ESA Climate Change Initiative (CCI+)
Essential Climate Variable (ECV)

Greenland_Ice_Sheet_cci+ (GIS_cci+)

End-to-end Uncertainty Budget (E3UB)

Prime & Science Lead: René Forsberg

DTU Space, Copenhagen, Denmark

rf@space.dtu.dk

Technical Officer: Marcus Engdahl

ESA ESRIN, Frascati, Italy

Marcus.Engdahl@esa.int

Consortium: Asiaq Greenland Survey (ASIAQ)

DTU-Space, Department of Geodynamics (DTU-S)

DTU-Space, Department of Microwaves and Remote Sensing (DTU-N)

Danish Meteorological Institute (DMI)

Environmental Earth Observation IT GmbH (ENVEO)

Science [&] Technology AS (S&T)

Technische Universität Dresden (TUDr)

The Geological Survey of Denmark and Greenland (GEUS)

The Niels Bohr Institute (NBI)

University of Leeds, School of Earth and Environment (UL)

Signatures page

Prepared by	Jan Wuite Lead Author, ENVEO		Date: 27 Apr 2022
Issued by	Daniele Fantin, Project Manager, S&T		
Checked by	Rene' Forsberg Science Leader, DTU-S		Date: 27 Apr 2022
Approved by	Marcus Engdahl ESA Technical Officer		Date:

Table of Contents

Change Log

Issue	Author	Affected Section	Change	Status
0.6	ENVEO	All	Draft version	
1.0	All	All	First version	
1.1	GEUS	6	MFID update	
2.0	S&T	8	CFL update	

Acronyms and Abbreviations

Acronyms	Explanation
ATBD	Algorithm Theoretical Basis Document
C3S	Copernicus Climate Change Service
CCI	Climate Change Initiative
CCI+	Climate Change Initiative Plus
CFL	Calving Front Location
CS2	CryoSat-2
CSR	Center for Space Research, University of Austin
DEM	Digital Elevation Model
DInSAR	Differential Interferometric Synthetic Aperture Radar
DMI	Danish Meteorological Institute
DTU-N	DTU Microwaves and Remote Sensing Group
DTU-S	DTU Geodynamics Group
E3UB	End-to-End ECV Uncertainty Budget
ECV	Essential Climate Variable
EGSIEM	European Gravity Emergency Service project
ENU	East North Up
ENVEO	ENVironmental Earth Observation GmbH
EO	Earth Observation
ESA	European Space Agency
GCOS	Global Climate Observation System
GCP	Ground Control Point
GEUS	Geological Survey of Denmark and Greenland
GFZ	Deutsche GeoForschungsZentrum
GIA	Glacial Isostatic Adjustment
GIS	Greenland Ice Sheet
GLL	Grounding Line Location
GMB	Gravimetry Mass Balance
GPS	Global Positioning Sstem
GRACE(-FO)	The Gravity Recovery and Climate Experiment (Follow On)
IMBIE	Ice Sheet Mass Balance Inter-Comparison Exercise
InSAR	Interferometric Synthetic Aperture Radar
IPP	Interferometric Post-Processing
IV	Ice Velocity
IW	Interferometric Wide Swath
MAI	Multiple Aperture Interferometry
MCF	Minimum Cost Flow
MEaSUREs	Making Earth System Data Records for Use in Research
MFID	Mass Flow Rate and Ice Discharge
NBI	Niels Bohr Institute, University of Copenhagen
NEGIS	North East Greenland Ice Stream

OG	Outlet Glaciers
OIB	Operation Ice Bridge
PROMICE	Danish Program for Monitoring of the Greenland Ice Sheet
RA	Radar Altimetry
RMS	Root Mean Square
S&T	Science and Technology AS
S2	Sentinel-2
SA	Solution Area
SAR	Synthetic Aperture Radar
SEC	Surface Elevation Change
SLR	Satellite Laser Ranging
SMB	Surface Mass Balance
SOW	Statement of Work
TEC	Total Electron Content
TOA	Top of Atmosphere
TOPS	Terrain Observing by Progressive Scans
TPROP	Technical Proposal
TUDr	Technische Universität Dresden
UL	University of Leeds
URD	User Requirement Document

1 Introduction

1.1 Purpose and Scope

This document contains the End-to-end Uncertainty Budget for the Greenland_Ice_Sheet_cci (GIS_cci) project for CCI+ Phase 1, in accordance to contract and SoW (AD1 and AD2). The central aim is to ascertain error characteristics that permit identification of climate change over natural variability. The E3UB describes the end to end errors of ECV improvements, proposed for CCI+, and builds on the Phase 2 Comprehensive Error Characterisation Report (CECR) document (RD1) of the 'Greenland_Ice_Sheet_cci project.

The overall error and uncertainty budget for new and existing products will be provided or re-assessed and updated where needed, considering errors induced by new sensors, models, corrections, technical developments and continued validation/inter-comparison efforts, including Round Robin outcomes. It describes the best current understanding of the sources of errors, and uncertainties for the retrieval algorithms of the parameters: 'Surface Elevation Change (SEC)', 'Ice Velocity (IV)', 'Gravimetric Mass Balance (GMB)', 'Mass Flux and Ice Discharge (MFID)' and 'Supraglacial Lake (SL)'.

1.2 Document Structure

This document is structured into an introductory chapter followed by 5 chapters focussed on End-to-end Uncertainty Budget for each of the CCI+ parameters:

- Surface Elevation Change (SEC)
- Ice Velocity (IV-SAR)
- Ice Velocity (IV-Optical)
- Gravimetric Mass Balance (GMB)
- Mass Flux and Ice Discharge (MFID)
- Supraglacial Lake (SL)

1.3 Applicable and Reference Documents

Table 1.1: List of Applicable Documents

No	Doc. Id	Doc. Title	Date	Issue/ Revision/ Version
AD1	ESA/Contract No. 4000126023/19/I-NB + App. 1	CCI+ PHASE 1 - NEW R&D ON CCI ECVS, for Greenland_Ice Sheet_cci	2019.04.01	
AD2	ESA-CCI-EOPS-PRGM-SOW-18-0118 - Appendix 2 to contract.	Climate Change Initiative Extension (CCI+) Phase 1, New R&D on CCI ECVs - SOW	2018.05.31	Issue 1 Revision 6

Table 1.2: List of Reference Documents

No	Doc. Id	Doc. Title	Date	Issue/ Revision/ Version
RD1	ST-DTU-ESA-GISCCI-CECR-001	CECR for the Ice_Sheets_cci project of ESA's Climate Change Initiative	2016.09.09	2.1
RD2	ST-DTU-ESA-GISCCI-ATBD-001-v3.2	Greenland_Ice_Sheet_cci Algorithm Theoretical Baseline Document (ATBD)	2017.11.24	3.2
RD3	ST-DTU-ESA-GISCCI+-ATBD-001	Greenland_Ice_Sheet_cci+ Algorithm Theoretical Baseline Document (ATBD)	2021.10.28	1.4

RD4	ST-DTU-ESA-GISCCI+-PUG-001	Product User Guide (PUG)	2021.02.25	2.0
RD5	ST-DTU-ESA-GISCCI+-PVIR+001	Product Validation and Intercomparison Report (PVIR)	2022.04.27	2.1

2 End-to-end Uncertainty Budget for SEC

2.1 Introduction

This section describes the uncertainty budget and error sources for the SEC data products within the CCI+ Greenland Ice Sheet project. This document builds on the CECR of the CCI Greenland Ice Sheet project (RD1) and will not in detail repeat what is written there. We refer to (RD1) when relevant.

2.2 Sources of error/uncertainty

For RA the sources of errors and uncertainties in the individual height measurements and the derived SEC are summarized in the following:

An inherent source of uncertainty of a RA-derived surface height over ice sheets is the firn penetration into the snow/firn is unknown. The penetration depth is furthermore dependent on which retracker is applied, and the retracker which most accurately tracks the snow-air surface is not necessarily the one that proves most robust, which is important for SEC estimates.

Potential instrument biases will off course also impact the error budget, but such biases are often minimized by calibrations over flat terrain elsewhere on Earth where no surface penetration is expected. Therefore, these will not dominate the error budget.

From systematic intercomparisons within each RA data set and with alternative datasets such as airborne LiDAR (Sørensen et al., 2018; Simonsen et al. 2017; Nilsson et al., 2016) it is found that the agreement between the RA and the validation data is topography-related, and that RA generally maps higher elevations than the validation data - most likely due to the large footprint of conventional RA that favours mapping topographic highs (Schröder et al., 2019). This kind of topographic error cancels out when comparing repeating measurements, and hence is not a problem in SEC measurements.

The precision of the RA measurements is the interesting parameter for SEC and can be assessed through comparing each mission's crossover statistics: how well do ascending and descending measurements compare if they are collected close in time. In Schröder et al., (2019) they compare all Envisat crossovers within 31 days over Antarctica and their results are summarized in Figure 2.1 . From this figure it is clear that the standard deviation of elevation differences in crossovers increases with surface slope. Further, Schröder et al. (2019) has performed an analysis of a range of RA missions, and their results are provided in Table 2.1.
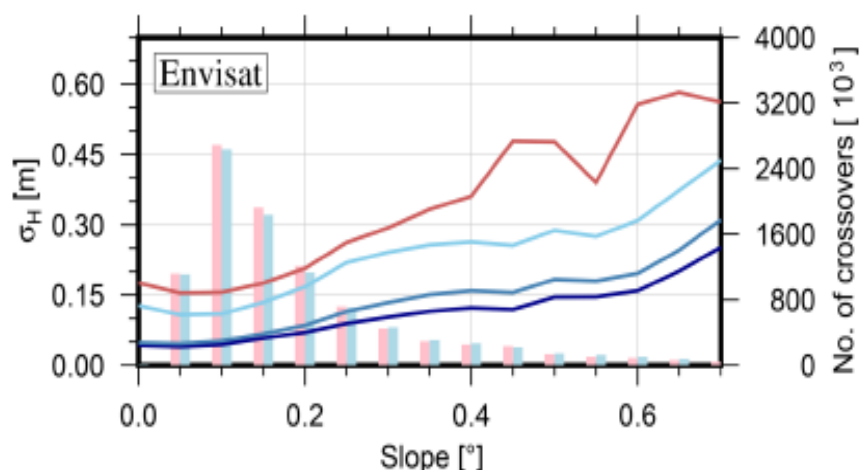

Figure 2.1: Crossover statistics of Envisat over Antarctica from Schröder et al. (2019). Standard deviation as a function of slope. The curves represent different retrackers employed to the data and the vertical bars indicate number of crossovers.

Table 2.1: Noise level and slope-related component of the measurement's precision. Adapted from Schröder et al. (2019). Units are in m and s in degrees.

Data set	Data center
Seasat	$0.21 + 1.91s^2$
Geosat	$0.17 + 0.86s^2$
ERS-1 (ocean)	$0.25 + 0.90s^2$
ERS-1 (ice)	$0.36 + 2.37s^2$
ERS-2 (ocean)	$0.23 + 0.75s^2$
ERS-2 (ice)	$0.38 + 2.57s^2$
Envisat	$0.17 + 1.03s^2$
ICESat	$0.05 + 0.25s^2$
CryoSat-2 (LRM)	$0.18 + 2.46s^2$
CryoSat-2 (SARIn)	$0.38 + 2.01s^2$

2.3 Methodology for determination of error and uncertainty

For each RA data point used to estimate the SEC, we assume that the uncertainty is defined by the relation provided in Table 2.1; hence it is a function of mission[1] and slope. These uncertainties are propagated through the least squares fitting in the SEC processor (σ_p). Furthermore, we acknowledge that an additional error source for the SEC is the fact that the surface might not change in accordance to the prescribed model. Therefore, we add also the RMSE of the model-observations for each SEC estimate and add this to the propagated measurement precision (σ_m) (Sørensen et al., 2018b):

$$\sigma\_total = \sigma\_m + \sigma\_p \qquad (2.1)$$

We should note here that neither of these error sources compensate for changes in elevation due to changes in penetration depth of the RA. Therefore, the RA-derived SEC will differ from a SEC derived from laser data from e.g. OIB. This effect is documented in Simonsen and Sørensen (2017) and shown in Figure 2.2.

σ_total defines the uncertainty of each SEC estimate and is used as input in the Kriging interpolation scheme that is used to inter/extrapolate SEC to cover the entire ice sheet. The error of course increases with distance to nearest SEC point, in which the error is exactly σ_total.

