ESA Climate Change Initiative (CCI)

Greenland Ice Sheet (GIS) Essential Climate Variable (ECV)

System Specification Document (SSD)

Science &Project 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	Daniele Fantin, S&T Lead Author	Digitally signed by Daniele S. M. Fantin Date: 2021.07.09 10:28:33 CEST	Date: 6 July 2021
Issued by	Daniele Fantin Project Manager, S&T		Date: 6 July 2021
Checked by	René Forsberg Science Leader, DTU-S	R. Forsberg	Date: 6 July 2021
Approved by	Marcus Engdahl ESA Technical Officer		

Table of Contents

Change Log

Issue	**Author**	**Affected Section**	**Reason**	**Status**
1.0	S&T	All	Draft version	
1.1	GEUS	5	Updated MFID	
1.2	ENVEO, S&T	3	Updated IV	Reseased to ESA on 6 July 2021

Acronyms

Acronym	Explanation
ASIAQ	Asiaq Greenland Survey
AWS	Amazon Web Services
CCI	Climate Change Initiative
DMI	Danish Meteorological Institute
DTU-N	DTU-Space, Department of Microwaves and Remote Sensing
DTU-S	DTU-Space, Department of Geodynamics
ECV	Essential Climate Variable
ENVEO	ENVironmental Earth Observation GmbH
ESA	European Space Agency
ESP	ENVEO SAR Software Package
GEUS	Geological Survey of Denmark and Greenland
GIS	Greenland Ice Sheet
GIS_cci	Greenland_Ice_Sheet_cci
GMB	Gravimetry Mass Balance
GPS	Global Positioning System
IV	Ice Velocity
IW	Interferometric Wide
MFID	Mass Flow Rate and Ice Discharge
NBI	Niels Bohr Institute, University of Copenhagen
NDWI	Normalised Difference Water Index
OPT	Optical
PSTG	Polar Space Task Group
S&T	Science and Technology AS
SAR	Synthetic Aperture Radar
SEC	Surface Elevation Change
SL	Supraglacial Lake
SLC	Single Look Complex
SSD	System Specification Document
TOPS	Terrain Observation with Progressive Scans
TUDr	Technische Universität Dresden
UL	University of Leeds

1 Introduction

1.1 Purpose and Scope

This document contains the System Specification for the Greenland_Ice_Sheet_cci (GIS_cci) project for CCI+ Phase 1, in accordance to contract and SoW [AD1 and AD2].

This document is part of Task 3 Systems Evolution within the GIS_cci project, as part of ESA Climate Change Initiative (CCI) program.

This is the System Specification Document (SSD), which specifies the characteristics of the ESA GIS_cci ECV processing system.

This SSD document builds on the Phase 2 SSD [RD1] of the 'Greenland_Ice_Sheet_cci' project.

1.2 Document Structure

This document is structured into an introductory chapter followed by five chapters describing the processing system for the CCI+ ECV parameters, which are:

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

Each ECV chapter provides:

- A specification of the purpose of an operational ECV production system and its intended use.
- An overview of the context of the system, defining all significant interfaces among system components and crossing the system's boundaries.
- A definition of the fundamental operations to be performed within the system to accept and process the inputs and to process and generate the outputs.
- A description of major constraints of the system.
- A description of operational scenarios for the system including data sources, valid ranges of values, timing considerations, operator requirements, and special interfaces.
- Specification of the environmental characteristics of where the system will be installed.
- Specification of the growth, expansion, and capability characteristics of the system.
- Description of the life cycle sustainment activities to be executed during the life cycle of the system.

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, and its Appendix 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 Statement of Work	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-SSD-001	System Specification Document	2017.11.08	1.1

Note: If not provided, the reference applies to the latest released Issue/Revision/Version

2 Surface Elevation Change (SEC)

2.1 System overview

The SEC operational ECV production system takes as input all available elevation data from the Level-2 products of radar altimetry missions since 1991 and processes these to produce 5-year (and for CS2 2-year) mean rate of surface elevation change products over the Greenland ice sheets. Input data for the SEC production system consist of multi-mission radar altimetry data (Table 2.1), auxiliary correction data, ice sheet DEM and ice mask. The SEC system is described in detail in the DPM and the GIS CCI project and any updates are also reported in the current ATBD.

Table 2-1: Current Radar Altimetry Data Sets

Mission	Product Level	Baseline	Source	Operational Dates
CCI+ GIS				
ERS-1	L2	REAPER v1.0	ESA	02-Aug-91 to 02-Jun-96
ERS-2	L2	REAPER v1.0	ESA	13-May-95 to 04-Jul-03
ENVISAT	L2	GDR v2.1	ESA	09-Apr-02 to 18-Oct-10
CryoSat-2	L2i SIN	Baseline-C	ESA	16-Jul-10 to 28-Apr-19
CryoSat-2	L2i LRM	Baseline-C	ESA	16-Jul-10 to 28-Apr-19
Sentinel-3B	L2	Baseline-003	ESA	01-Mar-16 to present
Sentinel-3B	L2	Baseline-003	ESA	01-May-18 to present

2.2 Operational scenarios

The SEC processor works in four distinct modes: True repeat track, along track, cross-over and Plane-fit. The choice of mode depends on the data used as input. We refer here to the ATBD for further information.

