step
- Other uncertainties

The first version of the EUB contained rather general, qualitative uncertainty descriptions, based on insights from previous studies. In the meantime, various products have been generated and validated for the two use case studies (see [RD3]). In this updated version 2 we summarize at the end of each subsection the main results obtained for the datasets derived for the two use cases Eastern Arctic and High Mountain Asia. The new sections are 2.2.4, 2.3.4, 2.4.4 and 2.5.4 for glacier area, 3.2.4, 3.3.4 and 3.3.5 for elevation change and 4.2.4, 4.3.4, 4.3.5 and 4.4.4 for velocity. For those products not being covered, we refer to related assessments from the literature.

2. Glacier extent

2.1 Ocean water

2.1.1 Major error sources

The presented method to map glaciers surrounded by ocean water is an extension of the standard method to map glaciers using a band ratio and a threshold value. It uses an additional threshold in the blue band to mask the comparably dark ocean water. The major error source is thus the selection of the correct threshold value for this band.

2.1.2 Uncertainties per processing step

It could be well possible that the selected threshold value is masking most of the surrounding ocean water but not all of it (e.g. when a different value also removes correctly classified glaciers). The resulting remaining water bodies with an often highly irregular shape can at best be selected and removed in the vector space. If the water is partly covered by sea ice, the method will only remove the clear water. The polygons of the wrongly classified sea ice are also at best selected and deleted in the vector space. Alternatively, the method presented in Section 2.2 can be used. In effect, this additional processing step introduces no additional uncertainties as possible wrong assignments will be manually identified and removed. Instead, the method decreases uncertainty of the product as a large number of calving glacier fronts have not to be digitized manually, but are separated pixel sharp from the water body. Manual digitizing would generalize the boundary across many pixels in an inconsistent way, introducing higher uncertainties than resulting from the raw classification.

2.1.3 Other uncertainties

There are no remaining quantifiable uncertainties in this method.

2.1.4 Use Case Eastern Arctic: Glacier mapping Franz-Josef-Land (Sentinel-2)

A carefully selected additional threshold in the blue band was applied to re-classify clean ocean water that was wrongly classified as glacier ice when using the standard band ratio method [RD3]. The required value was easy to find and removed clean ocean water completely and pixel sharp, i.e. no further manual corrections had to be performed in these regions. Accordingly, the uncertainty (here due to generalization by manual editing) decreased considerably. Hence, to obtain the best quality product in regions where glaciers have limited debris cover, it is highly recommended to use automated glacier mapping methods and apply manual digitizing only for remaining local corrections.

2.2 Removing clouds, sea ice and seasonal snow

2.2.1 Major error sources

When using cloud-based processing tools such as Google Earth Engine, one has to consider:

- Geolocation issues
- Image collection filter
- Statistical filtering
- Threshold selection
- Salt and pepper effect.

2.2.2 Uncertainties per processing step

Pre-processing
Geolocation: It is well known that Sentinel-2 images so far suffered from a geolocation problem due to a coarse DEM used during orthorectification Kääb et al. (2016). Since glaciers are often located in steep and rough terrain, imprecise and variable geolocation is a major problem for time series analysis (e.g. Sentinel-2A and -B images are shifted).

Image collection filter (from - to dates): The period allowed for image collection critically impacts on the quality of the results. The longer the period is, the higher is the chance to get a good results. However, glacier extents are quickly changing so a few years is a maximum duration of the period considered. The maximum number of images in an image collection has also to be found as this impacts on the processing time. Threshold values have to be defined for each region individually as also cloud frequency plays a role.

Main Processing
Statistical filtering: The standard statistical filtering (median and min) used is rather simple but effective. It will work over a wide range of conditions but maybe not for all. The result might be improved by using a larger image collection, but this comes at the expense that the date of the glacier outline is increasingly blurred. As individual pixel values are not altered by the method, the resulting image contains the digital numbers (DNs) from many individual images, acquired under different atmospheric conditions. In effect, methods that are based on reflectance calculation and consider an atmospheric / topographic correction might not work as good as the simple band ratio applied to the raw DNs (as these are pixel-wise consistent).

Threshold selection: Thresholds for the mapping of snow or ice are selected manually or based on values reported in the literature. They are thus subjective and if not reported, difficult to trace back. There might be more objective or automated ways for threshold selection but the sensitivity of the results to the threshold values is in general very small except for regions with snow and ice in cast shadow that should be analysed to find the 'best' value (e.g. Fig. 3b/c of doi.org/10.5194/essd-12-1805-2020). 'Best' means here that the workload for manual corrections is minimized.

Post-processing
Salt and Pepper effect: Due to strong changes in contrast (e.g. shadow on snow) and the geolocation problem, individual pixels with very different values might stand out in the final classification (called salt and pepper effect). Majority/minority analysis functions allow reducing or reclassifying such spurious pixels but their impact on the result needs to be further investigated.

Overall, there is no negative impact of the processing line on product quality (as remaining obvious errors are corrected manually), but a high positive impact on processing time and dataset completeness. The main drawback is that glacier outlines are no longer from a single date but refer to a shorter or longer date range.

2.2.3 Other uncertainties

As GEE is provided by a private company, long-term availability of this application is not guaranteed. Moreover, the 'Image Collection' filter is provided as an internal function in GEE and information about how it works and which scenes have been selected is difficult to obtain.

One might thus use the output rather as supplemental information to cover also regions missing in a more specific collection of scenes that is used as a base for the classification. Finally, it has to be mentioned that caution is necessary when working over large areas. If the region of interest exceeds a certain maximum amount of pixels, the result will not be exported. In effect, some experimenting is required to obtain the best possible result.

2.2.4 Use Case Eastern Arctic: Glacier mapping Svalbard (Sentinel-2)

For the year 2017/18 glacier inventory we have created for Svalbard we used manually selected scenes with near optimal glacier mapping conditions. Wrongly classified sea ice and lakes were removed manually, glaciers under clouds and remaining seasonal snow in the north of Svalbard were corrected using additional scenes (from September of the same year). Apart from the classic processing, we used GEE to export a mosaic of Sentinel-2 tiles for the September scenes. The lesson learned for the ECV Uncertainty budget is that manually selected scenes can be used to obtain best mapping results when a unique set of optimal scenes can be found and manual corrections for clouds are manageable. Otherwise (e.g. when a large number of scenes have to be mosaicked) the processing and filtering with GEE is preferable. It comes at the expense of a blurred time stamp, i.e. the outline of a glacier refers to a date range rather than a specific date.

2.3 Mapping debris-covered glaciers

2.3.1 Major error sources

Major errors in the data product (coherence images) are data voids resulting from layover and radar shadow (see also Section 4.2). Major errors in data interpretation result from moving objects that are not debris-covered glaciers (e.g. instable slopes or rock glaciers) or other regions connected to glacier fronts that lost coherence (e.g. water bodies).

2.3.2 Uncertainties per processing step

Data processing

The uncertainties introduced by the interferometric processing are described in Section 4.2 are not repeated here. It is important to note though, that debris-covered glaciers are comparably flat and are much less impacted by data voids than the surrounding terrain. Moreover, it has been shown [see RD2] that not all sensors are affected by data voids in the same way as orbiting geometries, SAR bands and repeat intervals differ.

Data Interpretation

The information provided by coherence images is used as a supplemental data source to more accurately map debris-covered glaciers. The datasets are thus used to increase product accuracy rather than decreasing it. However, as mentioned above, visual inspection of further datasets (e.g. false-colour composites or DEM products) is required for a correct interpretation as also other changing non-glacier objects have a low coherence. Additionally, the acquisition date of the coherence images has to be considered as also the extents of debris-covered glaciers change over time (though often not very strongly). There is thus no 1:1 transfer of glacier extents (in particular near the terminus) as visible in the coherence images and the datasets have to be interpreted carefully by the analyst.

In combination, data voids from the processing directly impact on the ability for the analyst to correct glacier outlines, but the impact might in most cases be small as additional data sources can be used for the interpretation (see Section 2.3.1 in [RD2]).

2.3.3 Other uncertainties

The part of the uncertainty that is difficult to quantify is the analyst-specific interpretation rule-set for debris-covered glaciers. In many cases there is a range of possibilities for interpretation and neither one nor the other might be fully right or wrong. As a measure to quantify the uncertainty in interpretation, we recommend performing a multiple digitizing of a couple of glaciers by different persons. The resulting variability in glacier extent gives a good measure of the uncertainties.

2.3.4 Use Case HMA: Glacier mapping in the Karakoram

We have not specifically investigated other methods for large-scale debris-cover mapping as we focussed for this use case on the investigation of specific glaciers. The study by Mölg et al. (2018) presents a detailed assessment of uncertainties for the existing debris-cover mapping methods.

