



# Sea Ice CCI+



**ESA CCI+ CLIMATE CHANGE INITIATIVE**  
PHASE 1: NEW R&D ON CCI ECVs

Contract number:

4000126449/19/I-NB



## CCI+ Sea Ice ECV

# SEA ICE THICKNESS ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD)

Reference: D2.1

Issue: 3.1

Date: 15 July 2021



**FMI**



Max-Planck-Institut  
für Meteorologie

**DTU**







**Norwegian  
Meteorological  
Institute**

**The Norwegian Meteorological Institute  
(METNO)**

Henrik Mohns Plass 1

N-0313 Oslo

Norway

Phone: + 47 22 96 30 00

Fax: + 47 22 96 30 50

E-Mail: [thomas.lavergne@met.no](mailto:thomas.lavergne@met.no)

<http://www.met.no>

**Contract**

PHASE 1 OF THE CCI+ CLIMATE CHANGE  
INITIATIVE NEW R&D ON CCI ECVs

SEA ICE ECV

**Deliverable**

D2.1 Sea Ice Thickness Algorithm Theoretical  
Basis Document

**CLIENT**

European Space Agency

**CLIENT REFERENCE**

4000126449/19/I-NB

**Revision date:**

09 June 2021

**Approval date:**

15 July 2021

**Principal Authors**

*Heidi Sallila, Finnish Meteorological Institute*

*Eero Rinne, Finnish Meteorological Institute*




*Stefan Hendricks, Alfred Wegener Institute*

*Stephan Paul, Alfred Wegener Institute*

## Change Record

Issue	Date	Reason for Change	Author(s)
1.0	29 August 2019	First version	H. Sallila, E. Rinne, S. Hendricks
2.0	29 May 2020	Second version	H. Sallila, E. Rinne, S. Hendricks
2.1	25 June 2020	Response to the RID_D2.1_SIT_v2.0	H. Sallila, E. Rinne, S. Hendricks
3.0	09 June 2021	Third version	S. Paul, H. Sallila, S. Hendricks, E. Rinne
3.1	17 July 2021	Response to the RID_D2.1_SIT_v3.0	S. Paul, H. Sallila, S. Hendricks, E. Rinne

## Document Approval

Role	Name	Signature
Written by:	S. Paul, H. Sallila, E. Rinne, S. Hendricks	
Checked by:	M. A. Killie	
Approved by:	A. M. Trofaier	



## Contents

0.1 List of Figures	7
0.2 List of Tables	9
<b>1 INTRODUCTION</b>	<b>10</b>
1.1 Purpose	10
1.2 Scope	10
1.3 Document Status	10
1.4 Acronyms and Abbreviations	10
1.5 Executive Summary	12
<b>2 INPUT AND AUXILIARY DATA</b>	<b>13</b>
2.1 Overview	13
2.2 Primary Altimeter Data Sets	15
2.3 Auxiliary data	16
<b>3 OVERVIEW OF THE SIT PROCESSING CHAIN</b>	<b>18</b>
<b>4 PRE-PROCESSING AND PRIMARY DATA (LEVEL-1 PRE-PROCESSING)</b>	<b>19</b>
4.1 General Filtering	19
4.2 Region Filtering	19
4.3 CryoSat-2 Radar Modes	20
4.4 Orbit Merging	20
<b>5 GEOPHYSICAL RETRIEVAL (LEVEL-2 PROCESSING)</b>	<b>21</b>
5.1 Surface-Type Classification	21
5.1.1 Procedure Description	21
5.1.2 Results	29
5.1.3 General Remarks	33
5.2 Dual Mission Orbit Cross Overs	33
5.3 Retracking	36
5.3.1 Procedure Description	36
5.3.2 Results	40
5.3.3 Envisat Backscatter Drift Correction	43
5.3.4 Pulse Deblurring	44
5.4 Geophysical Range Correction	45
5.5 Radar Freeboard and Sea-Surface Height	45
5.5.1 Sea-Surface Height Uncertainty	48
5.5.2 Radar Freeboard Uncertainty	49
5.6 Snow on Sea Ice	50
5.6.1 Snow Depth	50
5.6.2 Snow Depth Uncertainty	53

5.6.3 Snow Density	53
5.6.4 Snow Density Uncertainty	53
5.7 Sea-Ice Freeboard	54
5.7.1 Radar Freeboard	54
5.7.2 Sea-Ice Freeboard	54
5.7.3 Sea-Ice Freeboard Uncertainty	55
5.8 Sea-Ice Thickness	55
5.8.1 Freeboard to Thickness	55
5.8.2 Sea Ice Density Uncertainty	56
5.8.3 Sea Ice Thickness Uncertainty	56
5.8.4 Sea Ice Type (MYI Fraction) Uncertainty	56
<b>6 COLOCATION ON SPACE-TIME GRID (LEVEL-3 PROCESSOR)</b>	<b>58</b>
6.1 Grid Temporal Coverage	58
6.2 Grid Spatial Definition	58
6.3 Parameter Gridding	59
6.4 Level-3 Gridded Uncertainties	59
<b>7 GAP INTERPOLATION (LEVEL-4 PROCESSOR)</b>	<b>61</b>
<b>8 SEA ICE VOLUME COMPUTATION</b>	<b>63</b>
<b>9 REFERENCES</b>	<b>64</b>

## 0.1 List of Figures

Figure 2-1: Average winter (October to March) Arctic sea ice thickness in meters from October 1993 to March 2001 computed from pulse-limited ERS satellite altimeter measurements

Figure 2-2: Monthly gridded sea ice thickness data in the northern hemisphere with orbit coverage limits for March 2011 (top panel): Envisat (left) and CryoSat-2 (right) and in the southern hemisphere for September 2011 (lower panel)

Figure 2-3: Sea-ice thickness product level examples in both hemispheres. top: Daily orbit trajectories (l2p), bottom: monthly data on space-time grid with different resolution for northern and southern hemisphere

Figure 3-1: Flow chart for the Sea Ice Thickness Processor

Figure 5-1: Flowchart for the process of deriving thresholds for the new surface-type classification

Figure 5-2: Time-series of surface-type fractions for the sensor overlap period between CryoSat-2 (CS2) and Envisat RA-2 (ENV) for the Arctic

Figure 5-3: Time-series of surface-type fractions for the sensor overlap period between CryoSat-2 (CS2) and Envisat RA-2 (ENV) for the Antarctic

Figure 5-4: Visualizations of monthly sea-ice fraction, lead fraction, and valid fraction benchmarks for the Arctic (March 2012)

Figure 5-5: Visualizations of monthly sea-ice fraction, lead fraction, and valid fraction benchmarks for the Antarctic (September 2011)

Figure 5-6: Time-series of surface-type fractions for the sensor overlap period between Envisat RA-2 (Envisat) and ERS-2 RA (ERS-2) for the Arctic

Figure 5-7: Visualizations of monthly ERS-2 sea-ice fraction, lead fraction, and valid fraction benchmarks for the Arctic (March 2003)

Figure 5-8: Distribution of orbit crossovers between ERS2 and Envisat for February 2003 (left) and CryoSat-2 and Envisat for February 2011 (right) for various crossing time differences.

Figure 5-9: Waveform stack for an example XO between ERS-2 and Envisat in February 2003.

Figure 5-10: Clustered ERS-2 waveform stack of the same example XO between ERS-2 and Envisat in February 2003 as shown in Figure 5-9 (left). Top row shows the raw waveforms; bottom row waveforms aligned by their detected first-maximum index.

Figure 5-11: Visualizations of two monthly sets of figures (from left to right): Freeboard difference between Envisat RA-2 and CryoSat-2, the best achievable freeboard difference using an optimal retracker threshold, the sea-ice backscatter, the leading-edge width, and the iteratively estimated optimal threshold for November 2011 (top row) and March 2012 (bottom row)

Figure 5-12: Visualizations of averaged binned optimal threshold values on an x-y plane of leading-edge width and sea-ice backscatter for the Arctic. The blue plane is the 3rd order polynomial fit through all data points

Figure 5-13: As Figure 2-12 but for the Antarctic showing May 2011 (top row) and September 2011 (bottom row)

Figure 5-14: As Figure 2-13 but for the Antarctic and captured from two different viewpoints

Figure 5-15: Mean freeboard for each month of the sensor overlap period (top) for Envisat RA-2 (red) and CryoSat-2 (blue) and the corresponding mean freeboard difference between both sensors in centimetres (middle) and percent with reference to CryoSat-2 (bottom) for the Arctic

Figure 5-16: Setup as in Figure 2-16 but for the Antarctic

Figure 5-17: Scatterplot of all gridded freeboard estimates of CryoSat-2 (y-axis) vs. Envisat RA-2 (x-axis) for the Arctic (left) and the Antarctic (right)

Figure 5-18: Visualizations of the monthly averaged sea-ice backscatter reduction between 2002 and 2012 over ocean-type waveforms obtained between 70°N-75°N and 40°E-50°E

Figure 5-19: Hillshaded sea surface height of the DTU15 global mean sea surface height product for the SIT ECV target region in the northern and southern hemisphere

Figure 5-20: Computation of radar freeboard

Figure 5-21: Example from along-track CryoSat-2 freeboard retrieval. a) Orbit location, b) surface type classification with classes unknown, lead and sea ice with percentage and geographical location along track c) Lead detections, mean sea surface height and sea surface height anomaly d) unfiltered radar freeboard

Figure 5-22: Example of Surface Height Uncertainty

Figure 5-23: Steps for creating the monthly merged snow depth climatology. This example is for April, from left to right: 1) Warren snow depth climatology, 2) Monthly snow composite from daily AMSR2 data, 3) Low-pass filtered composite and 4) Merged Warren/AMSR2 with regional weight factor applied

Figure 5-24: Regional weight factor for the W99 snow depth climatology

Figure 5-25: Performance example of sea ice thickness with the merged W99/AMSR2 snow product. Upper panel is the AWI CryoSat-2 v2.0 sea ice product with W99 snow and lower panel the sea ice thickness with merged W99/AMSR2 snow depth climatology. The improvements are most drastic in areas outside the domain (marked with green rectangles) of the W99 climatology (marked with purple polygon)

Figure 5-26: Computation of sea-ice thickness

Figure 6-1: Gridded uncertainties (Example CryoSat-2 March 2015 Arctic data)

Figure 7-1: Evaluation of CryoSat-2 SIT area averages and their anomalies

Figure 7-2: CryoSat-2 SIT (left); Distance along iceline (DAL; right) for April 2016.

## 0.2 List of Tables

Table 1-1: Acronyms and Abbreviations

Table 2-1: Summary of used auxiliary data sets

Table 4-1: The table lists the latitudinal boundaries for the northern and southern hemisphere used for the region filtering

Table 5-1: Metrics for ocean surface-type classification of waveform data for all sensors, hemispheres, and radar modes

Table 5-2: Metrics for lead surface-type classification of Envisat RA-2, CryoSat-2 SAR mode, and CryoSat-2 SIN mode waveform data for the Arctic

Table 5-3: Metrics for sea-ice surface-type classification of Envisat RA-2, CryoSat-2 SAR mode, and CryoSat-2 SIN mode waveform data for the Arctic

Table 5-4: Metrics for lead surface-type classification of Envisat RA-2, CryoSat-2 SAR mode, and CryoSat-2 SIN mode waveform data for the Antarctic

Table 5-5: Metrics for sea-ice surface-type classification of Envisat RA-2, CryoSat-2 SAR mode, and CryoSat-2 SIN mode waveform data for the Antarctic

Table 5-6: Reference dates for the monthly snow climatology used for the estimation of linear interpolated snow depth with daily resolution

Table 6-1: Temporal definition for Level-3 products.

Table 6-2: Projection definition for Level-3 products.

Table 6-3: Grid extent and spacing for Level-3 products.

## 1 INTRODUCTION

### 1.1 Purpose

This document is the Algorithm Theoretical Basis for the Sea Ice ECV within CCI+ PHASE 1 - NEW R&D ON CCI ECVs, which is being undertaken by a METNO-led consortium. This document is based on the work of Phase 2 of the ESA CCI project and includes the new developments for the Sea Ice Thickness (SIT) aspects.

This document also contains preparation and documentation of the ongoing work to extend the CRDP to cover SIT from ERS-1 and ERS-2 satellites as well as for future level-4 products and computation of sea-ice volume. At the time of writing, these development items are not complete and these sections will be updated in the future versions of this document.

### 1.2 Scope

The scope of the document is to describe elements of the algorithms that are chosen for implementation during the third year of the CCI+ Phase 1, towards the production of the CDR during Year 3. The selected algorithms are presented and justified, but the document does not contain the results of research leading to the selection of these algorithms.

### 1.3 Document Status

This is the third issue of the ATBD document for the Sea Ice CCI+ project. The document describes the algorithms aimed for the final processing, which, however, will need added details and steps before finalizing the document. In addition, the auxiliary data sets used in the initial processing are introduced.

The description of the SIT retrieval algorithm is a combination of the state at the end of the Phase 2 of the SICCI project and the knowledge gained during CCI+. This version differs from the SICCI Phase 2 in that this document contains the added knowledge from ERS-1 and ERS-2 studies, including pulse deblurring correction, as well as the exception of novel snow estimates for the Arctic in areas where Warren climatology is considered outdated. Additionally, the introduction of dual mission orbit cross overs and the added benefit for the current and future CDR production is included in this document. This document will be iterated after the SIT algorithm development has been finalized and the CDR is produced and published.

### 1.4 Acronyms and Abbreviations

Table 1-1 below lists the acronyms and abbreviations used in this volume.

**Table 1-1: Acronyms and Abbreviations. Acronyms for the deliverable items (URD, etc...) and partner institutions (AWI,...) are not repeated.**

Acronym	Meaning
AMSR-E / AMSR2	Advanced Microwave Scanning Radiometer (for EOS / #2)
C3S	EU Copernicus Climate Change Service
CCI	Climate Change Initiative
CDR	Climate Data Record
CS2	ESA's CryoSat-2
DAL	Distance along iceline
DMSP	Defense Meteorological Satellite Program
DMXO	Dual-mission orbit crossover
EASE grid	Equal-Area Scalable Earth Grid
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
ENVISAT	ESA's Environmental Satellite
EO	Earth Observation
ERS	European Remote Sensing Satellite
ESA	European Space Agency
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
FYI	First Year Ice
ICDC	Integrated Climate Data Center
ICDR	Interim Climate Data Record
L1b, L2, L3C, ...	Satellite data processing Level (Level-1b, ...)
LEW	Leading-edge width
MSSH	Mean sea surface height
MYI	Multi-Year Ice
OSI SAF	EUMETSAT Ocean and Sea Ice Satellite Application Facility
PP	Pulse peakiness
RA	Radar Altimeter
RRDP	Round Robin Data Package
S0	Sea-ice backscatter
SIC	Sea Ice Concentration
SIT	Sea Ice Thickness
SAR	Synthetic Aperture Radar
SGDR	Sensor Geophysical Data Record
SIRAL	Synthetic Aperture Radar (SAR) Interferometer Radar Altimeter
SMMR	Scanning Multichannel Microwave Radiometer
SMOS	Soil Moisture and Ocean Salinity
SSH	Sea surface height

SSHA	Sea surface height anomaly
SSM/I	Special Sensor Microwave/Imager
SSMIS	Special Sensor Microwave Imager/Sounder
W99-AMSR2	The merged Warren-AMSR2 snow climatology
XO	Orbit crossover

## 1.5 Executive Summary

This document presents the algorithms for producing the Sea Ice Thickness Climate Data Record (SIT CDR) in CCI+. This document can be understood as a recipe book for a software engineer wanting to build a working SIT processor. It also gives the background of used algorithms and data for anyone wanting to understand the CRDP better.

The document includes all the necessary steps for converting altimeter waveforms into sea ice thickness in along-track (L2), monthly gridded (L3) format, and the gap-free gridded data (L4):

- Filtering data based on latitudes, possibly removing data points based on flags provided with the data
- Introduction and description of the dual-mission orbit crossovers between sensors in their respective mission-overlap periods
- Surface-type classification algorithms based on waveform parameters for differentiating between ocean, lead, sea ice, and ambiguous
- Waveform retracking of the different surface types to obtain ice elevations and sea surface height tie points over leads between ice floes.
- Application of geophysical range corrections including tidal correction
- Estimation of radar freeboard and along-track sea surface height from ice surface elevations and interpolated SSH tie points utilizing a mean sea surface
- Radar freeboard to sea-ice freeboard conversion by applying geometric corrections, with snow information
- Sea ice thickness calculation based on sea-ice freeboard and auxiliary parameters and the assumptions of hydrostatic equilibrium

Along with steps and algorithms, all primary and auxiliary data products used to create the SIT CDR are introduced.

Although the algorithms for the final production are chosen, there are some details that are necessary to be added to the ATBD in its final iteration.

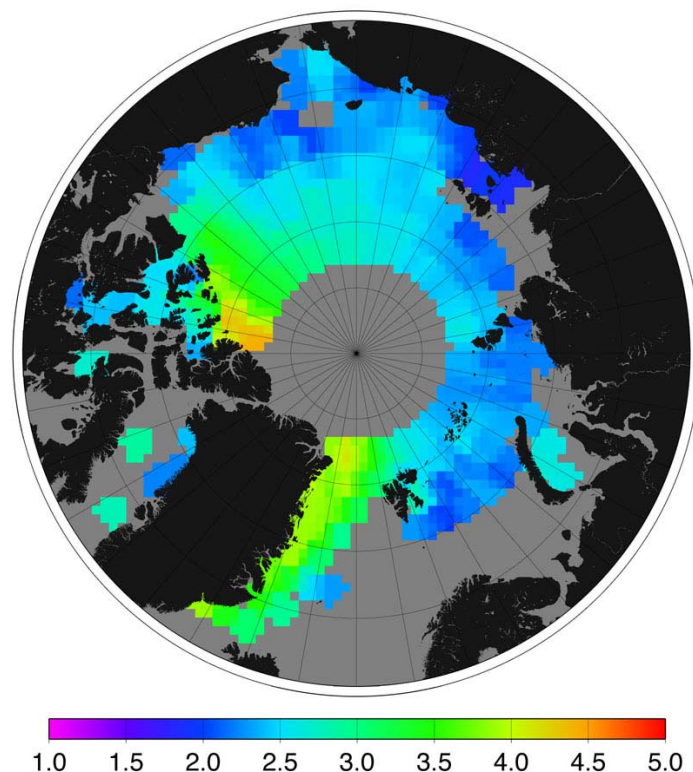


## 2 INPUT AND AUXILIARY DATA

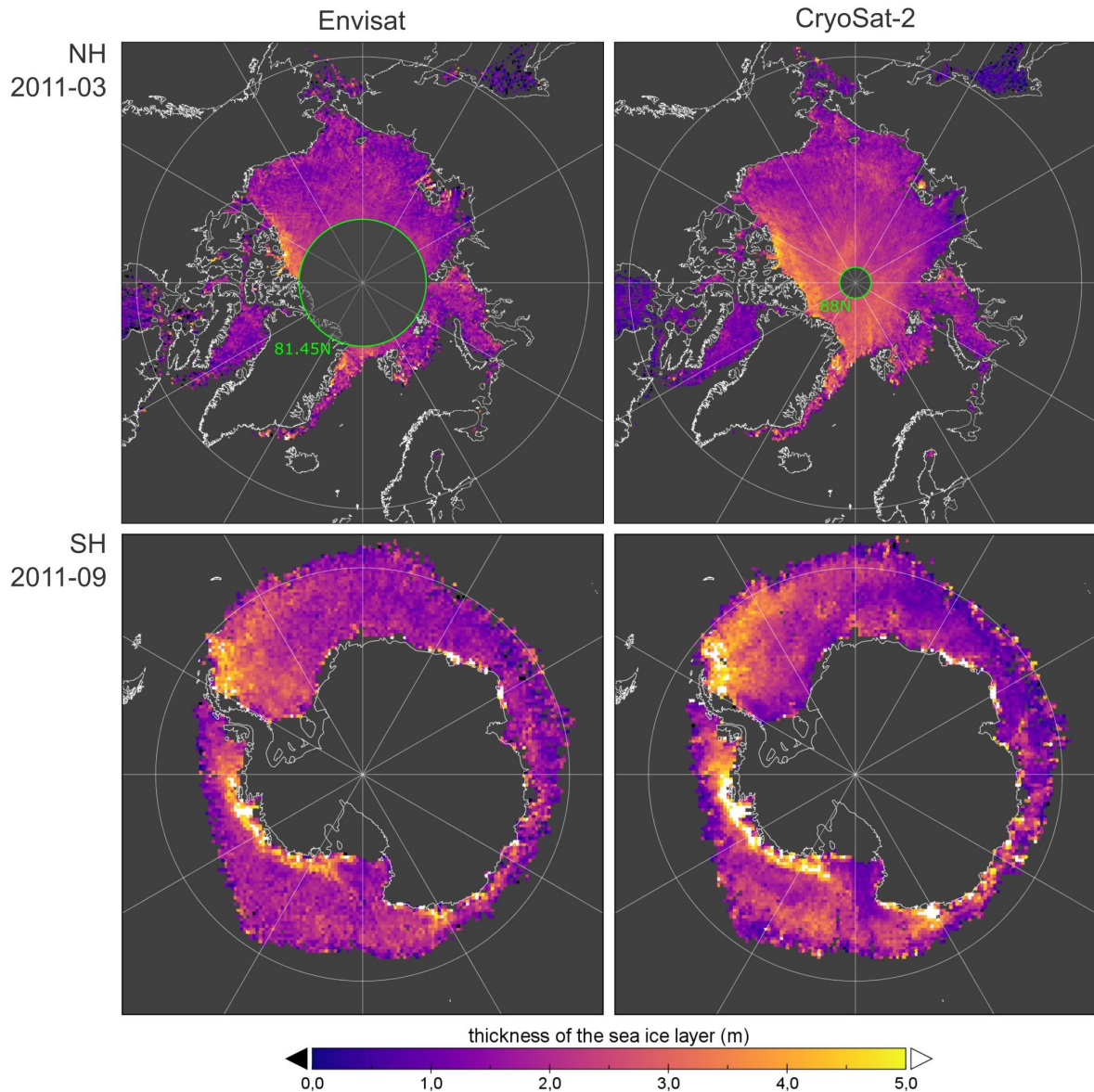
### 2.1 Overview

This part of the document is intended as a generic guide to setting up a sea ice thickness processing system for any polar orbiting satellite radar altimeter. The general method is described and specific examples are given. The general processing system is identical for pulse-limited as well as for SAR altimetry. Any sensor type specific differences are stated.