Table 2-2: User-defined input for the SEC system

User input	Format	Note
START TIME	ddmmyyyy	Start date of the chosen SEC period
END DATE	ddmmyyyy	End date of the chosen SEC period
GRID SIZE	X [km]	Grid cell size of final SEC grid
CORR. LENGTH	Y [km]	Correlation length of dh/dt data. This controls smoothness of SEC grid
MIN #DATA	Z [number]	Number of data points required in each grid cell/segment.
DEM	Name of DEM	Choice of DEM to use in the relocation
Model	#	Defines which model to solve for in the least squares regression
Flags	binaries	1 for use this parameter. 0 do not use this parameter. Number of flags relevant depends on mission data.

2.3 Hardware and software platform

The SEC operational ECV production system runs on most UNIX servers with adequate memory, and disk storage space. The current SEC processing system is developed and test in Mac-OS, and is transferred for operational-runs to a server system running ax86_64 GNU/Linux UNIX operating system.

2.3.1.1 Hardware

Table 2-3: Processing Hardware at for the SEC processor

Model	Supermicro 848A-R1K62B
Processor	4 X AMD Opteron 6376 (64 cores), 64 bit
Memory (RAM)	256 GB
Local Hard Drive	30TB
Network Attached Storage	>100TB

2.3.1.2 Operating system

The system currently runs on the UNIX operating system x86_64 GNU/Linux. The system has been tested and runs on other Unix/Linux systems as well, including Mac OS X.

2.3.1.3 Tools and libraries

The SEC processor is written in a mixture of different languages, which includes Fortran, C++, Python and R. The main operational kernel is written in Python, whereas R is used for Kriging.

Python 3.5 packages: numpy, tables, scipy, matplotlib, os, sys, multiprocessing, shapefile, netCDF4.

R packages: sp, gstat, ncdf4, mapplots, rgdal, scales, ggplot2.

2.4 Future concerns and developments

We are currently working on porting all code to Python version 3.5 and R version 3.6.0. Some parts of the SEC processor were previously written in C++. We expect no degradation in functionality of the SEC processor caused by the change to Python.

3 Ice Velocity

3.1 SAR Ice Velocity (SAR IV)

3.1.1 System overview

In this section we describe the processing system for deriving ice velocity maps from repeat pass SAR data applying both offset tracking and SAR interferometry (InSAR) techniques. The ENVEO SAR Software Package (ESP) Version 2.1 is used for generation of ice velocity maps. ESP 2.1 applies incoherent and coherent offset tracking for mapping displacements in range and azimuth on repeat pass SAR SLC data. It also measures line-of-sight displacements with InSAR and combines these to derive ice velocity from crossing-orbit measurements. The IV processor supports the common space-borne SAR sensors including Sentinel-1 IW, Radarsat-2, ERS 1/2, ENVISAT ASAR, ALOS PALSAR, TerraSAR-X, and TanDEM-X. Auxiliary data needed are a DEM and (optional) ocean/ice/land mask. For interferometry, an external a-priori ice velocity map can also be used for coregistration and calibration. Figure 3-1 shows the high-level processing line for the IV production. It includes the following three modules:

- MODULE IV: within this module SAR data are imported into the system and velocity maps are generated for pairs of repeat pass SAR data of the same track. This module corresponds either to offset tracking or InSAR processing. The output are IV maps in map projection as defined by the user. Each IV product includes a quality map.
- MODULE MERGE: this module combines the IV products from various tracks and image pairs. The output is a merged large/regional scale ice velocity map and a quality map.
- MODULE VAL: the validation module is an associated module, which enables the inter-comparison and validation of the generated ice velocity map with in-situ velocity measurements. The output is statistical information on the inter-comparison and scatterplots compiled in a quality assessment report.