2.4 Processing of Corona images

2.4.1 Major error sources

The processing of Corona images has two parts, DEM generation and orthorectification of the related satellite image. Any errors in the DEM impact on the geolocation of individual pixels. As the latter is a general issue for all satellite images (incl. Sentinel-2) we do not discuss details here but refer to Kääb et al. (2016) for related effects. Apart from the standard uncertainty sources in DEM generation (e.g. accuracy of GCPs used, stereo model parameters, handling of data voids, e.g. resulting from low contrast or radar shadow and layover) that are already described in the literature (e.g. Toutin and Grey 2000, Toutin 2004), there are some special error sources to be considered for Corona processing. These include:

- missing stereo coverage along the image boundaries
- an only roughly documented sensor model
- a highly variable, non-nadir image view geometry, making GCP identification difficult
- artefacts from the raw data (e.g. scanner and image acquisition issues)

2.4.2 Uncertainties per processing step

The uncertainties introduced by each processing step cannot be quantified, as only the final result of the DEM generation is accessible. We thus only get to know the combined uncertainties from all processing steps. The first two possibilities to determine the quality of the resulting DEM is creation of a hillshade and subtraction from another (reference) DEM. The hillshades will primarily reveal where the regions of uncertainty are located whereas the difference DEM will reveal how large they are. To use this information in a quantitative sense, it is important to mask glaciers and all other possibly changing objects and exclude major artefacts by applying a slope and/or magnitude threshold (e.g. exclude all terrain off-glaciers that is steeper than 30 degrees and DEM differences that are larger than 150 m). Such stable terrain differences give a first quantitative indication of DEM accuracy. Other reference DEMs might be applied for this estimation, but they might struggle with the same problems (e.g. artefacts).

When using the DEM to obtain elevation differences over glaciers, it might be sensible to further filter and reduce artefacts found on glacier surfaces. For the surge type glaciers we intend analysing, this is rather challenging as real changes can be in the same order of magnitude as the threshold value for the filter (or even higher). So this has to be carefully tested and evaluated for individual glaciers and might not be 'a one-threshold fits all' solution.

2.4.3 Other uncertainties

Usually the subsequent orthorectification is done per image stripe. If these are later mosaicked to cover a larger region, the missing regions between stripes are problematic. To get a complete mosaic, it is recommended to fill the data voids between the DEM stripes beforehand with data from another DEM. For glaciers that are in a different position then, the elevation differences will result in a geolocation shift, but this might be small and acceptable compared to the no-data option. But in general the Corona image series completely cover the earth surface for a certain area and single image stripes of one series are overlapping over the full width.

2.4.4 Use Case Eastern Arctic: Orthorectification Franz-Josef-Land (Corona)

For the use case eastern Arctic we orthorectified 12 non-stereo Corona KH-4 stripes acquired on 12.9.1962 over Franz-Josef-Land using an external DEM (ArcticDEM resampled to 10 m resolution) and ground control points (GCPs) collected from the Sentinel-2 scenes described in Section 2.1.4. Average root mean square errors of the location in x and y are 6.2 and 5.1 m, respectively, so about 1/2 pixel. During orthorectification the Corona stripes were resampled to a 10 m resolution and mosaicked for later glacier mapping. The elevation uncertainties of DEMs derived from Corona stereo pairs depend strongly on image conditions (snow cover, shadow) and imaging geometry (e.g. scene overlap). A discussion of related uncertainties is provided by Goerlich et al. (2017) and Kääb et al. (2021) for glaciers in the Tian Shan.

2.5 Detection of glacier surges

2.5.1 Major error sources

Our detection of glacier surges is based on operator-detection and interpretation of annual radar backscatter changes from stack statistics images, for now implemented in GEE. The major error sources stem from

- The data available within GEE (e.g., Sentinel-1 data holdings not complete within GEE)
- Operator-selection of processing parameters (e.g., ascending vs. descending orbits; polarization; definition of winter season; details see Section 2.5.2 'Data processing')
- Interpretation of backscatter changes
- Natural uncertainties (e.g., changes other than full surges; surges without pronounced crevassing).

2.5.2 Uncertainties per processing step

Data processing

A number of variations are possible during the data processing, all influencing the final stack statistics image. Most influence comes from the choice of orbit direction – ascending or descending, and season choice. For best results, both ascending and descending orbits are processed and investigated. However, if for some reason an operator choses only one orbit direction, or not both orbit directions are acquired or available in GEE, some surges might not be

recognized optimally, for instance due to radar shadow or overlay. The choice of optimal season on which the stack statistics is based might vary from region to region. If a general season is chosen for an entire hemisphere, sub-optimal conditions might be met for some regions. Test with varying season definitions help to understand this uncertainty, or to avoid it. For instance, varying the season as Nov to Mar, or Jan-Apr, or other periods. The goal is to choose a time period with as few short-term backscatter disturbances as possible, for instance from melt events. At the same time, a too short time period (e.g. only 1-2 acquisitions) has a limited impact on suppressing radar speckle effects.

Data interpretation
While for clear and full surges, interpretation of the backscatter changes between seasons seems robust, small surges or other glacier instabilities might not display clear backscatter changes. Comparison to optical images (also available in GEE) and comparison of results from different years are able to solve many of these uncertain cases, but several unclear cases certainly remain. The ultimate solution should thus be comparison to Sentinel-1 or Sentinel-2 / Landsat 8 derived velocity time series (see Section 4).

A further uncertainty is that a strict and unique relation between surging and backscatter increase through enhanced crevassing does not exist. Some surges might not display significantly enhanced crevassing or crevasses are covered by snow (undiscovered surges), or processes other than surging lead to backscatter increase ('false alarms'). It should be noted, though, that the latter type of cases leaves typically backscatter fingerprints that are clearly different from surges, or interesting anyway (e.g. avalanches, glacier lake changes, large calving events, strong glacier retreats/advances).

2.5.3 Other uncertainties

Similar issues as described in Section 2.2.3 for the Sentinel-2 image collection in GEE also applies to Sentinel-1. We discovered in rare cases Sentinel-1 scenes in GEE with georeference offsets. Such scenes lead to duplicate patterns in the stack results and are thus easily detected. In such cases, the offset scene(s) have to be detected and excluded from the stack.

Data holdings
If no Senetinel-1 data are available in GEE for a certain region and winter season, results will be empty and thus easy to detect. If data are available for a certain region and winter, though not complete, this might be more difficult to detect and deteriorate results in a hidden way. Additional statistics such as number of scenes and their temporal distribution could help to avoid such problems.

Above undiscovered surges also represent in principle an unknown uncertainty, unless an independent check of a region is done, not least based on velocity time series.

2.5.4 Use Cases Eastern Arctic and HMA: Detection of glacier surges (Sentinel-1)

The above method was applied globally for 2018-2019, and for other years for selected cases (Leclercq et al. 2021). The results were validated against optical and radar-derived velocity series using offset tracking for Svalbard and Alaska, which showed very high agreement (Leclercq et al. 2021). In addition, the method was used for a large number of cases in the Eastern Arctic and HMA where Sentinel-1 based velocity time series (processed by GAMMA Remote Sensing) suggested significant speed changes, and for the Chongtar glacier study (Fig. 2.1). In all cases, the Glaciers_CCI backscatter-change based method clearly confirmed

the glacier surges and enabled to trace back when first signs of instability/acceleration appeared. We conclude that the method, even if only qualitative and so far based on human interpretation, provides a very simple, quick and robust way of detecting glacier surges over large scales.

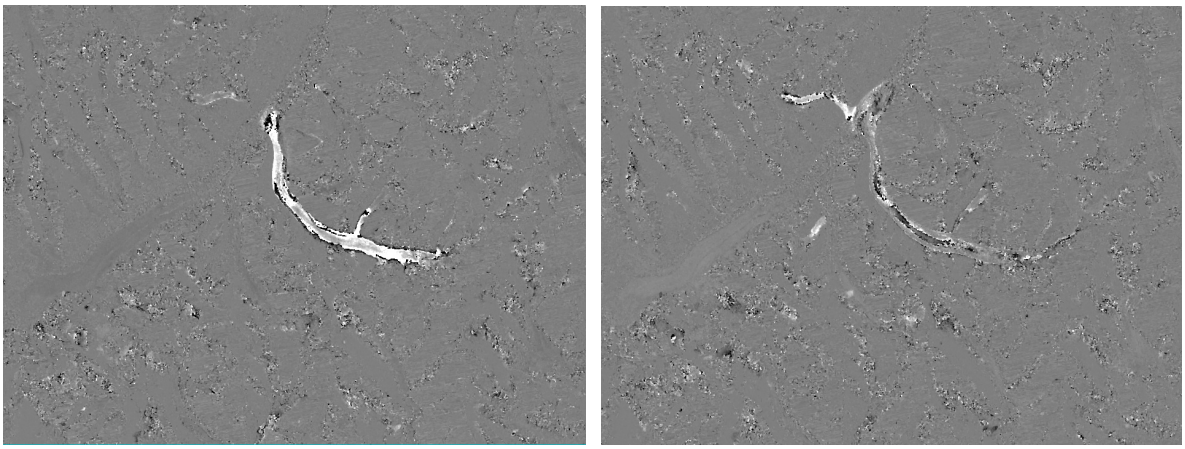

Figure 2.1: Sentinel-1 backscatter changes over Chongtar Glacier, left: 2020-2021, right 2021-2022. Bright zones indicate strongly increased crevassing, dark areas decreased crevassing, over the respective time period.

