



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 CONCENTRATION ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD)

Reference: D2.1

Issue: 2.1

Date: 22 June 2020




FMI



Max-Planck-Institut
für Meteorologie






 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 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 Concentration Algorithm Theoretical Basis Document
CLIENT European Space Agency	CLIENT REFERENCE 4000126449/19/I-NB
Revision date: 22 June 2020	Approval date: 20 July 2020
Principal Authors <i>Thomas Lavergne, Norwegian Meteorological Institute</i> <i>Atle M. Sørensen, Norwegian Meteorological Institute</i> <i>Rasmus Tonboe, Danish Meteorological Institute</i> <i>Leif Toudal Pedersen, Technical University of Denmark - DTU Space</i>	

Change Record

Issue	Date	Reason for Change	Author(s)
1.0	29.08.2019	First version	T. Lavergne, A. Sørensen, R. Tonboe
2.0	29.05.2020	Second version	T. Lavergne, R. Tonboe, A. Sørensen
2.1	22.06.2020	Minor updates addressing RIDs from ESA TO	T. Lavergne

Document Approval

Role	Name	Signature
Written by:	T. Lavergne, A. Sørensen, R. Tonboe	
Edited by:	M. A. Killie	
Approved by:	A. M. Trofaier	

Contents

1 INTRODUCTION	6
1.1 Purpose	6
1.2 Scope	6
1.3 Document Status	7
1.4 Acronyms and Abbreviations	7
1.5 Executive Summary	8
2 INPUT AND AUXILIARY DATA	8
2.1 Satellite data	8
2.1.1 SMMR data	8
2.1.2 SSM/I and SSMIS data	9
2.1.3 AMSR-E and AMSR2 data	9
2.2 Atmosphere reanalysis data (ERA 5)	9
2.3 Binary Land Mask	10
2.4 Sea Ice Extent Climatology	10
3 OVERVIEW OF THE SIC PROCESSING CHAIN	10
4 PRE-PROCESSING ON SWATH PROJECTION (L1P)	11
4.1 Data volume reduction	11
4.2 Apply inter-platform calibration	11
4.3 Collocate channels	12
4.4 Land spill-over correction	12
5 GEOPHYSICAL PROCESSING ON SWATH PROJECTION (L2)	12
5.1 Several algorithms, combining different microwave channels	12
5.2 A hybrid, self-tuning, self-optimizing sea-ice concentration algorithm	14
5.2.1 Merging equation for the hybrid algorithm	14
5.2.2 Applying the SIC algorithms	14
5.2.3 Tuning of the algorithm at 0% and 100% SIC	15
5.2.4 Dynamical selection of the SIC training data	16
5.3 Radiative transfer modelling for correcting atmospheric influence on TB	17
5.4 The distance-along-the-line	18
5.5 Open-water filtering	18
5.6 Reducing systematic errors at high-concentration range	20
5.7 Uncertainties at L2	20
5.7.1 Algorithm and tie-point uncertainty	21

5.7.2 Uncertainty for the hybrid algorithm	21
6 GRIDDING AND DAILY COMPOSITING (L3)	21
6.1 Gridding and daily averaging	21
6.2 Gridding and smearing uncertainty	22
7 FINAL ANALYSIS AND FILTERING (L4)	23
7.1 Apply masks and filters	23
7.1.1 Masking for land regions	23
7.1.2 Climatological maximum extent masking	23
7.1.3 Prepare cleaned SIC maps	23
7.1.4 Additional land spill-over correction (coastal regions)	23
7.2 Gap filling by interpolation	23
7.2.1 Temporal interpolation	24
7.2.2 Spatial interpolation	24
7.3 Compute total uncertainty	24
7.4 Possible melting or high T2m flag	25
8 TEMPORAL AGGREGATION	25
9 ALGORITHMS FOR ESMR (1972 - 1977)	26
10 REFERENCES	28

1 INTRODUCTION

1.1 Purpose

This document is the Algorithm Theoretical Basis Document for the Sea Ice Concentration ECV product within CCI+ PHASE 1 - NEW R&D ON CCI ECVs. It documents the scientific background and algorithm details of the methodologies implemented to generate Sea Ice Concentration Climate Data Records.

1.2 Scope

The scope of the document is to describe elements of the algorithms that are deemed mature enough for implementation during the 2nd year of the CCI+ Phase 1 project, towards production of test products during Year 2. The selected algorithms are presented and justified, but the document does not hold results of research work leading to the selection of these algorithms (some of which will be in D2.3 Update to ADP).