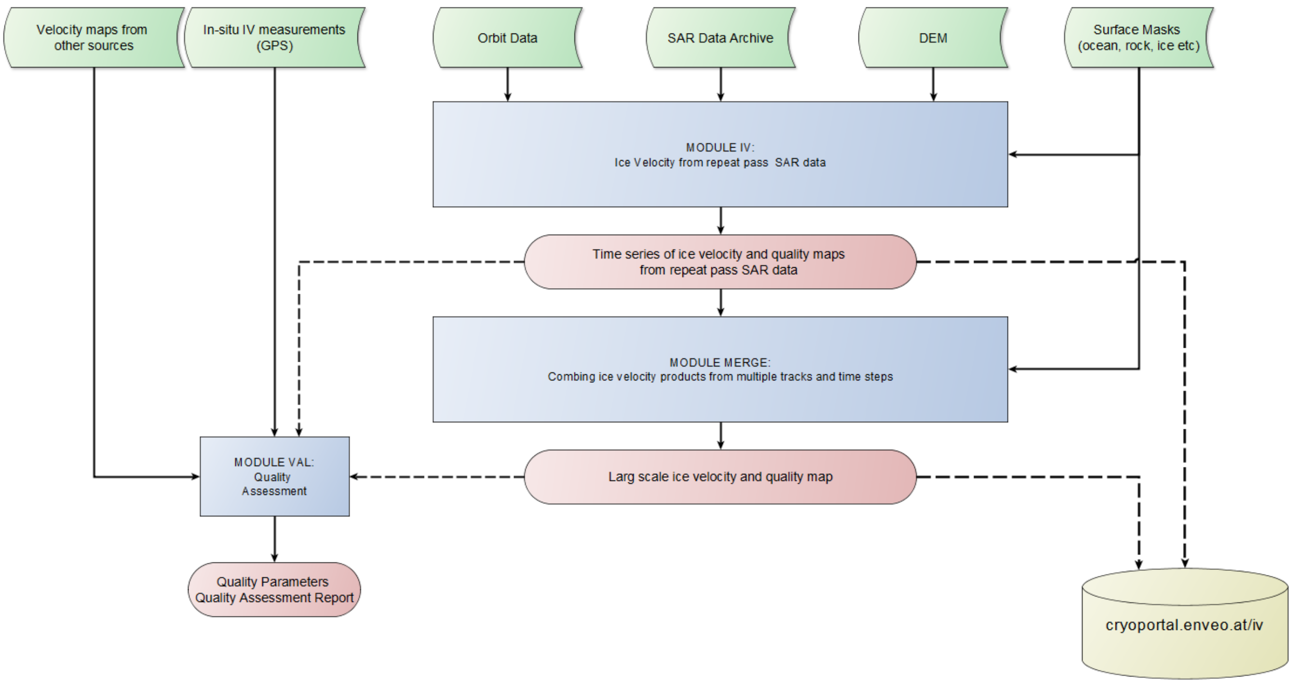

Figure 3-1: High-level flow chart of the IV processing system. Green – input data, Blue – processing modules, red - product and intermediate products.

MODULE IV

Offset-tracking

Figure 3-2 shows a high-level flow line of the ice velocity generation module with offset tracking. Regarding Sentinel-1 Interferometric Wide (IW) Single Look Complex (SLC) data the processing is done on the burst level. In the case of Stripmap mode data the processing is done on scene / frame level, which is practically very similar. Here we describe the Sentinel-1 IW SLC burst processing.

The module has access to the SAR data archive and orbit data, as well as the DEM and optionally the surface masks. The processing is done track by track, for each track image bursts are selected according to the time step. For Sentinel-1 processing we select the two shortest repeat pass periods (either 6, 12 or 24 days). Using image geometry and a digital elevation model the local shift between two bursts is calculated which is considered in the displacement calculation using incoherent (or coherent) offset tracking. Beside the displacements in slant range and azimuth the quality of the matching is calculated. Debursting of all processed burst pairs forms the displacement map in SLC geometry. After outlier removal the velocity of the 3 components is calculated using a DEM as input. Optionally, the surface masks are applied to mask ocean and un-glaciated areas.

The output of this module are velocity maps (E, N, ΔZ) in m/d per track and time step, and the quality map.

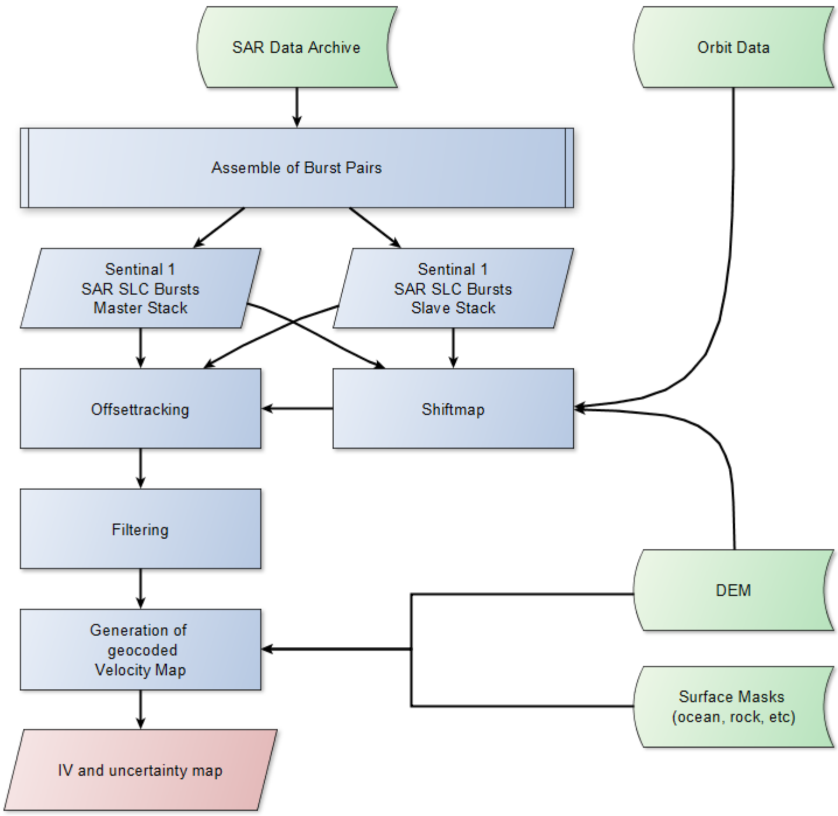

Figure 3-2: Processing steps for ice velocity generation using offset-tracking applied to Sentinel-1 IW TOPS SLC data.

InSAR

Figure 3-3 describes the flowline of InSAR processing for ice velocity. This processing line is dedicated to Sentinel-1 acquisitions in TOPS mode. Similarly, the offset-tracking module, the InSAR processing on Sentinel-1 data is performed at the burst level up to the phase unwrapping step. Burst-level processing includes co-registration, interferogram formation, flat-earth and topographic phase removal and coherence estimation. If working with Stripmap acquisitions instead, these processing steps are directly applied to the whole scene and the processing line is adapted by simply skipping the debursting step.

Burst interferograms are mosaicked at the debursting step to make a single large scale interferogram for each master-slave pair. Phase unwrapping and phase calibration are performed at the debursted level. Once calibrated, the phase can be converted into line-of-sight (LOS) velocity and are then geocoded.