---

[1] The Sentienl-3 noise level and the slope-related component of the measurement's precision will for this first iteration be mirrored by the CryoSat-2 SARIn estimates. This is due to the sub-optimal processing-chain currently used Sentinel-3 limits the number of crossovers at higher slopes areas. This hampers the derivation of an estimate for Sentinel-3, however, when a dedicated land-ice processor is implemented for Sentinel-3 the study by Schröder et al., 2019 be revised to include Sentinel-3.
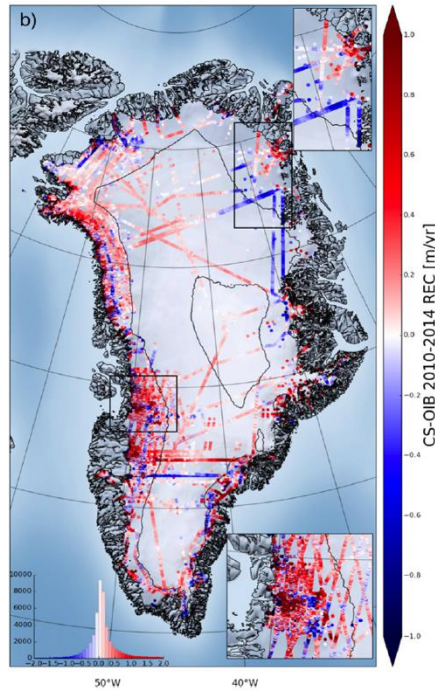
Figure 2.2: Differences between CryoSat and OIB observed elevation changes. From Simonsen and Sørensen (2017).

2.4 Error and uncertainty documentation

As described above the errors associated with the SEC grids is a function of satellite mission, surface slope, goodness of fit (of the least squares regression), and distance to nearest measurement point. Figure 2.3 shows examples of SEC grids and their associated error grids.
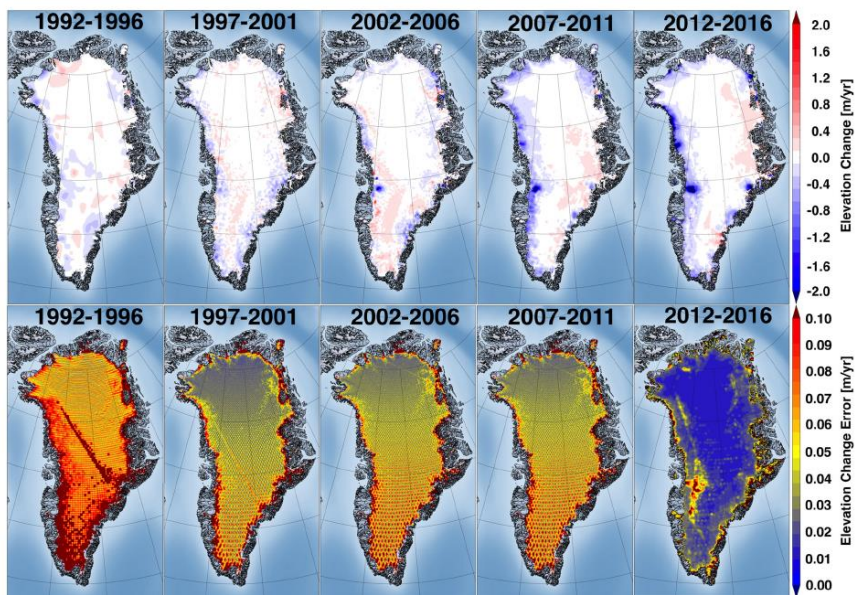
Figure 2.3: Examples of SEc grids and their associated error grids. from Sørensen et al., (2018b).

2.5 Guideline for using the product

Each SEC grid is provided by an associated error grid, so that the user can directly assess the validity of each SEC grid point.

2.6 References

Nilsson, J., Gardner, A., Sørensen, L. S., & Forsberg, R. (2016). Improved retrieval of land ice topography from CryoSat-2 data and its impact for volume-change estimation of the Greenland Ice Sheet. The Cryosphere, 10(6), 2953-2969.

Schröder, L., Horwath, M., Dietrich, R., Helm, V., Van Den Broeke, M. R., & Ligtenberg, S. R. (2019). Four decades of Antarctic surface elevation changes from multi-mission satellite altimetry. The Cryosphere, 13(2), 427-449.

Simonsen, S. B., & Sørensen, L. S. (2017). Implications of changing scattering properties on Greenland ice sheet volume change from Cryosat-2 altimetry. Remote Sensing of Environment, 190, 207-216.

Sørensen, L. S., Simonsen, S. B., Langley, K., Gray, L., Helm, V., Nilsson, J., ... & Davidson, M. W. (2018). Validation of CryoSat-2 SARIn Data over Austfonna Ice Cap Using Airborne Laser Scanner Measurements. Remote Sensing, 10(9), 1354.

Sørensen, L. S., Simonsen, S. B., Forsberg, R., Khvorostovsky, K., Meister, R., & Engdahl, M. E. (2018b). 25 years of elevation changes of the Greenland Ice Sheet from ERS, Envisat, and CryoSat-2 radar altimetry. Earth and Planetary Science Letters, 495, 234-241.

3 End-to-end Uncertainty Budget for IV (SAR)

3.1 Introduction

The main upgrade of the IV processing chain introduced in this cycle of the project is the application of TOPS interferometry for Sentinel-1 data. The processing chain developed in the first two CCI phases is based on offset-tracking and documented in (RD2) and will not be further discussed in this revision of the document. The interferometric processing chain being developed in the CCI+ project is described in (RD3).

3.2 Sources of error

3.2.1 Baseline errors

The nominal orbit accuracy for the precise Sentinel-1 orbits, generally available 21 days after acquisition, is 5 cm 3D RMS. For interferometry, it is the relative orbit error (baseline error), projected on to the line-of-sight direction that affects the IV measurement. Orbital errors are generally of low frequency nature and will therefore to a large extent be corrected when a low-order polynomial calibration of the interferometric displacement measurement is carried out. The method is detailed in (RD3) and Mohr and Boncori (2008).

3.2.2 IW coregistration errors

The major challenge in employing interferometry on Sentinel-1 data over ice sheets is the azimuth-dependent squint angle inherent in the beam steering employed in the Interferometric Wide (IW) swath mode. An azimuth shift, Δx, not accounted for in the coregistration will lead to a phase shift, $\Delta\phi$ (De Zan et al., 2014):

$$\Delta\phi = \frac{4\pi}{\lambda}\Delta x \sin\beta \tag{3.1}$$

where λ is the radar wavelength and β is the squint angle (the angle off broadside at which the target is in the centre of the antenna beam). The azimuth shift can arise both from global coregistration errors (i.e. orbital errors), from ionospheric artefacts, or from actual horizontal azimuth motion associated with the ice flow between the two acquisitions. The phase sensitivity to azimuth shift is much smaller than for shifts in the range direction, since the variation of β along a burst is on the order of $\pm 1°$. At burst boundaries, however, the squint angle changes abruptly, which can cause problems for the phase unwrapping, leading to burst-to-burst discontinuities. The azimuth shift necessary to produce a phase jump of π across a burst boundary with squint angle $+\beta$ and $-\beta$ at the two burst edges is:

$$\Delta \mathbf{x} = \frac{\lambda}{8\sin\beta} \tag{3.2}$$

which for typical values of β=1° for Sentinel-1 is 0.4m. Orbital errors will typically not be of this magnitude but could be observed for ionosphere-induced azimuth shifts. For real ice flow in the azimuth direction, this corresponds to an ice velocity of just 7 cm/day (25 m/y), assuming a 6-day temporal baseline. Employing an external ice velocity map derived from a multi-year average of offset-tracking measurements (or other sources), the bulk of the azimuth motion can be compensated for, as described in (RD3).

3.2.3 Decorrelation

The interferometric phase quality is affected by various contributions, and is measured by the interferometric coherence, which is a number between 0 and 1. The contribution of the different decorrelation effects on the coherence can be summed up by Scheiber et al. (2000):

$$\gamma = \gamma_{SNR} \cdot \gamma_{spatial} \cdot \gamma_{misreg} \cdot \gamma_{temporal} \tag{3.3}$$

At the frequency of Sentinel-1, the ice generally has a high backscatter (and hence SNR), so that the decorrelation due to thermal noise, γ_{SNR}, is small. Spatial decorrelation is also limited in Sentinel-1 data due to the small orbital tube. Since an ice sheet presents a non-stationary scenario, misregistration can

occur due to ice motion between acquisitions, even if the sub-resolution structure remains stable. The impact on the coherence of a range or azimuth misregistration, Δ, is given by:

$$\gamma_{misreg} = \text{sinc}\frac{\Delta}{\rho}, \quad |\delta| < \rho \tag{3.4}$$

where ρ is the resolution in the relevant dimension, and sinc(x)=sin(πx)/(πx). For a misregistration larger than one resolution cell, the coherence drops to zero. Sentinel-1 has the highest resolution in the range direction of approximately 2.5 m, which means that for a 6-day interferogram, an ice flow velocity projected in the range direction of 2.5 m/6 days = 0.44 m/d (160 m/y) leads to complete decorrelation if not accounted for. The use of a multi-year averaged external ice velocity map derived from offset-tracking (see (RD3)) in the coregistration procedure can account for the bulk of this effect.

Finally, temporal decorrelation occurs when the scene scattering properties change between image acquisitions, for example due to melt, snowfall or ice dynamics, and is the major contributor to coherence loss in IV estimation.

A low coherence leads to noise on the phase estimate, which again leads to noise on the displacement estimate. The phase variance associated with a given coherence level is given by Rodriguez and Martin (1992):

$$\sigma_\varphi^2 = \frac{1}{2N_L}\frac{1-\gamma^2}{\gamma} \tag{3.5}$$

3.2.4 GCP uncertainty

The interferometric displacement is a relative measurement, since the unwrapped phase is referenced to a starting point, for which the phase is only known modulo 2π; it is furthermore affected by contributions from orbit and TOPS coregistration effects as described in 3.2.1 and 3.2.2. A bias on the displacement of the phase unwrapping reference point will be propagated to the entire image, thus a calibration using ground control points of known displacement is required. For the ice sheet scenario, it may in many scenes be impossible to find non-moving points on bedrock which are phase-connected to the ice sheet. In this case, moving ground control points must be used. As described in (RD3), we use GCPs in slow moving areas, ideally on ice divides, as these will tend to show the least variation. An uncertainty on the GCP velocity will affect the calibration, and must be accounted for, as described in 3.3.

3.2.5 Phase unwrapping errors

Phase unwrapping errors can arise if steep phase gradients are present in the interferograms, if regions in the image are disconnected phase-wise, or at burst/swath boundaries in the presence of phase discontinuities caused by azimuth misregistration (see 3.2.2). The phase unwrapping is based on a minimum cost flow (MCF) algorithm, with weights for the phase unwrapping are generated from coherence values and intensity edge-detection. The MCF algorithm can to some extent exploit the reduced coherence at burst boundaries to unwrap "around" these discontinuities. A segmentation mask indicating disconnected segments in the unwrapped phase field is generated and output, and absolute phase estimation is first carried out on a main segment (by default the largest segment from the unwrapping segmentation mask), by averaging the difference of the calculated and measured path lengths to each of the GCP's in this segment having a valid interferometric phase. Absolute phases are then estimated for each of the remaining segments. Segments without valid GCP's are discarded. If the segmentation is done incorrectly, this could cause unwrapping errors within improperly connected regions. The error prediction module attempts to model this error, as described in section 3.3.

3.2.6 Velocity inversion errors

Since interferometry measures only displacement in the radar line-of-sight direction, each point for which a horizontal velocity is desired must be imaged from different directions, i.e. from ascending and descending tracks covering the same region. At high latitudes, the tracks are nearly orthogonal, whereas at lower latitudes, the crossing angle difference becomes smaller, and the velocity is poorly resolved in the North-South direction, since both line-of-sight vectors will be primarily oriented in the East-West direction.

With an uncertainty estimate available for each line-of-sight observation, the resulting 2D velocity errors can be estimated as part of the velocity inversion.