3. Elevation change

3.1 Radar penetration collection

This Glaciers_CCI+ product is mainly a collection of elevation differences between radar DEMs and reference data with no/little penetration. As such the uncertainty budget follows the Glaciers_CCI Phase II uncertainty budget [RD1] for DEM differencing, with the exception that radar penetration, which is part of the DEM differencing uncertainty budget, is treated as unknown, or target of the collection.

3.2 Seasonal and historic dh/dt from altimetry sensors

3.2.1 Major error sources

There are three major error sources:

- input data – including data corrections from auxiliary sources supplied within the input datasets
- cross-calibration
- modelling

The Antarctic mass balance trends reported in Shepherd et al (2019) used a similar data processing chain, and a similar error analysis will be performed.

3.2.2 Uncertainties per processing step

Dh aggregation from input, including surface-fitting
The input data uncertainty depends on the distribution of elevation measurements within each grid cell/epoch. The standard deviation of these measurements does not formally account for all uncertainty sources, but will include residual errors from radar penetration and volume scattering that are not removed by retracking and backscatter power correction, and factors such as radar speckle, satellite location uncertainty and atmospheric attenuation uncertainty which de-correlate within the cycle period (Wingham et al, 1998).

When calculating rates of surface elevation change, ie using a single cell and multiple epochs, the input error component is formed from the individual errors on each data point used. It is taken as the root mean square of the data point standard deviations, divided by the length of the time window.

dh time-series multi-mission cross-calibration
The cross-calibration uncertainty accounts for errors in the biases calculated between missions. For a given period, data from one or more missions may be used. If only one mission is used then no cross-calibration is necessary and the cross-calibration uncertainty contribution is zero. If more missions are used then the root mean square of the cross-calibration uncertainties of the missions involved is converted to an uncertainty on the rate of change, by dividing by the time period over which the rate is calculated

dh/dt modelling
The modelling uncertainty is the standard deviation of the model fit.

Uncertainty combination
The three uncertainties listed above are summed in quadrature to give a total uncertainty.

3.2.3 Other uncertainties

Overview
The input dataset includes corrections provided from many auxiliary datasets, e.g. atmospheric grids from weather-monitoring sites. Another possible source of uncertainty is in failure of the filtering (e.g. removal of outliers) applied during processing steps.

Estimation of Other uncertainties
Uncertainties from input data, even when not explicitly evaluated, are folded in via the statistical analysis of the surface-fitting during the first processing step, as above. Extreme outliers in altimetry measurements are an expected problem, given the complex environment from which they are derived. Outlier removal is applied to ensure a representative dataset at each step is used, since other statistical techniques (such as resampling from the same population) are not available. The uncertainty from any data point, where filtering failed, will be retained within the calculated standard deviations at each step.

As part of validation, error budgets will be compared for consistency.

3.2.4 Use Case Eastern Arctic: Radar Altimeter

For each of the four regions in the East Arctic use case, dh/dt was derived wherever possible on a 1 km by 1 km grid, in a moving time-window that was 5 years wide and stepped by 1 year to cover the period 1991 to 2021 inclusive. Figure 3.1 provides an overview of the data density per square kilometre in each of the four study regions. Table 3.1 details the mean values for each component of the error as well as the total error and the number of dh/dt values derived. However, the error distributions are very skewed, due to sampling issues and underlying glaciology, so a histogram of the total error distributions is shown in Figure 3.2. This error distribution is expected, and is similar to, e.g., that for the Antarctic ice sheet surface elevation change rates in the Copernicus Climate Data Store, shown in the pixel-level accuracy histogram in Gilbert (2021). The dominant error component is the uncertainty in the observations, reflecting the difficulty of radar altimetry over complex terrain.

Table 3.1: Summary of dh/dt total and component errors in each East Arctic region.

Region	Mean total error (m)	Number of dh/dt values	Mean input error (m)	Mean cross-calibration error (m)	Mean modelling error (m)
Svalbard	1.15	672149	0.86	0.13	0.62
Franz Josef Land	0.60	298177	0.44	0.08	0.34
Novaya Zemlya	0.84	374686	0.57	0.10	0.51
Severnaya Zemlya	0.84	322362	0.60	0.09	0.47

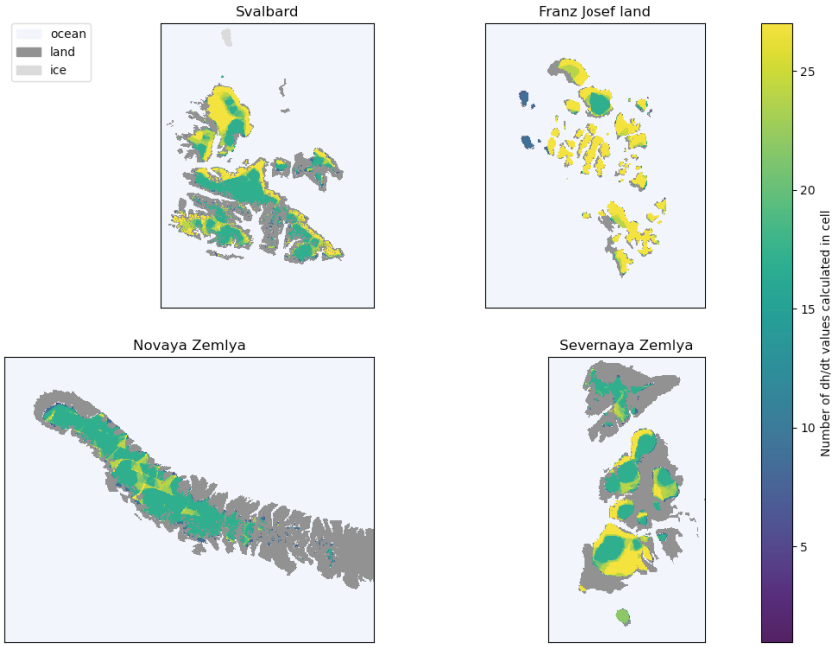

Figure 3.1: Data density in the glacierized areas of each East Arctic region. In each grid cell the maximum number of data points is 27.

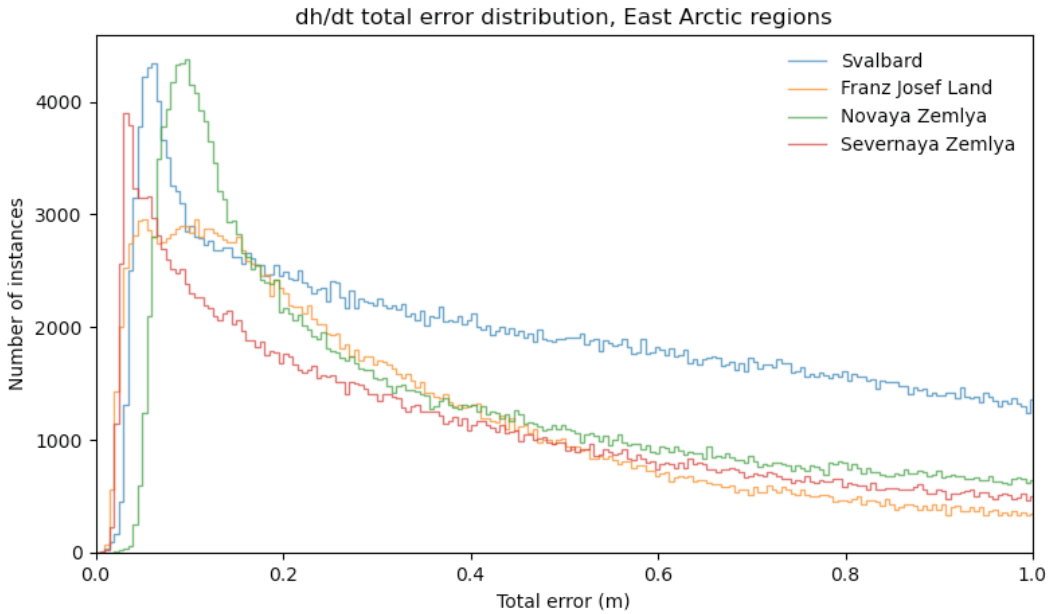

Figure 3.2: Histograms of the dh/dt total error distribution in each East Arctic region, excluding the distribution tails for clarity.

3.3 Seasonal dh/dt from ICESat in steep topography

The uncertainties in the High Mountain Asia surface elevation change datasets are similar to those in the Russian Arctic, whether the gridded dh dataset is derived by the surface fitting or the crossover method. However, in High Mountain Asia only a single mission is used, so cross-calibration is unnecessary.

3.3.1 Major error sources

There are two major error sources

- input data – including data corrections from auxiliary sources supplied within the input datasets
- modelling

3.3.2 Uncertainties per processing step

Dh aggregation from input, including surface-fitting
The input data uncertainty depends on the distribution of elevation measurements within each grid cell/epoch. The standard deviation of these measurements does not formally account for all uncertainty sources, but will include residual errors from factors such as satellite location uncertainty, elevation estimation from the ICESat-2 photon cloud, atmospheric attenuation, or uncertainties in a reference DEM if used to aid the surface-fitting or cross-over interpolation for instance for topographic normalisation (Treichler and Kääb 2016).
When calculating rates of surface elevation change, i.e. using a single cell and multiple epochs, the input error component is formed from the individual errors on each data point used. It is taken as the root mean square of the data point standard deviations, divided by the length of the time window.