The method used to extract sea ice thickness from radar altimetry data is based on the pioneering work of Peacock and Laxon, 2004; Laxon et al., 2003 for the ERS-2 mission. The method involves separating the radar echoes returning from the ice floes from those returning from the sea surface in the leads between the floes. This step of a surface-type classification is crucial and allows for a separate determination of the ice floe and sea surface heights. The freeboard that is the elevation of the ice upper side (or ice-snow interface) above the sea level can then be computed by deducting the interpolated sea surface height at the floe location from the height of the floe. Sea-ice thickness can then be calculated from the sea-ice freeboard with the additional information of the snow load. Figure 2-1 shows an example of the earliest results of Laxon et al. (2003) with aggregated ERS data. Figure 2-2 and Figure 2-3 show the progression in terms of spatial and temporal resolution of altimeter-based sea ice thickness information with the extension to the Envisat and CryoSat-2 platforms in the ESA CCI project since the first application of the method.

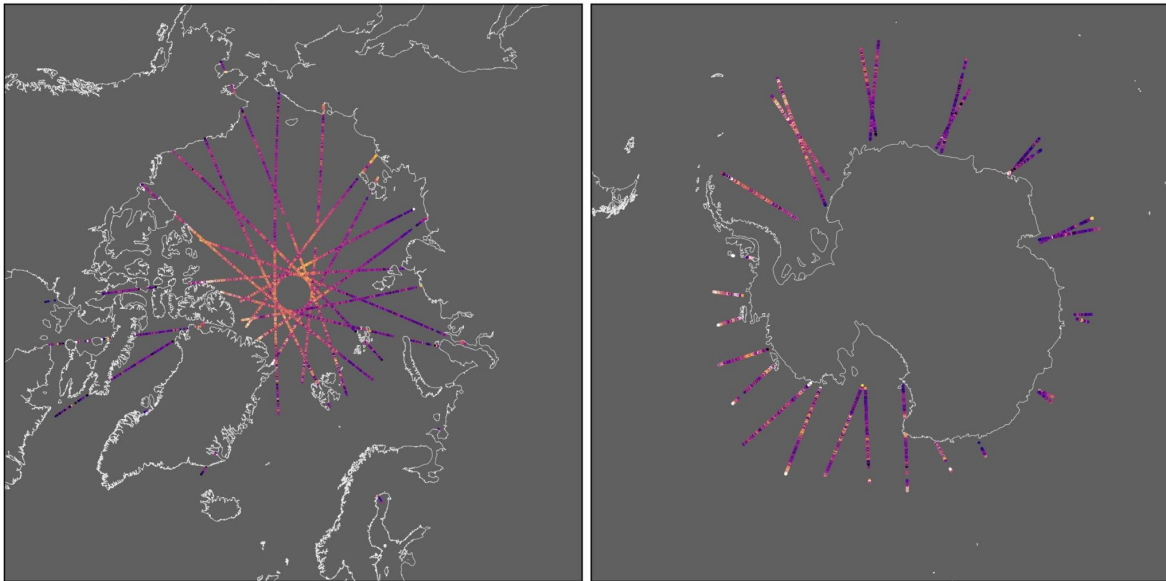


**Figure 2-1: Average winter (October to March) Arctic sea ice thickness in meters from October 1993 to March 2001 computed from pulse-limited ERS satellite altimeter measurements.**

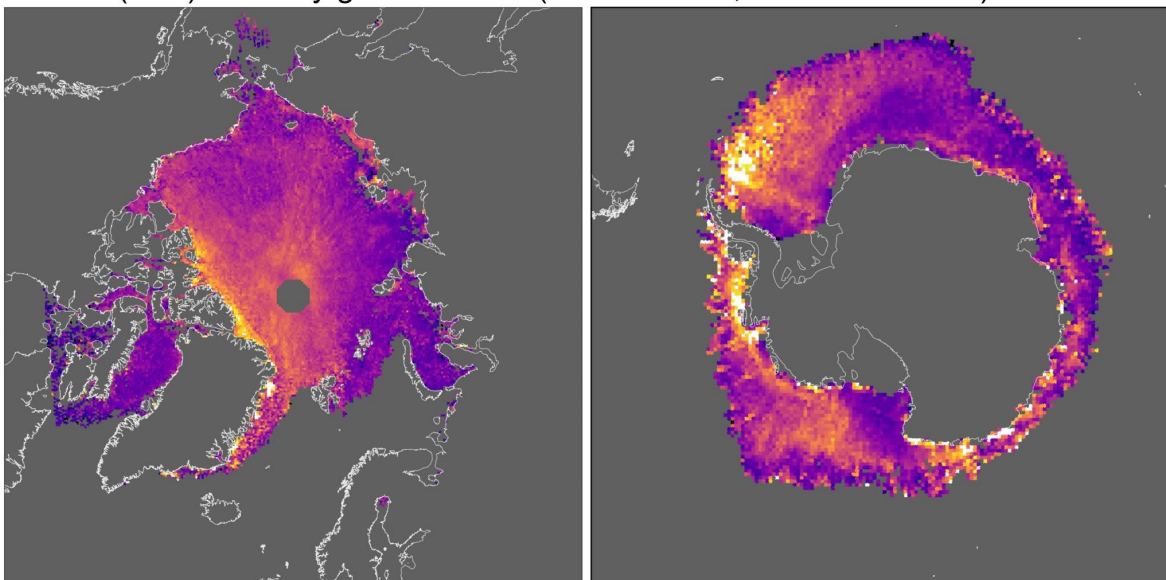


**Figure 2-2: Monthly gridded sea ice thickness data in the northern hemisphere with orbit coverage limits for March 2011 (top panel): Envisat (left) and CryoSat-2 (right) and in the southern hemisphere for September 2011 (lower panel)**

## Level 2 (L2P): Daily orbit trajectories



## Level 3 (L3C): Monthly gridded fields (Arctic: 25 km, Antarctic: 50 km)



**Figure 2-3: Sea-ice thickness product level examples in both hemispheres. top: Daily orbit trajectories (l2p), bottom: monthly data on space-time grid with different resolution for northern and southern hemisphere.**

## 2.2 Primary Altimeter Data Sets

The input data set must contain the radar echo waveforms and all other fields mentioned in this document such as altitude, range, atmospheric corrections and geophysical corrections. Figure 3-1 shows a flow chart for each step of the sea ice thickness processor. Each step is

explained in detail in the sections below. For ERS-1 and ERS-2 RA the REAPER Sensor Geophysical Data Record (SGDR) data (Brockley et al., 2017) is used. In the case of Envisat RA-2, the input for the sea ice thickness processor is version 3.0 of the Envisat SGDR data (ESA, 2018). The SGDR data contains the waveforms as well as all other required fields. The ERS data is provided per cycle files or daily files. For Envisat data, each orbit is stored in two data files. The earlier of the two data files contains the data for the ascending arc from -81.5 latitude up to +81.5 latitude, and the later of the two the descending arc from +81.5 latitude back down to -81.5 latitude. These files are read sequentially and the output split at appropriate points to make continuous Arctic and Antarctic passes.

For CryoSat-2, the version 1.0 of Baseline D orbit data files are used and separated into sections of different instrument modes by the processor. CryoSat-2's SIRAL altimeter is operated in two different modes over sea ice: a) In synthetic aperture radar (SAR) off-coast and b) in synthetic aperture radar interferometric (SIN) mode to enable more accurate land ice altimeter measurements with higher surface slopes. For the product generation both radar modes are used, but the processing does not utilize the interferometric information in SIN mode. In addition to the different altimeter type that improves the spatial resolution, the higher orbit inclination of CryoSat-2 allows sea ice thickness measurements in the Arctic up to 88N.

### **2.3 Auxiliary data**

The conversion into sea-ice freeboard requires either the use of auxiliary input data or a parametrization of snow depth. For the Arctic, where in SICCI Phase 2 only Warren climatology (W99, Warren et al., 1999) was applied, we now use a merged Warren-AMSR2 (W99-AMSR2) snow climatology for all the instruments, further described in Section 5.5. One main reason for the change is that the Warren climatology is based on data sets obtained from Arctic drift stations in regions of multi-year sea ice (MYI), snow depth values are suspected to be biased high over first-year sea-ice (FYI).

In order to discriminate between FYI and MYI in the Arctic, we resort to Copernicus Climate Change Service (C3S) Climate Data Record (CDR)/interim-CDR (ICDR). In SICCI Phase 2 a MYI fraction data set based on the Special Sensor Microwave Imager (SSM/I)/Special Sensor Microwave Imager Sounder (SSMIS) sensors on-board of the Defense Meteorological Satellite Program (DMSP) satellites provided by the Integrated Climate Data Center (ICDC) was used. The C3S CDR is produced with an algorithm that is optimized to produce consistent CDRs based on time series of passive microwave data of the above-mentioned instruments, in addition with SMMR and ECMWF ERA-Interim data.

**Table 2-1: Summary of used auxiliary data sets.**

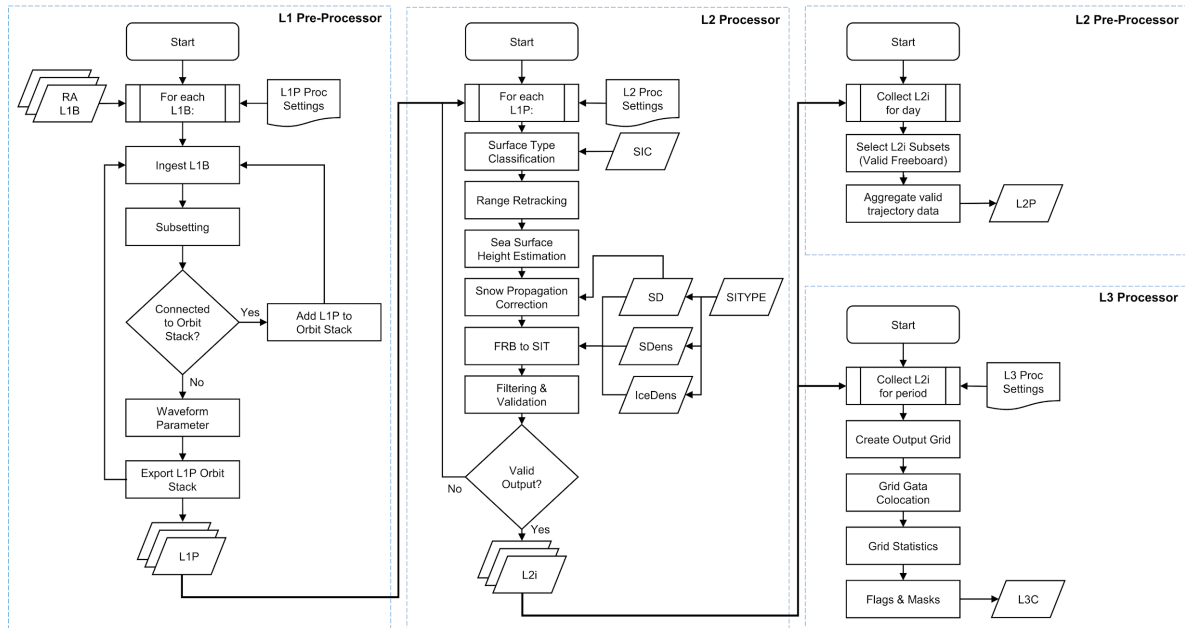
Parameter	ERS-1 /2 Arctic	ERS- 1 / 2 Antarctic	Envisat RA-2 Arctic	Envisat RA-2 Antarctic	CryoSat-2 Arctic	CryoSat-2 Antarctic
SIC	C3S CDR	C3S CDR	C3S CDR	C3S CDR	C3S CDR/ICDR	C3S CDR/ICDR
SIType	C3S CDR	Single Ice Type	C3S CDR	Single Ice Type	C3S CDR/ICDR	Single Ice Type
Snow Depth	Merged W99-AMSR2 climatology	AMSR-e climatology	Merged W99-AMSR2 climatology	AMSR-E/2 climatology	Merged W99-AMSR2 climatology	AMSR-E/2 climatology
Snow Density	Mallett et al., 2020	fixed/clim	Mallett et al., 2020	fixed/clim	Mallett et al., 2020	fixed/clim
MSS	DTU15	DTU15	DTU15	DTU15	DTU15	DTU15

For the Antarctic, we assume only a single sea-ice type being present. As the Warren climatology is only available for the Arctic, we use a snow-depth climatology derived from the Advanced Microwave Scanning Radiometer-EOS (AMSR-E) and AMSR-2 data for the Antarctic. This data set is based on a revised version of the approach described by Cavalieri et al. (2014) and provided by the ICDC.

Other required auxiliary input data sets for the estimation of sea ice freeboard and sea ice thicknesses comprise the use the sea-ice concentration (SIC) data obtained from the C3S CDR for both hemispheres, in contrast to the SIC product from Ocean and Sea Ice Satellite Application Facility (OSISAF) used in SICCI Phase 2. For mean sea-surface (MSS) height the product provided by the Danish Technical University (DTU) in its 2015 version is used.

A summary of all used auxiliary data sets for the production of the sea-ice thickness climate data record is presented in Table 2-1.

### 3 OVERVIEW OF THE SIT PROCESSING CHAIN



**Figure 3-1: Flow chart for the Sea Ice Thickness Processor**

Figure 3-1 presents an overview about the sea-ice thickness processing chain detailed into defined processors for the successive product data levels. The structure of the following sections is modeled after these processors that include details for each sensor. The geophysical retrieval starts with the surface-type classification, with the corresponding parametrization for ERS-1 and ERS-2, Envisat and CryoSat-2. This continues with a thorough description of the range retracking procedure and a necessary Envisat RA-2 backscatter correction. Furthermore, the processing chain of radar freeboard and sea-surface height derivation, the estimation of sea-ice freeboard, and the estimation of sea-ice thickness are described. The subsections contain the computation of the geophysical parameters as well as the corresponding uncertainties.

While the geophysical retrieval is implemented at full sensor resolution, the aggregation of the parameter to space-time grids is described in the following sections of this document.

The processors are implemented in the python sea-ice radar altimetry toolbox (pysiral). This open source software project is hosted at Github (<https://github.com/shendric/pysiral>) and allows the inspection of the actual implementation of all algorithm components.



## **4 PRE-PROCESSING AND PRIMARY DATA (LEVEL-1 PRE-PROCESSING)**

The main purpose of the pre-processing of the primary level-1 data is to provide a unified input format and data conventions for the following geophysical retrieval.

### **4.1 General Filtering**

For ERS-1 / 2 the filtering is yet to be confirmed.

There is some sensor-specific general filtering applied, which follows UCL's implementation of the Envisat algorithm used during SICC Phase 1. This filtering is based on the available flags in the Envisat data indicating any significant problems with any record. In UCL's implementation of the filtering for Envisat, the Measurement Confidence Data Flags (MCD Flags) in the SGDR data are examined for problem records. We remove records where the following flags are raised: 0 (Packet Length Error), 1 (OBDH invalid), 4 (AGC Fault), 5 (Rx Delay Fault) and 6 (Waveform Fault).

For CryoSat-2 level-1b data, no general filtering mechanisms are necessary.

### **4.2 Region Filtering**

The latitudinal boundaries within which Arctic and Antarctic sea ice is found are listed in the Table 4-1. The latitude values in the satellite data are examined and any data points outside these regions are rejected from the processing. The surface type flag in the data is also examined and any data not flagged as over ocean is also rejected.

**Table 4-1: The table lists the latitudinal boundaries for the northern and southern hemisphere used for the region filtering**

Area	Minimum Latitude	Maximum Latitude
Arctic	45.0	90.0
Antarctic	-90.0	-45.0

The data is cropped to the two latitude ranges and data over land masses are excluded, except if the orbit segment over land is shorter than 300 km. Else, the orbit is split into two segments.

### 4.3 CryoSat-2 Radar Modes

Only for CryoSat-2, the altimeter data is divided into orbit segments in three radar modes: LRM, SAR and SARIn. Of these we disregard LRM due to its little relevance over sea ice, thus using only SAR and SARIn modes. These come in separate product files, which are merged in the pre-processor. The merging process requires reducing the SIN waveforms from 512 to 256 range bins of the SAR waveforms. This can be done without losing waveform information as the sea ice waveforms are narrow and defined. The step is unique to CryoSat-2, as other platforms (ERS-1/2, Envisat) provide data only with a single radar mode.

### 4.4 Orbit Merging

Adjacent neighbouring orbit segments are merged into a single orbit segment over the polar regions whenever possible to enable consistent sea-surface height estimation across the Arctic polar basin. Due to the geography in the Antarctic, the descending and ascending orbit segments over the ocean will always be separated by the Antarctic continent.



## 5 GEOPHYSICAL RETRIEVAL (LEVEL-2 PROCESSING)

The level-2 processing step includes the retrieval of the geophysical variables from the pre-processed radar measurements, with the use of auxiliary data listed in Table 2-1.

### 5.1 Surface-Type Classification

The surface-type classification is a crucial part in the processing chain of deriving sea-ice freeboard (and therefore sea-ice thickness), as the detection of leads is pivotal for determining the sea-surface height. The sea-surface height in turn is used as the reference from which the sea-ice freeboard is calculated. Additionally, a clear distinction between leads, sea ice and ambiguous mixed signals (which will be excluded from the actual freeboard retrieval) helps to improve the quality and accuracy of resulting sea-ice freeboard estimates. In other words, a surface-type selection bias is very likely to also have an impact on the resulting sea-ice freeboard and hence also the sea-ice thickness and subsequent products.

In general, with smaller instrument footprint sizes, less surface-type mixing occurs. However, leads often dominate acquired waveforms due to their specular reflection, and therefore act as sources of strong off-nadir backscatter signals. These off-nadir leads can substantially decrease the quality of the range retracking and increase the sensors' footprint. This is especially true for pulse-limited radar altimeters such as ERS-1/2 RA and Envisat RA-2. In case of Envisat RA-2, the nominal circular footprint of 2 km in diameter (Connor et al., 2009) can increase to up to 10 km (Chelton et al., 2001) for strong off-nadir backscatter sources. Despite its much smaller footprint (1.65 km × 0.30 km), CryoSat-2 can also be affected by off-nadir leads, which will result in erroneous freeboard estimates (Armitage and Davidson, 2014).

#### 5.1.1 Procedure Description

The surface classification algorithm is based on a multi-parameter classification based on a consistent set of parameters for ERS-1, ERS-2, Envisat RA-2 and CryoSat-2. The set of classifiers is defined by the sea-ice backscatter (SIG0), the leading-edge width (LEW) and the pulse peakiness (PP) as classifiers to positively identify between lead-type and sea-ice-type from otherwise ambiguous-type waveforms.

The pulse peakiness is defined as in Ricker et al. (2014):

$$PP = \sum_{i=1}^{N_{WF}} \frac{\max(WF)}{WF_i} \cdot N_{WF}$$

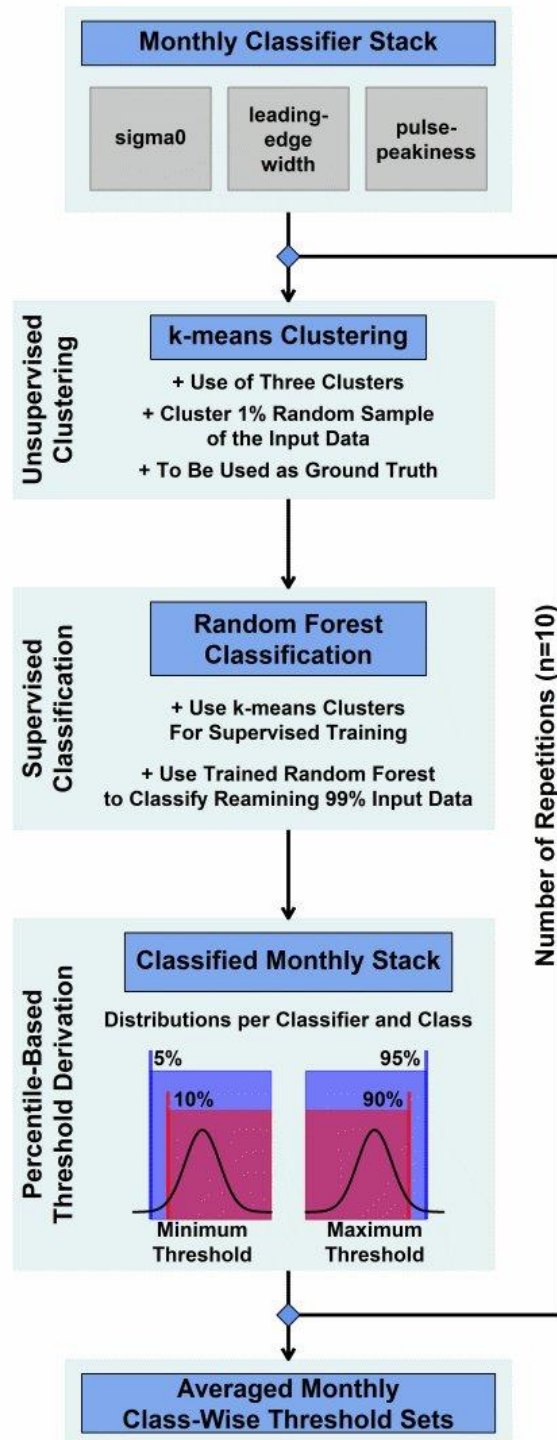
The leading-edge width is defined as the width in range bins along the power rise to the first maximum between 5 % and 95 % of the first-maximum peak power while using a ten-times oversampled waveform.