3.3 Methodology for determination of error and uncertainty
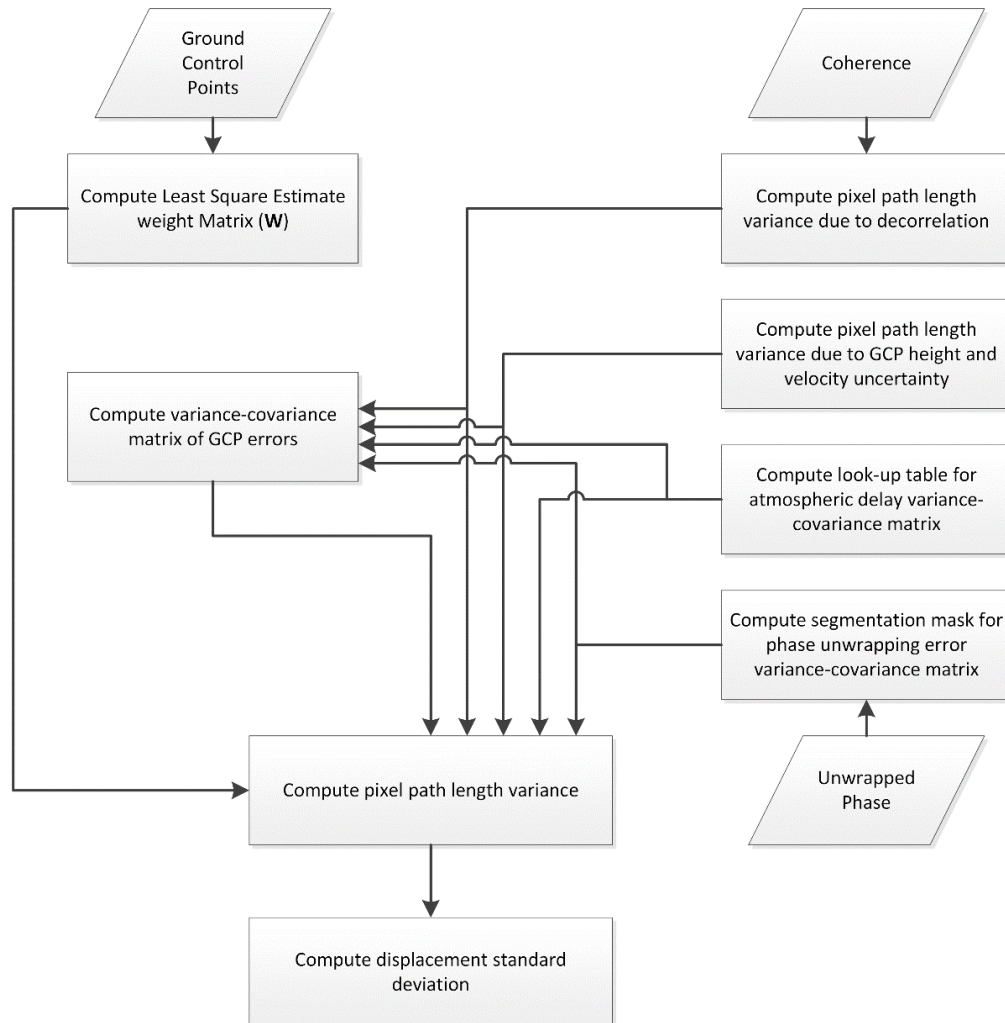
Figure 3.1: Error prediction framework for interferometric displacement measurement.

Error prediction is based on a mathematical framework for baseline error calibration, described in detail in Mohr and Boncori (2008) and Mohr and Boncori (2008). Estimation of the baseline errors is modelled as a least square problem in the presence of correlated noise. The error sources considered are interferometric decorrelation, GCP height and velocity uncertainty, atmospheric propagation and phase unwrapping. Phase unwrapping errors between any two pixels within the unwrapped phase field are considered to be a zero-mean random variable and an attempt is made to model their covariance. This is done in two steps. A segmentation mask is derived at first, based on the residue density and jumps greater than π radians in the unwrapped phase file, and subsequently this is used to assign the error covariance between pixel pairs. The segmentation mask serves the purposes of identifying consistently unwrapped regions.

As shown in Figure 3.1, statistical models for each of the main error sources are used to compute a variance-covariance matrix of the electrical path-length errors affecting the GCPs used for baseline refinement. Together with the weight matrix used during interferometric processing for baseline estimation, this allows computation of the error variance of each calibrated pixel path-length. Path-length error variances are finally converted to height and displacement error standard deviations. These

estimates are propagated in the velocity inversion to errors on Easting and Northing velocities when combining line-of-sight measurements from crossing tracks. In areas where multiple tracks are combined, the error estimates are used to do a weighted average of the displacement measurements, and the error estimate of the output product is updated to reflect this.

3.4 Error and uncertainty documentation

The output from the error prediction described in the previous section is a pixelwise estimate of the standard deviation and is provided with the NetCDF file containing the IV product.

3.5 Guideline for using the product

The IV product is distributed in NetCDF files following the conventions described in the Greenland_Ice_Sheet_cci Product User Guide (PUG). The estimated error standard deviations are provided in the variables land_ice_surface_easting_velocity_std and land_ice_surface_northing_velocity_std, on the same grid as the ice velocity estimates. An example from the North Greenland Ice Stream (NEGIS) is shown in Figure 3.2. This example is based on 13 different pairs, acquired on three descending and two ascending tracks. The standard deviation is seen to be lower for pixels which are observed on multiple tracks, which is to be expected due to the increased amount of averaging that can be done. Also visible is the difference in Northing versus Easting velocity standard deviation. This difference arises because the ascending/descending tracks are oriented primarily in the North-South direction, giving a higher sensitivity in the East-West direction, where the line-of-sight is primarily oriented.
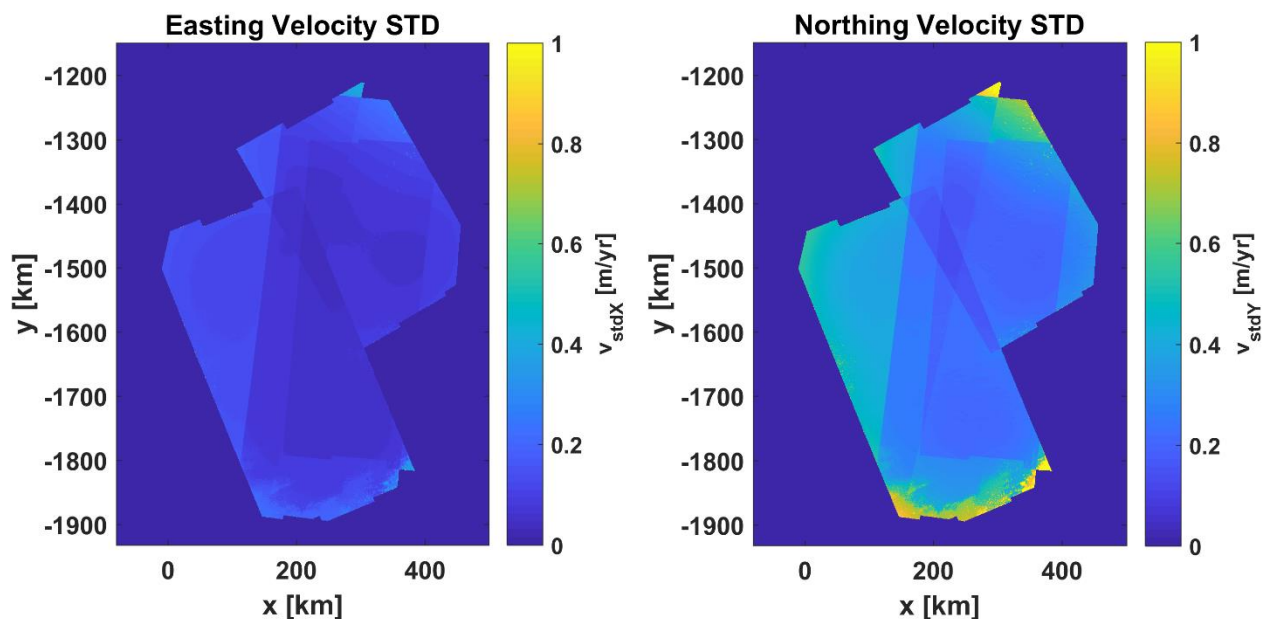
Figure 3.2: Example error estimates for interferometric IV product, for the NEGIS area.

3.6 References

Mohr, J.J.; Boncori, J.P.M.: An Error Prediction Framework for interferometric SAR Data. IEEE Transactions on Geoscience and Remote Sensing (2008), Vol 46, pp. 1600-1613.

De Zan, F; Prats-Iraola P.; Scheiber, R.; Rucci,A..: Interferometry with TOPS: coregistration and azimuth shifts, Proceedings of EUSAR 2014, June 2014, pp. 949-952.

Scheiber, R.; Moreira, A.: Coregistration of interferometric SAR images using spectral diversity, IEEE Transactions on Geoscience and Remote Sensing (2000), Vol 38, no. 51, pp. 2179 –2191.

Rodriguez, E.; Martin, J.M.: Theory and design of interferometric synthetic aperture radars, IEE Proceedings-F (1992), Vol 139, no. 2, pp.147-159.

Boncori, J.P.M. and Mohr, J.J.: A tuneable closed form model for the structure function of tropospheric delay, IEEE Geoscience and Remote Sensing Letters (2008) , Vol. 5, pp. 222-226.

4 End-to-end Uncertainty Budget for IV (Optical)

4.1 Introduction

In this section the error sources, uncertainties and methodology for characterisation of the errors of the derived optical IV ECV will be outlined.

The three main techniques used to derive ice velocity from satellite image data:

1. Differential SAR Interferometry (DInSAR)
2. Multiple Aperture Interferometry (MAI)
3. Offset tracking

These methods have different error characteristics depending on the different velocity regime (fast or slow), and several other aspects, such as the asperity of the terrain and the time of the imagery acquisition (applicable to optical imagery only). The optical IV ECV produced within the CCI project and continued during its extension CCI+, is based on an offset tracking technique, and more specifically on a feature-tracking method. Feature-tracking has been selected because of the robustness of this type of technique. Both SAR and optical imagery can be used as input for feature-tracking approaches, but SAR data have the advantage of an active sensor. This sensor is not affected by solar illumination (day/night) or cloud coverage, two aspects particularly relevant at high latitudes and known as sources of errors when feature-tracking techniques are applied to optical imagery. Nevertheless, optical IV is important because it allows filling gaps in SAR IV products during summer in areas with surface melt, particularly along the margins.

A general overview of the expected accuracy of the three methods aforementioned can be found in (RD3). Table 4.1 gives some general expected estimates, but individual errors will vary widely.

Table 4.1: Typical expected errors of the ECV IV product.

Area of ice sheet	Accuracy	Sensors	Method
Ice sheet interior	<10 m/y	ERS SAR	Calibration with balance velocities D-InSAR and MAI techniques
Ice sheet margin	10 to 30 m/yr	ERS SAR, ASAR, PALSAR	Calibration with stationary control points on bedrock. Speckle-tracking (highest accuracy) and feature-tracking (lowest accuracy)
	10 to 30 m/yr	Optical (Sentinel-2)	Feature tracking

4.2 Sources of error

The main sources of error for offset tracking techniques applied to optical imagery are:

- Sun angle variations across the image (δ_{sun}). These variations reflect in several solar illumination effects, which increase the noise in the imagery and consequently in the displacement retrieval.

- Cloud coverage (δ_{cloud}). The presence of clouds over ice and/or snow is hardly distinguishable. Consequently, it is one of the main sources of noise/failures for offset-tracking-based algorithms.

- Image misalignment between two acquisitions (δ_{misal}). This error may have a contribution in both x and y coordinates. An accurate multi-temporal co-registration of optical image time series is an indispensable pre-requisite for many analysis techniques, including feature-tracking.

- Presence of bed-related topographic features (δ_{topo}). Those features remain stationary as the ice flows over them. Consequently, they may distort the shape of the features, causing a misdetection

or an error in the measurement of the retrieved displacement. The topographic component impacts mainly the offset measurements along both x and y components.

- Intrinsic error of the offset-tracking algorithm selected for optical IV retrieval (δ_{noise}).

The sources of errors listed above lead to:

$$\delta(x, y) = \delta_{sun} + \delta_{misal} + \delta_{topo} + \delta_{cloud} + \delta_{algo} \qquad (4.4)$$

During the generation of the optical IV ECV, a series of pre- and post-processing steps were developed to reduce the effect of the various error components listed in Equation 4.1. These pre- and post-processing steps are described in more detail in Chapter 4.4. Before describing these steps, one of the sources of error listed above will be analysed more in depth: image misalignment.

4.2.1   Image misalignment

A very common source of error is the one caused by a misalignment between two image acquisitions. Sentinel-2, the high-resolution optical data source used for retrieving the optical IV ECV, is not exempt from image misalignment. Indeed, products processed before 15 June 2016 were affected by a yaw bias correction anomaly leading to low co-registration accuracies. These accuracies translate in a misalignment of up to 2 pixels (20 m).

The long-term performance for products processed after the yaw bias correction anomaly was solved (15 June 2016), shows an increased accuracy of 11-12 m at 95% for both Sentinels. Even if excellent, this value is still distant from Sentinel-2 multi-temporal geolocation target (0.3 pixel at 95% confidence level).

More recently, the geolocation performance of Sentinel-2 A has improved significantly after the calibration performed in December 2018. On the contrary, Sentinel-2 B's geolocation performance degraded again in October 2019. This degradation led to a new calibration on 27 November 2019.

The most recent multi-temporal performance statistics for the Sentinels are shown in Table 4.2.

Table 4.2: Multi-temporal performance statistics for Sentinel 2A and B (Source: L1C Data Quality Report).