The InSAR module uses the SAR data archive in SLC format as input. For Sentinel-1 data, only data with 6-day temporal baseline are used to avoid aliasing and decorrelation caused by fast ice motion or changing surface conditions due to melt or snowfall. The InSAR module uses precise orbits provided by ESA ~20 days after the acquisition. A digital elevation model must also be given as input to the module for co-registration, estimating the topographic phase and geocoding.

Optionally, an external a-priori ice velocity map (from e.g., offset tracking) can be inserted in the module. It is then used for the co-registration and calibration steps: at the co-registration step, it is used to update the co-registration map with local displacements; at the calibration step, it is used for selecting slow-motion ground control points which are used as reference values.

The output of this module are geocoded maps of LOS velocity. These maps are produced for 6-day and 12-day repeat pass data from all available tracks. In the MERGE module, corresponding to the inversion step, these images are used to derive a velocity map. The inversion step combines LOS velocities from crossing orbits to calculate the ice flow velocity components. The merged output is a 3-D velocity map (X, Y, Z) in meters per day.

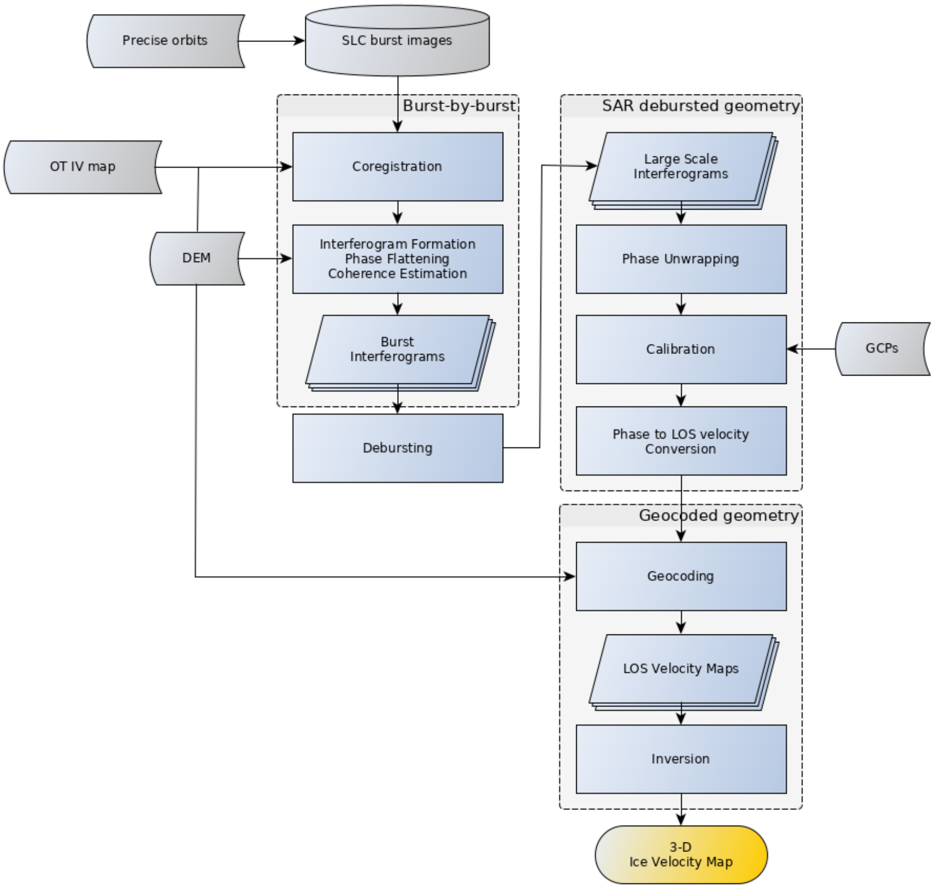

Figure 3-3: Processing steps for ice velocity generation using InSAR applied to Sentinel-1 IW TOPS SLC data. SLC – Single-Look Complex. OT – Offset Tracking. LOS – Line-of-sight. GCP – Ground Control Point.

MODULE MERGE

This module aims to produce a large/regional scale ice velocity map using the output of the MODULE IV as input. The generation of a regional ice velocity map requires the combination of IV results, from several tracks. The long tracks are geocoded into a common map projection of the output grid, using the annotated imaging and orbit parameters and a digital elevation model as input. The merging combines the velocity components separately in m/d. The output of this module is a large scale, regional or ice sheet wide flow velocity and quality map.

MODULE VAL

The validation module aims for an independent validation of the ice velocity product using in-situ ice velocity data, as provided by GPS. The module inter-compares the velocity components separately but can be also applied the velocity magnitude. Currently, the module requires manual interaction, mainly in the selection and pre-processing of the in-situ ice velocity data. Output are statistical parameters of the inter-comparison and scatter plots for visualisation.

This module allows also to inter-compare ice velocity products from other sources. The module takes different map projections into account and provides statistical parameters of the pixel-by-pixel inter-comparison, histograms, and spatial maps of the differences.

MODULE IV-DATABASE

The ice velocity products are linked to the IV database at cryoportal.enveo.at. The web-based portal allows easy access and simple analysis of all ice velocity products (single maps, monthly & annual). The current versions support the following analysis:

- Ice velocity visualization and download of data:
  - Visualize the full time series of ice velocity profiles along pre-defined central flow lines of all major outlet glaciers.
  - Visualize the full time series of velocity on fixed points along the flow line.
- Mass flux calculation:
  - For single glaciers
  - Sub basins
  - Basins
  - Ice sheet wide

3.1.2 Operational scenarios

Operational observations are mainly planned using Copernicus Sentinel-1 A/B satellites. There are generally four processing scenarios applied:

- Continuous observations of margins and the ice sheet interior (6 or 12-day repeat) with offset-tracking.
- Production of winter campaigns ice velocity map of the ice sheet interior (6 -day repeat) with InSAR.
- Production of annual ice sheet wide ice velocity maps (also within C3S).
- Production of monthly ice sheet wide ice velocity maps.