1.3 Document Status

This is the second version of the ATBD of the Sea Ice Concentration variable within CCI+ Phase 1. The content of the ATBD is largely based on the ATBD from Year 1 (itself based on the ATBD from CCI Phase 2).

1.4 Acronyms and Abbreviations

The table below lists the acronyms and abbreviations used in this volume.

Table 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)
CCI	Climate Change Initiative
CDR	Climate Data Record
DAL	Distance Along the (sea-ice) Line
DMSP	Defence Meteorological Satellite Program
EASE grid	Equal-Area Scalable Earth Grid
ECMWF	European Centre for Medium-Range Weather Forecasts
ESA	European Space Agency
ESMR	Electrically Scanning Microwave Radiometer
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
FoV (<i>alt</i> FOV)	Field-of-View
FYI	First Year Ice
ICDR	Interim Climate Data Record
L1B, L2, L3, ...	Satellite data processing Level (Level-1b, ...)
EPS, EPS-SG	EUMETSAT's Polar System, EPS Second Generation
MWI	MicroWave Imager (EPS-SG)
MYI	Multi-Year Ice
NSIDC	US National Snow and Ice Data Centre
OSI SAF	EUMETSAT Ocean and Sea Ice Satellite Application Facility
OWF	Open Water Filter
PMR	Passive Microwave Radiometer
PMW	Passive Microwave
RTM	Radiative Transfer Model
SIC	Sea Ice Concentration
SMMR	Scanning Multichannel Microwave Radiometer

SSM/I	Special Sensor Microwave/Imager
SSMIS	Special Sensor Microwave Imager/Sounder

1.5 Executive Summary

The suite of algorithms for producing the Sea Ice Concentration Climate Data Records in CCI+ are presented. They span all processing levels, from pre-processing of raw satellite data to monthly aggregated products:

- L1P “Level-1 Pre-Processing” are all the steps applied to prepare microwave brightness temperature data (L1B T_B) for processing of Sea Ice Concentration;
- L2 “Level 2” are the core of the geophysical algorithms. SIC and its uncertainties are indeed computed on swath-projection (L2). The algorithms cover the extraction and preparation of dynamic tie-points, the tuning of the algorithms to these tie-points, the tuning and evaluation of Open Water Filters, and of algorithm uncertainties.
- L3 “Level 3” are mostly algorithms to prepare daily gridded maps of SIC and associated parameters (uncertainties, filters, flags...)
- L4 “Level 4” are the final analysis where cleaned SIC fields are prepared, uncertainty contributions added, and remaining data gaps are filled with spatio-temporal interpolation.

In addition, the algorithm to prepare monthly averaged aggregated SIC products are covered in a final section.

This ATBD is a living document that will be edited regularly to reflect the algorithm development work in the project.

2 INPUT AND AUXILIARY DATA

2.1 Satellite data

2.1.1 SMMR data

The Scanning Multichannel Microwave Radiometer (SMMR) instrument on board the Nimbus 7 satellite operated from October 1978 to 20th August 1987 (Gloersen et al., 1992). For most of the period, the instrument was operated only every second day, due to power supply limitations. The instrument had 10 channels, from six Dicke radiometers, at five frequencies (6.6, 10.7, 18.0, 21.0, 37.0 GHz) and vertical and horizontal polarization. The scanning across track was ensured by tilting the reflector from side to side while maintaining constant incidence angle on the ground of about 50.2° . The scan track on the ground formed a 780 km wide arc in front of the satellite (Gloersen and Barath, 1977). Because of the satellite orbit inclination and swath width there is no coverage poleward of 84° .

2.1.2 SSM/I and SSMIS data

The Special Sensor Microwave/Imager (SSM/I) sensors on board the Defence Meteorological Satellite Program (DMSP) started its record with the F08 satellite on 9th July 1987, shortly before the SMMR ceased to operate. The SSM/I is a total power radiometer, with a conical scan measuring the upwelling radiation from the Earth at a constant incidence angle of about 53.1° at four frequencies (19.3, 22.2, 37.0, 85.5 GHz). The swath width is about 1400km and the polar observation hole extends to 87° .