Over the course of a winter season, ice conditions can change substantially. Similar to leads, young and thin-ice areas feature rather specular reflections compared to other ice types. Furthermore, the amount of leads varies seasonally and regionally. Based on fixed thresholds for a whole winter season, these changes are difficult to capture and the rejection

rate is increased unnecessarily. Hence, we decided on using monthly thresholds to improve the overall results and data quality.

There is a general lack of ground-truth data as collocated measurements of the same sea-ice situation are very difficult due to sea-ice drift and therefore rare. However, received waveforms feature very distinct characteristics and are well described in literature for sea ice and leads. These characteristics can also be deduced from the chosen set of classifiers. In order to bypass the lack of ground-truth, we decided to use a combination of unsupervised clustering and supervised classification.

Based on this combination, we are able to determine suitable thresholds for data acquired by ERS-1 and ERS-2, Envisat RA-2 as well as CryoSat-2. The workflow of how we derived the surface-type thresholds is summarized in Figure 5-1.



**Figure 5-1: Flowchart for the process of deriving thresholds for the new surface-type classification**

In a first step, the three classifiers are computed for all available L1b data per sensor and month in the sensor overlap period from January 1995 to June 1996 for ERS-1 and ERS-2, from May 2002 to July 2003 for ERS-2 and Envisat and from November 2010 to March 2012 for Envisat and CryoSat-2. We only use waveforms that are located between 70°N and 81.5°N for the Arctic, are obtained over the ocean, and feature a minimum sea-ice

concentration of 70%. The northern limit of 81.5°N was chosen to assure a maximum of consistency between Envisat RA-2 and CryoSat-2. In order to retrieve thresholds for the “Wingham Box” and other Arctic areas that are covered while CryoSat-2 operates in SIN mode, all waveforms above 70°N were used. For the Antarctic the same parameters apply, but waveforms are geographically limited to an area south of 65°S.

Next, 1 % of this monthly data is sampled randomly. This data sample is then separated into three clusters using k-means clustering (MacQueen et al., 1967; Hartigan et al., 1979). This methodology is widely used to separate input data of  $N$  observations into  $K$  clusters of equal variance, whereby the within-cluster sum-of-squares are minimized (MacQueen et al., 1967; Hartigan et al., 1979).

Generally, the preselection of the number of clusters can be a problem when utilizing k-means clustering. However, while we also tested a higher number of initial clusters with the perspective of later reunion of very similar clusters, a separation into just three classes turned out to be sufficient. Overall, lead waveforms account for a smaller fraction of the total measurements compared to sea-ice waveforms. Because of that and the fact that k-means clustering generally tends toward generating equal-size clusters (this is generally a presumption of k-mean algorithms), sole use of k-means clustering for the complete data set was not feasible.

This information in mind, the clustered 1 % data sample is therefore used as training data to train a random forest (Breiman, 2001). Random forests are an ensemble machine learning methods used for classification and are based on a large number of single decision trees that are fitted to randomized sub samples of the given training data set (Breiman, 2001). After initial training, the random forest can then be used for classification of the remaining data. Each tree in the trained forest then does a classification and casts a unique vote. In the end, the majority decides the resulting class. Each decision tree is thereby grown following certain rules: First, from the training data of size  $N$ ,  $N$  cases are sampled randomly with replacement as specific training data set for each single tree. Second, for  $M$  input parameters (in our case sea-ice backscatter, pulse peakiness, and leading-edge width), a fixed number  $m \ll M$  of the given input parameters is specified and randomly selected out of  $M$ . The best split on these selected parameters  $m$  is then used to split the node. Throughout the growing of the forest, the value of  $m$  is held constant. Third, each tree is grown out fully, i.e., to its largest possible extent. No pruning is applied. In contrast to single decision trees that tend to overfit, random forests do not overfit and are also capable of dealing with unbalanced data sets (Breiman, 2001).

For the here-used classification problem, we always grow a total number of 500 decision trees per training. Due to the small number of input parameters ( $M=3$ ), we set  $m$  to one.

The trained random forest for each month is then used to classify the remaining 99 % of the corresponding monthly data. From this classified data set, distributions for each of the three classifiers for each month in the sensor overlap period are obtained. These distributions feature clear distinctions along each classifier's respective total range for each surface-type class (leads, sea ice, and ambiguous). Sea-ice backscatter is on average in the upper part of the range for the lead class and in the lower for the sea-ice class. Similar observations are apparent for pulse peakiness (upper part for leads, lower for sea ice) and leading-edge width (lower part for leads, upper part for sea ice). In other words, leads feature higher sea-ice backscatter and pulse peakiness as well as shorter leading-edge widths. The opposite is seen in the sea-ice class. The class of ambiguous signals is placed in between.

Thresholds are then obtained from the resulting classifier distributions by using either the 5 % or 10 % percentile for a minimum threshold, or the 90 % or 95 % percentile in case of a maximum threshold. The choice of using the more strict (10 %/90 %) or less strict (5 %/95

%) percentile thresholds depends on the sensor. Due to its larger footprint and therefore an expected higher degree of surface-type mixing, we chose the more strict thresholds for ERS-1/2 RA as well as Envisat RA-2, and the less strict thresholds for CryoSat-2 due to its smaller footprint. For example, in order to derive thresholds for the detection of leads, the 5%/10% percentiles of the sea-ice backscatter and pulse-peakiness distributions would be used alongside the 90%/95% percentile of the leading-edge-width distribution.

The whole procedure, starting with randomly sampling 1% from the initial monthly stack, is then repeated ten times. In the last step, the average minimum/maximum thresholds for each classifier, surface-type class, and month in the sensor overlap period are estimated. These thresholds are summarized in Table 5-1 through Table 5-5 and are used for all months in the complete climate data record. Note the threshold units are in % for sea ice concentration (SIC), dB for backscatter (SIG0), and unitless for pulse peakiness (PP) and leading edge width (LEW) due to the last two being calculated over range bins.

**Table 5-1: Metrics for ocean surface-type classification of waveform data for all sensors, hemispheres, and radar modes**

Metric	Min	Max
Ocean waveforms are characterized by medium to low <b>pulse peakiness (PP)</b> values.		5
Only regions of very low ice concentration (SIC in %) values are suitable for the ocean surface type flag		5

**Table 5-2: Metrics for lead surface-type classification of Envisat RA-2, CryoSat-2 SAR mode, and CryoSat-2 SIN mode waveform data for the Arctic**

Metric	Month	ERS-1 RA		ERS-2 RA		Envisat RA-2		CryoSat-2 SAR		CryoSat-2 SIN	
		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Lead waveforms are characterized by strong <b>pulse peakiness (PP)</b>	JAN	16.36		20.96		46.90		67.30		264.30	
	FEB	16.72		20.51		46.40		66.30		257.90	
	MAR	17.02		20.55		46.20		66.60		253.60	
	APR	17.17		20.73		48.40		69.90		264.60	
	OCT	16.68		23.25		52.90		76.00		291.80	
	NOV	18.22		22.32		51.00		73.80		288.80	
	DEC	16.97		21.45		47.70		68.60		272.60	
Lead waveforms are also characterized by high <b>backscatter</b> values due to specular reflection ( <b>SIG0</b> , in dB)	JAN	24.95		27.45		28.80		23.80		24.90	
	FEB	24.35		26.93		28.60		23.20		25.00	
	MAR	25.10		27.10		28.50		23.30		24.10	
	APR	24.78		27.92		28.40		23.40		24.50	
	OCT	28.52		32.75		32.80		28.00		29.00	
	NOV	26.78		29.68		30.80		25.80		27.40	
	DEC	25.21		27.99		29.30		24.10		25.80	
Lead	JAN		1.18		0.97		0.82		0.77		1.10

waveforms feature a very steep increase in echo power and therefore short <b>leading-edge widths (LEW)</b>	FEB				0.98		0.82		0.78		1.11
	MAR		1.17		0.98		0.82		0.78		1.13
	APR		1.16		0.99		0.82		0.76		1.09
	OCT		1.18		0.94		0.82		0.72		1.02
	NOV		1.22		0.95		0.82		0.73		1.03
	DEC		1.15		0.96		0.82		0.76		1.07
				1.17							
Only lead classifications by waveform are expected that fall into regions of sufficient sea-ice cover (checked with <b>sea-ice concentration ; SIC</b> in %)	JAN	70		70		70		70		70	
	FEB	70		70		70		70		70	
	MAR	70		70		70		70		70	
	APR	70		70		70		70		70	
	OCT	70		70		70		70		70	
	NOV	70		70		70		70		70	
	DEC	70		70		70		70		70	

**Table 5-3: Metrics for sea-ice surface-type classification of Envisat RA-2, CryoSat-2 SAR mode, and CryoSat-2 SIN mode waveform data for the Arctic**

Metric	Month	ERS-1 RA		ERS-2 RA		Envisat RA-2		CryoSat-2 SAR		CryoSat-2 SIN	
		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Sea-ice waveforms shouldn't be peaky and therefore have a low <b>pulse peakiness (PP)</b>	JAN		7.56		12.34		16.00		30.50		99.40
	FEB		7.66		11.69		14.80		28.70		94.20
	MAR		7.80		11.75		14.10		28.10		89.90
	APR		8.06		12.38		14.20		28.50		90.00
	OCT		6.72		15.04		19.40		35.40		114.40
	NOV		8.32		13.62		19.30		34.90		113.90
	DEC		7.77		12.89		16.90		31.90		103.80
Sea-ice waveforms are also characterized by low <b>backscatter</b> values due to diffuse reflection ( <b>SIG0</b> in dB)	JAN	9.69	17.75	14.68	22.87	2.5	22.50	2.5	20.80	2.5	21.40
	FEB	9.70	17.76	13.98	22.23	2.5	21.80	2.5	19.90	2.5	20.90
	MAR	9.43	17.20	14.19	22.13	2.5	21.30	2.5	19.60	2.5	20.10
	APR	9.03	16.75	14.09	22.52	2.5	20.40	2.5	19.00	2.5	19.10
	OCT	10.02	18.85	17.94	27.09	2.5	25.90	2.5	25.70	2.5	24.30
	NOV	10.23	18.90	15.82	24.65	2.5	24.60	2.5	23.20	2.5	23.70
	DEC	9.90	17.99	15.22	23.57	2.5	22.80	2.5	21.10	2.5	22.00
Sea-ice waveforms feature a less steep increase in echo power and	JAN	0.90		0.85		0.81		1.02			1.55
	FEB	0.90		0.85		0.83		1.08			1.58
	MAR	0.90		0.85		0.83		1.10			1.62
	APR	0.90		0.85		0.83		1.11			1.64
	OCT	0.93		0.83		0.78		0.91			1.44
	NOV	0.89		0.83		0.78		0.90			1.44
	DEC	0.89		0.84		0.80		0.97			1.51

therefore longer <b>leading-edge widths (LEW)</b>										
Only sea-ice classifications by waveform are expected that fall into regions of sufficient ice cover (checked with <b>SIC</b> in %)	JAN FEB MAR APR OCT NOV DEC	70 70 70 70 70 70 70		70 70 70 70 70 70 70		70 70 70 70 70 70 70		70 70 70 70 70 70 70		70 70 70 70 70 70 70

**Table 5-4: Metrics for lead surface-type classification of Envisat RA-2, CryoSat-2 SAR mode, and CryoSat-2 SIN mode waveform data for the Antarctic**

Metric	Month	ERS-1 RA		ERS-2 RA		Envisat RA-2		CryoSat-2 SAR		CryoSat-2 SIN	
		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Lead waveforms are characterized by strong <b>pulse peakiness (PP)</b>	JAN	19.57		23.41		56.60		80.70		307.40	
	FEB	20.48		22.77		53.20		75.10		300.70	
	MAR	22.23		22.09		51.90		73.20		291.70	
	APR	18.67		22.35		50.70		69.50		288.50	
	MAY	18.66		22.21		50.10		69.70		283.70	
	JUN	17.87		22.53		49.30		69.30		284.20	
	JUL	18.67		22.44		49.50		69.20		276.90	
	AUG	18.90		22.34		49.10		69.50		284.40	
	SEP	18.57		23.06		49.30		69.70		278.90	
	OCT	19.51		22.92		51.60		71.70		289.40	
	NOV	21.23		22.62		53.90		76.00		299.40	
	DEC	20.04		23.69		55.10		78.10		307.70	
Lead waveforms are also characterized by high <b>backscatter</b> values due to specular reflection ( <b>SIG0</b> in dB)	JAN	28.21		32.98		33.20		28.50		29.20	
	FEB	28.96		32.85		32.10		26.80		29.00	
	MAR	28.59		31.22		31.80		26.20		28.50	
	APR	23.42		29.19		30.80		24.60		27.80	
	MAY	24.11		29.20		29.40		23.40		26.90	
	JUN	24.25		28.42		28.60		22.80		26.50	
	JUL	22.82		27.28		28.60		23.00		26.30	
	AUG	23.31		27.28		28.40		23.00		27.00	
	SEP	24.15		28.00		28.50		23.20		26.20	
	OCT	24.47		28.91		29.50		24.00		27.20	
	NOV	25.81		30.22		31.10		25.90		27.50	
	DEC	26.92		31.80		32.10		27.30		28.40	
Lead waveforms feature a very steep increase in echo power and therefore short <b>leading-edge widths (LEW)</b>	JAN		1.10		0.95		0.82		0.71		1.00
	FEB		1.06		0.96		0.82		0.73		1.01
	MAR		0.97		0.96		0.82		0.74		1.03
	APR		1.06		0.95		0.82		0.77		1.04
	MAY		1.09		0.95		0.82		0.77		1.06
	JUN		1.13		0.95		0.82		0.77		1.05
	JUL		1.08		0.95		0.82		0.78		1.07
	AUG				0.95		0.82		0.77		1.05
	SEP				0.94		0.82		0.77		1.07
	OCT				0.95		0.82		0.76		1.05

	NOV DEC				0.94 0.93		0.82 0.82		0.74 0.72		1.02 1.00
			1.08 1.12 1.08 1.03 1.07								
Only lead classifications by waveform are expected that fall into regions of sufficient ice cover (checked with <b>SIC</b> in %)	JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC	70 70 70 70 70 70 70 70 70 70 70 70		70 70 70 70 70 70 70 70 70 70 70 70		70 70 70 70 70 70 70 70 70 70 70 70		70 70 70 70 70 70 70 70 70 70 70 70		70 70 70 70 70 70 70 70 70 70 70 70	

**Table 5-5: Metrics for sea-ice surface-type classification of Envisat RA-2, CryoSat-2 SAR mode, and CryoSat-2 SIN mode waveform data for the Antarctic**

Metric	Month	ERS-1 RA		ERS-2 RA		Envisat RA-2		CryoSat-2 SAR		CryoSat-2 SIN	
		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Sea-ice waveforms shouldn't be peaky and therefore have a low <b>pulse peakiness (PP)</b>	JAN		5.21		14.34		24.60		40.10		138.40
	FEB		4.78		12.86		20.70		35.30		126.10
	MAR		4.61		13.65		19.60		32.90		124.90
	APR		5.73		14.05		18.80		30.20		127.30
	MAY		6.98		13.78		17.50		28.70		122.20
	JUN		7.29		13.89		16.90		28.90		121.00
	JUL		6.92		13.91		16.60		28.10		114.90
	AUG		7.03		13.71		16.10		28.00		115.80
	SEP		8.22		13.92		16.30		28.40		114.30
	OCT		7.20		14.13		18.10		29.60		121.20
	NOV		7.31		14.89		20.70		34.10		126.50
	DEC		6.16		15.13		22.80		36.60		135.20
Sea-ice waveforms are also characterized by low <b>backscatter</b> values due to diffuse reflection ( <b>SIG0</b> in dB)	JAN	8.98	16.17	13.54	26.86	2.5	27.20	2.5	26.30	2.5	26.40
	FEB	9.14	15.20	12.87	25.61	2.5	25.40	2.5	24.10	2.5	25.10
	MAR	8.94	15.63	14.20	25.40	2.5	26.70	2.5	25.10	2.5	27.60
	APR	8.57	15.15	15.47	25.03	2.5	27.20	2.5	26.20	2.5	27.30
	MAY	8.61	16.54	15.93	25.07	2.5	24.60	2.5	23.10	2.5	24.90
	JUN	8.44	16.88	15.25	24.33	2.5	23.10	2.5	20.90	2.5	24.20
	JUL	8.00	15.73	14.58	23.05	2.5	22.50	2.5	20.20	2.5	24.10
	AUG	8.19	15.97	14.96	23.22	2.5	21.70	2.5	19.10	2.5	24.90
	SEP	8.78	17.45	15.03	23.61	2.5	22.30	2.5	20.00	2.5	23.70
	OCT	8.11	16.53	15.65	24.33	2.5	23.30	2.5	20.60	2.5	25.00
	NOV	8.02	17.02	16.28	26.05	2.5	25.20	2.5	22.90	2.5	25.20
	DEC	8.69	16.82	16.57	28.01	2.5	26.10	2.5	23.90	2.5	25.00
Sea-ice waveforms feature a less steep increase in echo power	JAN	0.94		0.85		0.78		0.87		1.31	
	FEB	0.96		0.86		0.80		0.95		1.40	
	MAR	0.98		0.85		0.80		0.98		1.37	
	APR	0.95		0.84		0.80		1.02		1.34	
	MAY	0.90		0.84		0.81		1.07		1.37	
	JUN	0.91		0.84		0.80		1.07		1.38	



and therefore longer leading-edge widths (LEW)	JUL	0.91		0.84		0.80		1.12		1.41	
	AUG	0.91		0.84		0.81		1.13		1.41	
	SEP			0.84		0.81		1.11		1.42	
	OCT	0.90		0.84		0.80		1.08		1.38	
	NOV	0.92		0.84		0.79		0.95		1.36	
	DEC	0.93		0.84		0.78		0.92		1.33	
			0.94								
Only sea-ice classifications by waveform are expected that fall into regions of sufficient ice cover (checked with SIC in %)	JAN	70		70		70		70		70	
	FEB	70		70		70		70		70	
	MAR	70		70		70		70		70	
	APR	70		70		70		70		70	
	MAY	70		70		70		70		70	
	JUN	70		70		70		70		70	
	JUL	70		70		70		70		70	
	AUG	70		70		70		70		70	
	SEP	70		70		70		70		70	
	OCT	70		70		70		70		70	
	NOV	70		70		70		70		70	
	DEC	70		70		70		70		70	

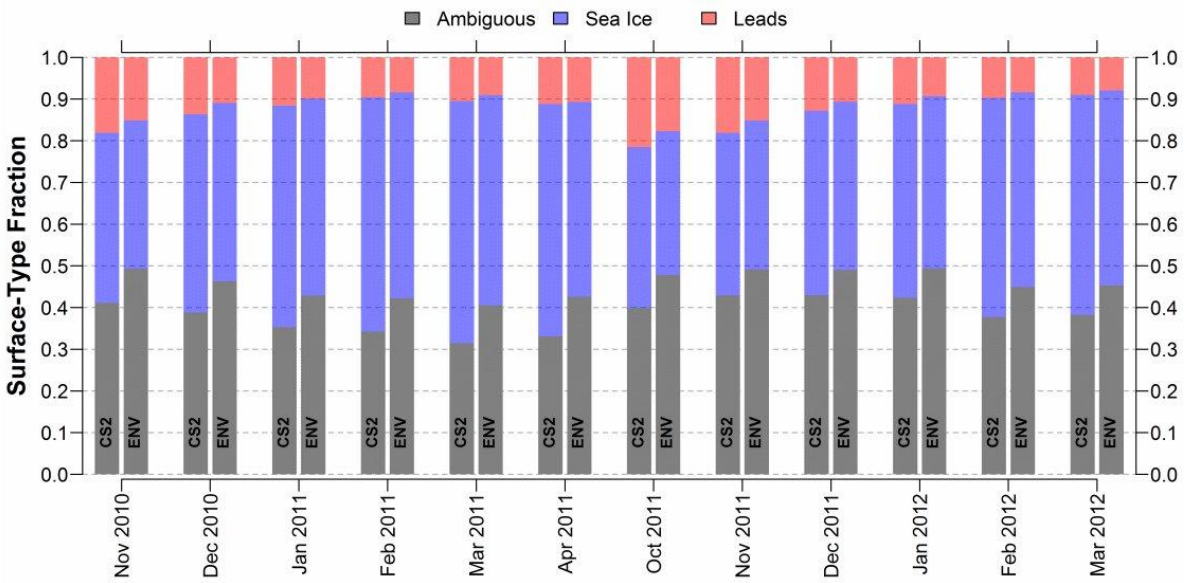
### 5.1.2 Results

Utilizing this new and sensor-consistent surface-type classification scheme results in overall much better agreement between ERS-2 RA, Envisat RA-2, and CryoSat-2 for typical benchmarks in their respective mission-overlap periods.