The track-by-track offset-tracking processing is applied on regions were continuous acquisitions of S1 IW SAR data are acquired. This processing scenario enables monitoring of ice velocity variations with short time intervals, e.g., 6 days in the case of Sentinel-1A and 1B. This observation scenario has been proposed by ENVEO and implemented by ESA for the Greenland Ice Sheet margins with continuous acquisitions Sentinel-1A and 1B since June 2015 (Figure 3-4).

A similar observation scenario is planned for track-by-track InSAR processing. InSAR processing will rather focus on mapping the ice sheet interior, where ice motion is slower, and conditions are stable from one acquisition to the other. For this purpose, Sentinel-1A and 1B data with 6-day time interval acquired during the winter campaign will be used. Data in wintertime are expected to have better coherence and the coverage of overlapping ascending and descending tracks over the ice sheet interior is improved during the winter campaign.

Ice sheet wide velocity maps for Greenland are generated on a monthly and annual basis and rely on the acquisition campaigns defined by ENVEO in agreements with CCI project partners, ESA and members of the PSTG group. They are performed additionally to the continuous acquisitions of S1.

Currently annual acquisition campaigns for Greenland Ice sheet are planned, with 4-6 repeat acquisitions for each track, in the period December to March (Greenland). In order to perform a complete map, ENVEO operates IV processing within 10 days after image acquisition and provides ESA feedback on further needed acquisitions to close gaps in the IV product with low coherence.

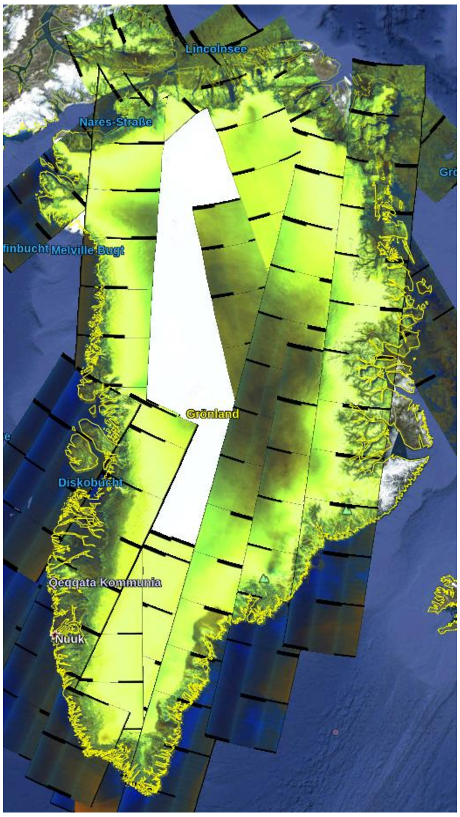

Figure 3-4: S1 tracks for continuous monitoring of Greenland ice velocity (Status: October 2019).

3.1.3 Hardware and software platform

The main processing at ENVEO is performed on the processing system described below.

3.1.3.1 Hardware

The main processing is done on 3 server machines and 18 virtual machines at the IKB cluster, which are connected to a mass storage of about 800 TB. The system applies OPENMP to support multiple CPUs and Cores.

Table 3-1: Processing Hardware for the ENVEO SAR IV processor.

	Processing	**Development, Visualization, Quality Control**
Model	GNU/Linux Centos 7	GNU/Linux Fedora
Number of machines	3	8
Processor	Intel Xeon CPU E5-2650 v2 @ 2.60GHz 16 cores	Intel Core i7-2600 CPU @ 3.40GHz
Memory (RAM)	128 GB	16 GB
Local Hard Drive	300 GB	1 TB
Network	Ethernet 10000baseT/Full	Ethernet 1000baseT/Full
Network Attached Storage	Ca 800 TB network storage	Ca 800 TB network storage

	Processing	
Model	GNU/Linux Centos 7	
Number of nodes	18	
Processor	Intel Xeon CPU-X5670 2.93GHz 2 cores	
Memory (RAM)	10 GB	
Local Hard Drive	-	
Network	Ethernet 10000baseT/Full	
Network Attached Storage	Ca 800 TB network storage	

3.1.3.2 Operating system

The ESP-IV processing system runs on common Linux operating systems. Currently the tested systems are CentOS release 6/7, Fedora 30 or later releases, but the software will also work on other common Linux/Unix systems.

3.1.3.3 Tools and libraries

Tools & Libraries

- GCC / OpenMP http://gcc.gnu.org
- cmake http://cmake.org
- PROJ.4 https://trac.osgeo.org/proj/
- Python (numpy, scipy, etc) https://www.python.org/
- GDAL (latest version) http://www.gdal.org/
- FFTW 3 http://www.fftw.org/
- Libxml http://www.xmlsoft.org/
- NetCDF, HDF http://www.unidata.ucar.edu/
- wget https://www.gnu.org/software/wget/
- gzip http://www.gzip.org/

Visualisation Tools

- QGIS http://qgis.org
- Cryoportal Ice Velocity (developed by ENVEO) https://cryoportal.enveo.at

3.1.4 Future concerns and developments

Further improvements of the software are planned. The ESP software has been connected and tested with Cluster systems utilizing several hundreds of Cores. This is especially of interest for campaign processing of big data sets as it occurs for polar ice sheets.