The Special Sensor Microwave/Imager (SSM/I) data set used for this reprocessing was prepared by EUMETSAT CM SAF and covers the period of available DMSP satellites instruments from 1987 to 2008 (F08, F10, F11, F13, F14, F15). Some SSM/I instruments continued their mission further than 2008, but these data are not included in the CM SAF FCDR.

The SSM/I instruments have five low frequency channels that are mostly similar to some of those on SMMR. In addition, two higher frequency channels at 85GHz, with twice the sampling rate and better spatial resolution, are available on the SSM/I starting with DMSP F10 (the 85 GHz channels had a malfunction on F08).

The SSMIS instruments are a slight evolution of the SSM/I concept, and most characteristics that drive the design of SIC CDRs are similar to SSM/I. Noticeable differences are the size of the polar observation hole (89°), and the center frequency of the high-frequency channels (91.1 GHz). The SSMIS instruments were also on board DMSP satellites, and we use F16, F17, and F18 missions (F19 was a short-lived mission, and F20 was never launched). DMSP F18 is thus the last available SSMIS instrument.

2.1.3 AMSR-E and AMSR2 data

The AMSR-E instrument on board the Earth Observing System (EOS) satellite Aqua recorded passive microwave data from 1st June 2002 until 4th October 2011. This instrument measured vertically and horizontally polarized brightness temperatures at 7 frequencies (6.9, 7.2, 10.7, 18.7, 23.8, 36.5 and 89 GHz), thus 14 channels in all. Thanks to a larger antenna reflector, AMSR-E had significantly better spatial resolution than SSM/I or SSMIS. It also had a wider swath, and thus a smaller polar observation hole (89.5°).

The AMSR2 instrument on board the Global Change Observation Mission – Water (GCOM-W1) satellite provides similar data to the AMSR-E instrument (no 7.2 GHz channels), with even slightly better resolution. The first data for AMSR2 are from 23rd July 2012.

2.2 Atmosphere reanalysis data (ERA 5)

The microwave radiation emitted by the ocean and sea ice travels through the Earth's atmosphere before being recorded by the satellite sensors. Scattering, reflection, and emission in the atmosphere add or subtract contributions to the radiated signal, and challenge our ability to accurately quantify sea-ice concentration.

An central step in our Level-2 processing is thus the explicit correction of the T_b for the atmospheric contribution to the top of the atmosphere radiation (see Sect. 5.2). For this purpose, we access global hourly fields from C3S ERA5 reanalysis (produced by ECMWF). Fields of 10 m wind speed, 2 m air temperature, and total column water vapour are used. ERA5 reanalysis currently starts in January 1979 and is available throughout the time period of our CDRs. ERA5 will be extended with the 1950-1978 period by summer 2020.

The ERA5 reanalysis is described in Hersbach et al. (2020).

2.3 Binary Land Mask

Land masks for the target 25×25 km grids (one for NH and one for SH) are computed based on the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) 0.05×0.05° land mask (Donlon et al., 2012). This mask was re-used in the ESA CCI sea surface temperature (SST) (Phase 1) L4 data records and was selected as the input mask for the v2 SIC CDRs to increase cross-ECV consistency. The masks are tuned to closely match that of the NSIDC SIC CDR (the NSIDC “SSM/I” 25 km Polar Stereographic mask). On average, in the NH, this corresponds to setting all 25×25 km grid cells with a fraction of land lower than 30 % to water (and these cells can thus potentially be covered with sea ice). There is no right or wrong binary land mask at such coarse resolution, and the decision to tune to the NSIDC SIC CDR land mask is to help an intercomparison of data records.

During CCI+, a new landmask might be prepared to better match those of the SST CCI and/or other ocean ECVs.

2.4 Sea Ice Extent Climatology

The monthly varying maximum sea-ice extent climatology implemented in Meier et al. (2017) was used as a basis for our own climatology. The modifications included manual editing of some single pixels based on US National Ice Center, Canadian Ice Service, and Norwegian Ice Service ice charts (e.g. along the coast of northern Norway, for some summer months in the vicinity of Nova Scotia). The climatology of peripheral seas and large freshwater bodies (e.g. Bohai and Northern Yellow Seas, Great Lakes, Caspian Sea, and Sea of Azov) was also revisited. The cleaned climatologies were then expanded with a buffer zone of 150 km in the NH and 250 km in the SH. The larger expansion in SH is to cope with the positive trends in the SH sea-ice extent (Hobbs et al., 2016).