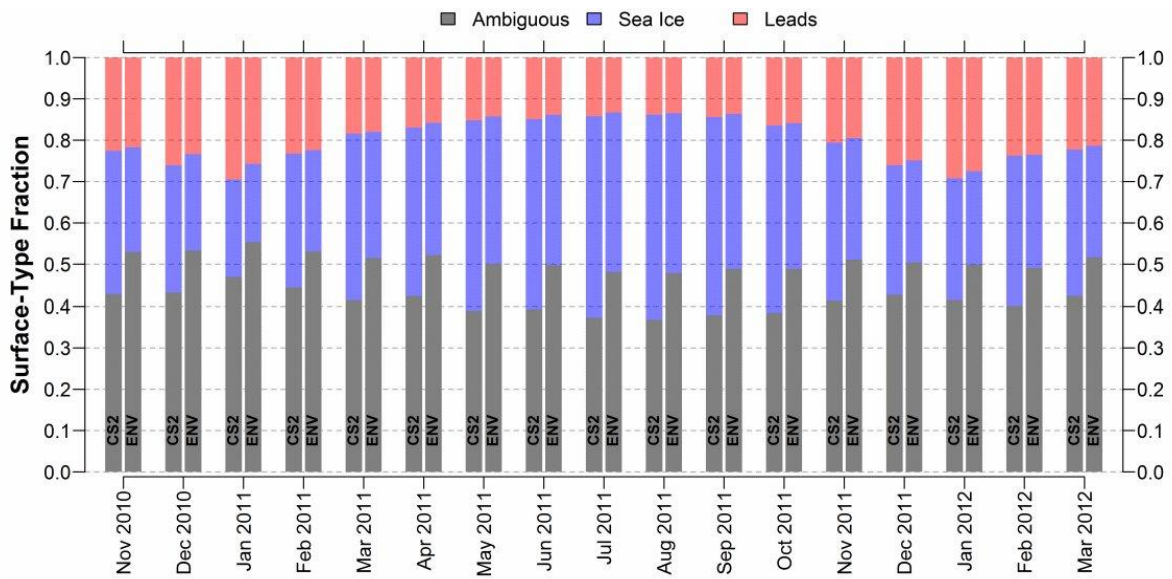
In general, this approach allows for substantially more wave forms being classified as either lead or sea-ice type that were otherwise rejected before. With respect to CCI Phase 1, where there was a very high fraction of lead detections compared to only a very small fraction of classified sea-ice type waveforms, the spatial patterns and distributions of these occurrences are now better in line with what one would expect for Envisat RA-2. Furthermore, the intermission consistency for the Arctic as well as the Antarctic has improved substantially (Figure 5-2 and Figure 5-3; Figure 5-4 and Figure 5-5). A similar effect can be seen for ERS-1/2 RA. However, there is a substantial drop in identified 'sea-ice'-type waveforms for ERS-2 RA (about a third compared to Envisat RA-2, Figure 5-6), leading to an overall still rather low valid fraction of 'valid waveforms (Figure 5-7). The fraction of 'sea-ice'-type waveforms appears to be lowest around the central Arctic, hinting on limitations of the current procedure on handling rougher multi-year sea-ice waveform types. This is still an ongoing task for a later and final iteration of this document.

The overall increased number of valid waveforms has an additional positive side effect on the overall data record: It allows for a much higher spatial resolution to be used in the final gridded Level 3 product without any compromises on overall coverage. Here, we are able to provide a 25 km resolution gridded data set for the Arctic and Antarctic for all sensors.

Direct comparisons of surface-type class fractions (i.e., either ambiguous, lead, or sea ice) over the course of the sensor overlap period reveal an overall very good agreement between CryoSat-2 and Envisat RA-2 (Figure 5-2 and Figure 5-3). While the fraction of lead- and sea-ice waveforms is on average slightly smaller for Envisat RA-2 than for CryoSat-2 (about 8 % for the Arctic and 10 % for the Antarctic), both sensors show a similar seasonal development in both hemispheres.



**Figure 5-2: Time-series of surface-type fractions for the sensor overlap period between CryoSat-2 (CS2) and Envisat RA-2 (ENV) for the Arctic**



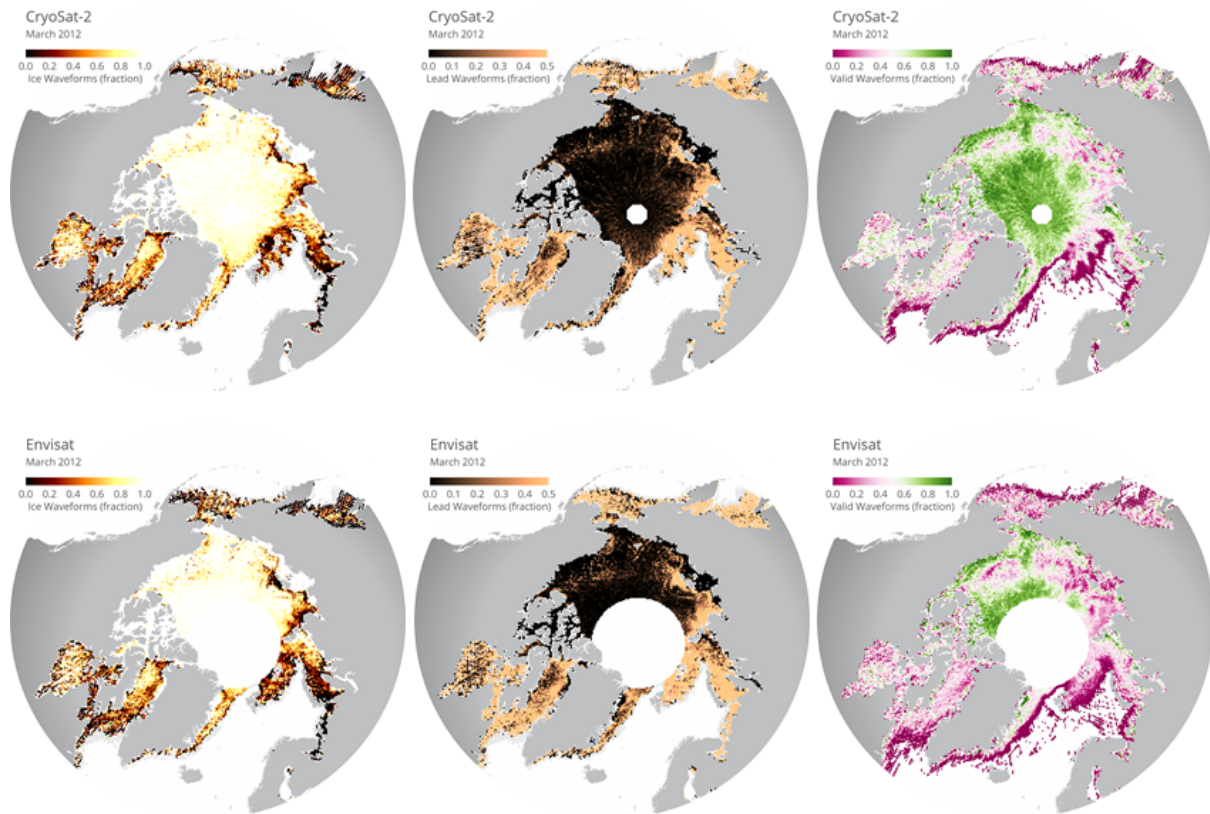
**Figure 5-3: Time-series of surface-type fractions for the sensor overlap period between CryoSat-2 (CS2) and Envisat RA-2 (ENV) for the Antarctic**

This seasonal change in the present sea-ice cover is also apparent from the derived surface-type class thresholds (Table 5-1 - Table 5-5). During summer months (Antarctic) and the early winter (Arctic), the number of lead waveforms is higher and returns from new and young ice tend to be more specular, which results in higher maximum thresholds in sea-ice backscatter and pulse peakiness. This observed seasonal shift in the distributions of the classifiers will also play an important role in the description of the new retracker scheme.

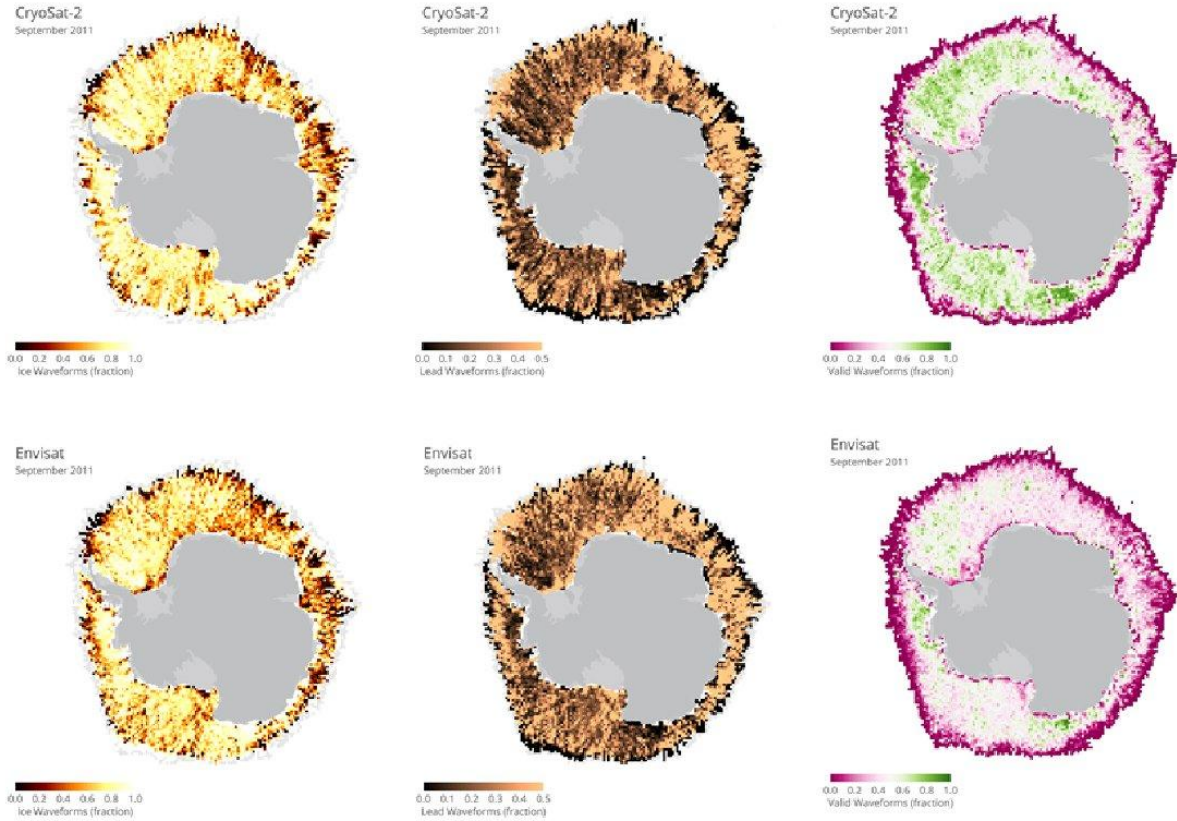
An exemplary visualization of monthly map-wise inter-comparisons between Envisat RA-2 and CryoSat-2 based on the benchmarks of valid-, lead-, and sea-ice fraction is shown in Figure 5-4 and Figure 5-5. In these gridded data sets, the overall good agreement is

confirmed. However, there are small differences, and as mentioned earlier, slightly smaller valid fractions for Envisat RA-2. This behavior is expected and results most likely from the much larger footprint of Envisat RA-2, especially in regions with high rates of sea-ice dynamics such as the Beaufort Sea, but also in the Laptev Sea. Here, the increased surface-type mixing likely prevents a clearer separation between waveform types.

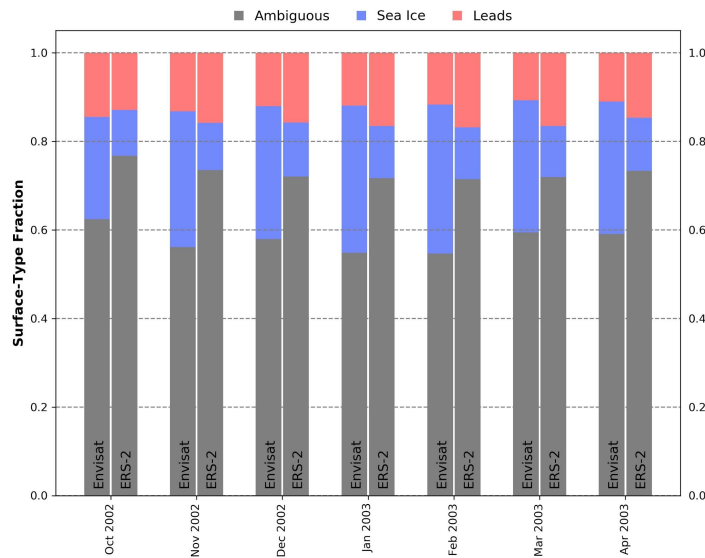
Nevertheless, both comparisons highlight the overall good agreement that could be achieved between both sensors with this new surface-type classification scheme and the chosen settings during the sensor overlap period. These results therefore lay the foundation for a proper inter-mission sea-ice freeboard and sea-ice thickness data record.



**Figure 5-4: Visualizations of monthly sea-ice fraction, lead fraction, and valid fraction benchmarks for the Arctic (March 2012)**

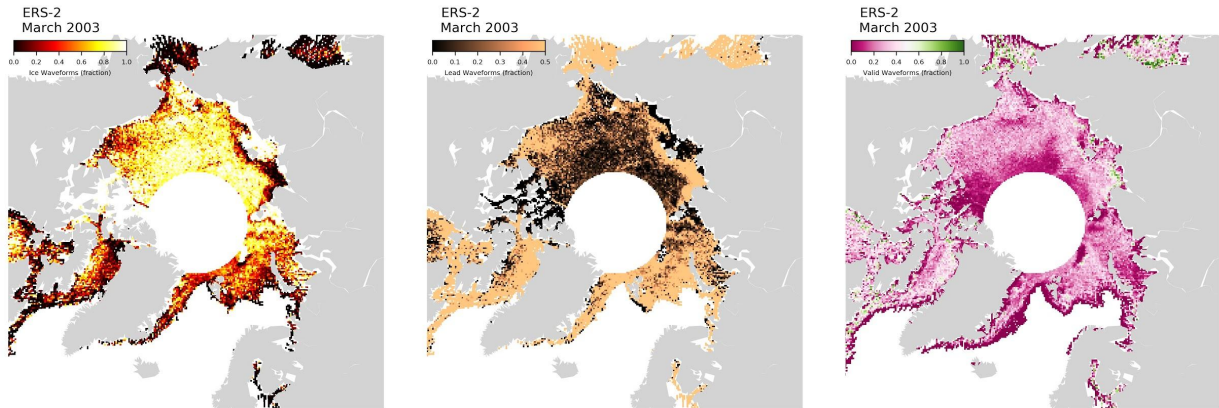


**Figure 5-5: Visualizations of monthly sea-ice fraction, lead fraction, and valid fraction benchmarks for the Antarctic (September 2011)**



**Figure 5-6: Time-series of surface-type fractions for the sensor overlap period between Envisat RA-2 (Envisat) and ERS-2 RA (ERS-2) for the Arctic**





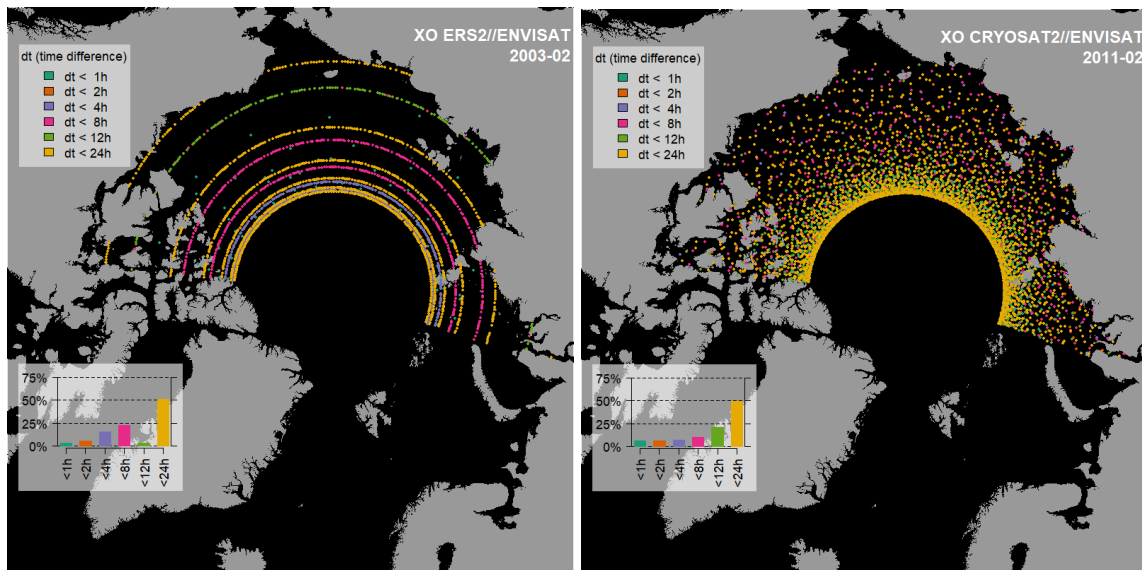
**Figure 5-7: Visualizations of monthly ERS-2 sea-ice fraction, lead fraction, and valid fraction benchmarks for the Arctic (March 2003)**

### 5.1.3 General Remarks

The implementation of ERS-1/2 RA data poses a more difficult task than Envisat RA-2 during CCI Phases 1 and 2 due to differences in instrument parameters (e.g., bin width and bin number, range of noise level, size of waveform artifacts), all for the worse. All of this led to the implementation and preliminary introduction of a more sophisticated approach to evaluate instruments and resulting waveform differences in a spatiotemporally more direct way using inter-mission orbit crossovers.

## 5.2 Dual Mission Orbit Cross Overs

The basis for the retracker tuning of ERS-2, and in the future also Envisat RA-2, as well as the surface-type classification of all available sensors, are dual mission orbit crossovers (DMXO) for their respective overlap periods. All DMXO utilize waveforms acquired per sensor within a 25km radius around the actual orbit cross over (XO) in the central Arctic and the marginal seas East of 70°E and West of 95°W and above a latitude of 70°N (Figure 5-8). For the best possible results all analysis is limited to XOs that occurred within a maximum time difference of two hours.



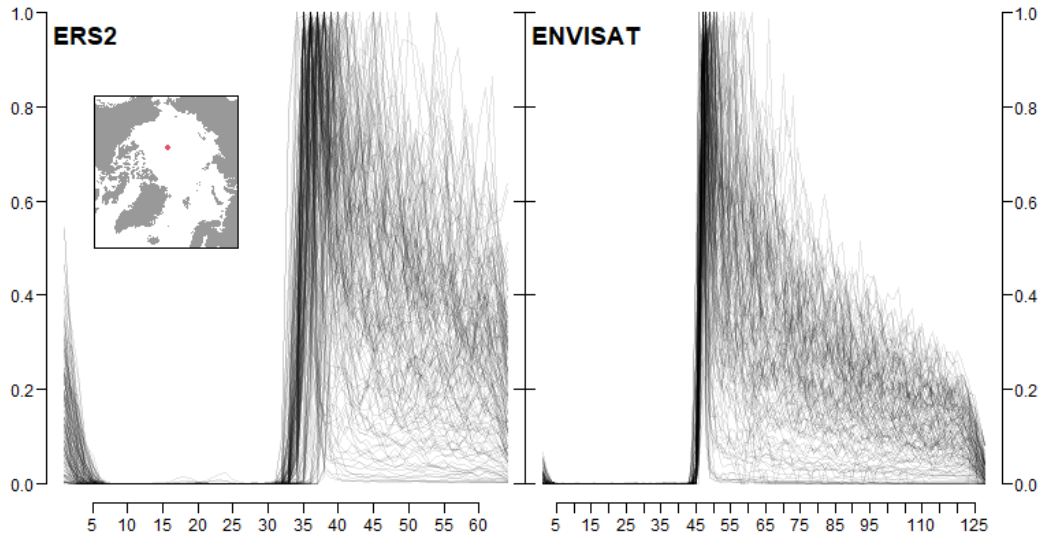
**Figure 5-8: Distribution of orbit crossovers between ERS2 and Envisat for February 2003 (left) and CryoSat-2 and Envisat for February 2011 (right) for various crossing time differences.**

Based on the examples shown in Figure 5-8, this results in 95 (ERS2/Envisat) and 274 (CryoSat-2/Envisat) XOs within 1 hour time difference, respectively and a total of 1065 (ERS2/Envisat) and 2577 (CryoSat-2/Envisat) for the respective mission overlap periods.