3.2 Optical Ice Velocity (Opt IV)

3.2.1 System overview

In this section we describe the processing system for deriving ice velocity maps from pairs of Sentinel-2 optical images applying an offset tracking technique. The optical ice velocity processor developed by S&T consists of 6 modules, which are listed in Table 3-2. To run these modules, we use an automated pipeline processing manager called Dagger developed by S&T, which strings together the IV modules, manages the data streams, and handles instantiation of the processing steps. The high-level processing line for the IV production is shown in Figure 3-5.

Table 3-2: Name and purpose of each module composing the optical IV processor

Module name	Purpose
Product download	Queries and downloads S2 products within a given AOI and time range.
Cloud masking	Creates a cloud mask of an S2 product.
IV_pair	Finds all available pairs of S2 products between 7 and 15 days apart within the queried time range.
IV	Creates an IV map from a given S2 pair.
IV_merge	Creates a single stacked or averaged IV product of all the IV maps. Mosaicking tiles if needed.
IV_release	Prepares the final IV product in a zip file with proper description, quick look etc.

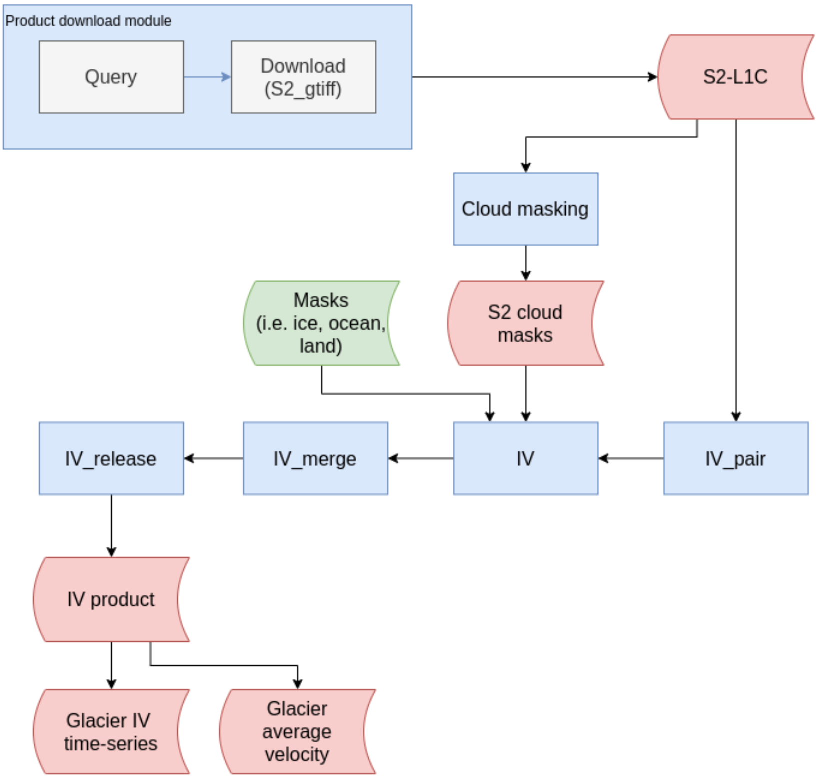

Figure 3-5: high level flow chart of the optical IV processor. Green identifies input data, blue processing modules, grey specific processing tasks within a module and red are intermediate and final products.

3.2.2 Product download module

The product download module takes as required input an AOI and a time range, from which it queries Sentinel-2 L1C images and downloads to a local folder. Additional criteria such as cloud coverage can be specified. The S2 product will be converted to the GeoTIFF format.

Cloud masking

At the cloud masking step, we will create a cloud mask for each downloaded S2 L1C product. These will then be utilized in the IV pipeline.

IV pair

All possible pairs of the retrieved S2 images that are at a minimum of seven and a maximum of 15 days apart, will be found. These pairs will then be passed on.

IV

For each pair of S2 images, we will apply a feature tracking algorithm to calculate the north and east ice velocity. Furthermore, the co-registration offset will be measured, and the root-mean-squared(rms) will be extracted from the ice velocities.

IV Merge

The IV maps generated in the previous step will be averaged, stacked and mosaiced together in the case where the AOI covers multiple tiles. Both a time series and an averaged IV map will be created.

IV Release

The final output will consist of the IV time series and the averaged IV map. These will be delivered in a zip file and NetCDF products together with a quick look image, comments, and product description as the CCI requirements have specified.

3.2.3 Operational scenarios

The operation scenario relies on the availability of Copernicus Sentinel-2 A/B satellites acquisitions. Firstly, the observations for 9 pre-selected glaciers are retrieved based on temporal and spatial criteria through a network interface (e.g., SciHub or AWS) and stored locally. The production of the seasonal IV map is planned at the end of the summer season, after the end of S2 satellites observational season over Greenland region.

Once the observational season is over, glacier based IV maps are generated according to two processing scenarios:

- Production of IV time-series using observations with 6-12 days repeat.
- Production of seasonal IV map.

The scenario has been proposed by S&T and approved by ESA.

3.2.4 Hardware and software platform

3.2.4.1 Hardware

Model	Custom-built
Processor	Intel Core i7 CPU @ 3.40GHz
Memory (RAM)	16 GB
Local Hard Drive	1 TB
Network Attached Storage	1x30TB

Figure 3-6: Processing Hardware for the IV processor

3.2.4.2 Operating system

The processor operates on any Linux system.