Dh/dt modelling
The modelling uncertainty is the standard deviation of the model fit.

Uncertainty combination
The two uncertainties listed above are summed in quadrature to give a total uncertainty.

3.3.3 Other uncertainties

Overview
The input dataset includes corrections provided from many auxiliary datasets, e.g. atmospheric grids from weather-monitoring sites. Another possible source of uncertainty is in failure of the filtering (e.g. removal of outliers) applied during processing steps.

Estimation of further uncertainties
Uncertainties from input data, even when not explicitly evaluated, are folded in via the statistical analysis of the surface-fitting during the first processing step, as above. Extreme outliers in altimetry measurements are an expected problem, given the complex environment from which they are derived. Outlier removal is applied to ensure a representative dataset at each step is used, since other statistical techniques (such as resampling from the same population) are not available. The uncertainty from any data point where filtering failed will be retained within the calculated standard deviations at each step.

As part of validation, error budgets will be compared for consistency.

3.3.4 Use Case HMA: ICESat-2 performance

For the Chongtar Glacier study, the ICESat-2 ATL03 photon-product was mainly used as the ATL06 elevation product did not give useful results (Paul et al. in review). Due to the systematic off-pointing at mid-latitudes, ICESat-2 tracks are not repeated exactly in the study area and the ATL06 data *alone* proved too sparse, both geographically and temporally, for

further analysis of the surges. The ICESat-2 surface elevations fall into the time gap of the DEMs available for Chongtar between 2015 and 2020, thus providing additional temporal information on the surge development (Fig. 3.4).

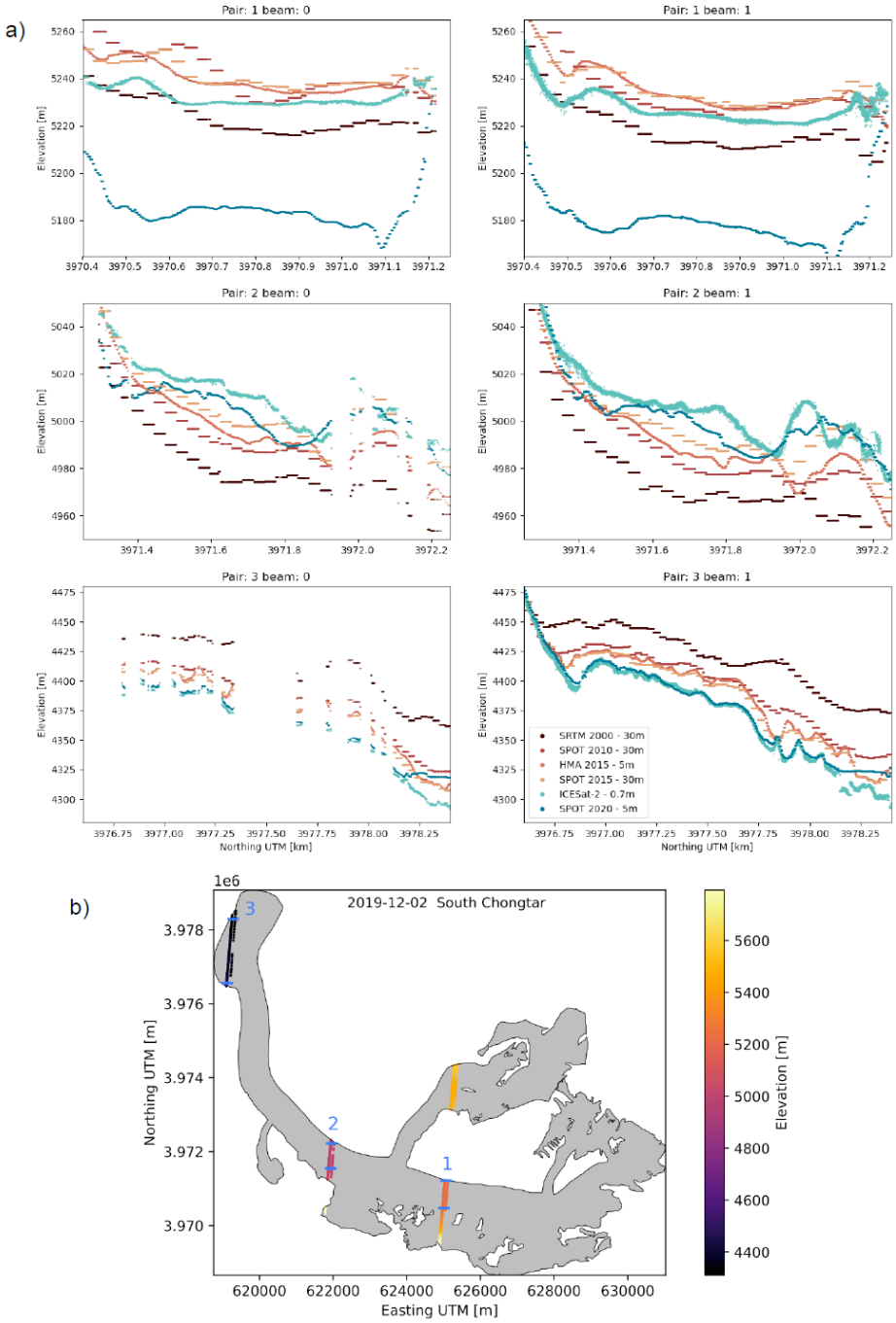

Figure 3.4: a) Elevation profiles of one ICESat-2 overpass over the tongue of South Chongtar glacier on 2.2.2019. Filtered ATL03 photon elevations are shown in blue, corresponding SRTM DEM elevations in black. The profiles are numbered from East to West (see map in panel b), i.e. profile 1 is highest up on the glacier. The left panels show the weak laser beam, the right panels the strong laser beam with c. four times as many photons per pulse shot. b) Locations of the six elevation profiles on South Chongtar glacier shown in a). The blue markers and numbers indicate which part of the glacier tongue is shown in a), and the profile number. The easternmost beam of each pair is the weak beam.

The surface elevation detail and accuracy of the freely available ICESat-2 ATL03 photon data surpasses all other datasets, including a SPOT 2020 high-resolution DEM. When combined with one or several DEMs, the higher temporal resolution of ICESat-2 provided additional information on how the elevation changed in-between DEM time stamps (Fig. 3.3.1). This may be very useful for slower changes or to further constrain the onset/end of a rapid change, such as a surge. However, ICESat-2 only provides elevation profiles with varying locations, which makes this data type more demanding to analyse. The footprints of the ICESat-2 ATL06 time series alone are too sparse to derive any useful trends in glacier surface elevation for smaller glaciers. Both the weak and the strong laser beams of ICESat-2's three beam pairs provide equally good data in the snow-covered accumulation areas. On darker and more rugged surfaces, the weak beam yields considerably fewer photon returns than the strong beam. The mean uncertainty of the ATL06 ICESat-2 data was ±5.37 m. However, we assume that ATL03 elevation uncertainties are in the order of decimetres on the relatively smooth glacier surface.

3.3.5 Use Case HMA: DEM differences

In the Chongtar study (Paul et al. in review) we quantify elevation changes of the glaciers before and during the surge using differences between multi-temporal DEMs from both optical and SAR sensors. Specifically, we used the SRTM DEM from February 2000, a SPOT5-HRS DEM from January 2010, a SPOT6 DEM from October 2015, and a SPOT7-derived DEM from October 2020. These DEMs were co-registered to the HMA DEM mosaic (Shean, 2017) used as a reference due to its superior spatial resolution and accuracy over stable terrain (off-glacier) compared to the other DEMs. For comparison, we also analysed elevation changes derived from the ASTER time series by Hugonnet et al. (2021). The co-registration accuracy of the DEMs was estimated from the DEM differences calculated on stable terrain (off-glacier) with slopes smaller than 40°. We found no indication of remaining horizontal shifts between the DEMs (Fig. 3.4).

On stable terrain, the mean elevation differences, which are sensitive to extreme values, are <1.4 m for all DEM difference pairs except for the SPOT 2020-SRTM2000 DEM pair (2.4 ±8.8 m). However, the differences between the DEMs are within the range of expected uncertainties considering the very steep and rugged terrain. After the co-registration, all DEMs provided useful results to track elevation and volume changes, independent of glacier size. The compared DEMs are of similar quality over glaciers, but the SPOT 2010 DEM used by Gardelle et al. (2013) suffered from strong artefacts at steep slopes. The elevation values of the SPOT 2015 and HMA DEM (which is also from 2015 in this region) are basically identical apart from individual raster cells. So elevation changes from 2000 (SRTM) to 2015 (HMA DEM) can also be derived from freely available DEMs. The SPOT 2020 DEM had impressive quality but the raw image pair had to be purchased.

The DEM time series from ASTER images derived by Hugonnet et al. (2021) shows the same trends as from the DEMs used here. However, the ASTER DEM time series missed detecting local changes of smaller glaciers. In addition, the strong spatial filtering inherent in the ASTER dataset, smoothens artefacts and data gaps off and to some degree also details on glaciers. Nevertheless, this dataset provides a larger overview and complementary information on cumulative elevation changes from 2000 to 2019 and the temporal evolution over several larger glaciers can be well followed.