3 OVERVIEW OF THE SIC PROCESSING CHAIN

Fig. 1 gives an overview of the processing chain for the SIC CDRs. The red boxes are data (stored in data files) and the blue boxes are processing elements that apply algorithms to the data. The whole process is structured into four chains, at Level 1P (left), Level 2, Level 3, and Level 4 (right). The input Level 1 (L1) data files hold the fields observed by the satellite sensors at the top of the atmosphere, in satellite projection: the brightness temperatures (TB) are structured in swath files. The Level 1 Preprocessing (L1P) prepares the L1B swath files for SIC processing. The Level 2 (L2) chain transforms these into the environmental variables

of interest, but still on swath projection: the SIC, its associated uncertainties, and flags. The L2 chain holds an iteration (marked by the “2nd iteration” grey box) similar to the workflow in Tonboe et al. (2016) and stemming from the developments of Andersen et al. (2006). This iteration implements two key correction schemes: the atmospheric correction algorithm at low-concentration range (Sect. 5.2) and a novel correction for systematic errors at high-concentration range (Sect. 5.4). The Level 3 (L3) chain collects the L2 data files and produces daily composited fields of SIC, uncertainties, and flags on regularly spaced polar grids. These fields can and will typically exhibit data gaps, e.g. in case of missing satellite data. The Level 4 (L4) chain fills the gaps, applies extra corrections, and formats the data files that will appear in the CDR.

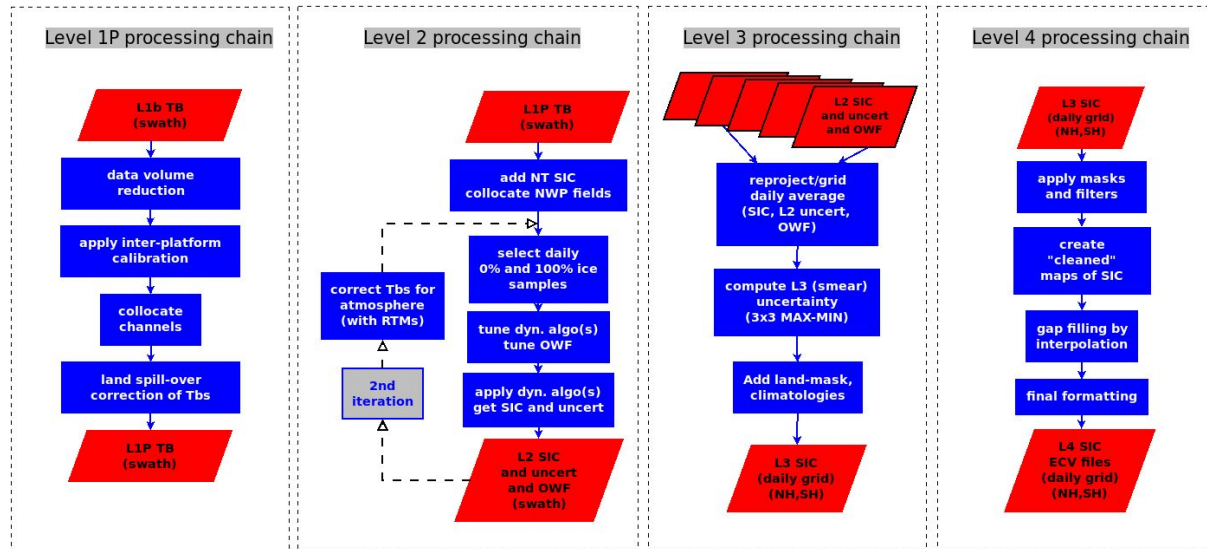


Figure 1: From left to right, the four main elements (Level 1P, Level 2, Level 3, and Level 4) in the sea-ice concentration (SIC) processing workflow. The red boxes depict data files, the blue boxes correspond to individual steps (a.k.a. algorithms) in the processing. The files that exit a processing chain (e.g. the “L2 SIC and uncert and OWF” at the bottom of the Level 2 processing chain) are the input for the next level of processing. Acronyms: NT is the Nasa Team algorithm, OWF is open-water filter, RTM is radiative transfer model, uncert stands for uncertainty.

The next four sections document the L1P, L2, L3, and L4 algorithms.

4 PRE-PROCESSING ON SWATH PROJECTION (L1P)

4.1 Data volume reduction

We focus our processing on the polar and sub-polar oceans. A first pre-processing step is thus to remove all instrument scans whose observations are all in the $[-35^{\circ}; +35^{\circ}]$ latitude band, as well as all scans well inside land and continents (200 km distance).