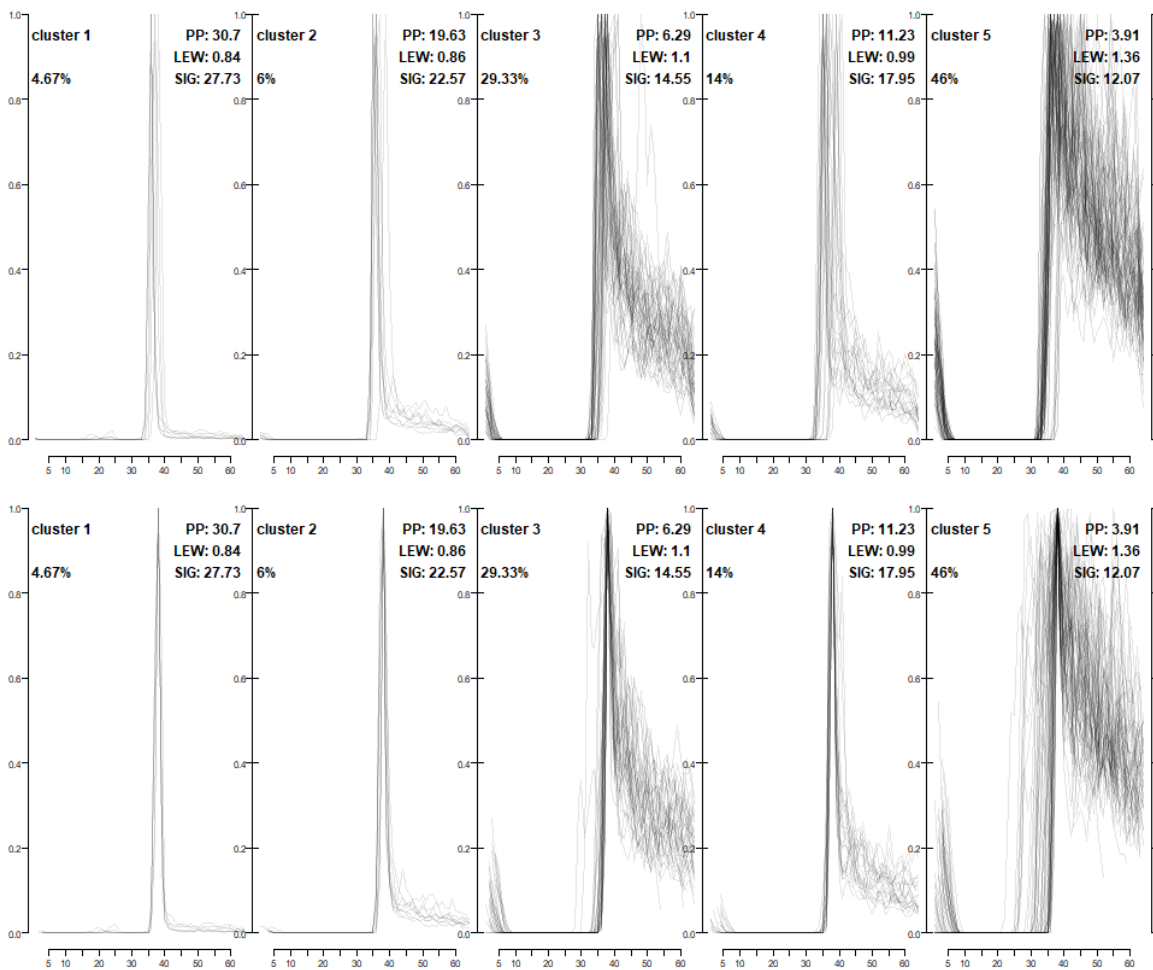
Working with these XOs allows for a i) more detailed investigation into waveform parameters and limitations of current implementations as well as ii) more direct possibility to tune (and retune) older sensors to CryoSat-2 whenever necessary as the data basis is reduced compared to the full available data record of all sensors.

Figure 5-9 features waveform stacks of ERS-2 and Envisat RA-2 for a single XO in the central Arctic. Here, the overall higher noise in the ERS-2 waveforms is apparent as well as the in-parts larger artifacts in the first few bins compared to those of Envisat RA-2.

Through clustering (example in Figure 5-10) and subsequent training of a machine-learning based classification algorithm (e.g., Random Forests; Breiman, 2001) and its direct implementation into the processing software we expect to achieve better surface-type classification results in the future for all sensors as well as making threshold-based classifications obsolete.



**Figure 5-9: Waveform stack for an example XO between ERS-2 and Envisat in February 2003.**



**Figure 5-10: Clustered ERS-2 waveform stack of the same example XO between ERS-2 and Envisat in February 2003 as shown in Figure 5-9 (left). Top row shows the raw waveforms; bottom row waveforms aligned by their detected first-maximum index.**

In the current implementation, our adaptive threshold retracker algorithm is trained based on the identified XOs for each month of the overlap period between ERS-2 and Envisat to achieve a better overall match between both sensors than the previous implementation of tuning based on the gridded data products. In the future, this approach shall be adopted for all available sensors. Details shall be provided in the respective sub section.

## 5.3 Retracking

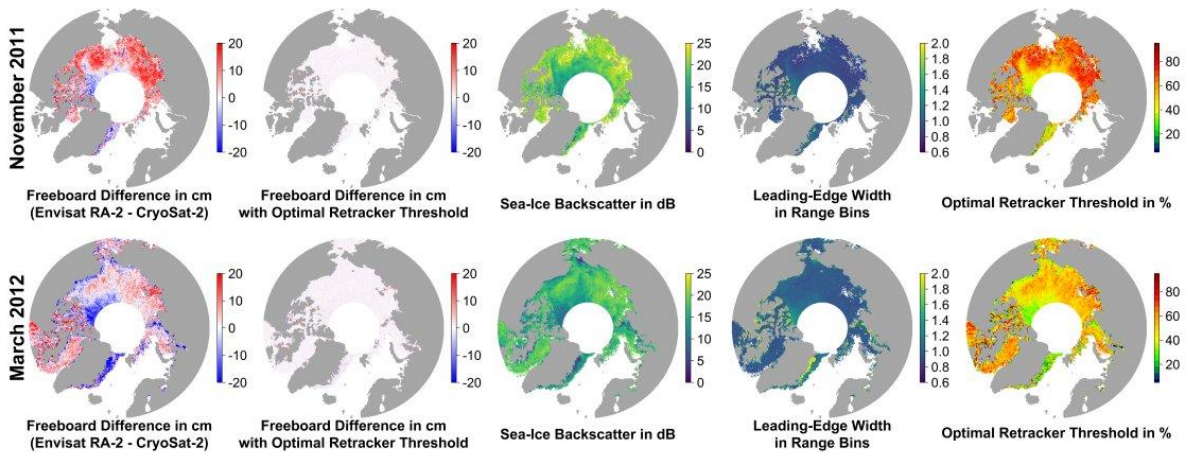
### 5.3.1 Procedure Description

The range retrieval algorithm for Envisat RA-2 and CryoSat-2 waveforms is identical for sea-ice and lead waveforms. Ocean waveforms are currently discarded. The used Threshold First Maximum Retracker Algorithm (TFMRA, Helm et al., 2014; Ricker et al, 2014) is based on the following steps:

- a) Estimate the noise level as the average of the first 5 bins of the waveform. However, in case of Envisat RA-2 we are following UCL's implementation and discard the counts in the first 5 bins of the echo as these just contain artefacts of the FFT.
- b) Oversampling of the echo waveforms by a factor of 10 using linear interpolation
- c) Smoothing of the oversampled waveform with a window filter size of 11 range bins
- d) Locating the first local maximum of the waveform: Must be higher than noise level + 15% of absolute peak power.
- e) Obtain the range value at a specified threshold of the power of the detected first maximum, by linear interpolation of the smoothed and oversampled waveform.

Continuing on the last point, the choice of retracker threshold is pivotal for the range estimation. Following AWI's implementation for CryoSat-2 (Ricker et al., 2014), we keep a consistent threshold of 50% from the first maximum peak power both for leads and sea-ice waveforms. For pulse-limited altimetry such as for Envisat RA-2, retracking near the maximum power for leads proved to be essential to retrieve reasonable freeboard estimates later on. Therefore, a threshold of 95% was chosen for leads from Envisat RA-2 waveforms. However, using a single fixed threshold of, e.g., 50% similar to CryoSat-2, results later on in sea-ice freeboard estimates that feature an overall smaller variation than CryoSat-2 estimates. Furthermore, expected thin-ice regions feature ice that is too thick and vice versa. We relate this behavior to the much larger footprint and therefore increased mixing of surface types of different surface-roughness scales in every Envisat RA-2 waveform.





**Figure 5-11: Visualizations of two monthly sets of figures (from left to right): Freeboard difference between Envisat RA-2 and CryoSat-2, the best achievable freeboard difference using an optimal retracker threshold, the sea-ice backscatter, the leading-edge width, and the iteratively estimated optimal threshold for November 2011 (top row) and March 2012 (bottom row)**

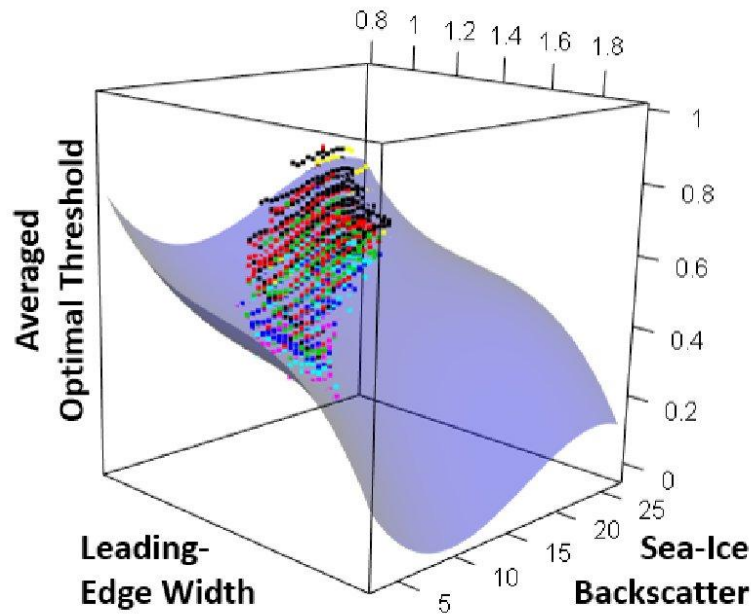
From Figure 5-11 it appears that differences in sea-ice freeboard are related to differences in the waveform parameters of sea-ice backscatter and leading-edge width (as well as pulse peakiness, which is strongly correlated with sea-ice backscatter, but is not shown here). Areas of potential multi-year ice near the Canadian Archipelago and areas influenced by multi-year ice export are in general substantially too thin (e.g., about 20 cm and more in March), whereas areas of predominantly first-year sea ice are in general too thick in the Envisat RA-2 data. However, the level of freeboard difference appears to be seasonal, where Envisat RA-2 appears to be unable to keep track of these seasonal changes.

As these differences in sea-ice freeboard between CryoSat-2 and Envisat RA-2 appear to be indeed strongly correlated to patterns in the sea-ice backscatter and the leading-edge width of Envisat RA-2 waveforms, we decided to apply a tuning scheme by computing an adaptive range retracker threshold as a function of sea-ice backscatter and the leading-edge width to mitigate the differences. Due to the already mentioned larger footprint of Envisat RA-2 and hence increased mixing of different surface types, it appears to be necessary to treat waveforms differently according to the waveform shape (and hence surface properties) by means of retracking the main scattering horizon.

In order to derive the functional relationship between threshold and sea-ice backscatter/leading-edge width, we first processed all Envisat RA-2 for the complete sensor overlap period. This processing was done using the TFMRA with a fixed threshold for leads of 95 % and a threshold for sea-ice waveforms that was changed in each run. This sea-ice threshold ranged between 5 % and 95 % in steps of 5 %. For example, in the first run the complete data set was processed using a retracker threshold of 5 % for sea-ice waveforms and the resulting sea-ice freeboard was calculated. In the next run, a fixed threshold of 10 % was used for all sea-ice waveforms and so on, until the last run with a sea-ice threshold of 95 % was computed and the resulting sea-ice freeboard was calculated.

From this data set, the optimal threshold, i.e., the threshold that yields the smallest difference in freeboard between Envisat RA-2 and CryoSat-2, was iteratively derived. An exemplary result is also shown in Figure 5-11. Again, also the optimal thresholds reflect the seasonal change in waveform parameters with a varying range of optimal threshold values that are in general higher for the early winter than the values in late winter.

● Oct ● Nov ● Dec ● Jan ● Feb ● Mar ● Apr



**Figure 5-12: Visualizations of averaged binned optimal threshold values on an x-y plane of leading-edge width and sea-ice backscatter for the Arctic. The blue plane is the 3<sup>rd</sup> order polynomial fit through all data points**

Next, average optimal threshold values were calculated for each 0.25 dB sea-ice backscatter and 0.025 leading-edge width bin on an x-y plane. A 3D visualization of this is shown in Figure 5-12. For months November through March both occurrences in the sensor overlap period were used. October and April, which were only covered once during the sensor overlap period were each added twice to circumvent issues of underrepresentation in their number of data values added to the total.

Through this compilation of monthly data points, three 3<sup>rd</sup> order polynomial planes were fitted based on different weighting schemes in order to maximize the adjusted  $R^2$ . As weights we used either the number of optimal threshold values per bin in the x-y plane, the inverse standard deviation of all optimal threshold values per bin ( $1/\sigma$ ), or no weights at all.

For the Arctic, the result shown in Figure 5-12 is based on the inverse standard deviation as weights and achieved an adjusted  $R^2$  of 0.94. All shown data points have a minimum of 50 occurrences and were obtained in the central Arctic only.

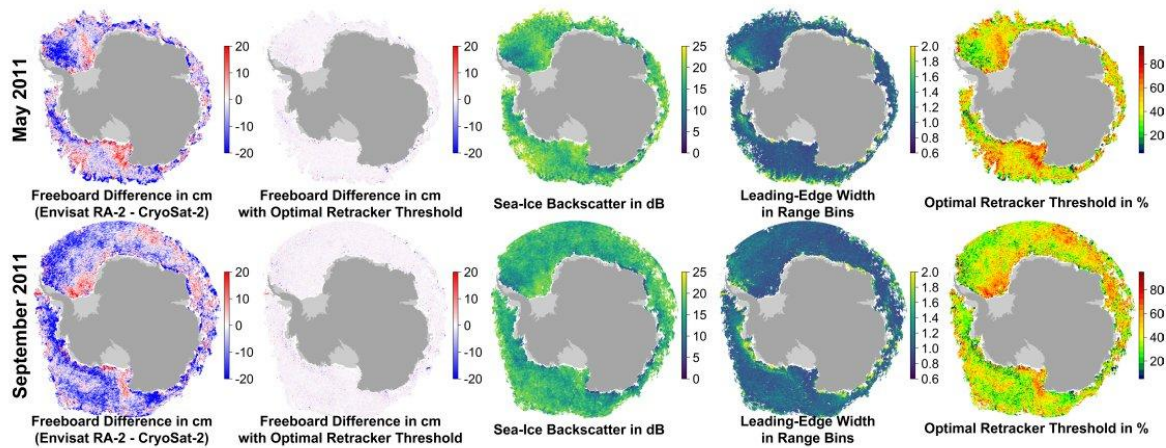
In Figure 5-12, the seasonal shift is also present: Early winter months tend towards shorter leading-edge widths and higher sea-ice backscatter values (October in yellow and November in Black), whereas late-winter months feature longer leading-edge widths and lower backscatter values.

The optimal threshold ( $th_{opt}$ ; in decimal values) to be used in the adaptive range retracking as a function of sea-ice backscatter ( $\sigma^0$ ) and leading-edge width ( $lew$ ) is given by the following equation:

$$th_{opt} = k_1 - k_2 \times lew + k_3 \times lew^2 - k_4 \times lew^3 - k_5 \times \sigma^0 + k_6 \times \sigma^2 - k_7 \times \sigma^3$$

Where  $k_1 = 3.4775697362$ ,  $k_2 = 5.9296875486$ ,  $k_3 = 4.3516498381$ ,  $k_4 = 1.0933131955$ ,  $k_5 = 0.0914747272$ ,  $k_6 = 0.0063983796$ ,  $k_7 = 0.0001237455$ .

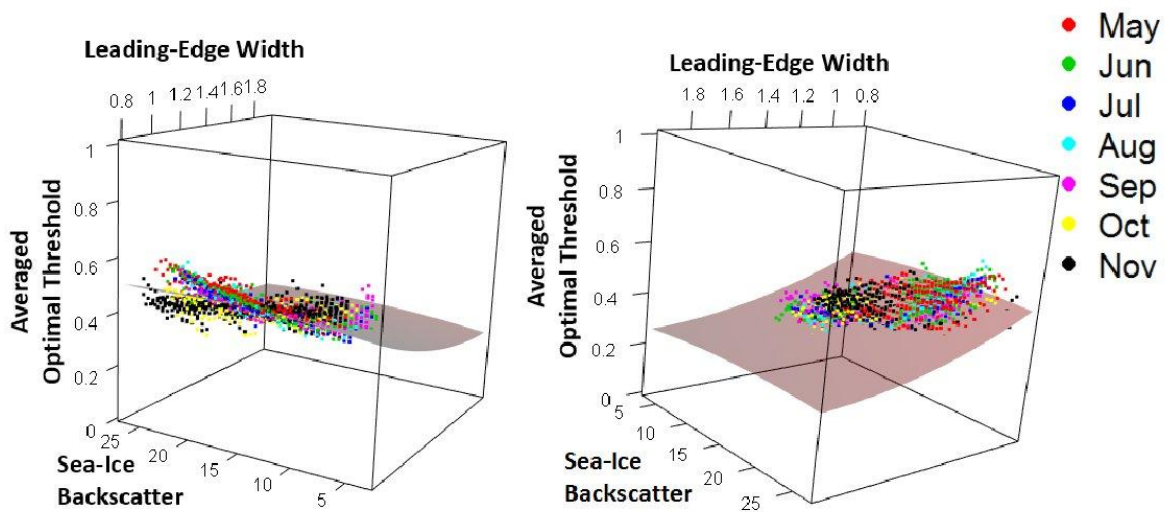
In a first attempt, we applied the same equation that was derived from the northern hemisphere data also to the southern hemisphere. However, this did not improve the results. The reason for that can partly be seen in Figure 5-13. In contrast to the Arctic, the differences between early and late winter is less prominent in the sea-ice freeboard differences as well as the optimal-threshold values. Additionally, patterns in sea-ice backscatter and leading-edge width are less correlated in some areas. This is potentially related to surface flooding and/or large fast-ice areas with a different snow stratigraphy.



**Figure 5-13: As Figure 2-12 but for the Antarctic showing May 2011 (top row) and September 2011 (bottom row)**

For the Antarctic, a 2nd order polynomial fit resulted in the best statistical result (adjusted  $R^2$  of 0.77) to describe the optimal threshold as a function of leading-edge width and sea-ice backscatter (Figure 5-13).

For the Antarctic, the result shown in Figure 5-14 is based on the number of optimal threshold values per bin as weights. All shown data points also have a minimum of 50 occurrences and were obtained by excluding the marginal ice zones of the Antarctic as well as the austral summer months. However, compared to the Arctic, there is a much larger spread between months.



**Figure 5-14: As Figure 5-12 but for the Antarctic and captured from two different viewpoints.**

The equation to be used for deriving the optimal threshold (in decimal values) in the Antarctic adaptive range retracking as a function of sea-ice backscatter ( $\sigma^0$ ) and leading-edge width ( $lew$ ) is stated below:

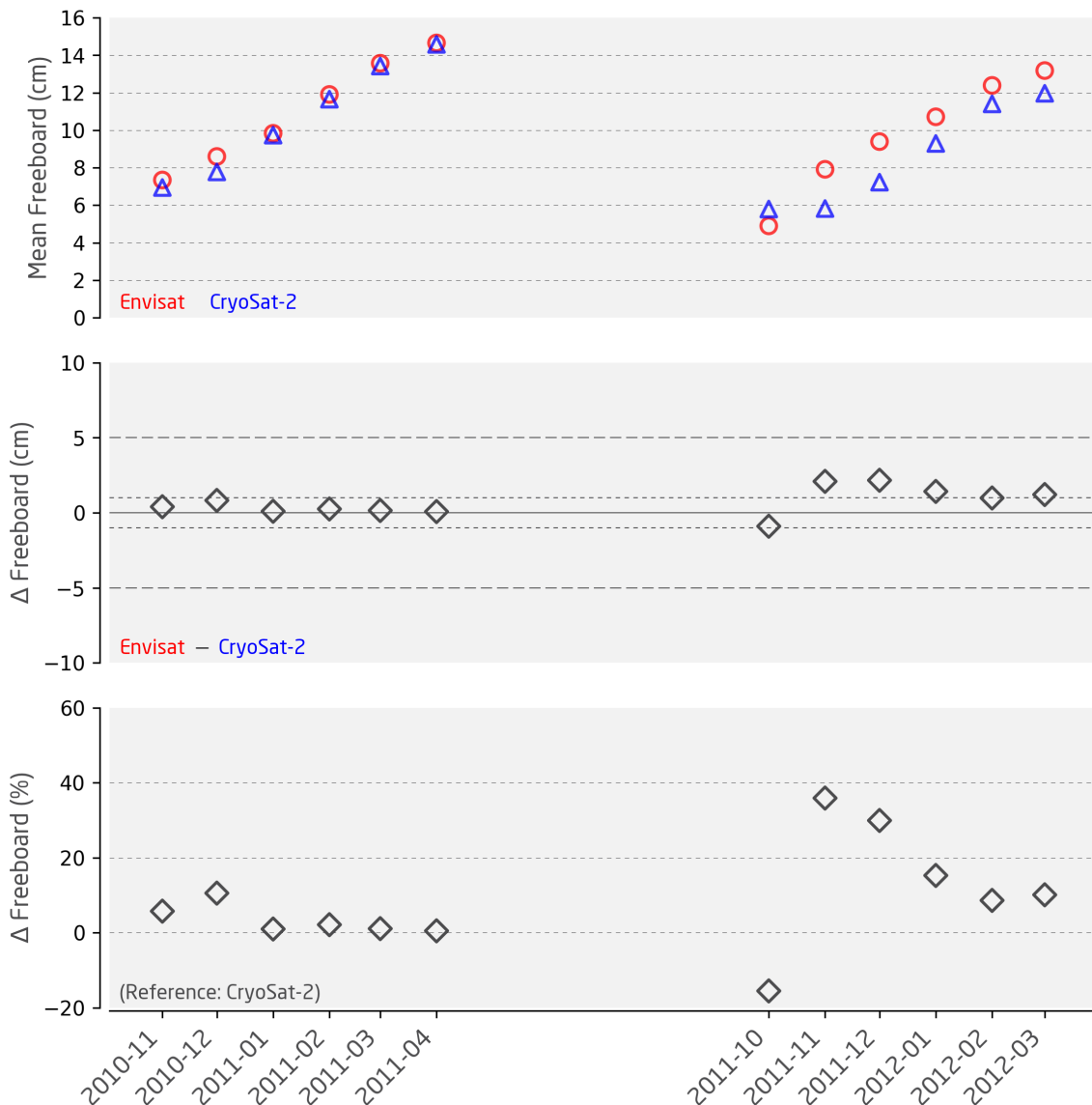
$$th_{opt} = k_1 - k_2 \times lew + k_3 \times lew^2 + k_4 \times \sigma^0 - k_5 \times \sigma^{0^2}$$

Where  $k_1 = 0.8147895184$ ,  $k_2 = 0.5555823623$ ,  $k_3 = 0.1347526920$ ,  $k_4 = 0.0055934198$ ,  $k_5 = 0.0001431595$

Utilizing both equations, for each retracking of each sea-ice waveform, the to-be-used threshold is calculated from the waveform-associated sea-ice backscatter and leading-edge width value. This threshold is then believed to yield the mean-scattering surface in accordance with CryoSat-2 measurements.