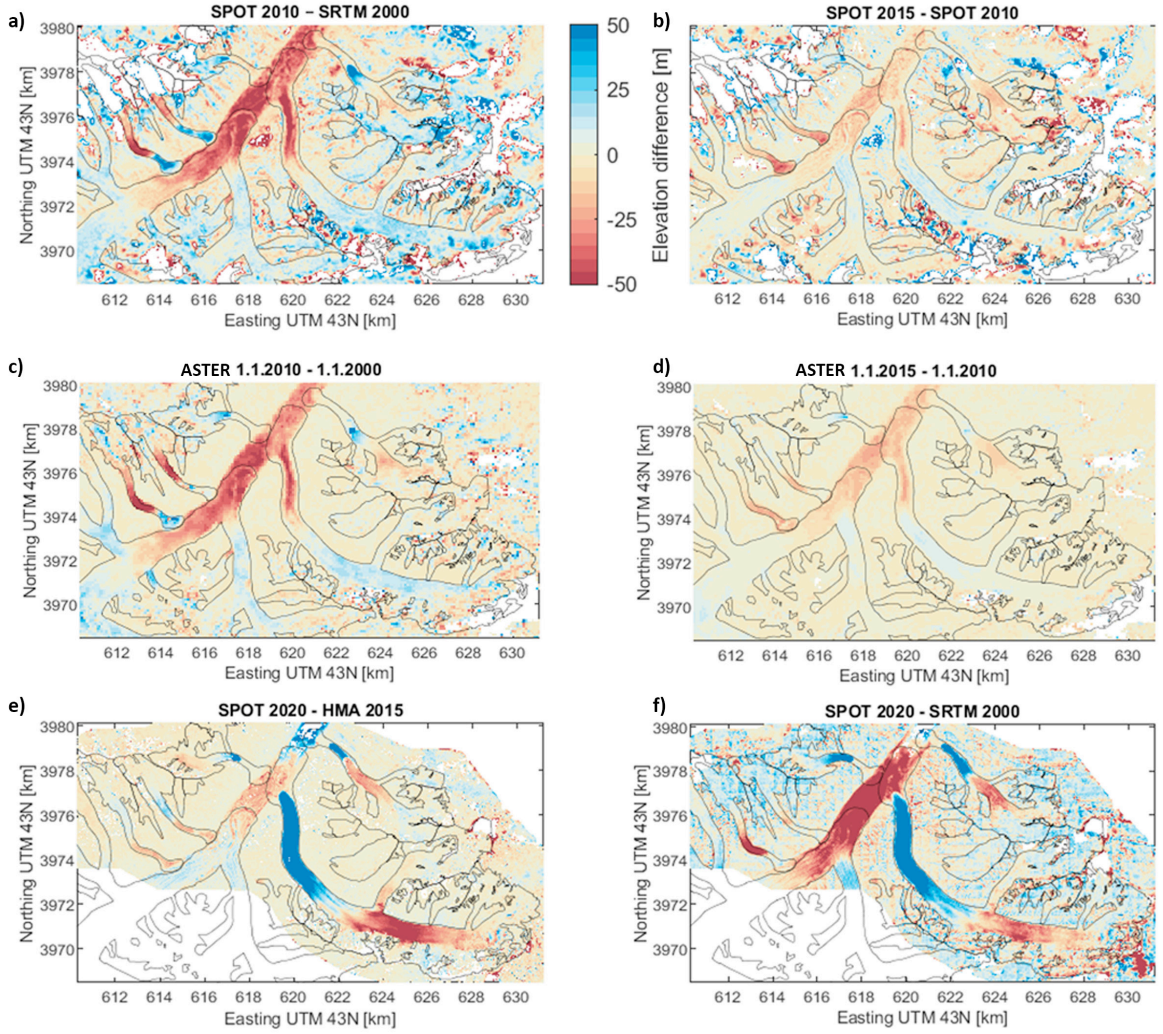

Figure 3.5: 2D elevation difference maps over the study region. a) SPOT 2010 – SRTM 2000, b) SPOT 2015 – SPOT 2010, c) ASTER time series (2010-2000), d) ASTER time series (2015-2010), e) SPOT 2020 – HMA 2015, f) SPOT 2020 – SRTM 2000.

4. Velocity

4.1 Introduction

In Phase 2 of Glaciers_CCI, intensity cross-correlation was selected as the primary algorithm for ice surface velocity estimation using both optical and SAR satellite images. Sources of errors and uncertainties, both internal (i.e. algorithm dependent) and external, as well as methods for accuracy determination and the accuracy to be reported were discussed in detail in [RD1].

As a general statement, the accuracy of individual glacier displacement measurements from repeat satellite optical data using offset tracking is on the order of one pixel. For a one year Landsat panchromatic image pair this corresponds to an accuracy of 15 m/yr. However, in case of good visual contrast, such as given for crevassed and snow-free glaciers, and satellite optical data of the latest generation (Landsat 8 and Sentinel-2) surface displacement accuracies of 10-20% of a pixel can be reached. This corresponds e.g. for Landsat 8 data (15 m) to 1.5-3 m or 24-48 m/yr for a time interval of 16 days.

For SAR sensors, we estimate the reliability of the cross-correlation algorithm to return co-registration parameters as accurate as 1/10th of an image pixel. This corresponds for the ALOS PALSAR and TerraSAR-X data separated by a temporal interval of 46 respectively 11 days to an accuracy of about 10 m/yr and for the Sentinel-1 IWS data separated by a temporal interval of 6/12 days to an accuracy of about 30 m/yr.

More specifically, the accuracy of offset-tracking between two optical or SAR images is reported in the velocity products as:

- a comment from an experienced operator based on visual inspection of the resulting displacements (consistency of ice flow field, sensor effects, etc.);
- characterisation of the matching quality of individual ice velocity estimates: map of correlation coefficient and/or signal-to-noise ratio, depending on the algorithm implementation, given for each displacement estimation;
- similarity between original measurements and low-pass filtered ones for each displacement, given as deviations in x- and y-offsets, or vector magnitude and direction;
- statistical measures for stable ground matches (mean, standard deviation, RMSE);
- if temporally consistent matches based on images with higher resolution or ground-based measurements of displacements exist, summary of deviations (mean, standard deviation, RMSE, min, max).

Though requiring little supervision and showing satisfactory performances, especially for fast ice flow, offset-tracking velocity products suffer from reduced performance in slow-moving areas with velocity below about 5 cm/day and their degradation by ionospheric streaks, which are frequent in the polar regions. SAR interferometry (InSAR), on the other hand, can reach a precision of one to two orders of magnitude better than offset-tracking, is especially suitable for slow moving ice and is capable to provide a higher spatial resolution. However, InSAR requires the preservation of the SAR signal between repeat observations, which can be disturbed by weather effects (accumulation, wind, surface melt) and by high velocity gradients in shearing zones and termini of outlet glaciers. Moreover, contrary to offset-tracking, InSAR is

only sensitive to the component of the motion along the line-of-sight (LOS) direction and therefore requires acquisitions from overlapping descending and ascending orbits to derive the 2-D surface velocity field. A systematic SAR acquisition planning is therefore needed from crossing orbits within a short period. Before the launch of Sentinel-1A/B, such data were not systematically acquired and not always with a satisfactory temporal baseline.

The unprecedented data set provided by Sentinel-1 satellites opens the possibility to improve the current ice velocity measurements over Arctic glaciers and ice caps by applying InSAR in a more systematic fashion. As there might be some regions not covered by both, ascending and descending passes, we derive in these regions the ice velocity by combining the LOS velocity from the available pass with a flow direction map calculated from the multi-annual OT velocity map. With the a-priori knowledge of the flow direction, the ambiguity of the velocity vector orientation is solved and a single viewing geometry is sufficient to determine the velocity magnitude.

The recent failure of Sentinel-1B has a strong impact on InSAR applications, because the temporal baseline of the mission is now reduced from 6 to 12 days. Longer revisit times mean a reduced coherence over glaciers due to snowfall, melt or wind and ice motion. In many cases, InSAR can therefore no more be used to monitor slow moving ice at high spatial resolution. The reduced revisit time also affects the uncertainty of OT products that are averaged over a longer time span (e.g. monthly or annual maps) as fewer data are included.

In this document, the ECV uncertainty budget for offset-tracking is not duplicated. Instead, we will concentrate the current discussion on SAR interferometric techniques and on sensor and method synergies.

4.2 Interferometric techniques

4.2.1 Major error sources

Errors in interferometric products can arise from several sources:

- Decorrelation and random phase noise
- Baseline errors
- Phase unwrapping errors
- Phase calibration errors

4.2.2 Uncertainties per processing step

Decorrelation and random phase noise

Interferometric phase quality can be affected by a number of time-, geometry- and processing-related factors. Interferometric phase accuracy is estimated with the coherence estimator, i.e. low coherence corresponds to large phase uncertainty. Coherence loss can be caused by temporal changes (snowfalls, melt or wind), geometrical decorrelation, mis-registration or thermal noise. For Sentinel-1, geometrical decorrelation is limited thanks to the small orbital tube. Typical values of SNR show that thermal noise is negligible. In practice, phase variance is approximated by the Cramer-Rao bound, using the coherence and the effective number of looks (Hanssen, 2001).