4.2 Apply inter-platform calibration

In some input data sources, such as the Fundamental Climate Data Record (FCDR) of SMMR, SSM/I, and SSMIS data, the coefficients to calibrate T_B data inside a series of

satellites are computed but not applied. A pre-processing step is thus to apply those coefficients.

This inter-calibration step will generally have minimal impact on the results from our SIC algorithms since these are designed to adapt their formulation and coefficients (the tie-points) to the T_B data (they are dynamic algorithms).

4.3 Collocate channels

T_B data are arranged as (scanline, scanpos) two-dimensional arrays in the L1B swath files. Due to the conical scan mechanism of the instrument, this indexing does not guarantee that observations at a given index for each of the different frequency channels are closest in space once projected onto the Earth's surface. This is especially true for the high-frequency channels (near 90 GHz) and at the edges of the swath.

Since our SIC algorithms combine T_B data from different frequency channels, it is key that they are first collocated in terms of Earth location. This can effectively be achieved with nearest-neighbour collocation.

4.4 Land spill-over correction

Due to the coarse resolution of the sensors used, especially SMMR, SSM/I, and SSMIS, the T_B data are influenced by land emissivity several tens of kilometres offshore from the coastlines. The microwave emissivity of land is comparable to sea-ice emissivity and much higher than that of ocean water. This means that sea-ice concentration will be consistently overestimated in coastal regions with less than 100% sea ice concentration.

We adopt and adapt the swath-based correction algorithms of Maass and Kaleschke (2010). The basic principle is that a fine-resolution land mask is used together with the antenna viewing geometry to simulate (and correct for) the contribution of land emissivity to the observed T_B . The algorithm of Maass and Kaleschke (2010) was adopted with some modification and tuning, including (a) the computation of the fraction of land in each FoV in the view geometry of the antenna (not after projection to a map), and (b) the approximation of the antenna pattern functions as Gaussian (normal distribution) shapes indexed on the aperture angle from the central view direction, instead of distance on a projection plane. At the end of this step, T_B in FoV that overlap land and ocean (coastal FoVs) are corrected for the contribution by land, and can enter the Level 2 SIC algorithms.

5 GEOPHYSICAL PROCESSING ON SWATH PROJECTION (L2)

5.1 Several algorithms, combining different microwave channels

The algorithm baseline described in the rest of this Level-2 chapter is generic enough to be applied to any combination of microwave frequency channels. Thus, the sections below should be understood as “algorithm-independant”. When we run the production system, several SIC algorithms are run in parallel, that combine different frequency channels.

Different frequency channels indeed have different characteristics that can be interesting for building a climate data record of sea-ice concentration. Generally, using lower frequency channels (e.g. 6.9 GHz) yield a higher accuracy than using high frequency channels (e.g. 89.0 GHz). This is because the contrast between open water and sea ice is largest with the lower frequencies, related to the fact that the atmosphere is more transparent. Another general rule is that lower microwave frequencies offer coarser spatial resolution than the higher frequencies. The choice of which microwave frequencies to combine in a SIC algorithm is thus a balance between accuracy and spatial resolution. Used in combinations, 19 and 37 GHz frequency channels provide a compromise accuracy/resolution and are also the only ones being available without interruption since the late 1970s. With now 30 years of near-90 GHz imagery, there is an interest in re-visiting SIC algorithms using these frequencies to improve the spatial resolution of the Climate Data Record.

The table below summarizes the algorithms (channel combinations) used so far, including in CCI+ Sea Ice Year 1, and CCI Phase 2.

Phase	Algorithm identification	Microwave channels
CCI+ Year 1&2 / OSISAF	SICCI3LF_corrSICCI3LF	(19V, 37V, 37H)
CCI+ Year 2	N90LIN_corrSICCI3LF	(91V, 91H)
CCI+ Year 1	SICCI3HF_corrSICCI3HF	(19V, 91V, 91H)
CCI+ Year 1	SICCI3K4_corrSICCI3K4	(19V, 19H, 37V, 37H)
CCI+ Year 1	SICCI3AF_corrSICCI3AF	(19V, 19H, 37V, 37H, 91V, 91H)
CCI Phase 2 / OSISAF	SICCI2LF_corrSICCI2LF	(19V, 37V, 37H)
CCI Phase 2	SICCI2VLF_corrSICCI2VLF	(6V, 37V, 37H)
CCI Phase 2	SICCI2HF_corrSICCI2HF	(19V, 91V, 91H)

Table 2: Summary of algorithms investigated in CCI+ Year 1 & 2, and CCI Phase 2 / OSI SAF. See text for more details.