3.2.4.3 Tools and libraries

The Opt-IV processor requires Python 3.8 or later for running support scripts.

Linux Libraries required: gdal

Python Libraries required(in addition to the Python standard library): click, h5py, loguru, netCDF4, numba, numpy, opencv, osr, pyproj, pytest, scipy, yaml, zipping,

3.2.5 Future concerns and developments

During the first 2 years of CCI+ project, the following improvements of the processor have been performed:

- Improvement of co-registration between input images.
- Improvement of IV_merge (mosaicking) module.

- Improve of masking module for removing artefacts, including interpolation using a distance threshold.
- Improvement of an independent module for validation and error analysis.

Further improvements of the software are currently ongoing within CCN1.

4 Gravimetric Mass Balance (GMB)

GMB with GRACE+FO using the same system as in the SSD from CCI. Refer to [RD2] for full description.

5 Mass Flux and Ice Discharge (MFID)

5.1 System overview

The flow rate and ice discharge algorithms have two primary steps

- Calculation of discharge gate locations.
- Calculation of discharge through the gates.

Gate locations are calculated once using a baseline velocity product and then remain fixed in space and time. Discharge is then calculated for each time when a velocity map exists.

Note - some of the text below comes from Mankoff et al. (2019) where this algorithm was first described.

Gates are algorithmically generated for fast-flowing ice (greater than 150 m yr-1) close to the ice sheet terminus determined by the baseline-period data. We apply a 2D inclusive mask to the baseline data for all ice flowing faster than 150 m yr-1. We then select the mask edge where it is near the BedMachine ice mask (not including ice shelves), which effectively provides grounding line termini. We buffer the termini 10 km in all directions creating ovals around the termini and once again down-select to fast-flowing ice pixels. This procedure results in gates 10 km upstream from the baseline terminus that bisect the baseline fast-flowing ice. We manually mask some land- or lake-terminating glaciers which are initially selected by the algorithm due to fast flow and mask issues.

We select a 150 m yr-1 speed cut-off because slower ice, taking longer to reach the terminus, is more influenced by SMB. The choice of a 10 km buffer follows from the fact that it is near-terminus and thus avoids the need for (minor) SMB corrections downstream, yet is not too close to the terminus where discharge results are sensitive to the choice of distance-to-terminus value (Mankoff et al. 2019), which may be indicative of bed (ice thickness) errors.

Discharge is calculated at each gate by multiplying the provided velocity by the thickness of the ice at the gate location.

5.2 Operational scenarios

The operational scenario for this product is to generate an estimate of total Greenland ice discharge for every timestep when any velocity product exits. If a velocity product does not cover all gates or all portions of all gates, velocity is linearly interpolated from the surrounding times, or forward- or backward-filled for the last and first time, respectively.

5.3 Hardware and software platform

5.3.1 Hardware

The development environment is a common laptop. It currently has 32 GB of RAM and 5 TB of internal storage, but can be run on a significantly smaller system. We suggest at least 8 GB of RAM and 1 TB of storage.

5.3.2 Operating system

We use the Linux operating system, specifically Ubuntu 18.04.3 LTS, with the following CPU and architecture:

```
Architecture:        x86_64
CPU op-mode(s):      32-bit, 64-bit
Byte Order:          Little Endian
```

```
CPU(s):                8
On-line CPU(s) list:   0-7
Thread(s) per core:    2
Core(s) per socket:    4
Socket(s):             1
NUMA node(s):          1
Vendor ID:             GenuineIntel
CPU family:            6
Model:                 142
Model name:            Intel(R) Core(TM) i7-8650U CPU @ 1.90GHz
Stepping:              10
CPU MHz:               758.158
CPU max MHz:           4200.0000
CPU min MHz:           400.0000
BogoMIPS:              4224.00
Virtualization:        VT-x
L1d cache:             32K
L1i cache:             32K
L2 cache:              256K
L3 cache:              8192K
NUMA node0 CPU(s):     0-7
```

5.3.3 Tools and libraries

We use the following tools for this workflow:

This work was performed using only open-source software, primarily GRASS GIS (Neteler et al. 2012) and Python (Van Rossum and Drake Jr 1995), in particular the Jupyter (Kluyver et al. 2016), pandas (McKinney 2010), numpy (Oliphant 2006), statsmodel (Seabold and Perktold 2010), x-array (Hoyer and Hamman 2017), and Matplotlib (Hunter 2007) packages. The entire work was performed in Emacs (Stallman 1981) using Org Mode (Schulte et al. 2012). The parallel (Tange 2011) tool was used to speed up processing. We used proj4 (PROJ contributors 2018) to compute the errors in the EPSG 3413 projection. All code used in this work is available in the Supplemental Material of Mankoff et al (2019).

5.3.4 Future concerns and developments

Development of this work is ongoing at http://github.com/mankoff/ice_discharge where we use additional IV products that support time series back to 1986, additional SEC products that cover the portions of the ice sheet where the gates are located, and other modifications.

Concerns for the future of this work are mitigated by the fact that the entire process is documented, and the code is available for anyone to replicate and improve the product. However, replicating the work has proven non-trivial so far.

Nonetheless concerns include the quality of the ice thickness data, and the ability to generate SEC maps at regular frequency and with enough spatial coverage to cover the gate locations.