Baseline errors

Inaccuracies in the annotated state vectors of the input SLC images introduce errors in the baseline calculation, which in turn leaves a residual phase contribution in the interferogram

after the reference phase removal (flat-earth and topographic phase). Such orbital errors can be corrected by using a low-order polynomial approximation of the interferometric phase. However, in the case of Sentinel-1, precise orbit vectors are available 3 weeks after the nominal acquisition with a 5 cm 3-D 1-sigma accuracy and make this correction unnecessary.

Phase unwrapping
Phase unwrapping errors occur in regions of low coherence, complex topography or phase aliasing and can propagate at a local scale or throughout the entire image. They can be detected to some extent, but hardly quantified. For this reason, a segmentation mask of the unwrapped phase is used for masking regions with clear phase unwrapping errors.

Phase calibration
In the case of ice velocity mapping, stable points, e.g. over bedrocks, should be preferred for phase calibration but they are often not connected to the ice surface in the interferogram image. For this reason, phase calibration is usually performed against slow-moving GCP measurements. These GCP measurements can come either from GPS data or another velocity map, e.g. from offset-tracking. In both cases, uncertainties of the GCP measurements propagates to the interferometric phase measurements during the phase calibration step. The uncertainty on the GCP velocity can be accounted for by projecting it on the sensor line-of-sight.

Uncertainty combination - Least-squares inversion for ice velocity retrieval
Interferometry provides only line-of-sight velocity measurements. In order to retrieve 2-D surface velocity, line-of-sight velocity maps from crossing orbits, i.e. with different viewing geometries, must be combined together. In the final step of the processing, line-of-sight velocity maps generated with SAR interferometry are combined together through a least-squares inversion.

To account for interferometric processing uncertainties, phase variances from random phase noise and phase calibration are summed up and used as weights of the input data in the least-squares inversion. Final uncertainties on the x- and y-components of the velocity vector are provided by the diagonal elements of the variance-covariance matrix of the least-squares parameters, which includes phase variance given as weights and errors introduced by the inversion.

4.2.3 Other uncertainties

In TOPSAR mode, co-registration errors may induce phase jumps in the along-track direction that are complicated to quantify. Although they can clearly be identified by eye in a TOPSAR interferogram, these phase jumps can be caused by a number of phenomena, like e.g. ionospheric scintillations, along-track ice motion or orbit inaccuracies.

In the interferometric processing, atmospheric and ionospheric phase contributions are not corrected. The uncertainties introduced by these effects would require ancillary data (e.g. maps of the Total Electron Content as provided by NASA/JPL) or modelling to be estimated.

Detection of phase unwrapping errors may fail in places and some phase unwrapping errors remain in the final LOS velocity products. However, these errors are not quantified and they are not included in the error budget.

4.2.4 Use Case Eastern Arctic: Interferometric velocities for Svalbard (Sentinel-1)

The interferometric Svalbard velocity map is generated from 27 Sentinel-1 individual 6-day pairs, along three different orbits, acquired between 1.10.2018 and 1.3.2019. The individual LOS velocity measurements are visually inspected for decorrelation, phase unwrapping errors and calibration biases, before the least-square inversion. Phase unwrapping errors are particularly likely due to the complex topography, the presence of narrow bridges connecting different regions of the archipelago island and the overall poor coherence. In case of large scale errors susceptible to introduce significant bias in the velocity measurements, the LOS velocity is not used for the least-square inversion. However, given the limited amount of data and the need for measurements along overlapping ascending and descending orbits, some LOS velocity maps tarnished with known errors had to be kept for insuring a satisfactory coverage. As a consequence, the interferometric velocity map of Svalbard exhibits obvious biases in places. To remove outliers, a median filtering with a window of 5x5 pixels is further applied to the velocity map. On average, the error estimates on the x- and y-velocity components are smaller than 1 cm/day, with values respectively as large as 5-10 cm/day and 10-20 cm/day.

4.3 Merging INSAR with offset-tracking flow direction

4.3.1 Major error sources

In the case of a combination of interferometric velocity measurements with offset-tracking flow directions, the main source of errors is the interferometric processing. If the flow direction is not known accurately, like e.g. in regions of diverging flow, this method suffers from artifacts and large errors may be introduced locally.

4.3.2 Uncertainties per processing step

Interferometric processing
Uncertainties caused by the interferometric processing have been described in Section 4.2. They include random phase noise, baseline errors, phase unwrapping errors and phase calibration errors.

Uncertainty combination - Linear regression
In the case of a combined use of interferometric velocity measurements with offset-tracking flow directions, the 2-D surface velocity vector is computed in two steps: first, a linear regression is applied to the line-of-sight measurements in order to retrieve the magnitude of the velocity vector; second, the estimated magnitude is projected on the x- and y- directions of the offset-tracking flow vector.

As for the least-squares inversion applied in the interferometric case (see Section 4.2.2), the interferometric processing uncertainties are accounted for by summing up random phase noise and phase calibration variances. The total variance is used to weight input line-of-sight measurements for the linear regression. The final uncertainty is obtained from the linear regression variance of the slope parameter and it is projected the flow direction to obtain uncertainties on the x- and y- velocity components.

4.3.3 Other uncertainties

Flow directions are derived from the multi-annual offset-tracking velocity map and are another possible source of uncertainties. However, the uncertainty on the flow direction is not in-

cluded in the error budget because they are used as deterministic mapping coefficients in the linear regression. Although this is not yet included in the error budget, the uncertainties on the flow direction could be partially taken into account for the projected velocity components by combining it with the slope parameter uncertainty. However, this would not include the error introduced in the estimation of the velocity magnitude at the linear regression step.

4.3.4 Use Case Eastern Arctic: merging offset tracking and interferometry?

Similar to the fully interferometric method (see Section 4.2), an ice velocity map of Svalbard is generated using the LOS velocity maps from 27 Sentinel-1 individual 6-day pairs combined with the flow direction derived from an OT multiannual map. The error mitigation and the choice of the pairs is described in Section 4.2.4. In the case of the Svalbard velocity map, the relevant error is the error on the velocity magnitude. Inaccuracies on the flow direction introduce in places magnitude errors up to several tens of centimetres per day. Thanks to their local nature, these errors can be mitigated by applying a median filter to remove outliers. The mean error amounts to almost 2 cm/day and is larger than the fully interferometric case.

4.3.5 Use Case HMA: comparing optical with radar velocities

In the Chongtar glacier study (Paul et al. in review), we compared glacier flow fields from Landsat ETM+/OLI. TerraSAR-X (TSX), Sentinel-1, Sentinel-2, and Planet cubesats. As can be seen in Fig. 4.1, the 2D surface flow velocities derived from the 15 m resolution Landsat 8 panchromatic band for the period July 2018 to July 2019 are the same as the 10 m resolution Sentinel-2 data for the period August 2018 to August 2019. The same applies to TSX velocities from April to May 2014 compared to the annual mean values from Landsat 8 over the period July 2013 to July 2014. The Planet cubesat images cover only the lower part of the glacier. Here, the Planet velocity (Fig. 4.1) reveals the same increase/decrease pattern as the Sentinel-2 velocity map. This is different when comparing velocities derived from Sentinel-1 (Fig. 4.2). The large image template sizes of 128 x 64 (450 m x 900 m) even for the largest glacier (tongue width 800 m) result in a strong underestimation of Sentinel-1 velocities with errors much greater than those reported in previous studies for larger Arctic glaciers (Paul et al., 2017; Strozzi et al., 2017). The poor performance of Sentinel-1 is mainly due to the glacier size situated between steep mountain flanks. Because of the relatively large size required for the matching window, too many non-moving pixels outside the glacier are included, which considerably affects the velocity estimation. The minimum width of a glacier to be reliably monitored with Sentinel-1 in the Himalayas is likely about 2 km.

We assessed the uncertainties of glacier flow velocities from stable terrain velocity observations, where flow velocities are supposed to be zero. The uncertainties in glacier flow velocity are mainly related to the co-registration accuracy, orthorectification, the time interval between image pairs (Fig. 4.1), surface conditions (shadow, snow, etc.), and the spatial resolution of the images. The larger the time window between two pairs, the smaller the uncertainty of the measured velocity. Despite the higher resolution, the uncertainty is higher for Planet than for Sentinel-2. For Sentinel-2, the orthorectification error is minimized because the imagery comes from the same relative orbit (Kääb et al., 2016). On the contrary, we have different orbital paths between Planet image pairs and thus further geometric corrections may be needed to minimise this error, as also suggested by Kääb et al. (2017) and Millan et al. (2019). In addition, the very small uncertainties in the TSX stable terrain are likely due to the accurate co-registration of the image pairs (Fig. 4.3).