The algorithm identification is in the form: <ALGO2>_corr<ALGO1>, where <ALGO2> might be the same as <ALGO1>. As was introduced in section 3, the Level-2 chain has an iteration through an atmospheric correction step. This atmospheric correction step requires a “first guess” SIC value, which is computed with <ALGO1> (1st iteration). Once the brightness temperature are corrected for the atmosphere, they are input to <ALGO2> (2nd iteration of the Level-2 chain).

Thus “SICCI3LF_corrSICCI3LF” uses the same algorithm (the same frequency channels) for the 1st and 2nd iterations, while “N90LIN_corrSICCI3LF” uses the SICCI3LF (19V, 37V, 37H) algorithm for the first-guess value of the atmospheric correction step, but applies the

N90LIN (“near-90GHz linear” algorithm, with channels 91V and 91H) in the second iteration. The capability to mix algorithms in this manner is new in CCI+ Year 2.

5.2 A hybrid, self-tuning, self-optimizing sea-ice concentration algorithm

A new sea-ice concentration algorithm formulation was developed during the ESA CCI Sea Ice projects (2008-2017) and is used for the SIC CDRs. It is an evolution of the algorithms used in Tonboe et al. (2016). In this section, we describe both how the algorithm is trained to T_B training data sets, and how it is then applied to actual T_B measurements recorded by satellite sensors.

5.2.1 Merging equation for the hybrid algorithm

We call the SIC algorithm a hybrid algorithm because it combines two other SIC algorithms: one that is tuned to perform better over open-water and low-concentration conditions (named B_{OW} for “best open water”), and one that is tuned to perform better over closed-ice and high-concentration conditions (named B_{CI} for “best closed ice”). The combination equation is quite simply a linear weighted average of B_{OW} and B_{CI} results, where w_{ow} is the open-water weight and SIC is expressed as sea-ice fraction [0; 1]:

$$w_{OW} = 1 \text{ if } B_{OW} < 0.7;$$

$$w_{OW} = 0 \text{ if } B_{OW} > 0.9;$$

$$w_{OW} = \frac{B_{OW} - 0.7}{0.2} \text{ if } B_{OW} \in [0.7 - 0.9];$$

$$SIC_{hybrid} = w_{OW} \times B_{OW} + (1 - w_{OW}) \times B_{CI}$$

5.2.2 Applying the SIC algorithms

Be \vec{T} a triplet of brightness temperatures measured from the satellite. For the SICCI3LF algorithm (see Table 2), the triplet holds the Ku V-pol channel (19v), Ka V-pol channel (37v) and Ka H-pol channel (37h). The SIC (noted C) is computed as:

$$C(\vec{T}) = \frac{\vec{v} \cdot (\vec{T} - \langle \vec{T}^W \rangle)}{\vec{v} \cdot (\langle \vec{T}^I \rangle - \langle \vec{T}^W \rangle)}$$

where $\langle \vec{T}^W \rangle$ (resp $\langle \vec{T}^I \rangle$) is the tie-point for 0% SIC (resp 100% SIC) and $\vec{v} = (v_{19v}, v_{37v}, v_{37h})$ is the vector of the algorithm coefficients.

The equation above is a generic form for applying B_{OW} and B_{CI} algorithms, once the two tie-points are computed (see Sect. 5.2.4), and vectors \vec{v}_{OW} and \vec{v}_{CI} are known (see Sect. 5.2.3).

5.2.3 Tuning of the algorithm at 0% and 100% SIC

In the case of SICCI3LF (19V, 37V, 37H), the tuning of B_{OW} and B_{CI} is an evolution from the tuning of the Bristol (BRI, Smith and Barrett, 1994) algorithm. The Bristol algorithm is tuned to a training data set of purely 100% SIC and purely 0% SIC T_B samples by selecting a unit vector $v_{Bristol}$ that is orthogonal to the 100% SIC line (itself sustained by vector u). Vector $v_{Bristol}$ is selected so that the data plane ($u, v_{Bristol}$) holds the open water tie-point (the average point of all 0% SIC samples). This is illustrated in Fig. 2.