### 5.3.2 Results

Here, we want to show and discuss some of the results using the adaptive threshold retracker for Envisat RA-2 in the sensor overlap period. For the Arctic, Figure 5-15 shows the average freeboard in centimeters per month, the average freeboard difference in centimeters as well as percent during the sensor overlap period for Envisat RA-2 and CryoSat-2. While in the first winter season, the match is nearly perfect with absolute average freeboard differences below one centimeter, the second winter season shows larger differences.

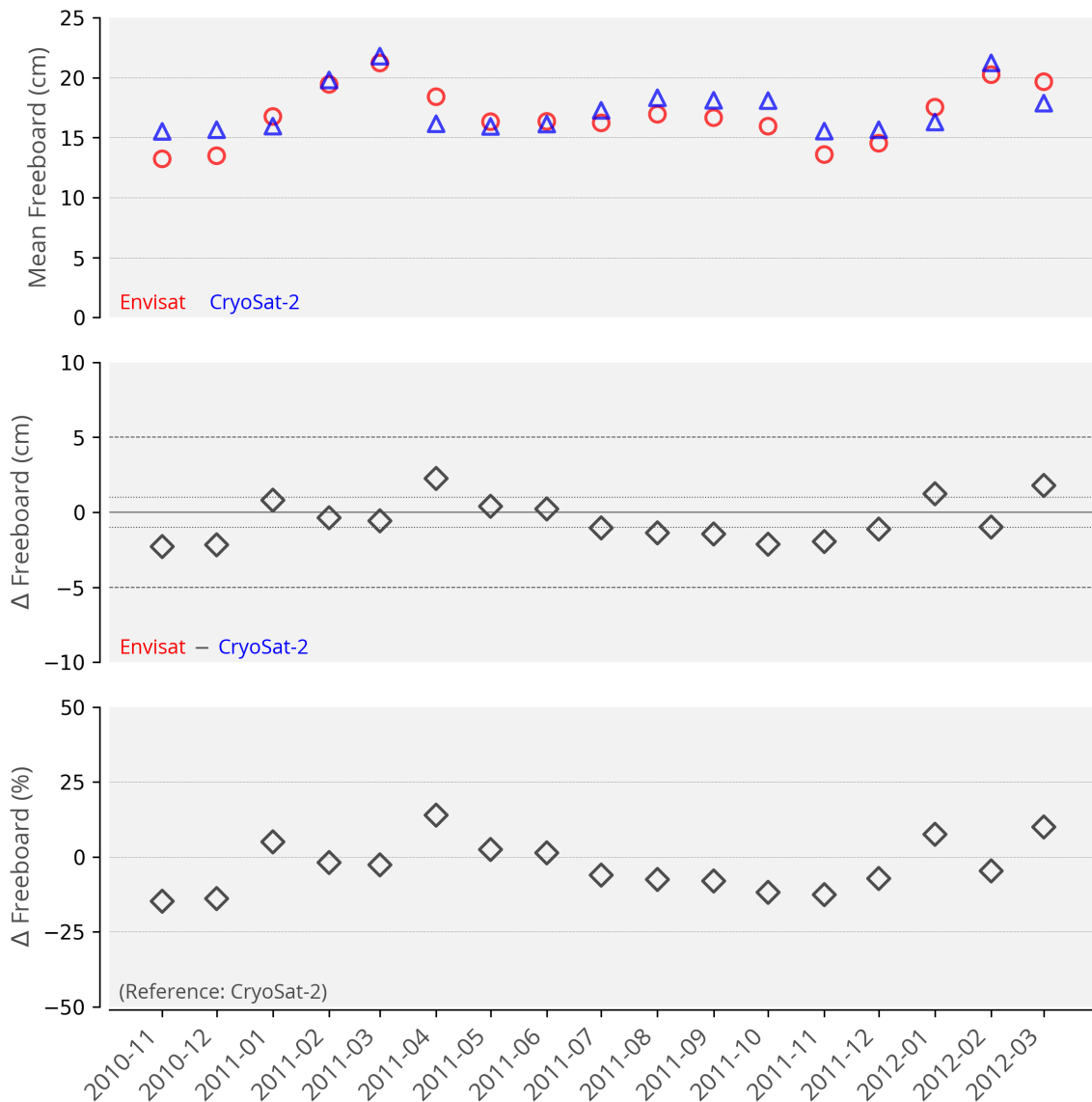


**Figure 5-15: Mean freeboard for each month of the sensor overlap period (top) for Envisat RA-2 (red) and CryoSat-2 (blue) and the corresponding mean freeboard difference between both sensors in centimetres (middle) and percent with reference to CryoSat-2 (bottom) for the Arctic**

However, these differences are still below three centimeters, which is a significant improvement over phase 1. Especially for the Arctic spring period (March & April), differences in average freeboard are 1.2 cm or better. The stability, i.e., the range of monthly differences, is 3.1 cm.

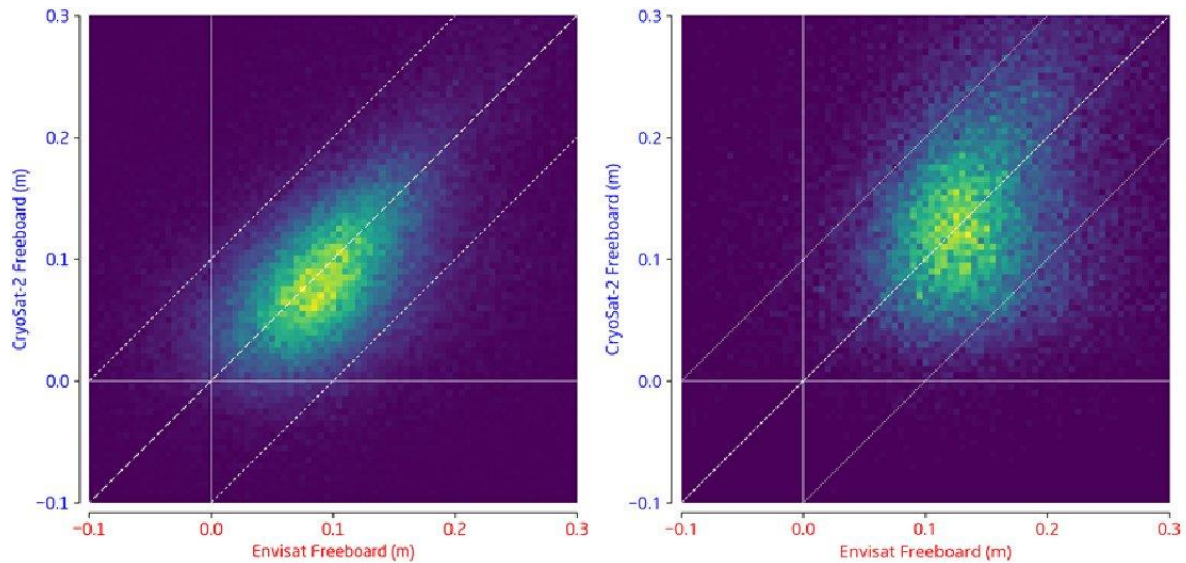
For the Antarctic (Figure 5-16), results are not as good as for the Arctic. Overall the algorithm has less skill to match Envisat RA-2 freeboards to the ones of CryoSat-2. This is very likely related to other physical processes such as more prominent snow stratigraphy and surface flooding. However, issues causing these differences are subject to further investigation. Overall, there is a stronger seasonality in the differences between summer and winter, which also leads toward a higher range of monthly differences of 4.6 cm.





**Figure 5-16: Setup as in Figure 5-15 but for the Antarctic**

Putting all gridded freeboard values of Envisat RA-2 and CryoSat-2 against each other underlines these observations (Figure 5-17). While the algorithm is able to achieve very good agreement for the Arctic (Figure 5-17, left), the results are slightly more diffuse for the Antarctic (Figure 5-17, right).



**Figure 5-17: Scatterplot of all gridded freeboard estimates of CryoSat-2 (y-axis) vs. Envisat RA-2 (x-axis) for the Arctic (left) and the Antarctic (right).**

Due to the overall changes to the processing scheme compared to CCI Phase 2 to implement the ERS-1/2 RA sensor family based on DMXOs, the adaptive retracker parameters  $k_0 - k_6$  (for the calculation of optimal threshold) will be derived the the resulting evaluations incorporated into the next and final iteration of this document.

### 5.3.3 Envisat Backscatter Drift Correction

Over the course of Envisat's life span, it appears that the RA-2 instrument has been degraded (Helm, 2017, pers. comm.). This results in a slight linear reduction in received backscatter over the years (Figure 5-18). As this can affect both the surface-type classification as well as the range retracking (as both are dependent on the received sea-ice backscatter), a correction had to be applied.

The monthly degradation factor of -0.003269253 was derived from the monthly averages of ocean-type waveforms in the Barents Sea (70°N-75°N and 40°E-50°E). Ocean-type waveforms are derived independent from the sea-ice backscatter classifier and we assume the surface roughness sufficiently random compared to ice-type waveforms for our analysis.

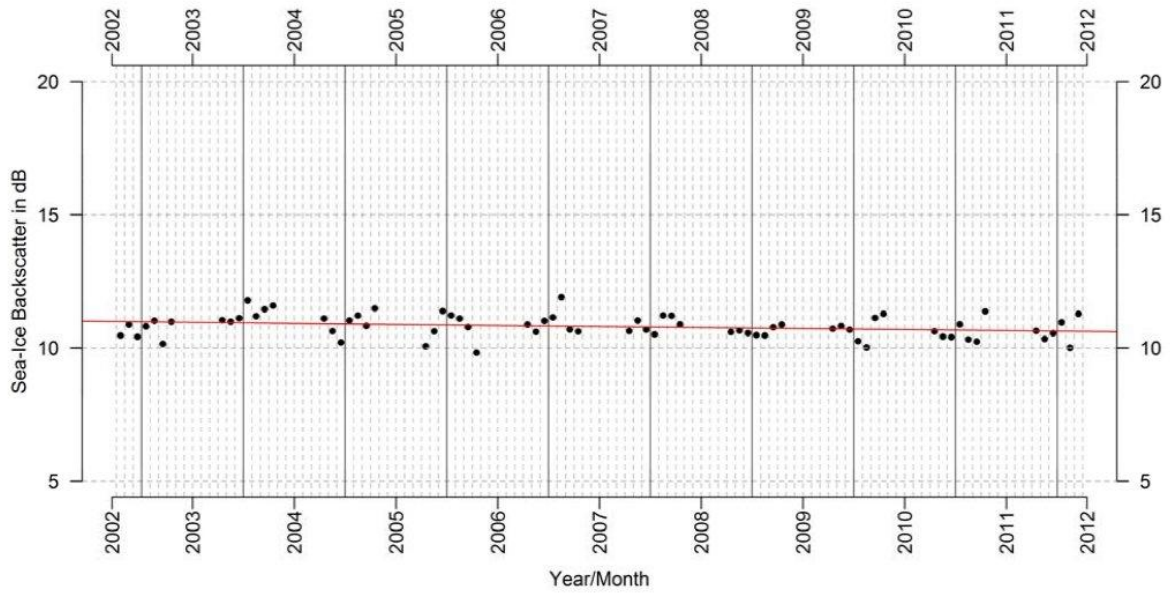
As the surface-type classification as well as the range retracking was derived from data in the sensor overlap period (November 2010 to March 2012), all backscatter values had to be corrected towards this base period. In order to accomplish that we picked June 2011 as a reference point.

Using the below given formula, we calculated the necessary backscatter drift correction:

$$t_{sift} = 12 \times (\alpha_{ref} - \alpha) + (m_{ref} - m)$$

$$\sigma_{drift}^0 = -0.003269253 \times t_{sift}$$

Here,  $t_{sift}$  is the time shift factor in months between the reference year ( $a_{ref}$ ) and month ( $m_{ref}$ ) and the currently processed year ( $a$ ) and month ( $m$ ). The resulting backscatter drift correction  $\sigma_{drift}^0$  is then added to the sea-ice backscatter before the surface-type classification and the range retracking. By doing so, the in general slightly higher backscatter values during earlier years of Envisat's lifespan are reduced to the level during the sensor overlap period.



**Figure 5-18: Visualizations of the monthly averaged sea-ice backscatter reduction between 2002 and 2012 over ocean-type waveforms obtained between 70°N-75°N and 40°E-50°E**

### 5.3.4 Pulse Deblurring

A significant challenge in ERS retracking is the pulse blurring which results from the range window moving during waveform averaging sequence, causing short wavelength noise in the retracked elevations. For a detailed description of pulse blurring, see Peacock and Laxon (2004). Due to a strong linear trend between the height error signal  $\varepsilon$  and the retracked elevation, a correction can be applied to the elevation measurements as a function of  $\varepsilon$ . A linear correction to the retracked range is applied if the height error signal  $\varepsilon < 0$ . We utilise the `htl_disc_out` variable from the REAPER SGDR files, convert this into  $\varepsilon$  in meter and apply the correction as introduced in Peacock (1998), and revisited in ESA (2021):

$$h_{corr} = h_{rtk} - \frac{\varepsilon}{m}$$

where  $h_{corr}$  is the corrected height calculated from the retracked height  $h_{rtk}$ , HTL error signal  $\varepsilon$  and slope of the trend  $m$ . The slope term is assumed to be constant for all  $\varepsilon$ , and based on simulations of Peacock (1998) it is -3.5 for ERS-1 and -5.0 for ERS-2.



## 5.4 Geophysical Range Correction

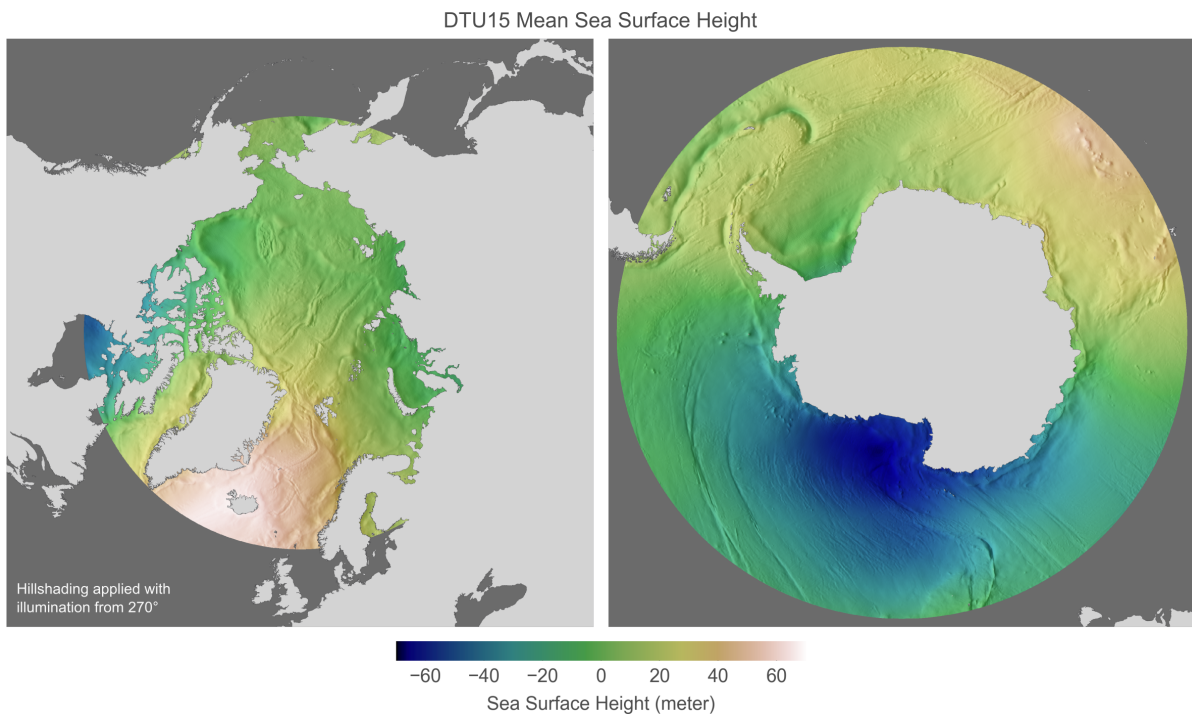
The range is corrected for the changes in sea level due to tides and atmospheric pressure. The specific geophysical range corrections are:

- Elastic ocean tide
- Geocentric polar tide
- Long-period ocean tide
- Solid earth tide
- Inverse barometric correction

## 5.5 Radar Freeboard and Sea-Surface Height

The vast majority of the signal seen in the floe and lead elevations retracked in the last section is caused by unevenness in the Earth's gravity field and mean circulation of the ocean currents. This fixed signal known as the mean sea surface must be removed before any interpolation of the sea surface heights is attempted. Many models of the mean sea surface are available and there will almost certainly be one present in the satellite data product. It is however advisable to use a consistent mean sea surface height product based on altimeter data from the target period (1993-2020).

An example for such a global mean sea surface height product is DTU15 (Table 2-1), which is based on radar altimeter data from ERS-1 to CryoSat-2 and thus spans the target SIT ECV period and region. DTU15 mean sea surface height is visualized in Figure 5-19.



**Figure 5-19: Hillshaded sea surface height of the DTU15 global mean sea surface height product for the SIT ECV target region in the northern and southern hemisphere**

With the mean sea surface height removed from the sea surface heights in the leads, the remaining signal will be due to time variant changes in sea surface height caused by variability in the magnitude and direction of ocean currents; the dynamic topography and long wavelength errors in tides and atmospheric corrections. This signal varies on a scale of a few hundred kilometres. The ice freeboard, or the height the ice floe protrudes above the sea surface, is determined by interpolating the sea surface height beneath the floe location and subtracting it from the height of the floe. Figure 5-20 and Figure 5-21 illustrate this calculation. Practically, the residual of the lead elevations with respect to the mean sea surface height (MSSH) yield the sea surface height anomaly (SSHA). The instantaneous sea surface height (SSH) is then defined by

$$SSH = MSSH + SSHA$$

The sea surface height measurements are then linearly interpolated and smoothed by a box filter using a window size of 25 km. A minimum of one lead must exist for each orbit to allow a proper estimation of the instantaneous sea surface height. The result from subtracting the interpolated and smoothed sea-surface height from all retracked sea-ice elevations yields the radar freeboard (RFRB). Radar freeboard in contrast to the sea-ice freeboard is not corrected for the slower wave propagation speed in the snow layer and therefore biased low.

$$RFRB = ELEV_{ice} - SSH$$

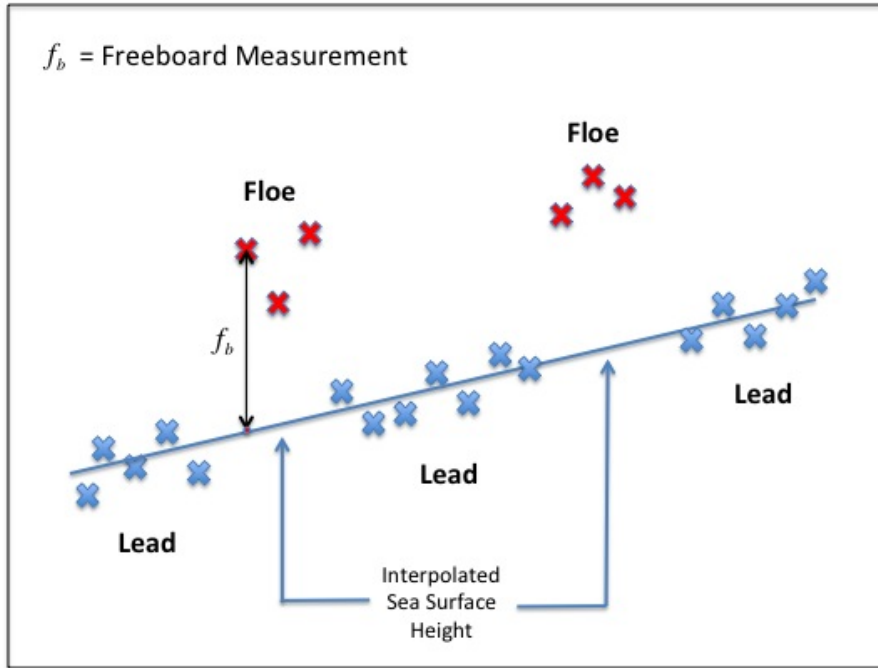
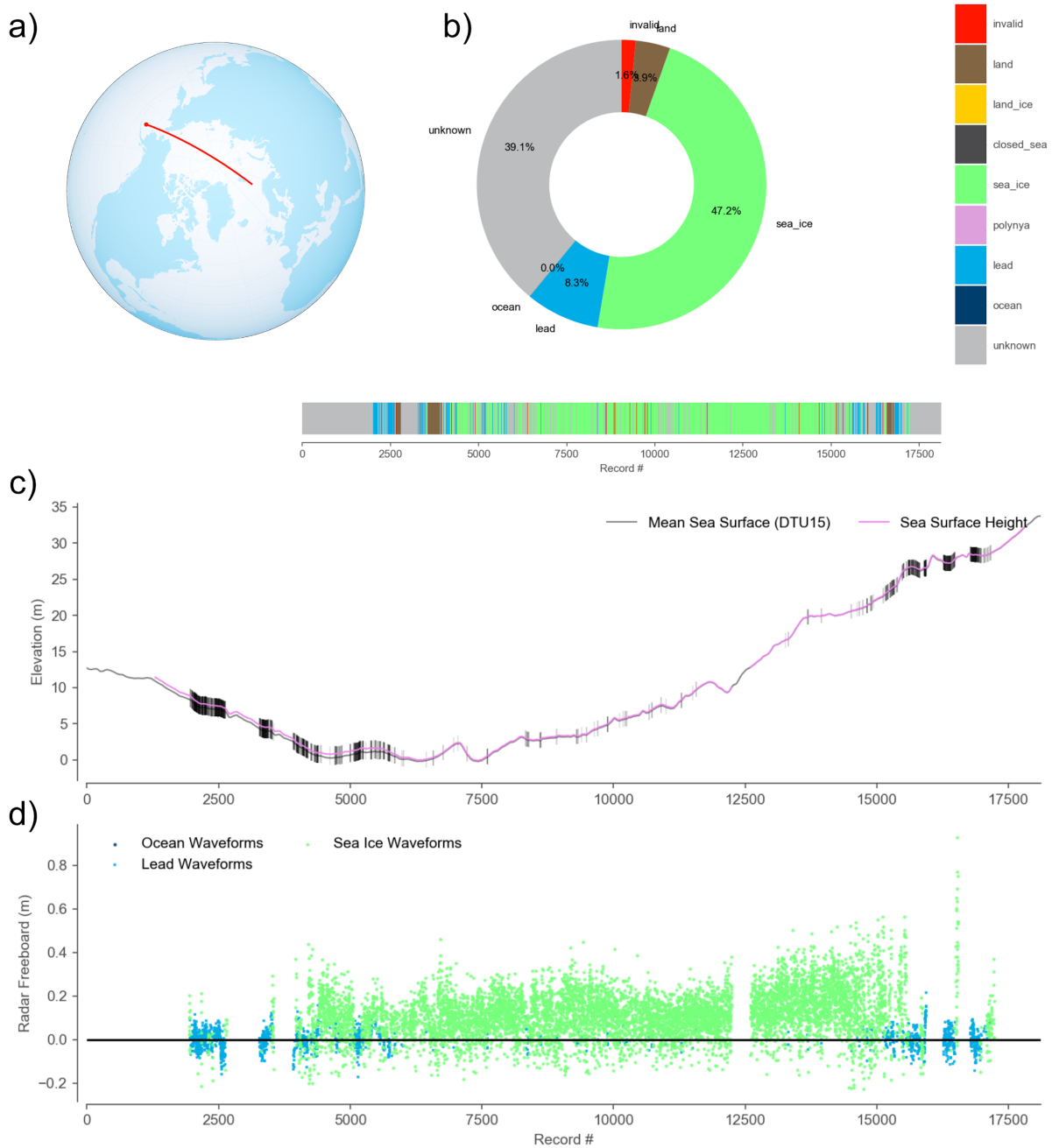