Co-registration error	0<X<0.5 pixels	0.5<X<1 pixels	1<X<1.5 pixels	>1.5 pixels
S2A % of products	58%	35%	6%	1%
S2B % of products	40%	43%	16%	1%
S2A/S2B % of products*	33%	31%	34%	2%

The figures listed in Table 4.2 are not yet compliant with the mission specifications of 0.3 pixel at 95% confidence level, which are expected to be met after the activation of the global reference image. ESA has reported that the final version of the Global reference image has been completed in March 2019 and that it is currently under validation.

Finally, in 2019 ESA has started a global reprocessing (still ongoing) of Sentinel-2 archive. The reprocessing includes ortho-rectification with GCPs obtained via automatic matching.

To correct the displacement between Sentinel-2 images, a pre-processing step has been included in the optical IV CCI processing chain. The step is described more in detail in the next Chapter.

4.3 Methodology for determination of error and uncertainty

Determining the uncertainty correlated to each source of error listed in Chapter 4.2 requires different methodologies.

4.3.1 Sun variations

Sun variations reflect in solar illumination effects, which increase the noise in the imagery and consequently in the displacement retrieval.

The sun-target-sensor geometry is often defined by two angles:

1) The incidence angle as determined by the solar elevation angle (the complement of the solar zenith angle).

2) The viewing angle, defined as the deviation of the sensor from the nadir position.

When the observed surface is not flat and presents irregularities, the incidence angle is modified by the slope and aspect of the target surface, impacting the calculation of IV with optical imagery as input. A typical example are the shadows generated by steep mountain sides in glacial valleys. Specifically, such shadows can cause matches which are displaced relative to the probable direction of the flow.

Several mitigation actions may be taken for minimizing the effect caused by Sun variations:

1) Select input images taken at a similar time of day.

2) Calculate the exact location and orientation of the shadows from the sun angle and other satellite parameters.

3) Apply a topographic correction to account for changes in illumination-shadows caused by slopes.

4) Filtering out pixels with unrealistic displacements (shadow will be mainly stable) in a post-processing step.

Actions 1) and 4) are included in the current version of the optical IV processing chain. The implementation of 2) and 3) is currently not foreseen.

4.3.2 Cloud coverage

Cloud detection is one of the most crucial steps during the pre-processing of optical satellite images. Indeed, failure to mask out the clouds from the image can have a significant negative impact on any subsequent analyses, including displacement detection via feature-tracking.

Multiple algorithms are currently in use to identify pixels contaminated by clouds. They can be roughly divided into three groups, depending on whether they use a single acquisition, multiple acquisitions (acquired by a satellite at different dates), or machine learning techniques.

Unfortunately, the cloud masks currently available are not very reliable when applied to imagery dominated by ice and snow, independently on the technique used. Sentinel-2 cloud masks, which are provided together with the Level-1C products, were tested but the results were not considered satisfactory and its inclusion as a pre-processing step of the optical IV processing chain discarded.

To overcome the unavailability of a reliable cloud mask, a simple cloud filtering was implemented as a post-processing step. During this post-processing step, the Root Mean Square (RMS) error of the easting-northing velocity for every pixel of an IV map previously generated is computed. Being the standard deviation a measure of how spread out the residuals are, after the calculation of the RMS error a threshold is set and pixels with RMS larger than the threshold masked out. Even if very basic, this approach works reasonably well because the RMS error of the velocity of cloudy pixels is often much higher than the one of free-clouds ones.

4.3.3 Image misalignment

A pre-processing step determining the co-registration has been developed and included in the optical IV processing chain. The algorithm computes the offset between two rasters, with one of the two taken as reference, and returns the offset in both x and y directions in sensor geometry.

The arguments of the co-registration algorithm are:

- r1, r2: Rasters 1 and 2
- N: Number of samples
- window_size: Window size in pixels
- search_radius: Search radius in pixels
- metric: Similarity metric function. It takes two arrays as positional argument
- threshold: If the output of metric is less than threshold, the sample is disregarded
- Mask (optional): Mask to be applied to raster 2

The script performs the following steps:

1. Obtain N random points.
2. Exclusion of mask-pixels with value below the defined threshold from the offset calculation.
3. Search matches in the search window of search_radius.
4. Return the average pixel offset over all samples.
5. Subtract the average pixel offset from raster 2.

Note: The two raster images must reside on the same grid and have the same extent.

An alternative, more conventional approach is the development of a method based on a set of pre-defined GCPs. Even if more reliable, at high latitudes some a potential issue needs to be highlighted. The GCPs could potentially be not identifiable or misidentified at different times of the year, sometimes even within the same month. This is due to the extreme conditions present at those latitudes, which can cause the GCPs to be partially or fully covered by snow/ice, but also to the absence of structures with sharp edges, such as buildings or road crossing. In such cases, the identification of the GCPs can be problematic, leading to errors.

Validation activity on the method described in this chapter is still ongoing. If results were not satisfactory, it would be foreseen to test a GCP-based approach.

4.3.4 Bed-related topographic features

The identification of bed-related topographic features is quite challenging. One of the biggest challenges is the presence of supraglacial lakes. Neither their emergence nor their location is easily predicable. Moreover, their number increases year after year, mainly because of the thinning of the ice sheet and of the melting of the superficial snow happening earlier and earlier during the summer season. Another factor to consider is that supraglacial lakes can sometimes be confused with clouds.

Finally, other bed-related topographic features can be a source of noise if those features remain stationary as the ice flows over them.

A processing step fully dedicated to the identification of bed-related topographic features is currently not implemented in the optical IV processing chain.

Nevertheless, lakes are partially filtered out during the cloud filtering post-processing step. This post-processing step has been already described in the section "Cloud coverage".

Identification of supraglacial lakes using ML-based approaches or using a complete and up to date database could lead to a drastic reduction of this source of errors, but its inclusion is not foreseen during the time of this project.

4.3.5   Algorithm

Feature-tracking accuracies are approximately 2 times worse than algorithms based on other types of methods, as for example speckle-tracking (De Lange et al., 2007). Compared to interferometry-based methods, they are not affected by low or absence of coherence, and they do not have a limit on the maximum measurable velocity magnitude, at least when applied to the range of velocities of the Greenland ice-sheets (< 13 km/y).

As documented in (RD3), the feature-tracking algorithm used to produce the optical IV ECV calculates a correlation value for every location at which the reference chip will entirely fit within the search chip. The best integer-pixel match is taken, and the algorithm interpolates a peak correlation location. In the following step, the correlation strength is determined. This parameter is a combination of peak height of the correlation map, height of peak to second-highest peak, and height of peak to background value of correlation index.

The algorithm then returns a flag parameter indicating if a good match was found, or if not giving an indication of what was the reason of the failure. The flag parameter can assume the following values:

1. Success.

2. Failure: correlation peak too near the edge of search chip.

3. Failure: subsidiary peak comparable in strength to the main peak.

4. Failure: diagonal displacement from nominal location exceeds the maximum specified by the user.

In the final step, the x and y error estimates are derived from a peak-height-to-peak-width comparison.

4.4   Error and uncertainty documentation

Each generated map of a Cartesian velocity component will be accompanied by its associated error. The error will be provided as a map, in the same geometry as the associated measurement, providing a measure of uncertainty on a per-pixel basis.

Product's metadata contain information on both the offset correction (in both x and y directions) computed during the pre-processing steps and on the RMS threshold used during the cloud-filtering post-processing step.

4.5   Guideline for using the product

It is expected that most estimated errors will be in the range of 10-30 m/yr, as outlined in (RD3). The selected, high-resolution (50m) gridded representation in north- and east-velocity components ensure a simple-to-use product, easily accessible for non-experts.

As final remark, it should be kept in mind that the technique employed tends to degrade when not applied to features rich and/or coherent areas.

For further guideline refer to ATBD (RD3) and PUG (RD4).

4.6   References

De Lange, R., A. Luckman, T. Murray Improvement of satellite radar feature tracking for ice velocity derivation by spatial frequency filtering, IEEE Trans. Geosc. Rem. Sens. 45, 2309-2317 (2007).

S2 MPC, L1C Data Quality Report, Reference: S2-PDGS-MPC-DQR, Issue: 47, Date: 06/01/2020

5 End-to-end Uncertainty Budget for GMB

5.1 Introduction

Errors in mass changes derived from GRACE and GRACE-FO monthly solutions have several origins. The three major contributions arise from:

1. Errors in the monthly gravity field solutions.

2. Leakage errors due to the limited spatial resolution achieved by both missions.

3. Errors in models used to reduce superimposed mass signals.

In the following we will give an overview on the different sources of errors, the applied methodologies for their assessment and a description of the accuracy measures provided along with the GMB products. More detailed discussions on selected aspects have been given previously, e.g., by Wahr et al. 2006; Horwath and Dietrich, 2009; Velicogna and Wahr, 2013; Barletta et al. 2013 and more studies cited in what follows.

5.2 Sources of error

In the following subsections the error components are discussed.

5.2.1 Monthly gravity field solutions

Monthly GRACE/GRACE-FO solutions (usually represented by a set of spherical harmonic Stokes coefficients) are significantly afflicted by errors. The errors have a pronounced covariance structure, induced by the different sensitivities of the satellite observations to different parts of the spherical harmonic spectrum and by the different susceptibility of parts of the spectrum to modelling errors involved in the GRACE processing procedure.

In geographic representations of gravity field, or mass, changes from GRACE/GRACE-FO, the typical north-south striping is a representation of the mission's error covariance structure. One reason is the along-track (hence mainly north-south) direction of the ranging observations between the twin satellites, giving higher sensitivity to this direction (e.g. Schrama et al. 2007). A second reason arises from temporal gravity field changes within a month. Actual temporal variations may alias into striping patterns of spatial variations (Seo et al. 2008). The Level-2 processing involves a de-aliasing procedure, where high-frequency tidal and non-tidal mass variations in atmosphere and ocean are reduced from the observations using Atmosphere and Ocean De-aliasing Level-1B (AOD1B) products (Dobslaw et al., 2016). Nevertheless, limitations in the utilized background models will lead to aliasing in monthly gravity field solutions caused by residual mass variations which have not been reduced. Processing centres provide calibrated errors, derived by a degree-dependent scaling of formal errors for each coefficient, as an estimate for the correlated GRACE errors.

Algorithms for the determination of mass changes intend to minimize GRACE error effects, e.g. by means of a suitable filter. Since available approaches are not able to entirely remove correlated errors, residual errors will inevitable be propagated to regional mass change estimates.

The magnitude of monthly solution errors varies over time. An increased error level is visible towards the end of the GRACE mission lifetime, caused by the ageing instruments, peaking in solutions following August 2016 when the accelerometer on board of one of the spacecrafts had to be switched off (Dahle et al., 2019b). GRACE-FO monthly solutions are derived from observations of one accelerometer only from the very beginning. Nevertheless, the error level of GRACE-FO solutions is lower than for GRACE single accelerometer solutions and is comparable to solutions of Release 05 from the middle of the GRACE era (Dahle et al., 2019a).

5.2.2 Signal leakage

The limited spatial resolution of GRACE-type missions, caused by the attenuation of small-scale gravity changes at satellite altitude, induces signal leakage. Signals from adjacent regions cannot be precisely

separated. In this way, signals from outside a region (e.g. a drainage basin) may leak into the regional estimate (leakage-in). Likewise, mass signals within the region under investigation may not be completely recovered, but under- or overestimated (leakage-out). Leakage errors are most pronounced for the investigation of small regions. Since leakage errors are related to actual mass changes, they are not uncorrelated from month to month. Their temporal correlation behaviour is instead inherited from the respective behaviour of the underlying mass changes.

Leakage-out originates from mass changes of the GIS itself. Variations in SMB cause leakage on inter-annual and long-term scales. Changes in ice dynamics, i.e. ice discharge, induce leakage mainly affecting long-term signal components (e.g. linear trends). Mass changes causing leakage-in may be subdivided into near-field and far-field signals. Ice mass changes in neighboring drainage basins are one distinct source of leakage-in. However the most important source of leakage-in is due to the nearby Ellesmere Island (white area in Figure 5.1) and to the outer glaciers surrounding the ice sheet (dark violet area in Figure 5.1).
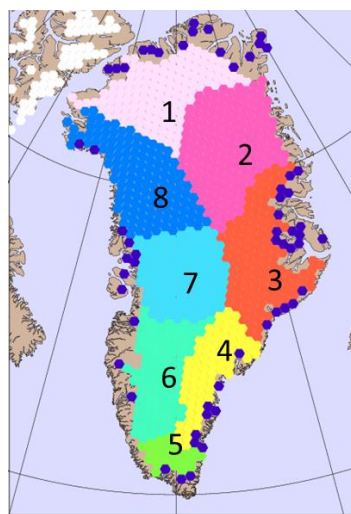

Figure 5.1: Overview of the drainage basin definition and the solution area. Ellesmere Island is shown in white and GIS surrounding glaciers and ice caps in dark violet.