When tuning BOW (resp BCI) we rather find the unit vector v_{OW} (resp v_{CI}) that minimizes the spread of the retrieved SIC at 0% SIC (resp 100% SIC) when the SIC algorithm is evaluated against the training data sets. The vectors v_{OW} (resp v_{CI}) are generally not aligned with $v_{Bristol}$ (Fig. 2 left panel).

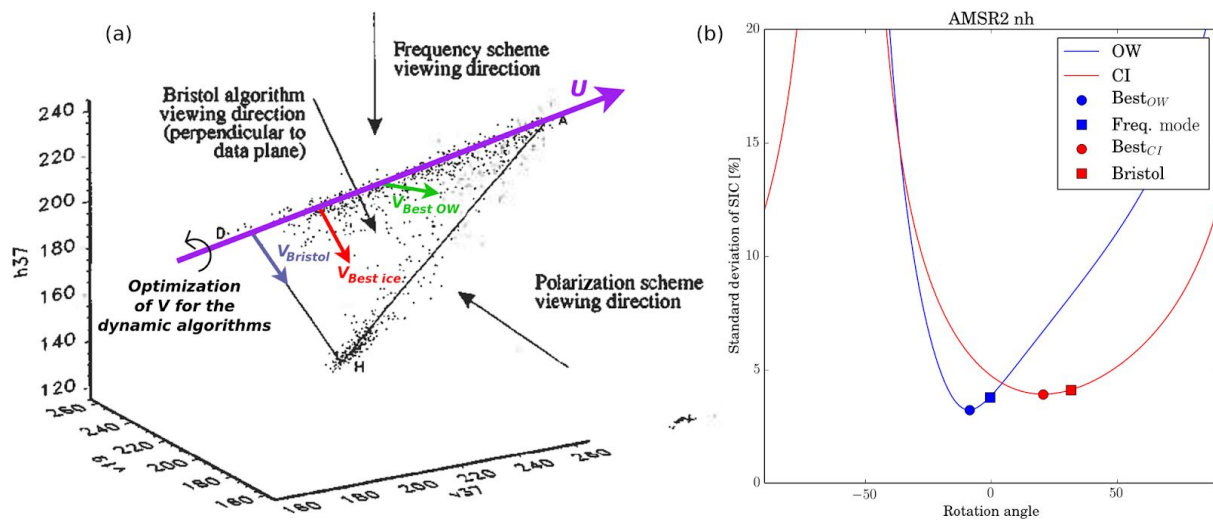


Figure 2 (a) Three-dimensional diagram of open-water (H) and closed-ice (ice line between D and A) brightness temperatures in a 19V, 37V, 37H space (black dots). The original figure is from Smith (1996). The direction U (purple) is shown, and vectors $v_{Bristol}$ (blue), $v_{Best-ice}$ (red), and $v_{Best-OW}$ (green) are added, as well as an illustration of the optimization of the direction of V for the dynamic (self-optimizing) algorithms. **(b)** Evolution of the SIC algorithm accuracy for open-water (blue) and closed-ice (red) training samples as a function of the rotation angle θ in the range $[-90^\circ; 90^\circ]$. Square symbols are used for the BFM (Bootstrap frequency mode, Comiso, 1986) and BRI (Bristol, Smith and Barrett, 1994) algorithms. Disk symbols locate the new, self-optimizing algorithms.

The optimization of vectors v_{OW} (resp v_{CI}) is best implemented as a brute-force optimization over a 1 dimensional axis representative of the rotation angle vector \vec{v} has wrt. a reference direction. In Fig. 2 (b), this optimization process is illustrated, with choosing 0° rotation angle as the Bootstrap Frequency Mode (BFM, Comiso, 1986) algorithm, in a case using AMSR2 data from the Northern Hemisphere. The solid lines plot the variation in the accuracy (measured as standard deviation of SIC, on the y axis) of the SIC algorithms defined by the rotation angle (x axis) against the 0% SIC (OW, blue) and 100% SIC (CI, red) training T_B data. The minimum of the blue (resp red) curves is reached at angle θ_{OW} (resp θ_{CI}) that defines v_{OW} (resp v_{CI}).