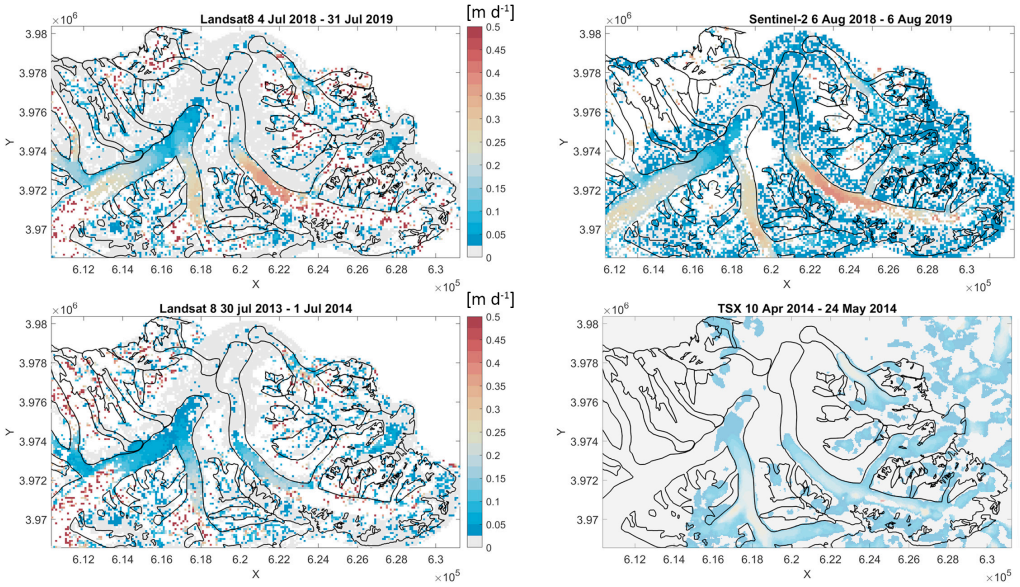

Figure 4.1: Comparison of 2D surface flow velocities [m d^{-1}] from Landsat, TSX and Sentinel-2.

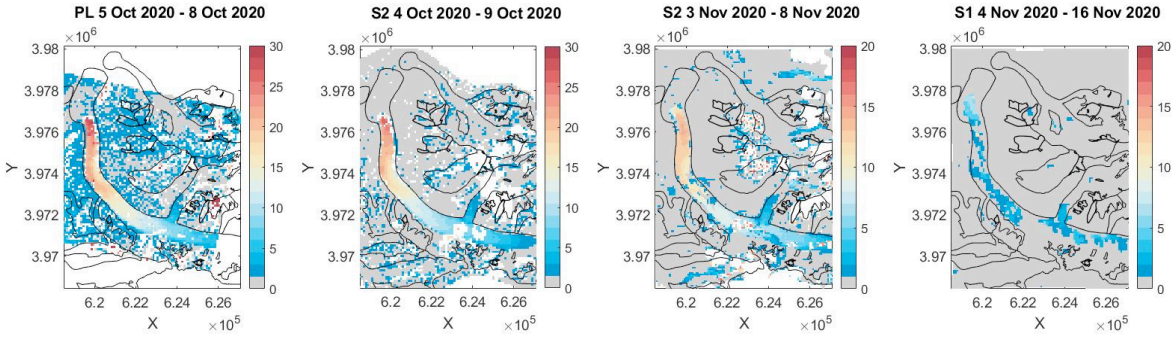

Figure 4.2: Comparison of 2D surface flow velocities [m d^{-1}] from Planet (PL), Sentinel-2 (S2) and Sentinel-1 (S1).

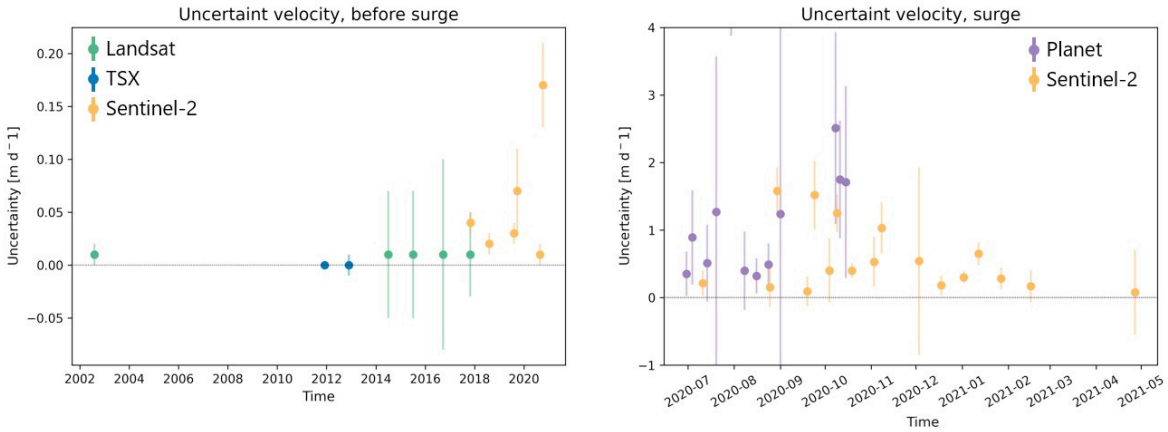

Figure 4.3: A comparison of the uncertainty in the flow velocity estimation derived from optical and SAR data over stable terrain (area outside the glacier) for scenes with one-year intervals (before the surge) and monthly intervals of up to 3 days (during the surge).

4.4 Drainage divides from SAR interferometry

4.4.1 Major error sources

Fringe patterns from wrapped differential SAR interferograms are proposed to manually correct the outlines of drainage basins derived from semiautomatic algorithms, e.g. using a DEM, outlines of glacier complexes, and hydro-logical modeling tools (Rastner et al., 2017). The fringe patterns are particularly helpful in detecting false divides on flat glaciers or icecaps, i.e. exactly where optical DEMs have problems due to saturation, while over mountainous areas they are much less useful. Fringe patterns alone cannot be employed to draw drainage divides everywhere, because they are limited by loss of coherence, sensitivity to the satellite line-of-sight direction only, and atmospheric disturbances. Fringe patters are thus not straightforward to be interpreted and only when superimposed to previously derived drainage basins, i.e. when combined with other methods, they indicate where ice divides are wrong. Ideally, the objective would be to find an automatic way to improve the divides by means of the fringe patterns, but so far only manual correction of drainage divides from fringe patters is considered. In this document we differentiate between the error sources arising from the SAR interferometric processing and those arising from the manual correction of drainage divides. Errors of deriving outlines of drainage basins from semiautomatic algorithms are not considered here. An extended discussion on this topic can be found in Kienholz et al. (2013).

4.4.2 Uncertainties per processing step

Interferometric processing
Uncertainties caused by the interferometric processing have been described in Section 4.2. Here, we recall the major consequences for deriving drainage divides from SAR interferometry.

Decorrelation and random phase noise are a severe limitation to SAR interferograms. Whereas coherence is often low for the lower sections of outlet glaciers, it is usually sufficiently high over the slower moving interior parts of the ice caps to obtain clearly visible fringes. With the increasingly availability of 6/12-day Sentinel-1 and 14-day ALOS-2 PALSAR-2 image pairs, there is nowadays a growing number of coherent winter SAR interferograms to update drainage divides over all major Arctic ice caps. As mentioned above, the Sentinel-1B failure will strongly limit the availability of this product, as the loss of coherence over a longer period is much more likely.

The accuracy and acquisition date of the DEM used to remove the topographic related phase from the SAR interferograms is crucial for the computation of the fringe patterns. The relationship between a change in the topographic height z and the corresponding change in the interferometric phase ϕ is given by (Bamler and Hartl, 1998):

$$\frac{\partial \phi}{\partial z} = \frac{4\pi}{\lambda} \frac{B_{\perp}}{R \sin \theta}$$

Here R is the distance from the SAR to the scatterer, θ the incidence angle, and B⊥ is the component of the baseline perpendicular to the line-of-sight direction. For Sentinel-1, with a wavelength of 5.55 cm, a nominal incidence angle of 35° and a nominal slant range of 693 km, an error of 10 m in the estimation of the topographic height (e.g. as a consequence of an inaccurate or aged DEM) results for a perpendicular baseline component of 50 m in a phase error of 0.09π. For ALOS-2 PALSAR-2, with a wavelength of 22.9 cm, a nominal incidence

angle of 35° and a nominal slant range of 628 km, an error of 10 m in the estimation of the topographic height results for a perpendicular baseline component of 100 m in a phase error of 0.05π. Hence, phase signals on fringe patterns of Sentinel-1 or ALOS-2 PALSAR-2 interferograms can be indeed interpreted as ice surface displacement in the satellite line-of-sight direction with possible atmospheric disturbances.

In the interferometric processing, tropospheric phase contributions are not corrected and the introduced uncertainties are thus casual in the fringe image. However, changes in atmospheric water content are expected to be small during high latitude winter, so that the resulting uncertainty in phase is expected to be at most ±0.25 cycle for C-Band (Gray et al., 1997) and maybe half of this value for L-Band.

To mitigate the ionospheric phase screens for differential SAR interferograms, the range split-spectrum technique has proven very effective (Gomba et al. 2016) and can be employed to improve the quality of the fringe images. Finally, SAR interferometry provides only line-of-sight velocity measurements. If the ice flow direction approaches an angle perpendicular to the line of sight, SAR interferograms are not sensitive to the motion. The availability of the interferograms from ascending and descending orbits might mitigate this problem.