Figure 5-20: Computation of radar freeboard

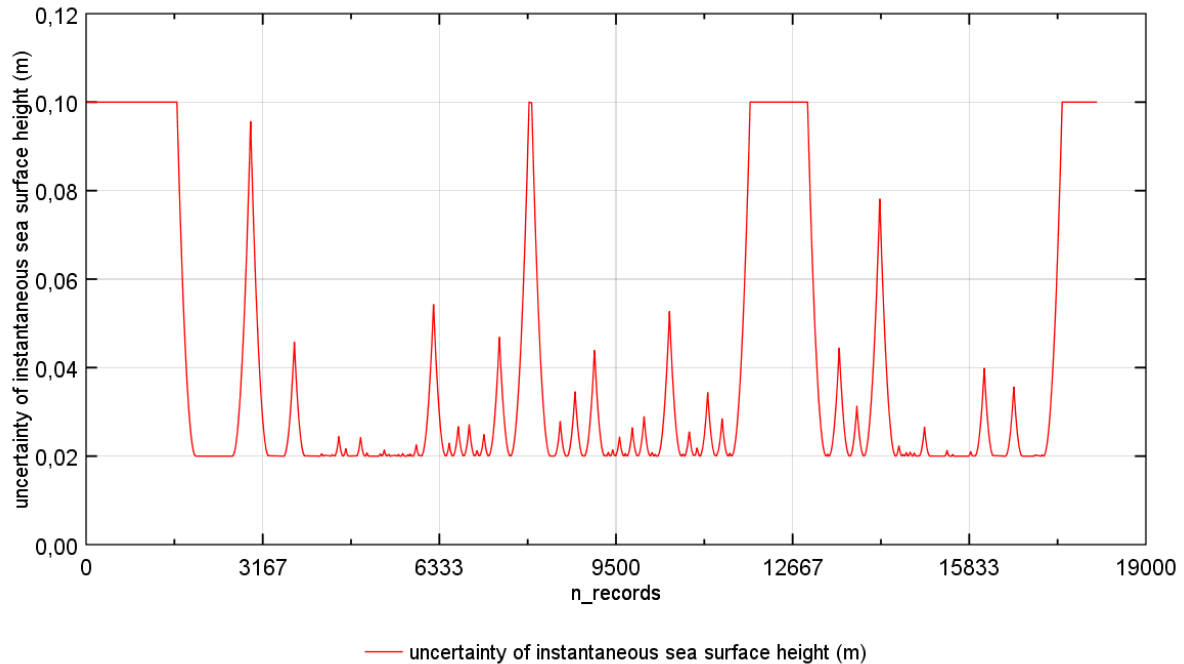


**Figure 5-21: Example from along-track CryoSat-2 freeboard retrieval. a) Orbit location, b) surface type classification with classes unknown, lead and sea ice with percentage and geographical location along track c) Lead detections, mean sea surface height and sea surface height anomaly d) unfiltered radar freeboard**

### 5.5.1 Sea-Surface Height Uncertainty

The uncertainty of the sea surface height depends on the base SSH uncertainty and the distance to the closest sea surface height tie point. The values for base SSH uncertainty is assumed to be 2 cm to include effects such as leads covered with thin ice and the maximum uncertainty is assumed as 10 cm (example of the value range in Figure 5-22) based on

investigations of the typical variation of the anomaly between the instantaneous sea surface height and mean sea surface along polar crossing orbits.



**Figure 5-22: Example of Sea Surface Height Uncertainty**

The sea surface height uncertainty is computed as

$$\sigma_{ssh} = \begin{cases} 0.02 \text{ m} + 0.1 \text{ m} \times \left( \frac{d_{tp}}{100 \text{ km}} \right)^2, \wedge d_{tp} < 100 \text{ km} \\ 0.1 \text{ m}, \wedge d_{tp} \geq 100 \text{ km} \end{cases}$$

With  $d_{tp}$  as the distance to the next sea surface height tie point.

### 5.5.2 Radar Freeboard Uncertainty

The radar freeboard uncertainty is computed by error propagation of the range or elevation uncertainty and the sea surface height uncertainty. For the simple case of radar freeboard being the difference between elevation and sea surface height, the radar freeboard uncertainty is given by:

$$\sigma_{rfrb} = \sqrt{\sigma_{elev}^2 + \sigma_{ssh}^2}$$

The elevation uncertainty  $\sigma_{elev}$  are fixed assumptions based on noise estimations for ERS-1/2, Envisat RA-2, and CryoSat-2 SIRAL sensors.

## 5.6 Snow on Sea Ice

The sea-ice thickness retrieval from altimeter data critically depends on the knowledge of snow (depth and density) information. In absence of a basin-scale observational data set, we utilized climatological information from the Warren et al. 1999 (W99) snow climatology of Arctic sea ice in SICCI Phase 2. For the Arctic, this has been replaced with a merged climatology created by AWI. This new snow product merges the monthly Warren snow climatology with daily snow depth from AMSR2 data, provided by the Institute for Environmental Physics of the University Bremen, over first-year sea ice, creating monthly snow depth fields of snow depth and density parametrizations.

### 5.6.1 Snow Depth

For the merging of the two data sets, monthly composites of the AMSR2 snow depth fields are created to match the monthly resolution of the W99 climatology for the months from October to April. After that a Gaussian low pass filter with the size of 8 grid cells is applied on the AMSR2 snow depth composite, negative snow depths are removed and upper range limit is set to 60 cm. Then a regional weight factor  $w$  is created to ensure a smooth transition between the inner Arctic Basin domain and the area where AMSR2 is used. The merged snow depth ( $sd_{merged}$ ) is computed as:

$$sd_{merged} = w \cdot sd_{W99} + (1 - w) \cdot sd_{AMSR2}$$

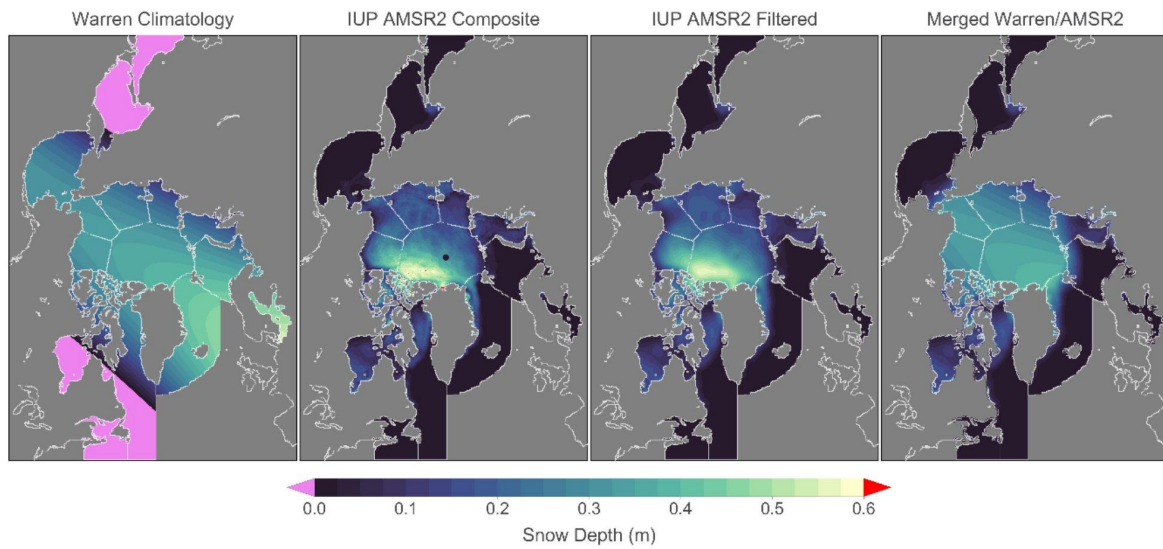
Figure 5-23 contains examples of the merging steps and Figure 5-24 for the regional weight factor.

Following the common practice to modify the W99 snow climatology by reducing the values by 50 % over first-year sea ice in the central Arctic (Tilling et al., 2018), the reduction is applied based on the ice type information for the particular orbit. This correction stems from Kurtz & Farrell (2011), which showed IceBridge measured snow thicknesses on FYI to be about 50% of the W99 estimates that are based on measurements made on MYI. Note that this scaling is applied only on the W99 snow, not on the AMSR2 snow depth. The scaled snow depth is:

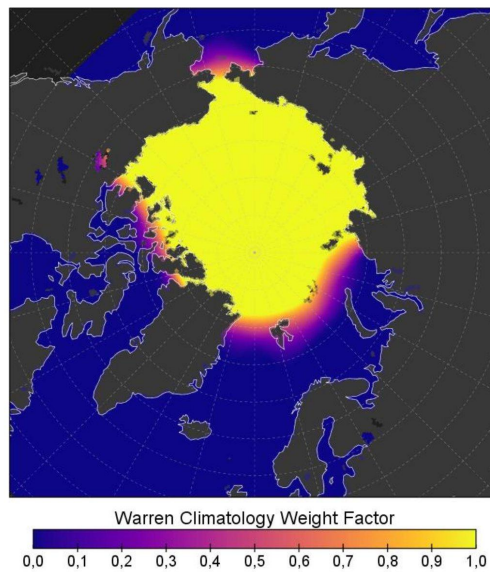
$$c = (1 - f_{myi}) * c_{fyi} * w$$

$$sd = sd_{merged} - c \cdot sd_{merged}$$

Where  $c_{fyi} = 0.5$  is the W99 scaling over first-year sea ice,  $c$  the total scaling factor including multiyear sea ice fraction  $f_{myi}$  and the weight factor.

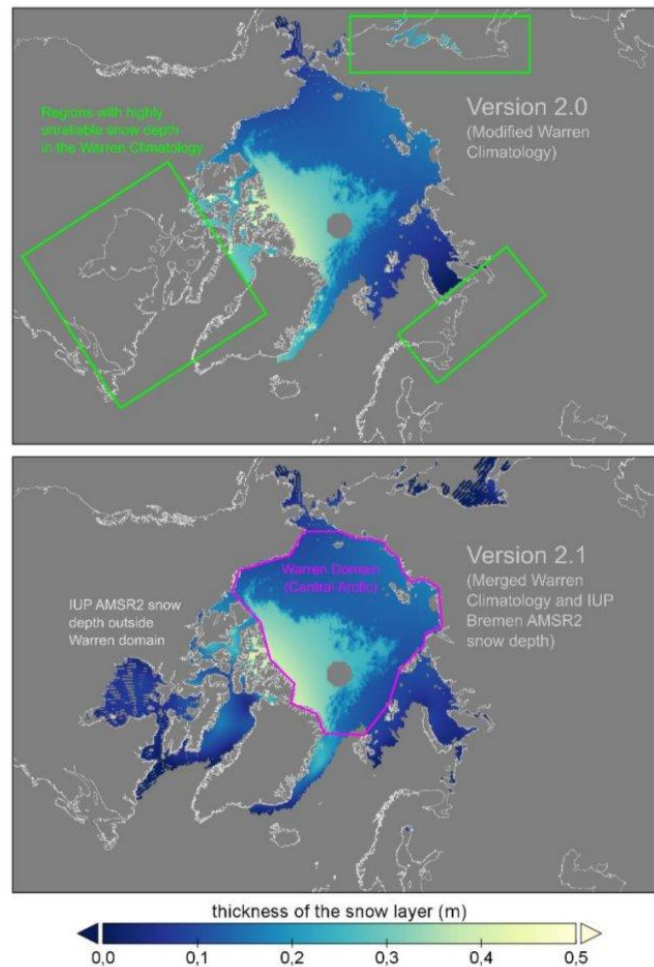


**Figure 5-23: Steps for creating the monthly merged snow depth climatology. This example is for April, from left to right: 1) Warren snow depth climatology, 2) Monthly snow composite from daily AMSR2 data, 3) Low-pass filtered composite and 4) Merged Warren/AMSR2 with regional weight factor applied**



**Figure 5-24: Regional weight factor for the W99 snow depth climatology**

An example of the results with merged snow is in Figure 5-25. There are significantly less data gaps outside the central Arctic Basin, while retaining the W99 information on areas potentially covered with multiyear sea ice, areas where AMSR2 lacks sensitivity.



**Figure 5-25: Performance example of sea ice thickness with the merged W99/AMSR2 snow product. (Top) AWI CryoSat-2 v2.0 sea ice product with W99 snow. (Bottom) sea ice thickness with merged W99/AMSR2 snow depth climatology. The improvements are most drastic in areas outside the domain (marked with green rectangles) of the W99 climatology (marked with purple polygon)**

Utilization of the monthly fields lead to unrealistic jumps in daily sea-ice freeboard and thickness values between the last day of a month and the first day of the next. Therefore, the monthly climatology is attributed to a reference day and linear interpolation is used between these days.

The reference day for the monthly climatology is the center of the month except for October and April. The reference day for these months is set to be the beginning and end respectively, as the merged snow climatology does not exist for September and May and extrapolation proved unreliable.

The linear interpolation only affects the snow depth values before the ice-type based 50% correction.



**Table 5-6: Reference dates for the monthly snow climatology used for the estimation of linear interpolated snow depth with daily resolution**

Month	Oct	Nov	Dec	Jan	Feb	Mar
Reference Day	1.10.	15.11.	15.12.	15.1.	15.2.	30.3.

For the Antarctic, with only a single ice type, a simpler approach is taken by applying the AMSR-E/2 snow-depth climatology provided by the ICDC. The climatology is based in averages for each calendar day of the daily data, available at the ICDC University Hamburg: (<https://icdc.cen.uni-hamburg.de/en/esa-cci-sea-ice-ecv0.html>).

### 5.6.2 Snow Depth Uncertainty

The uncertainty of the merged snow depth is derived in a similar fashion. We merge the uncertainty provided by the W99 climatology and the AMSR2 snow depth using the regional weighting factor.

$$\sigma_{sd}^{merged} = w \cdot \sigma_{sd}^{W99} + (1 - w) \cdot \sigma_{sd}^{AMSR2}$$

The uncertainty of snow depth (sd) is represented as the scaled uncertainty plus an uncertainty term for the scaling itself:

$$\sigma_{sd} = \left( \sigma_{sd}^{merged} - c \cdot \sigma_{sd}^{merged} \right) + \left( sd \cdot c \cdot \sigma_{fyi} \cdot c_{fyi} \right)$$

In the southern hemisphere the field `mediansnowdepth\_filtered100\_variability` of the snow depth climatology product is used as an uncertainty estimate.

### 5.6.3 Snow Density

Here, we use a linear increasing snow density over the winter season following Mallett et al., 2020. This replaces using the average snow density in the central Arctic Basin in SICCI Phase-2. Similar to snow depth, the increase of snow density is computed with daily resolution.

$$p_s = 6.5 \times t + 274.51$$

With  $t$  representing the time in fractional month since October 15.

### 5.6.4 Snow Density Uncertainty

In the Arctic the snow density uncertainty ( $\sigma_{\rho}^s$ ) is provided by the Warren climatology as well. The difference in sea ice density between FYI and MYI is small, therefore the snow density and its uncertainty are assumed to be independent from the myi fraction.

In the Antarctic, we assume a fixed uncertainty of 20 kg/m<sup>3</sup>.

## 5.7 Sea-Ice Freeboard

Sea-ice freeboard is the height of the sea ice surface above the instantaneous sea surface height. Its estimation requires obtaining the range from the satellite to the snow/ice interface, since it is safe to assume that sea ice in the northern hemisphere is covered with snow in the winter month. Here, the assumption is made that the influence of cold and dry snow on Ku-Band radar backscatter and thus the radar range is negligible. There are several studies shedding doubt on this assumption, however in the absence of a meaningful temporal and regional parametrizations of snow backscatter effects on radar range, this assumption is kept for the time being.

A certain impact by the snow is the slower wave-propagation speed of the radar signal in the snow layer. We therefore make the distinction between the terms radar freeboard and sea-ice freeboard, depending on whether any snow related correction has been applied.

### 5.7.1 Radar Freeboard

The initial radar-derived freeboard is then obtained by subtracting the sea-surface height from the sea-ice elevation. At this stage, the ice elevation is based on a conversion of the two-way travel time into range  $r$  with the vacuum light speed and the altitude  $alt$  of the satellite:

$$elev_{sea\ ice} = r - alt$$

$$rfrb = elev_{sea\ ice} - ssh$$

### 5.7.2 Sea-Ice Freeboard

The final step on the freeboard retrieval is the application of a geometric correction that accounts for the slower wave propagation speed of the radar ( $\Delta r_{WP}$ ) signal in the snow layer. The correction is linear dependent on snow depth ( $sd$ ) and thus implemented as a fraction of snow depth and its value is based on the ratio of EM wave propagation speed in snow with average density ( $\rho_s$ ) and in vacuum.

$$frb = rfrb + \Delta r_{WP}$$

with

$$\Delta r_{WP} = \left( \frac{c}{c_s} - 1 \right) \cdot sd$$

$$c_s = c(1 + 0.51 \times p_s)^{-1.5}$$

Valid sea ice freeboards are assumed to range from 0 to 2 meter, while the range is extended by the range noise (0.25 meter) for individual footprint. Thus orbit data outside the range of -0.25 m to 2.25 meter are filtered.

### 5.7.3 Sea-Ice Freeboard Uncertainty

In addition to the radar freeboard uncertainty, the sea ice freeboard uncertainty needs to take the component introduced by the snow wave speed correction into account. While the wave speed reduction is assumed to be reasonably well known, the additional uncertainty is controlled by snow depth uncertainty.

$$\sigma_{frb} = \sqrt{(sd \times \sigma_{sd})^2 + \sigma_{rfrb}^2}$$

## 5.8 Sea-Ice Thickness

### 5.8.1 Freeboard to Thickness

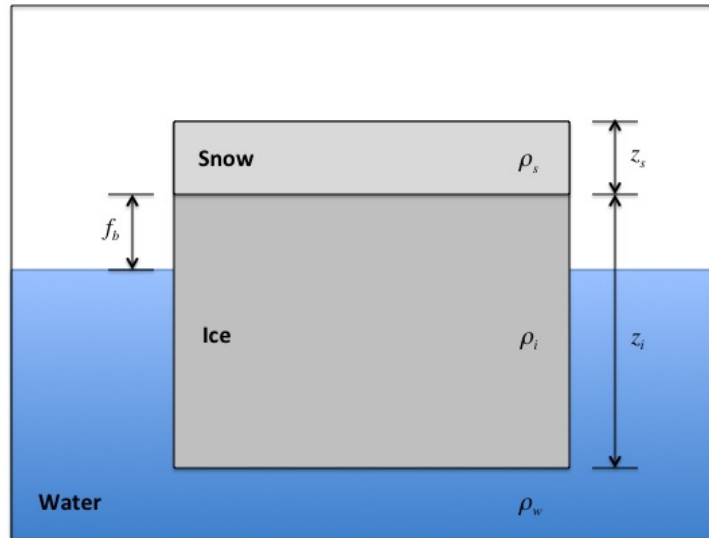
The final step in the processing is to convert sea-ice freeboard to sea-ice thickness. The ice floe may or may not be covered by snow, but field studies have shown that if the floe is indeed snow covered the radar reflection and hence height measurement relate to the snow ice interface. This however may not always be the case as was shown by the laser / radar altimeter study in Fram Strait during the RRDP exercise in CCI Phase 2. This most certainly is not the case for areas of seasonal sea ice, such as the Baltic Sea, for most of the winter. Thus freeboard values should be understood as "altimeter freeboard" values. That is, for the cold central Arctic they can be assumed to represent the ice freeboard, but for marginal areas the elevation measured is somewhere between the ice and snow freeboard. But since this effect cannot be parameterized with available EO data, it is always assumed in the processing that the dominant reflector is the snow/ice interface.