Since oceanic mass variations have been largely removed during the GRACE de-aliasing process using AOD1B products, only residual variations, due to uncertainties in the utilized background models, will possibly bias the ice mass estimates. The same applies to high-frequencies mass changes of the atmosphere. Leakage from far-field regions stems from continental hydrology, oceans and globally distributed ice masses such as the AIS. For a quantification of leakage errors from different near- and far-field source see Groh et al. (2019).

5.2.3 GIA

Glacial isostatic adjustment (GIA) has been recognized as the major cause of uncertainty in ice mass balance estimates from GRACE. GIA is the solid Earth phenomenon responsible for the mantle flow from the equatorial region towards the Pleistocene deglaciated areas. This mass flow within the Earth produces variations in topography and in gravity, and the latter is detected by GRACE mainly as a positive mass trend that, in many areas, cannot easily be distinguished from surface mass accumulation. While in Antarctica GIA has a very strong signature in the time variable gravity field, GIA accounts only for about the 2-5% of the total mass changes in Greenland, depending on the utilized GIA model.

5.2.4 Missing degree one information

Another source of uncertainty is related to geocenter motion and the different ways to infer it. Large mass variations, like the grounded ice melting in Greenland and Antarctica, generate not only gravity changes but also Earth surface displacements. These translate into a displacement of the center of mass (CM) with respect to the geometric center of the Earth (center of figure CF): this movement is known as the geocenter motion. Therefore, mass balance estimates should carefully take into account the geocenter motion. However, since the GRACE satellites move together with the center of mass of the Earth, in principle they cannot detect the geocenter motion, so this effect has to be recovered by other means. The choice for the correction for geocenter motion is still an open issue: there are different ways to calculate it and therefore more than one possible correction exist and this can be a source of variability in GRACE-derived mass balance estimates.

Different data sets can be used to add degree one information to the GRACE monthly solutions. Cheng et al. (2013a) use repeated SLR observations to derive degree one information from translations of the ground station network. Other approaches combine different types of data like GRACE and ocean models (Swenson et al., 2008; Sun et al., 2016) or GRACE, ocean models and GPS (Rietbroek et al., 2012). Time series from these approaches differ both in amplitude and in linear trend (Barletta et al., 2013). Differences in trend are mainly related to the nature of the time series. While Swenson et al. (2008) have reduced degree on trends caused by GIA, time series of Cheng et al. (2008) include the full degree one signal. All solutions exhibit errors caused by the limitations of the individual utilized techniques. One limitation of all geometric approaches (e.g. using SLR or GPS) is the sparse and irregular distribution of observing stations on the Earth's surface.

5.2.5 Different data source and methodologies

Another source of variability in mass balance computation arises from different strategies for processing the raw data by the GRACE/GRACE-FO project centers; thus discrepancies between different time-variable gravity field solutions are known.

In Barletta et al. (2013) the difference between the two official and most commonly used data sets (GFZ and CSR) are analyzed and used to derive uncertainties. Also an interesting overview of the difference between GFZ's old RL05 and the current RL06 release is shown in Dahle et al. (2019b), which confirms a lower noise content in the most recent releases.

For CCI+ we have selected the widely used CSR RL06 solution series (Bettadpur, 2018). In the course of the project CSR-based GMB products will be complemented by a product based on the solution series provided by TU Graz (ITSG-Grace2018, Kvas et al., 2019). The ITSG releases are known for their low noise level and form the core of a new European Gravity Emergency Service project (EGSIEM) and its successor COST-G, where the goal is to make new, even more accurate combined GRACE products (Meyer et al., 2019). Comparison of results based on different Level-2 data sets (e.g. CSR RL06 and ITSG-Grace2016) show only minor differences in the derived trends for GIS (Mottram et al., 2019).

Different approaches in extracting the mass balance from GRACE/GRACE-FO Level-2 data can be another source of variability and uncertainties. The impact of different methods for processing Level-2 data was investigated in the Round Robin Exercise conducted within the 2nd phase of the GIS_cci project. We found that GIS mass balance estimates for the period 2003-2013 vary between -252Gt/yr and -274Gt/yr (Groh et al., 2019). This spread is solely caused by different processing schemes applied to identical input data. In principle, independent methods should produce the same results if careful calibration and cross validation have been carried out.

5.3 Methodology for determination of error and uncertainty

5.3.1 Data error

We derive the uncertainties which are related to the data errors provided directly with the GRACE monthly models by using a Monte-Carlo-like approach in which 200 simulations are performed. The simulations are created from Stokes coefficients drawn from normal distributions with zero mean, and the standard

deviation provided with the GRACE/GRACE-FO Level-2 data. We deal with both precision and accuracy errors in our final results. The precision error accounts for the statistically distributed random error around one average value. The accuracy error accounts for how much the expected value deviates from the "true" value. For the precision error we provide the 95% confidence interval (2 sigma) propagated from the data.

Barletta et al. (2013) shows that the accuracy error associated with the differences of the methods is about 2%, which is smaller than the difference found by the use of different data sets. We test the accuracy error associated with the use of different data sets (CSR05 and ITSG16) and we find that the difference is small (see Figure 5.2, Figure 5.3 and Figure 5.4).

However, the precision error (2 sigma) propagated from the data is about 50% of the total error (Barletta et al. 2013). So the average (over the time series) of the precision error is an approximation that accounts for the accuracy error coming from uncertainty on degree one, from the use of different data set and different methods, and from the uncertainties on models used to produce the GRACE Level-2 data (as the ocean and atmospheric model). In our final estimate for each time series we provide a monthly error which is the sum of the precision error and the average of the precision error over the whole time series.
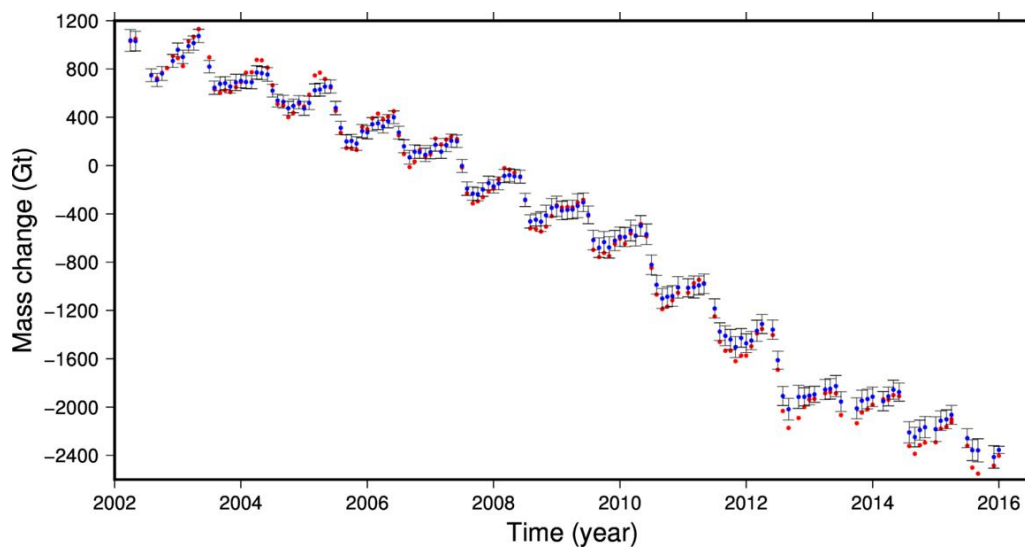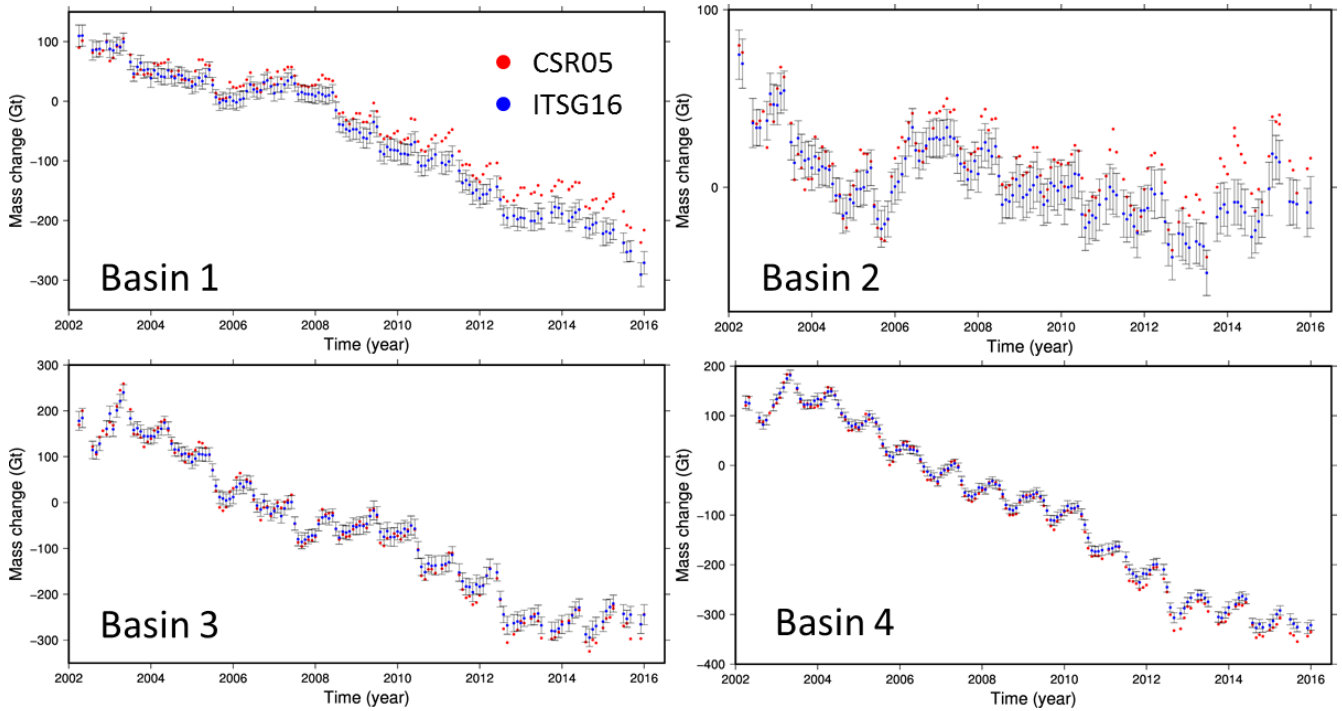

Figure 5.2: The whole Greenland time series.
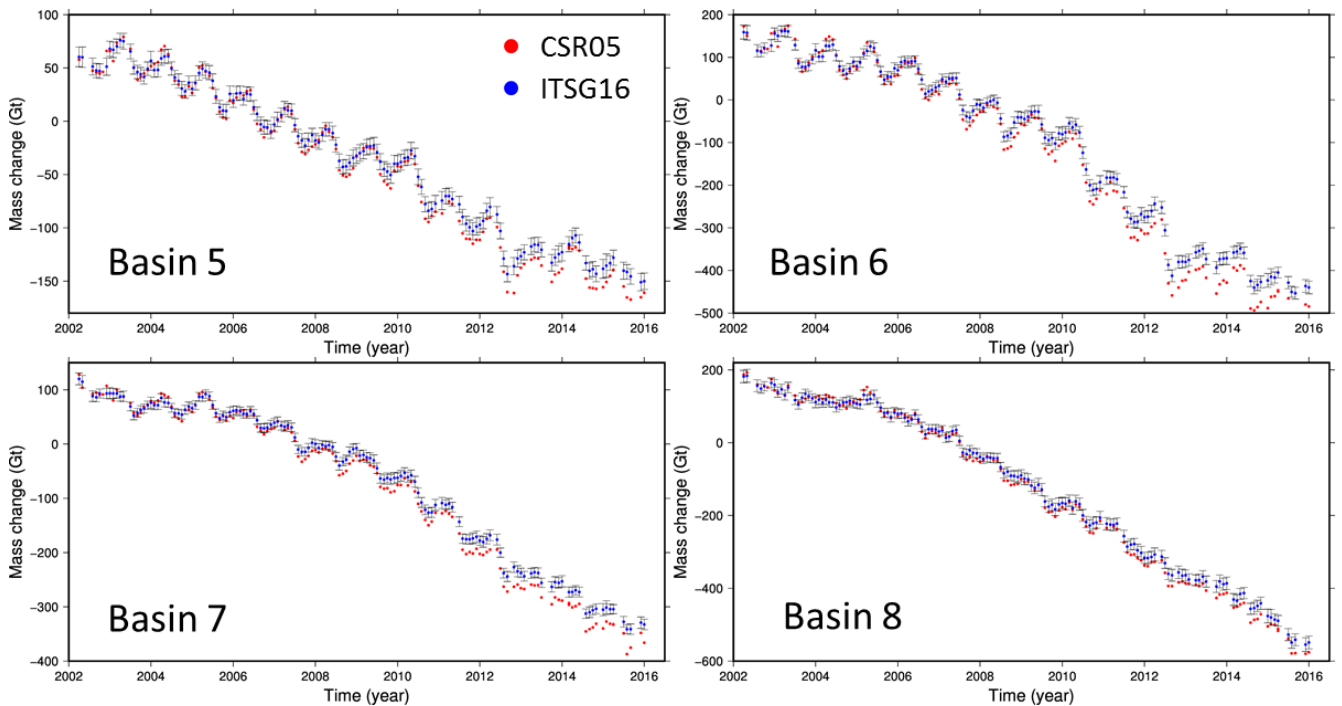
Figure 5.3: Time series for basin 1 to 4.