By summing up all the errors arising from the interferometric processing (in particular potential DEM errors and atmospheric artifacts), we have an expected uncertainty in phase of about one fourth of a phase cycle at C-Band and of about one eight of a phase cycle at L-Band. This corresponds to an expected uncertainty of about 20 cm/year for 12 days Sentinel-1 interferograms and of about 40 cm/year for 14 days ALOS-2 PALSAR-2 interferograms.

Manually correcting drainage divides
If the outlines of the drainage basins derived from semiautomatic algorithms (e.g. Rastner et al., 2017) do not follow the fringe patterns, they are manually edited to the extent possible. As any other similar mapping exercise, the required manual digitization is prone to subjective interpretation. Decorrelation and atmospheric artefacts might even make the interpretation more difficult, because making the fringe pattern less visible and not only related to ice surface velocity. As for glacier outlines, it would be possible to use the variability of repeated manual digitization as a measure of uncertainty in positioning the outlines (Paul et al., 2017).

4.4.3 Other uncertainties

Phase unwrapping errors and phase calibration errors do not apply here, because the fringe patterns arise from wrapped differential SAR interferograms. As previously stated, errors of deriving outlines of drainage basins from semiautomatic algorithms considering a DEM (e.g. Rastner et al., 2017) are not examined here.

4.4.4 Use Case Eastern Arctic: Drainage divides Austfonna

We have used the fringe pattern derived by interferometric processing of ALOS-1 PALSAR-1 to constrain ice divides calculated from different DEMs. The comparison revealed a good agreement of the main ice divides derived from the different DEMs and with the fringe pattern, in particular along the highest elevations where the ice cap is very flat. Despite the difficulties in interpreting the fringes, we conclude that they provide an interesting additional information layer when creating ice divides over very flat terrain.

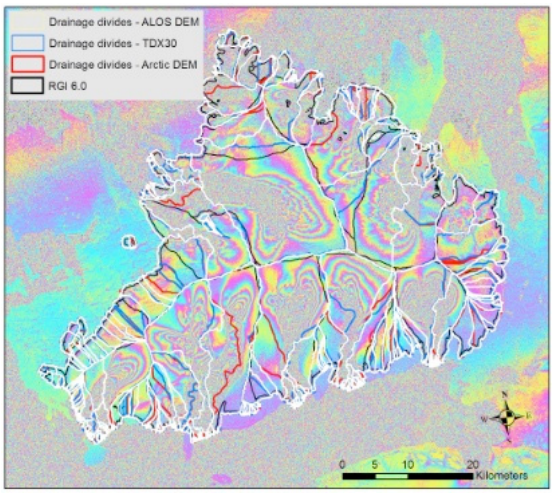

Fig. 4.4: Fringe pattern derived from ALOS Palsar for Vestfonna with drainage divides derived from three different DEMs (ALOS, TDX and Arctic DEM) and glacier outlines and divides from the RGI6 (black).

5. References

Bamler, R. and P. Hartl, 1998, Synthetic aperture radar interferometry. *Inverse Problems*, 14, R1-R54.

Gardelle, J., Berthier, E., Arnaud, Y., and Kääb, A. (2013): Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999–2011. *The Cryosphere*, 7, 1263–1286.

Gilbert, L., (2021): Product Quality Assessment Report, Cryosphere service: Ice Sheets and Ice Shelves ECV, https://datastore.copernicus-climate.eu/documents/satellite-GRACE/version3/C3S_312b_Lot4_D2.IS.6-v3.0_202103_PQAR_SEC_i1.0.pdf

Goerlich, F., T. Bolch, K. Mukherjee and T. Pieczonka (2017): Glacier Mass Loss during the 1960s and 1970s in the Ak-Shirak Range (Kyrgyzstan) from Multiple Stereoscopic Corona and Hexagon Imagery. *Remote Sensing*, 9 (3), 275; doi.org/10.3390/rs9030275.

Gomba G., A. Parizzi, F. De Zan, M. Eineder and R. Bamler (2016): Toward operational compensation of ionospheric effects in SAR interferograms: The Split-Spectrum Method. *IEEE Transactions on Geoscience and Remote Sensing*, 54(3), 1446-1461.

Gray, A. L., K.E. Mattar, D. Geudtner and P.W. Vachon (1997): *Experiments at CCRS using ERS tandem mode data*. In: 3rd ERS Scientific Symposium, 17 – 21 March 1997, Florence, Italy Proceedings, vol. 2, European Space Agency Special Publication, *ESA SP-414*, 1001-1006.

Hanssen, R.F. (2001): *Radar Interferometry: Data interpretation and error analysis*. Kluwer Academic Publishers, pp. 95-130.

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., and Kääb, A. (2021): Accelerated global glacier mass loss in the early twenty-first century. *Nature*, 592, 726–731.

Kääb, A., Winsvold, S.H., Altena B., Nuth, C., Nagler, T. and Wuite, J. (2016): Glacier remote sensing using Sentinel-2. Part I: Radiometric and geometric performance, and application to ice velocity. *Remote Sensing*, 8(7), 598; doi.org/10.3390/rs8070598.

Kääb, A., Strozzi, T., Bolch, T., Caduff, R., Trefall, H., Stoffel, M., and Kokarev, A. (2021): Inventory and changes of rock glacier creep speeds in Ile Alatau and Kungöy Ala-Too, northern Tien Shan, since the 1950s. *The Cryosphere*, 15, 927-949.

Kääb, A., Altena, B., and Mascaro, J. (2017): Coseismic displacements of the 14 November 2016 Mw 7.8 Kaikoura, New Zealand, earthquake using the Planet optical cubesat constellation. *Natural Hazards and Earth System Science*, 17, 627–639.

Kienholz C., R. Hock and A.A. Arendt (2013): A new semi-automatic approach for dividing glacier complexes into individual glaciers. *Journal of Glaciology*, 59 (217), 913-925.

Leclerq P., Kääb A., Altena B. (2021): Brief communication: Detection of glacier surge activity using cloud computing of Sentinel-1 radar data. *The Cryosphere*, 15, 4901-4907.

Millan, R., Mouginot, J., Rabatel, A., Jeong, S., Cusicanqui, D., Derkacheva, A., and Chekki, M. (2019): Mapping surface flow velocity of glaciers at regional scale using a multiple sensors approach. *Remote Sensing*, 11, 2498, https://doi.org/10.3390/rs11212498.

Mölg, N., Bolch, T., Rastner, P., Strozzi, T. and Paul, F. (2018): A consistent glacier inventory for the Karakoram and Pamir region derived from Landsat data: Distribution of debris cover and mapping challenges. *Earth Systems Science Data*, 10, 1807-1827.

Paul, F., T. Bolch, K. Briggs, A. Kääb, M. McMillan, R. McNabb, T. Nagler, C. Nuth, P. Rastner, T. Strozzi and J. Wuite (2017): Error sources and guidelines for quality assessment of glacier area, elevation change, and velocity products derived from satellite data in the Glaciers_cci project. *Remote Sensing of Environment*, 203, 256-275.

Paul, F., Piermattei, L., Treichler, D., Gilbert, L., Girod, L., Kääb, A., Libert, L., Nagler, T., Strozzi, T., and Wuite, J. (in review): Three different glacier surges at a spot: What satellites observe and what not. *The Cryosphere Discussions* [preprint], https://doi.org/10.5194/tc-2021-370.

Rastner P., T. Strozzi and F. Paul (2017): Fusion of multi-source satellite data and DEMs to create a new glacier inventory for Novaya Zemlya. *Remote Sensing*, 9(11), 1122; doi.org/10.3390/rs9111122.

Shean, D.E., Alexandrov, O., Moratto, Z., Smith, B.E., Joughin, I.R., Porter, C.C. and Morin, P. J. (2016): An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very high-resolution commercial stereo satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116, 101–117.

Shepherd, A., et al (2019): Trends in Antarctic Ice Sheet Elevation and Mass. *Geophysical Research Letters*, 46 (14) 8174-8183, doi.org/10.1029/2019GL082182.

Toutin, T. (2004): Geometric processing of remote sensing images: models, algorithms and methods. *International Journal of Remote Sensing*, 25(10), 1893-1924.

Toutin, T. and Gray, L. (2000): State-of-the-art of elevation extraction from satellite SAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55(1), 13-33.

Treichler, D. and Kääb, A. (2016): ICESat laser altimetry over small mountain glaciers. *The Cryosphere*, 10, 2129-2146.

Wingham, D.J., Ridout, A.J., Scharroo, R., Arthern, R.J., and Shum, C.K. (1998): Antarctic elevation change from 1992 to 1996. *Science*, 282, 456-458.

6. Acronyms

ALOS	Advanced Land Observing Satellite
ATBD	Algorithm Theoretical Baseline Document
DEM	Digital Elevation Model
DN	Digital Number
GCP	Ground Control Point
GEE	Google Earth Engine
ICESat	Ice, Cloud, and land Elevation Satellite
InSAR	Interferometric SAR
IWS	Interferometric Wide Swath
LOS	Line Of Sight
PALSAR	Phased Array type L-band Synthetic Aperture Radar
RADAR	Radio Detection and Ranging
RGI	Randolph Glacier Inventory
RMSE	Root Mean Square Error
SAR	Synthetic Aperture Radar
SLC	Single Look Complex
UCR	Uncertainty Characterisation Report