5.4 References

Hoyer, Stephan, and Joseph J. Hamman. 2017. "Xarray: N-D Labeled Arrays and Datasets in Python." *Journal of Open Research Software* 5 (April). https://doi.org/10.5334/jors.148.

Hunter, John D. 2007. "Matplotlib: A 2D graphics environment." *Computing in Science & Engineering* 9 (3): 90–95.

Kluyver, Thomas, Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, et al. 2016. "Jupyter Notebooks – a Publishing Format for Reproducible Computational Workflows." Edited by F. Loizides and B. Schmidt. IOS Press.

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.

McKinney, Wes. 2010. "Data Structures for Statistical Computing in Python." In *Proceedings of the 9th Python in Science Conference*, edited by Stéfan van der Walt and Jarrod Millman, 51–56.

Neteler, M., M. H. Bowman, M. Landa, and M. Metz. 2012. "GRASS GIS: A Multi-Purpose Open Source GIS." *Environmental Modelling & Software* 31: 124–30. https://doi.org/10.1016/j.envsoft.2011.11.014.

Oliphant, Travis E. 2006. *A Guide to Numpy*. Vol. 1. Trelgol Publishing USA.

PROJ contributors. 2018. *PROJ Coordinate Transformation Software Library*. Open Source Geospatial Foundation. https://proj4.org/.

Schulte, Eric, Dan Davison, Thomas Dye, and Carstin Dominik. 2012. "A multi-language computing environment for literate programming and reproducible research." *Journal of Statistical Software* 46 (3): 1–24.

Seabold, Skipper, and Josef Perktold. 2010. "Statsmodels: Econometric and Statistical Modeling with Python." In *9th Python in Science Conference*.

Stallman, Richard M. 1981. "EMACS the Extensible, Customizable Self-Documenting Display Editor." *Proceedings of the ACM SIGPLAN SIGOA Symposium on Text Manipulation* -. https://doi.org/10.1145/800209.806466.

Tange, O. 2011. "GNU Parallel - the Command-Line Power Tool.";*Login: The USENIX Magazine* 36 (1): 42–47. https://doi.org/10.5281/zenodo.16303.

Van Rossum, Guido, and Fred L. Drake Jr. 1995. *Python Reference Manual*. Centrum voor Wiskunde en Informatica Amsterdam.

6 Supraglacial Lake (SL)

6.1 System overview

This section describes the Python processing system for identifying and delineating supraglacial lakes from optical imagery. We do this using a normalised difference water index (NDWI), with a modified system based on Yang 2019 that comprises of 22 modules, in a workflow from satellite image retrieval to vector shapelife generation.

6.1.1 Data Collector

The data collector module downloads Sentinel-2 L1C images through SciHub. The downloaded zipped data is then sorted and stored locally in a pre-defined folder structure based on tile, relative-orbit, and acquisition date. At this stage, the data is unzipped, and file directory structure is flattened to avoid long filenames and paths, not compatible with Python.

6.1.2 Scene retrieval

Relevant scenes are retrieved from the folder structure based on tile information that cover the region of interest. The file paths are stored for subsequent processing stages.

6.1.3 NDWI Generator

This is one of two main processing modules, utilising the Arcpy package in python. The green and near infrared band are combined to form the NDWI raster. These are outputted as raster files to facilitate multiple threshold analysis in an efficient manner in the next module.

6.1.4 Thresholding and Lake Extraction

Thresholds ranging between a minimum and maximum determined by the acquisition date and/or the user to produce multiple lake classifications with varying sensitivity. If a mask is available, lakes are extracted from the NDWI within the masked area, e.g., ice extent and region of interest. These products are initially saved as a binary raster format. These binary raster datasets for varying thresholds (maximum, minimum) are subsequently vectorised to delineate supraglacial lake polygons.

6.1.5 Dataset compiler

Metadata for each vectorised polygon identified in each image are generated and assigned to each feature (supraglacial lake) in the shapefile. The metadata consists of area, perimeter, satellite, tile, relative orbit, and acquisition date. Following this, shapefiles for each individual image from the same acquisition date are merged and the original metadata is retained. Polygons delineated from the minimum and maximum thresholding are aggregated to only retain overlapping polygons identified by both thresholds, thereby retaining the largest extents, and removing false matches e.g., saturated snow. This stage is implemented to best identify lakes with some ice cover.

6.2 Operational scenarios

This is an R&D component and is therefore not an operational product. The processing system is automated for batch processing of regions of interest, e.g., the upstream catchments of Jakobshavn Isbræ, and Nioghavlfjerdsbræ.

6.3 Hardware and software platform

6.3.1 Hardware

In house server machine with the following specs:

Processor: Intel Xeon(R) 2.20GHz dual processor

RAM: 128GB

Available storage: 3.5 TB

Optional external storage also available.

6.3.2 Operating system

Windows 10 64bit operating system

6.3.3 Tools and libraries

The processor requires python 3 for running support scripts. The scripts dependencies are:

- Arcpy – for raster handling
- Gdal – for vector handling
- Geopandas – for data management
- Pathlib – for file directory management

Additionally, core python modules are needed which are available through most common python 3 distributions, e.g., os, sys.

6.4 Future concerns and developments

The processing is dependent on the availability of datasets through network interfaces such as sentinel-hub. The satellite scene download module is independent of the lake identification and classification procedures. Future opportunities could be foreseen in integrating the image download, sorting and storage with the processing modules. This would enable it to be operational. A more adaptive and automated selection of thresholds could be implemented, and a more efficient data storage would be a benefit for operational uses.

End of document