Figure 5.4: Time series for basin 5 to 8.

5.3.2 GIA

The GIA error is meaningful only for the linear trends in mass changes. For our best value we chose to use the Caron et al. (2018) model, which is a statistical model derived from a large ensemble of GIA models. Its contribution to the entire GIS mass change is about +10.6 Gt/yr. From a comparison with other models (Peltier et al., 2018; Fleming and Lambeck, 2004; A et al., 2013) we derived an uncertainty of about +/- 5.1 Gt/yr.

5.3.3 Missing degree one information

Usually uncertainties of the utilised degree one time series are based on the intercomparison of different available solutions. By comparing three independent degree one solutions Barletta et al. (2013) found that for certain drainage basins the error effect on a monthly solution can reach up to 40% of the total monthly error. We account for this error as explained in section 5.2.1.

The degree one contribution to the linear ice mass trend of the entire GIS was found to be -3.5 Gt/yr, mainly originating from GIA with an uncertainty of +/- 3 Gt/yr.

5.3.4 Leakage error

The mass estimate of each of the basins is obtained as the sum of the point mass changes within each basin mask definition (Figure 5.1).Each basin is colour-coded with dark violet being the surrounding glaciers and ice caps not connected to the ice sheet. The basin definition (Figure 5.1) covers only the ice sheet, but the outer glaciers contribute to the signal, and therefore have to be taken into account too.

In order to give an estimate at basin scale of the effect of the outer glaciers (OG) leakage we compute two solution which represents an upper and lower bound for the mass loss. One solution does not include the OG in the solution area (SA) so the signal of OG is partially included in the ice sheet and therefore we get an upper bound. The second solution includes the OG in the solution area, but we integrate only over the ice sheet (same area as in previous solution). The inversion method tends to place more mass on the edge and therefore if the outer glaciers are included in the solution area, part of the GIS signal will be pushed in the outer glaciers area. So, the second solution is a lower bound.

If the truth was in the middle we just take the average of those 2 values, however we believe that the second solution (with OG in SA) greatly underestimate the mass loss. We give both values in Table 5.1 and we also give a weighted average (W-Ave) and its standard deviation which represents the leakage error at basin scale. For the weighted average we use a weight of 3 for the first solution (No OG in SA).

This leakage error it's between 4% and 10% of the trend. The ratio error/trend in Table 5.1 can be used to get an extra contribution to the error in the monthly solution for applications that don't include the outer glacier signal.

Table 5.1: Trend at basin scale for 2 solutions without and with outer glaciers (OG) in solution area (SA).

	No OG in SA	OG in SA	W-Ave	st.dev
GIS01	-26.32	-21.95	-25.23	2.18
GIS02	-3.62	-2.91	-3.44	0.36
GIS03	-36.68	-30.59	-35.16	3.04
GIS04	-39.21	-35.62	-38.31	1.79
GIS05	-17.29	-15.15	-16.76	1.07
GIS06	-49.05	-45.11	-48.07	1.97
GIS07	-34.56	-31.36	-33.76	1.60
GIS08	-55.33	-51.28	-54.31	2.03
GIS	-262.05	-233.97	-255.03	14.04

Note that the total in Table 5.1 is only over the ice sheet. While for the whole Greenland we include the outer glaciers in the integration.

When we consider the whole Greenland we still have some leakage effect coming from Ellesmere Island. Based on the analysis of Table 5.1 we can assume that about 10% of Ellesmere Island signal leak into the GIS solution. If we integrate our solution over that region we obtain a mass change of -32.5 Gt/yr so for GIS we can assume a leakage error of about 3.2 Gt/yr.

5.4 Error and uncertainty documentation

Monthly mass change time series per basin will be provided with an average monthly error estimate. This is the error described in section 5.3.1.

Using the basin-scale mass change time series, mass balance (i.e. mass change per time) estimates, e.g. linear trends over the entire observational period, is also be provide in Table 5.1. Since effects on the long-term trends are the most critical errors, we will pay particular attention on the provision of comprehensive error estimates for long-term trends. These errors will contain the full range of uncertainties affecting the linear trend as described in the previous section. In particular, this comprises the formal error of the linear trend, derived by propagating the scaled error standard deviation, errors caused by GIA model uncertainties, leakage errors from different globally distributed sources, and errors in degree one.

Gridded GMB products showing 5-year mean trends of mm water equivalent per year (for detailed description see PSD) are considered to be used as a tool for visualising the spatial patterns. No error estimates are provided for every grid cell of the gridded product.

Table 5.2 summarizes the individual error components to be considered in the total error budget of the final GMB products through the example of the entire GIS. The estimation procedures as well as the expected ranges, which are either based on the Round Robin results or taken from the literature, are described in previous sections.

Table 5.2: Summary of errors on the trend for the whole Greenland.

Error source	Estimation procedure	Expected range
GRACE solutions	Propagation of the scaled standard deviation	3.2 Gt/yr
GIA model	Intercomparison of different models	5.1 Gt/yr
Leakage	Analysis of use of different SA	3.2 Gt/yr
Degree one	Intercomparison of different degree one time series	3 Gt/yr
Total	Individual components summed in quadrature	7 Gt/yr

5.5 References

Barletta, V. R., Sørensen, L. S., and Forsberg, R. (2013). Scatter of mass changes estimates at basin scale for Greenland and Antarctica. The Cryosphere, 7(5), 1411–1432.

Bettadpur, S. (2018). UTCSR Level-2 Processing Standards Document for Level-2 Product Release 0006. Austin: Center for Space Research, The University of Texas at Austin.

Caron, L., Ivins, E. R., Larour, E., Adhikari, S., Nilsson, J., & Blewitt, G. (2018). GIA Model Statistics for GRACE Hydrology, Cryosphere, and Ocean Science. Geophys. Res. Lett., 45. doi:10.1002/2017GL076644.

Cheng, M. K., Ries, J. C., &Tapley, B. D. (2013a). Geocenter Variations from Analysis of SLR Data. In Z. Altamimi, & X. Collilieux (Eds.), Reference Frames for Applications in Geosciences (pp. 19–25). International Association of Geodesy Symposia, 138.Springer Berlin Heidelberg.

Dahle, C., Murböck, M., Flechtner, F., and GFZ GRACE/GRACE-FO SDS Team (2019a). Level-2 GRACE/GRACE-FO RL06 Products from GFZ. Presentation, GRACE/GRACE-FO Science Team Meeting, JPL, Pasadena.

Dahle, C., Murböck, M., Flechtner, F., Dobslaw, H., Michalak, G., Neumayer, K., Abrykosov, O., Reinhold, A., König, R., Sulzbach, R., & Förste, C. (2019b). The GFZ GRACE RL06 Monthly Gravity Field Time Series: Processing Details and Quality Assessment. Remote Sens., 11(18), 2116.

Dobslaw, H., Bergmann-Wolf, I., Dill, R., Poropat, L., & Flechtner, F. (2016). Prodcut Description Document for AOD1B Release 06. Potsdam: Deutsches GeoForschungsZentrum GFZ.

Fleming, K., & Lambeck, K. (2004). Constraints on the Greenland Ice Sheet since the Last Glacial Maximum from sea-level observations and glacial-rebound models. Quat. Sci. Rev., 23(9-10), 1053–1077. doi:10.1016/j.quascirev.2003.11.001.

Groh, A., Horwath, M., Horvath, A., Meister, R., Sørensen, L. S., Barletta, V. R., Forsberg, R., Wouters, B., Ditmar, P., Ran, J., Klees, R., Su, X., Shang, K., Guo, J., Shum, C. K., Schrama, E., & Shepherd, A. (2019). Evaluating GRACE Mass Change Time Series for the Antarctic and Greenland Ice Sheet—Methods and Results. Geosciences, 9(10), 415.

Horwath, M., & Dietrich, R. (2009). Signal and error in mass change inferences from GRACE: the case of Antarctica. Geophys. J. Int., 177(3), 849–864.

Kvas, A., Behzadpour, S., Ellmer, M., Klinger, B., Strasser, S., Zehentner, N., & Mayer-Gürr, T. (2019). ITSG-Grace2018: Overview and Evaluation of a New GRACE-Only Gravity Field Time Series. J. Geophys. Res. Solid Earth, 124(8), 9332–9344.

Meyer, U., Jean, Y., Kvas, A., Dahle, C., Lemoine, J. M., & Jäggi, A. (2019). Combination of GRACE monthly gravity fields on the normal equation level. J. Geod., 93(9), 1645–1658.

Mottram, R., B. Simonsen, S., Høyer Svendsen, S., Barletta, V. R., Sandberg Sørensen, L., Nagler, T., Wuite, J., Groh, A., Horwath, M., Rosier, J., Solgaard, A., Hvidberg, C. S., & Forsberg, R. (2019). An Integrated View of Greenland Ice Sheet Mass Changes Based on Models and Satellite Observations. Remote Sens., 11(12), 1407.

Peltier, W. R., Argus, D. F., & Drummond, R. (2018). Comment on "An Assessment of the ICE-6G_C (VM5a) Glacial Isostatic Adjustment Model" by Purcell et al. J. Geophys. Res. Solid Earth, 123(2), 2019–2028. doi:10.1002/2016JB013844.

Rietbroek, R., Fritsche, M., Brunnabend, S. - E., Daras, I., Kusche, J., Schröter, J., Flechtner, F., & Dietrich, R. (2012). Global surface mass from a new combination of GRACE, modelled OBP and reprocessed GPS data. J. Geodyn., 59-60, 64–71.

Schrama, E. J., Wouters, B., &Lavallée, D. A. (2007). Signal and noise in Gravity Recovery and Climate Experiment (GRACE) observed surface mass variations. J. Geophys. Res. Solid Earth (1978–2012), 112(B8).

Seo, K. W., Wilson, C. R., Chen, J., &Waliser, D. E. (2008). GRACE's spatial aliasing error.Geophys. J. Int., 172(1), 41-48.

Sun, Y., Riva, R. E. M., & Ditmar, P. (2016). Optimizing estimates of annual variations and trends in geocenter motion and J2 from a combination of GRACE data and geophysical models. J. Geophys. Res. Solid Earth, 121(11), 8352–8370

Velicogna, I and J Wahr (2013). Time-variable gravity observations of ice sheet mass balance: Precision and limitations of the GRACE satellite data. Geophys. Res. Lett., 40(12), 3055-3063.

Wahr, J., Swenson, S., &Velicogna, I. (2006). Accuracy of GRACE mass estimates.Geophys. Res. Lett., 33, L06401.

6 End-to-end Uncertainty Budget for MFID

Note that much of the text below is from Mankoff et al. (2019) and is reproduced here either verbatim, or with minor changes to reflect minor changes in the work presented here.

6.1 Sources of error

Error sources include the following:

- Ice thickness
- Change in ice thickness
- Velocity
- Errors due to map projection warping
- Human error and software bugs

6.2 Methodology for determination of error and uncertainty

We determine error by using the upstream error provided by the input data products, considering the relative scale of those errors on the final product, considering whether or not errors are random or systematic, and whether or not our variables are truly independent.

6.3 Error and uncertainty documentation

6.3.1 Thickness

Thickness uncertainty is provided by the BedMachine data set (Morlighem, Williams, et al. 2017a, 2017b).

6.3.2 Velocity

Velocity standard deviation is provided by the PROMICE Sentinel 1 velocity product.

6.3.3 Discharge

Discharge is the combination of thickness and velocity. We assume ice thicknesses < 20 m are incorrect where ice speed is > 100 m yr-1. We adjust these thicknesses using a poor fit (correlation coefficient: 0.3) of the $\log_{10}$ of the ice speed to thickness where the relationship is known (thickness > 20 m). We set errors equal to one half the thickness (i.e. $\sigma_H = \pm 0.5\,H$). We also test the sensitivity of this treatment to simpler treatments, and have the following five categories:

Here we describe how we estimate the uncertainty related to the ice discharge following a simplistic approach. This yields an uncertainty of the total ice discharge of approximately 10 % throughout the time series.