Since the ice floe is in isostatic equilibrium, a simple calculation using freeboard and snow depth, and the densities of snow, sea ice and sea water, can be used to compute the thickness. Figure 5-26 illustrates this calculation. The final thickness is given by:

$$z_i = \frac{z_s \rho_s - f_b \rho_w}{\rho_w - \rho_i}$$

Where  $z_i$  is sea ice thickness,  $z_s$  snow depth,  $\rho_s$  snow density,  $f_b$  sea ice freeboard,  $\rho_w$  density of seawater and  $\rho_i$  density of sea ice.

For water density we use the fixed values of 1024 kg/m<sup>3</sup>. Direct measurements of sea ice density suggest that the density of multi-year ice is less than that of first-year ice. We therefore use a parameterization of the sea ice density that is scaled by the multi-year ice fraction between the density of multi-year ice (882 kg/m<sup>3</sup>) and first-year ice (916.7 kg/m<sup>3</sup>).



**Figure 5-26: Computation of sea ice thickness**

### 5.8.2 Sea Ice Density Uncertainty

Similar to snow depth, sea ice density ( $\rho_i$ ) is a parameter obtained by scaling between the values for FYI and MYI using the myi fraction. To estimate the uncertainty ( $\sigma_\rho^i$ ), we scale between the uncertainties of FYI ( $\sigma_\rho^{fyi}$ ) and MYI density ( $\sigma_\rho^{myi}$ ) and add a term for the scaling uncertainty.

$$\sigma_\rho^i = \sigma_\rho^{fyi} + f_{myi} \times (\sigma_\rho^{myi} - \sigma_\rho^{fyi}) + \sigma_{fmyi} \times (\sigma_\rho^{fyi} - \sigma_\rho^{myi})$$

### 5.8.3 Sea Ice Thickness Uncertainty

The sea ice thickness uncertainty is computed as the error propagation of the input uncertainties.

$$\sigma_{sit} = \sqrt{\left(\frac{\rho_w}{\rho_w - \rho_i} \sigma_{frb}\right)^2 + \left(\frac{f_{rb} \cdot \rho_w + s_d \cdot \rho_i}{\rho_w - \rho_i} \sigma_\rho^i\right)^2 + \left(\frac{\rho_s}{\rho_w - \rho_i} \sigma_{sd}\right)^2 + \left(\frac{s_d}{\rho_w - \rho_i} \sigma_\rho^s\right)^2}$$

### 5.8.4 Sea Ice Type (MYI Fraction) Uncertainty

In the Arctic the myi fraction uncertainty ( $\sigma_{fmyi}$ ) is taken directly from the MYI fraction product (field `my\_sea\_ice\_area\_fraction\_sdev`)

No sea ice type product is available in the Antarctic and the general assumption is that all sea ice can be described as FYI. Nevertheless we assume a static uncertainty of 10% for the MYI fraction to account for sea ice type based uncertainties.

## 6 COLOCATION ON SPACE-TIME GRID (LEVEL-3 PROCESSOR)

Level-3 sea ice thickness is processed by mapping the orbit-based Level-2 data onto a spatiotemporal grid. The temporal and spatial dimensions are described in the following subsections.

### 6.1 Grid Temporal Coverage

The data will be processed monthly for the winter season between October 1st and April 30th. Temporal specifics are described in Table 6-1:

**Table 6-1: Temporal definition for Level-3 products.**

	Monthly
Start of temporal coverage	First day of month 00:00:00 UTC
End of temporal coverage	Last day of month 23:59:59.999 UTC

### 6.2 Grid Spatial Definition

Data for both hemispheres will be gridded into the Equal-Area Scalable Earth Grid version 2 (EASE2-Grid) with 25km resolution. The projection is defined in Table 6-2 and grid extent and spacing in the Level-3 product are defined in Table 6-3.

**Table 6-2: Projection definition for Level-3 products.**

Property	Hemisphere	Value
false_easting	North/South	0.0
false_northing	North/South	0.0
grid_mapping_name	North/South	lamber_azimuthal_equal_area
inverse_flattening	North/South	298.257223563
latitude_of_projection_origin	North	90.0
	South	-90.0
longitude_of_projection_origin	North/South	0.0
proj4_string	North	+proj=laea +lon_0=0 +datum=WGS84 +ellps=WGS84 +lat_0=90.0
	South	+proj=laea +lon_0=0 +datum=WGS84 +ellps=WGS84 +lat_0=-90.0

semi_major_axis	North/South	6378137.0
-----------------	-------------	-----------

**Table 6-3: Grid extent and spacing for Level-3 products.**

Property	Value
Grid Dimension	(432, 432)
Grid Spacing (km)	25.0
Grid Notation	Center Coordinates
Grid x extent in projection coordinates (km)	(-5387.5, 5387.5)
Grid y extent in projection coordinates (km)	(-5387.5, 5387.5)

### 6.3 Parameter Gridding

Level-3 processing will grid Level-2 intermediate (I2i) files. All the Level-2 data points within the specific timeframe are transformed into projection coordinates and assigned an index of a corresponding grid cell in the target grid. Each target grid cell will then possess a dedicated parameter stack that contains all the geophysical variables from Level-2 data that were associated with that specific cell. There is no filtering applied at this stage, except for radar freeboard, where freeboard values in leads need to be set as NaN in the Level-3 processor. The parameter stack of Level-2 data ( $p_{i,L2}$ ) is used to compute the gridded parameter geophysical value  $p_{L3}$  as an arithmetic mean, ignoring non-numeric values:

$$p_{L3} = \frac{1}{n_{L2}} \cdot \sum_{i=0}^{n_{L2}} p_{i,L2} \quad \text{if } p_{i,L2} \neq NaN$$

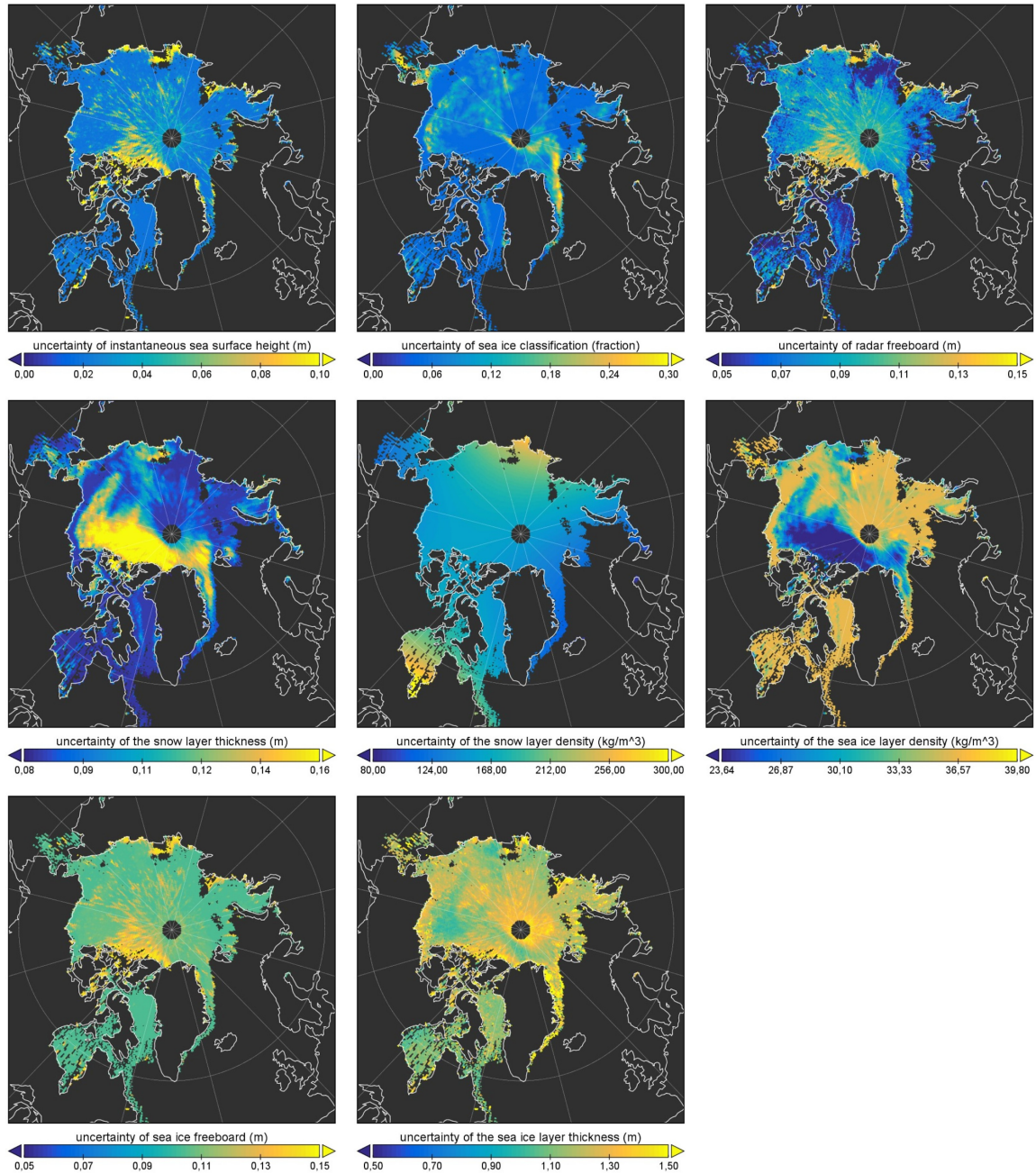
The geophysical parameters that will undergo gridding are:

1. radar freeboard
2. freeboard
3. sea ice thickness
4. snow depth
5. snow density
6. sea ice density
7. sea ice type
8. sea ice concentration

### 6.4 Level-3 Gridded Uncertainties

The Level-3 product contains the average uncertainties for freeboard/thickness respectively per grid cell to reflect that the biggest uncertainty components, e.g. snow depth, sea ice density, retracker biases, are not random uncertainties that would be reduced by averaging, examples in Figure 6-1. The uncertainties of the gridded radar freeboard, freeboard and sea ice thickness are therefore computed again with the error propagation functions, only that we

use the weighted mean error for uncertainties of random variables (radar freeboard) and the average uncertainty from the orbit data for variables with systematic error components (snow depth, sea ice and snow density). This approach introduced in CCI+ results in a more realistic uncertainty magnitude compared to the SIT CRDP v2.0.



**Figure 6-1: Gridded uncertainties (Example CryoSat-2 March 2015 Arctic data)**

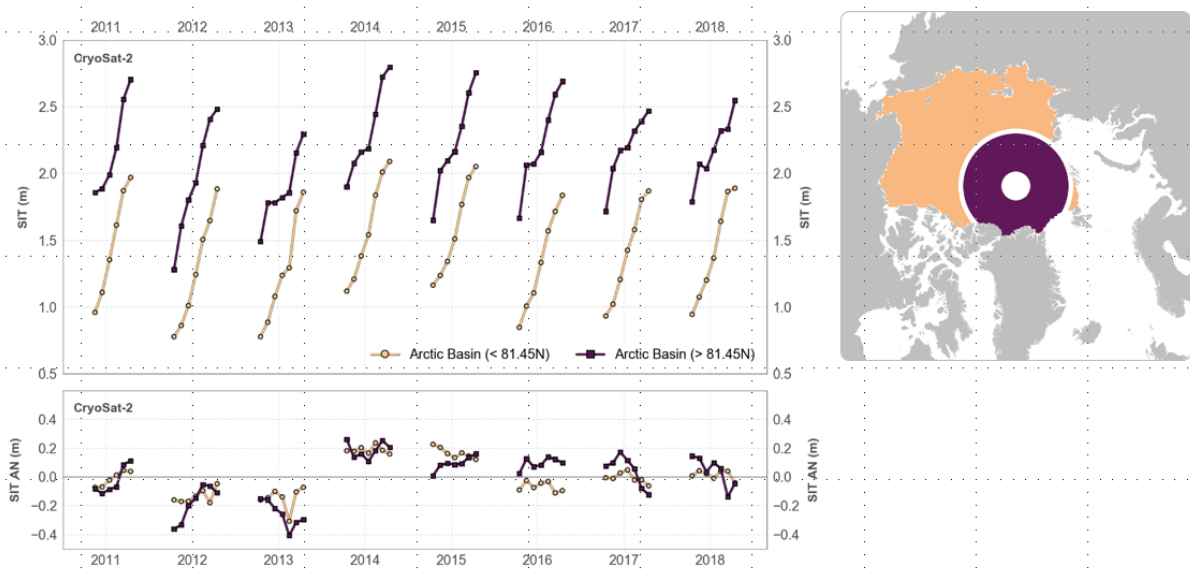


## 7 GAP INTERPOLATION (LEVEL-4 PROCESSOR)

Generally, the Level-4 data sets are created from lower level data which contains gaps due to lack of data coverage within the target period. Specifically, the Level-4 processor ingests sea-ice concentration data to determine where sea-ice thickness information needs to be available and computes an analysis of the available sea-ice thickness information from one or multiple platforms from Level-2 (trajectories) and Level-3 (space-time grids) data.

Areas that profit from the L4 product generation comprise the central Arctic basin “covered” by pole hole, the marginal seas with reduced spatiotemporal coverage, as well as topographical difficult areas such as the Canadian Archipelago. However, appropriate auxiliary data and methods are needed and tailored towards the specific regional challenges such as artifacts in the auxiliary data (e.g. from sea-ice concentration data) or varying interpolation window sizes.

Level-4 sea-ice thickness information for the ERS-1/2 and Envisat platforms in the northern hemisphere will rely on interpolation over significant distances, as these systems provide data only up to 81.5 deg north. As shown in Figure 7-1, at least the spatiotemporal evolution and variability within the Arctic Basin in- and outside the pole-hole area are rather consistent and co-varying.



**Figure 7-1: Evaluation of CryoSat-2 SIT area averages and their anomalies**

A suitable metric guiding the interpolation process using optimal interpolation could be the *distance along iceline* (DAL; Figure 7-2), as well as CCI+ sea-ice concentration data as background.

Results from and an analysis of the finalized Level-4 processor will be available with the finalized version of this document at a later stage.

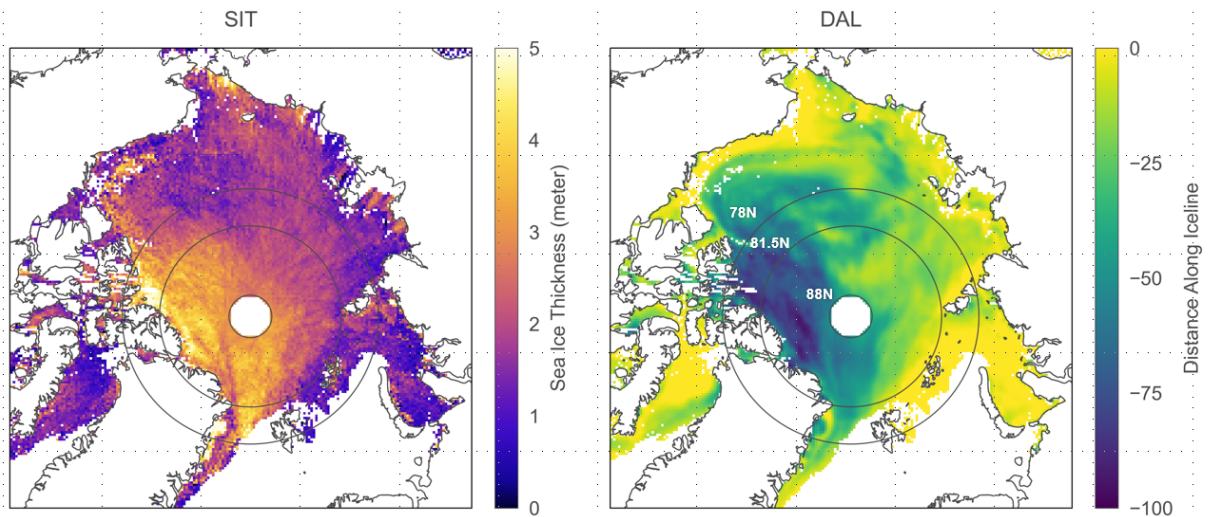


Figure 7-2: CryoSat-2 SIT (left); Distance along iceline (DAL; right) for April 2016.

## 8 SEA ICE VOLUME COMPUTATION

This section is to be completed in the future versions of the document.

For volume calculation, the individual sea ice thickness measurements are coupled with ice concentration values (C3S, Table 2-1). This is done using the gridded Level-4 product as gap-free data and its corresponding area (i.e., the area of a grid cell) is needed. Volume is calculated only where ice concentration is above 15%, so an ice extent mask is applied to rule out areas outside the 15% concentration. Sea-ice volume is then the sum product of sea ice thickness, concentration as well as the cell marine area of all grid cells covered.

$$SIV = \sum_{i=0}^n SIT_i \cdot SIC_i \cdot A_i^{marine}$$

The input for the marine area per grid cell is a dedicated land mask developed in CCI based on high resolution coastline data and generated for the target grid. The land mask is shared between all sea ice ECV variables. The mask contains the fraction of all grid cells covered by ocean domains and the marine fraction is then derived by multiplying this fraction with grid cell area (25 km<sup>2</sup>). The marine fraction is independent from the ice covered fraction of the grid cell, which is provided by the sea-ice concentration product.

The finalized section including uncertainty computation will be available with the data at a later stage.

## 9 REFERENCES

- Armitage, T. W. K. and Davidson, M. Using the interferometric capabilities of the ESA CryoSat-2 mission to improve the accuracy of sea ice freeboard retrievals *IEEE Trans. Geosci. Remote Sens.*, 52 (1), pp. 529-536, 2014.
- Breiman, L. (2001) Random Forests. *Machine Learning*, 45, 5-32. <http://dx.doi.org/10.1023/A:1010933404324>
- Brockley, D. J., Baker, S., Féménias, P., Martínez, B., Massmann, F-H., Otten, M., Paul, F., Picard, B., Prandi, P., Roca, M., Rudenko, S., Scharroo, R., and Visser, P.:REAPER: Reprocessing 12 Years of ERS-1 and ERS-2 Altimeters and Microwave Radiometer Data, *IEEE TGRS*, doi: 10.1109/TGRS.2017.2709343, 2017.
- Cavalieri, D. J., Markus, T., and Comiso, J. C. 2014. AMSR-E/Aqua Daily L3 12.5 km Brightness Temperature, Sea Ice Concentration, & Snow Depth Polar Grids. Version 3. June 2002 – October 2011. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: [http://dx.doi.org/10.5067/AMSR-E/AE\\_SI12.003](http://dx.doi.org/10.5067/AMSR-E/AE_SI12.003).
- Chelton, D. B., Ries, J. C., Haines, B. J., Fu, L.-L., and Callahan, P. S.: Satellite altimetry, *International Geophysics*, 69, 7–26, 2001.
- Connor, L. N., Laxon, S. W., Ridout, A. L., Krabill, W. B., andMcAdoo, D. C.: Comparison of Envisat radar and airborne laser altimeter measurements over Arctic sea ice, *Remote Sens. Environ.*, 113, 563–570, 2009.
- European Space Agency, 2018, RA-2 Sensor and Geophysical Data Record - SGDR. Version 3.0, <https://doi.org/10.5270/EN1-85m0a7b>
- European Space Agency, Fundamental Data Records for Altimetry: Detailed Processing Model CLS-ENV-NT-20-0424, 2021.
- Hartigan, J. A. and Wong, M. A.: Algorithm AS 136: A K-Means Clustering Algorithm, 28, 100–108, <http://www.jstor.org/stable/2346830>, 1979.
- Helm, V., Humbert, A., and Miller, H. (2014) Elevation and elevation change of Greenland and Antarctica derived from CryoSat-2, *The Cryosphere*, 8, 1539-1559, <https://doi.org/10.5194/tc-8-1539-2014>.
- Kurtz N., Farrel SL.: Large-scale surveys of snow depth on Arctic sea ice from Operation IceBridge. *Geophysical Research Letters*, Vol. 38, L20505 2011. doi:10.1029/2011GL049216
- Laxon, S. W., Peacock, N. R. & Smith, D. M. (2003) High interannual variability of sea ice thickness in the Arctic region. *Nature*, doi:10.1038/nature2050, 947-950.
- MacQueen, J. (1967) Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1: Statistics, 281--297, University of California Press, Berkeley, Calif., <https://projecteuclid.org/euclid.bsmmsp/1200512992>
- Mallett, R. D. C., Lawrence, I. R., Stroeve, J. C., Landy, J. C., and Tsamados, M.: Brief communication: Conventional assumptions involving the speed of radar waves in snow introduce systematic underestimates to sea ice thickness and seasonal growth rate estimates, *The Cryosphere*, 14, 251–260, <https://doi.org/10.5194/tc-14-251-2020>, 2020.
- Peacock, N. R. (1998) Arctic sea ice and ocean topography from satellite altimetry, Ph.D. thesis, University College London.

Peacock, N. R. & Laxon, S. W. (2004) Sea surface height determination in the Arctic Ocean from ERS altimetry. *Journal of Geophysical Research*, Vol. 109, No. C7, C07001 10.1029/2001JC001026.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., and Davidson, M. (2014) Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation, *The Cryosphere*, 8, 1607-1622, doi:10.5194/tc-8-1607-2014.

Tilling, R, Ridout, A., Shepherd, A.: Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data. *Advances in Space Research*, Volume 62, Issue 6, 15 September 2018, Pages 1203-1225, doi: <https://doi.org/10.1016/j.asr.2017.10.051>

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., and Colony, R. (1999) Snow depth on Arctic sea ice, *J. Climate*, 12, 1814–1829.