At each pixel we estimate the maximum discharge, $D_{\max}$, from

$$D_{\max} = \rho \, (V + \sigma_V) \, (H + \sigma_H) \, W, \tag{6.1}$$

and minimum discharge, $D_{\min}$, from

$$D_{\min} = \rho \, (V - \sigma_V) \, (H - \sigma_H) \, W, \tag{6.2}$$

where ρ is ice density, V is baseline velocity, σ_V is baseline velocity error, H is ice thickness, σ_H is ice thickness error, and W is the width at each pixel.

We use $\rho = 917$ kg m-3 because the gates are near the terminus in the ablation zone and ice thickness estimates should not include snow or firn, although regionally ice density may be < 917 kg m-3 due to crevasses. We ignore the velocity error σ_V because the proportional thickness error (σ_H/H) is an order of magnitude larger than the proportional velocity error (σ_V/V) yet both contribute linearly to the discharge. W is location-dependent due to the errors between our working map projection (EPSG 3413) and a more

accurate spheroid model of the earth surface. We adjust linear gate width by up to ~4% in the north and ~-2.5% in the south of Greenland (area errors are up to 8%). On a pixel by pixel basis we used the provided thickness uncertainty except where we modified the thickness (H < 20 m) we prescribe an uncertainty of 0.5 times the adjusted thickness. Subsequently, the uncertainty on individual glacier-, sector-, region-, or ice sheet scale is obtained by summing, but not reducing by the square of the sums, the uncertainty related to each pixel. We are conservative with our thickness error estimates, by assuming the uncertainty range is from $D_{\min}$ to $D_{\max}$ and not reducing by the sum-of-squares of sectors or regions.

6.3.4 Map projection

Our work takes place in a projected coordinate system (EPSG 3413) and therefore errors are introduced between the "true" earth spheroid (which is itself an approximation) and our projected coordinates system. We address these by calculating the area projection error due to EPSG 3413 which is approximately +8 % in Northern Greenland and -6 % in Southern Greenland, and multiplying variables by a scaling factor if the variables do not already take this into account. Velocities are "true velocities" and not scaled, but the gate width is scaled.

6.4 Guideline for using the product

The primary source of uncertainty is the ice thickness uncertainty, and that is systematic and fixed in time. Therefore, aggregating discharge spatially or temporally will not reduce that uncertainty. Even so, other sources of errors which are random do exist and will be reduced when aggregating results over large areas or times. We therefore suggest using GIS-wide discharge or regional discharge, and using caution if examining individual glaciers. Similarly, annual or monthly averaged or summed results will have a higher signal-to-noise ratio than individual time steps.

6.5 References

Mankoff, Kenneth D., William Colgan, Anne Solgaard, Nanna B. Karlsson, Andreas P. Ahlstrøm, Dirk van As, Jason E. Box, et al. 2019. "Greenland Ice Sheet Solid Ice Discharge from 1986 Through 2017." *Earth System Science Data* 11 (2): 769–86. https://doi.org/10.5194/essd-11-769-2019.

Morlighem, M., C. N. Williams, E. Rignot, L. An, J. E. Arndt, J. L. Bamber, G. Catania, et al. 2017a. "BedMachine V3: Complete Bed Topography and Ocean Bathymetry Mapping of Greenland from Multi-Beam Echo Sounding Combined with Mass Conservation." *Geophysical Research Letters*, September. https://doi.org/10.1002/2017gl074954.

Morlighem, M., C. Williams, E. Rignot, L. An, J. E. Arndt, J. Bamber, G. Catania, et al. 2017b. "IceBridge BedMachine Greenland, Version 3." NASA National Snow; Ice Data Center Distributed Active Archive Center. Boulder, Colorado USA: NASA National Snow; Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/2CIX82HUV88Y.

.

7 End-to-end Uncertainty Budget for SL

7.1 Introduction

Supraglacial lakes (SLs) have been identified in the Ilulissat region, West Greenland, for various periods between 2016-2019. Each SL was detected from freely available Sentinel-2 imagery, specifically for the region defined in Figure 7.1. The region consists of four Sentinel-2 tiles (T22WEB, T22WFB, T22WEC, and T22WFC), covering an area of approximately 40,000 km^2 that stretches from the coastline to the interior of the ice sheet. The region includes the major glacier outlet Jakobshavn Isbræ (also known as its Greenlandic name, Sermeq Kujalleq).
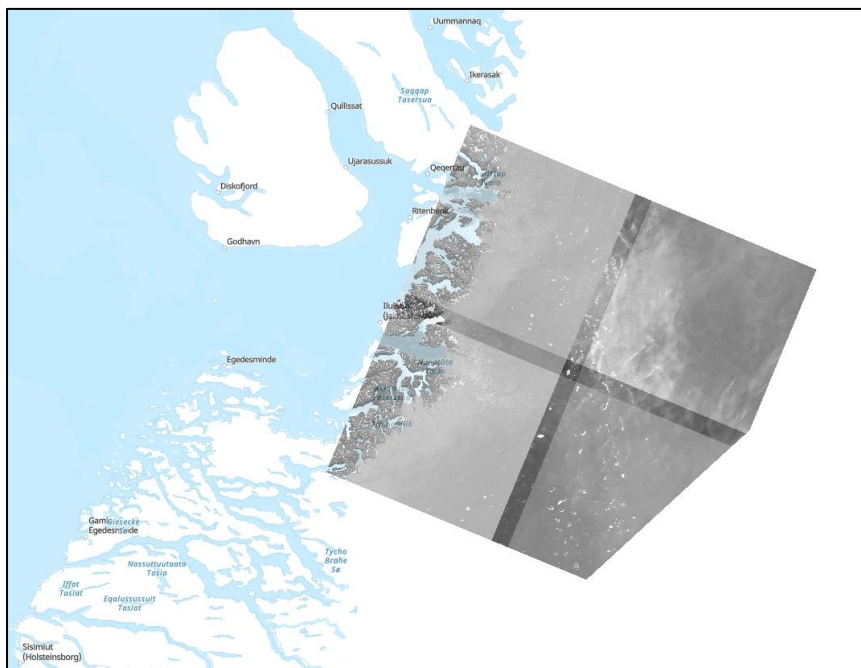

Figure 7.1: The study region for identifying supraglacial lakes, consisting of four Sentinel-2 tiles that cover the Ilulissat region in West Greenland. The example footprints shown are NDWI images, compiled from the green and NIR Sentinel-2 satellite images acquired on 25th August 2019. White regions on the ice sheet denote supraglacial lakes.

SLs were detected using an adaption of the method adopted by Yang and Smith (2013), recently translated into Python 2 (Yang, 2019). We adapted this method by translating it into Python 3, refining it for batch processing through one master script, and adjusting the file handling procedures so that they were in keeping with our standardised file structure system.

The SL detection method largely utilises the ArcGIS python package ArcPy for handling and manipulating raster data; for instance, using the 'ExtractByMask' algorithm for raster masking, and the 'Reclassify' algorithm for generating binary matrices from NDWI images. The workflow of the method can be simplified to the following key steps:

1. Create the NDWI raster files

2. Create the binary rasters, based on upper and lower threshold limits

3. Mask the binary rasters, using a shapefile denoting the ice extent (if provided)

4. Vectorise the binary rasters, only retaining positive matches

5. Filter vectorised polygons to remove false matches in a two-step process:

    a. SLs below a given area threshold (<0.01 sq km) are removed as they are deemed as misclassification, e.g. water filled crevasses and saturated snow.

    b. SLs within a mask denoting sinks in the ice surface are retained (if provided), and those outside the mask are filtered by a second threshold of <0.02 sq km

6. Merge overlapping polygons derived from upper and lower threshold limits

7. Assign information to each SL detected (i.e. area, perimeter, time slice, satellite tile detected from), and combine them from each time slice into individual outputted shapefiles for each unique acquisition

Statistics are then performed on the outputted shapefiles using the Geopandas package available with Python, producing general information such as the total SL area and area change with time, and SL count.

7.2 Sources of error

Sources of error are associated with each step in the workflow outlined in Section 7.1:

Source of error	Step that error is introduced	Approach for limiting error	Potential magnitude of error (0=insignificant, 3=large)
Scene does not cover the whole area	1	Do not use acquisition dates where tile extent is limited.	0
Ice-covered lakes	2	In the workflow, a dual binary threshold which identifies the upper and lower extent of each SL. If a SL is present for both thresholds then the upper extent is retained, which will best compensate for ice-covered lakes.	2
Saturated snow	2	The dual binary threshold identifies the upper and lower extent of each SL. If a SL is present for both thresholds then the SL is retained, reducing the probability of false matches.	1
Lakes containing sediment that changes their spectral signature	2	Similar to delineating ice-covered lakes, the dual binary threshold will limit the chances of misidentifications by only retaining lakes that are detected using the upper and lower threshold.	2
Mask is not up to date for the given time slice	3	Ensure the ice extent mask is up to date for each acquisition. We have ice extent masks for each year (i.e. 2016, 2018, 2019).	0
Vectorising process alters the extent and form of each SL	4	Ensure that the vectorised form purely follows each delineated raster pixel (i.e. un-simplified) rather than smoothing the feature (i.e. simplified).	0
Lake count and area information is not truly representative (e.g. in cases where one lake is represented as two shape features)	5	Lake features are merged if they are overlapping or grouped together if they are within the same sink.	0

As demonstrated, many of these errors were considered during the development of this workflow, and therefore errors introduced at steps 1, 3, 4 and 5 are insignificant. The remaining error, introduced at step 2, will now be the focus of the following sections (7.3 and 7.4).

7.3 Methodology for determination of error and uncertainty

Previous studies have determined the error associated with remotely sensed SL detection through two main approaches:

1. By comparing the number and size of lakes detected from different satellite sensors

Williamson et al. (2018) compared SLs detected from contemporaneous Landsat 8 and Sentinel-2 images (collected within 90 minutes of one another) to assess variations in detected SL areas, and defined error based on the differences in SL area between the two satellite sensors. Generally, there was high correlation between the two sets of SL areas (R^2=0.999), with an RMSE error estimation of 0.007 km^2 (equivalent to 7 Sentinel-2 pixels).

2. By comparing methodologies to assess variability in the number and size of lakes identified

Leeson et al. (2013) compared three separate SL detection methods (from Sundal et al., 2009, Selmes et al., 2011, and Johansson and Brown, 2013) to manual delineations of SLs in the region of Russell Glacier, West Greenland, to constrain error and evaluate the effectiveness of each method. In all, they found that the three methods evaluated generally underestimated the number of SLs, and generate a small number of false positive matches – 9% using the Sundal et al. (2009) method, 5% using the Selmes et al. (2011) method, and 33% using the Johansson and Brown (2013) method. SL areas are, on average, within 0.73 km^2 of those manually derived SLs, which was subsequently used as a measure of accuracy for the three methods.

For this study, we chose to adopt an error determination method similar to those described in the latter of the two approaches outlined above. A comparison of automatically derived and manually derived SLs was divided into two assessments: 1) an assessment of the number of SLs detected; and 2) an assessment in the accuracy of SL areas and forms. However, given the time constraints of the project it would take to do a total SL count and the given time constraints of the project, we largely focused on the second of these two assessments.

7.3.1 Determination of lake count error

To account for lake count error, we decided to adopt the error determination analysis from Yang and Smith (2013), from which the method employed in this study is adopted (Yang, 2019). This is a viable approach given that the method used is the same (merely translated from Matlab to Python for the purpose of this study), and the error determination in the Kang and Smith (2013) study is a suitable approach for this study.

Yang and Smith (2013) determined error in lake count by comparing the algorithm-derived water bodies to a set of manually derived water bodies. Overall, the algorithm captured 85% of the manually derived SLs. This percentage value was adopted for the error and uncertainty documentation as a flat error estimation for all SLs identified in this study. Therefore, all lake count results from this study have an error of ±15%.

7.3.2 Determination of lake area error

To account for lake area error, we chose all SLs from the algorithm-derived dataset for a select tile and time slice (T22WEC, 20/08/2019), and assessed the accuracy of their areas by comparing them to manually derived SLs.

A sensitivity test was performed to evaluate the accuracy of manually derived SLs, conducted in a similar manner to the method described by Leeson et al. (2013). Users delineated a selected SL over 10 iterations, from which a range of areas were computed to form an estimation of error (Table 7.1).

Table 7.1

User	Average area (km^2)	Maximum area	Minimum area	Error (km^2)

		(km²)	(km²)	max area – min area
1	0.1322	0.1399	0.1293	0.0106
2	0.1243	0.1362	0.1176	0.0186
3	0.1164	0.1249	0.1059	0.0190
AVERAGE	**0.1243**			**0.0161**

The average size of the manually delineated SL was 0.1243 km² (by all users). The variation in areas delineated by each user was used as a measure of error. Each user had slight differences in their range of SL areas, between 0.0106 km² and 0.0190 km², with an average of 0.0161 km². Based on the average SL area and the average error of each user, a percentage error of 13.0% was computed which represents the error of manually delineated SLs.

Following this, the accuracy of algorithm-derived SLs was evaluated by comparing their areas to manually delineated SLs. The SLs for this assessment represented a diverse range with an assortment of different sizes, and a selection of open water and ice-covered lakes. Figure 7.2 demonstrates the range of SLs within this tile, with three examples of simple classification (SL 1), minor ice cover and bed sediment causing slight misclassifications (SL 2), and major ice cover causing the most severe misclassification (SL 3).
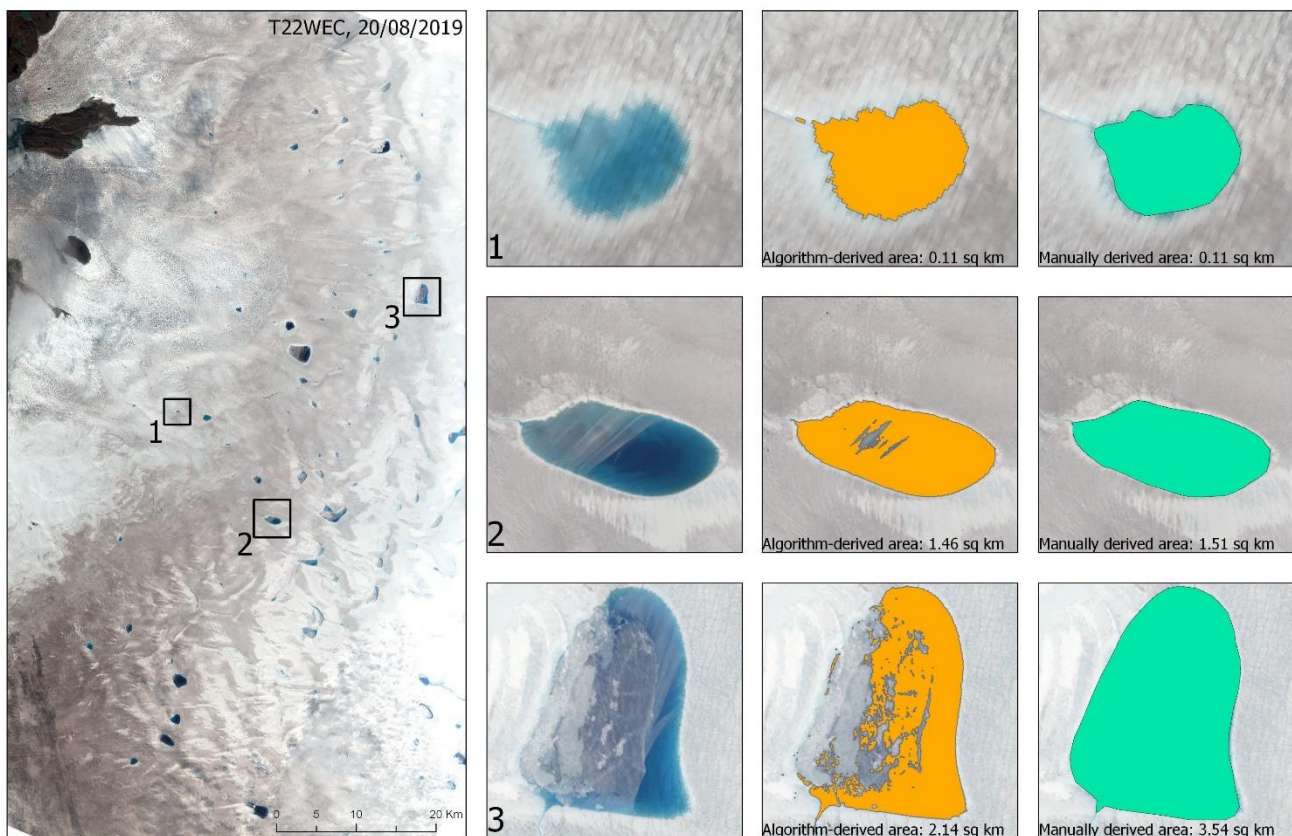

Figure 7.2: Examples of SLs in the area of interest (shown as an RGB composite) that are algorithm-derived (orange) and manually derived (green). SL 1 is a typical small lake in the Ilulissat catchment, for which calculating area algorithmically was the most accurate (i.e. within 5% of the manually derived areas). The algorithm detection of SL 2 exhibits minor problems due to ice-cover and bed sediment, which alter the spectral signature of the lake. This can lead to differences in area of up to 30%. SL 3 is an example of the most severe misclassification, with ice-cover drastically limiting the detection of lake area. In the case of

SL 3, there is a difference of 1.4 km² between the algorithm-derived area and the manually derived area, equating to an error of 40%.

By comparing the algorithm-derived SLs with the manually derived SLs, we computed an average difference of 0.229 km² equating to an average error (i.e. area difference divided by the total area of each SL) of 22%. This error surpasses the error introduced by manual delineations of SLs (13%), and therefore SL area error can be accounted for in this comparison

SL area error is skewed by the most severe misclassifications (as shown in the example of SL 3 in Figure 7.2). Error is significantly reduced in cases where SLs are simple classifications (e.g. SL 1 in Figure 7.2), with an average error of 15% in these instances. For the purpose of this study, where total SL area is used as a measure of SL extent, the average error of 22% will be used as the overall error estimation for SL area.

7.4 Error and uncertainty documentation

Error source	Error/uncertainty estimation	Notes
SL count	±15%	Adopted from Yang and Smith (2013), based on comparison between algorithm-derived SLs and manually derived SLs
SL area	±22%	Based on comparison of algorithm-derived SL areas and manually derived SL areas (as documented in Section 7.3)

7.5 References

Johansson, A. M., Brown, I. A., 2013. Adaptive classification of supraglacial lakes on the West Greenland ice sheet. IEEE J. Select. Topics Appl. Earth Obs. Remote. Sens., 6(4), 1998–2007. doi:10.1109/JSTARS.2012.2233722.

Leeson, A.A., Shephard, A., Sundal, A. V., Johansson, A. M., Selmes, N., Briggs, K., Hogg, A. E., Fettweis, X., 2013. A comparison of supraglacial lake observations derived from MODIS imagery at the western margin of the Greenland Ice Sheet. J. Glaciol., 59 (218), 1179-1188. doi:10.3189/2013JoG13J064.

Selmes, N., Murray, T., James, T. D., 2011. Fast draining lakes on the Greenland ice sheet. Geophys. Res. Lett., 38(15), L15501. doi: 10.1029/2011GL047872.

Sundal, A.V., Shepherd, A., Nienow, P., Hanna, E., Palmer, S., Huybrechts, P., 2009. Evolution of supra-glacial lakes across the Greenland ice sheet. Remote Sens. Environ., 113(10), 2164–2171. doi:10.1016/j.rse.2009.05.018.

Williamson, A. G., Banwell, A. F., Willis, I. C., Arnold, N. S., 2018. Dual-satellite (Sentinel-2 and Landsat 8) remote sensing of supraglacial lakes in Greenland. Cryosphere, 12, 3045-3065. doi:10.5194/tc-3045-2018.

Yang, K., 2019. Supraglacial river and lake analysis [software]. figshare. doi:10.6084/m9.figshare.9758051.v1.

Yang, K., Smith, L., 2013. Supraglacial streams on the Greenland Ice Sheet delineated from combined spectral-shape information in high-resolution satellite imagery. IEEE Geosci. Remote Sens. Lett., 10 (4), 801- 805.

8 End-to-end Uncertainty Budget for CFL

8.1 Introduction

In this section the error sources, uncertainties and methodology for characterisation of the errors of the Sentinel-2 based deep learning-based Calving Front Location (CFL) detection will be outlined.

The CFL localization algorithm has the following main steps to determine the front locations:

1. Semantic segmentation of the image into ice, land and sea classes.
2. Noise removal (small, misclassified regions and morphological smoothing).
3. End point localization (where ice, land and sea meet).
4. Shortest cost path detection.

8.2 Sources of error

The main sources of error for CFLs detection based on optical imagery are:

- Misclassification of regions primarily due to:
  - o Lack of textures (e.g., fresh snow cover over sea ice and land).
  - o Inaccuracies in deep learning-based segmentation.
  - o Presence of clouds.
- Noise removal modifying the front locations through smoothing.
- End point localization requires false positive suppression.
- Shortest path algorithm assumes minimal cost at the front line but that depends on the segmentation quality.

In the following, errors will be interpreted as line-to-line distances between predicted and real calving front locations. Segmentation error is an incorrectly classified pixel group. If these pixels are on the front line, then they also contribute to the line-to-line errors.

The sources of errors listed above lead to:

$$\delta(x,y) = \delta_{segmentation} + \delta_{endpoint\ localization} + \delta_{smoothing} + \delta_{shortest\ path} + \delta_{geolocation} + \delta_{time\ difference}$$

8.2.1 Segmentation error

Determining bounds for deep learning-based segmentation algorithms is still an open research topic. This segmentation error is the primary source of all the errors because all further processing steps rely on it. The segmentation assigns a single class to each pixel which has a 10m resolution. Subpixel localisation is impossible with this technique, which yields up to 5m error for discretization error.

Ambiguity: in the presence of clouds or if ice melange texture is indistinguishable from the ice sheet texture, then the deep learning segmentation might not be able to find the calving front location. This means that the algorithm doesn't detect the front line in the field of view.

The segmentation error can only be empirically determined, but it is expected that cloud cover, fresh snow and dirt on ice sheet worsen the predictability of the front lines as they lead to class confusion.

8.2.2 Noise Removal

The noise removal algorithm has two steps: first, the pixel probability predictions are converted to connected regions by selecting the most likely class. Floating ice and supraglacial lakes could be classified as ice and sea respectively, however, they should be classified as sea and ice, respectively. Objects below

threshold size are removed. These are only contributing to the front line endpoint localization algorithm, and should not contribute to front localization error, if the segmentation did not break down due to model ambiguity.

The noise removal algorithm also applies morphological opening and closing with 70m radius, which removes smaller than 70m structures from the front line.

8.2.3 Shortest path finding

The shortest path algorithm's primary aim is to prevent meandering structures to appear in the front lines. Such structures are unlikely to exist, but if they are present, then the algorithm might take a shortcut between two points. The algorithm developed during the project removes at least 1 pixel at sharp corners, leading up to 10m line to line error, and maximum as large as the structure itself. The most likely side effect is to ignore cracks deep into the ice sheet.

8.2.4 Geolocalization and time difference errors

Geolocalization and time difference errors are virtual errors originating from the possible misalignment of the validation imagery and possible time difference between the product and the validation measurement and the error of the algorithm producing the validation data. The effect of the time difference can be approximated by the product of the ice velocity and time difference if calving event hasn't occurred.

8.3 Methodology for determination of error and uncertainty

Due to the strong reliance of the segmentation algorithm which has an analytically non-quantifiable uncertainty, the errors are statistically determined during validation. This takes into account all the algorithmic errors but cannot take into account the uncertainty of the validation data. However, this validation data uncertainty is assumed to be insignificant compared to the calving front localization error. For more information see the PVIR [RD5].
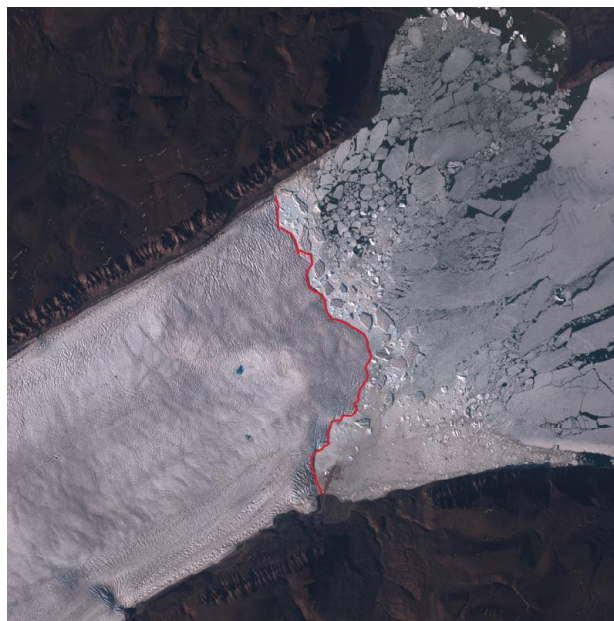

Figure 1-3: Example CFL of the Hagen glacier

8.4 Error and uncertainty documentation

Unfortunately, it is impossible to determine the error for a given input image in a non-empirical way. The product's metadata contains information on the empirical error estimation based on the PVIR [RD5].

This validation found that for the calving front locations have a maximum mean error of 80m, with local maximum error of up to 600m.

8.5 Guideline for using the product

It is expected that most estimated errors will be in the range of 80m.

It should be kept in mind that the employed automated front line extraction technique tends to degrade when not applied to features rich areas, in presence of clouds, dense ice melange and fresh snow.

For further guidelines refer to ATBD [RD3] and PUG [RD4].

End of document