At the end of the tuning (this is repeated on a daily basis, see Sect. 5.2.4), the tie-points $\langle \vec{T}^W \rangle$ and $\langle \vec{T}^I \rangle$ are stored, as well as vectors v_{OW} and v_{CI} . This is all what is needed to later apply the B_{OW} and B_{CI} on satellite swath data, and combine the results into the hybrid SIC value (Sect. 5.2.1).

The geometric description above holds for the other 3-channels algorithms (see Table 2). The algorithms with only two channels (e.g. N90LIN) do not give room for such an optimization. The algorithms with more than 3 channels have two or more degrees of freedom for the optimization, which cannot be as easily described geometrically.

5.2.4 Dynamical selection of the SIC training data

As described in the previous section, tuning the algorithms requires two sets of training data: one from OW areas (SIC = 0 %) and one from areas we assume have fully CI cover (SIC = 100 %). The training of the algorithms is performed separately for each instrument and for each hemisphere. In addition, the training is updated for every day of the data record and is based on a [-7; +7 days] sliding window worth of daily samples. The sliding window is relatively short so that tie points react more rapidly to seasonal cycles, e.g. onset of melting.

The dynamic training of the algorithms allows us to (a) adapt to interseasonal and interannual variations of the sea-ice and open-water emissivity, (b) cope with different calibration of different instruments in a series, or between different FCDRs, (c) cope with slightly different frequencies between different instruments (e.g. SMMR, SSM/I, and AMSR-E all have a different frequency around 19 and 37 GHz), (d) mitigate sensor drift (if not already mitigated in the FCDR), (e) compensate for trends potentially arising from the use of NWP reanalysed data to correct the T_B .

The CI training sample is based on the results of the NASA Team (NT) algorithm (Cavalieri et al., 1984): locations for which the NT value is greater than 95 % are used as a representation of 100 % ice (Kwok, 2002). Earlier investigations, e.g. during the ESA CCI Sea Ice projects, confirmed that NT was an acceptable choice for the purpose of selecting closed-ice samples. The tie points for applying the NT algorithm are from Appendix A in Ivanova et al. (2015). AMSR2 tie-points are not included in Ivanova et al. (2015), and we use the same tie points as for AMSR-E. To ensure temporal consistency between the SMMR and later instruments, the closed-ice samples for NH are only used for algorithm tuning if their latitude is less than 84° N, which is the limit of the SMMR polar observation hole.

NEW in CCI+ (Year 1): The selection of the OW tie-point samples was revised during CCI+. During CCI, the training areas varied on a monthly basis, following a monthly maximum ice extent climatology (itself derived from NSIDC data). In the forthcoming CDRs, the OW tie-point samples shall be selected in a 150 km wide belt that encompass the extent of the sea-ice cover on a daily basis. In practice, NT sea-ice concentration is computed and gridded on a daily basis, and a 150 km wide belt starting 150km from the ice edge is computed. All observations in the swath falling into this belt are selected for computing the OW tie-point.

5.3 Radiative transfer modelling for correcting atmospheric influence on TB

As described in Andersen et al. (2006) and confirmed in Ivanova et al. (2015), the accuracy of retrieved sea-ice concentration can be greatly improved when the brightness temperatures are corrected for atmospheric contribution by using a radiative transfer model (RTM) combined with surface and atmosphere fields from NWP reanalyses. The correction using NWP data is only possible in combination with a dynamical tuning of the tie points, so that trends from the NWP model are not introduced into the SIC data set. The correction scheme implemented in the CDRs is based on a double-difference scheme.

The scheme evaluates the correction offsets δTB (one per channel), the difference between two runs of the RTM: TB_{nwp} uses estimates from NWP fields (in our case ERA5), while TB_{ref} uses a reference atmospheric state with the same air temperature as TB_{nwp} , but zero wind, zero water vapour, and zero cloud liquid water. δ_{TB} is thus an estimate of the atmospheric contribution at the time and location of the observation.

$$TB_{nwp} = F(W_{nwp}, V_{nwp}, L_{nwp} = 0; T_S, SIC_{ucorr}, \theta_0)$$

$$TB_{ref} = F(0, 0, 0; T_S, SIC_{ucorr, instr})$$

$$\delta_{TB} = TB_{nwp} - TB_{ref}$$

$$TB_{corr} = TB - \delta_{TB}$$

For TB_{nwp} , the RTM function F simulates the brightness temperature emitted at view angle θ_0 by a partially ice-covered scene with sea-ice concentration SIC, and with surface and atmospheric states described by W_{nwp} (10 m wind speed, m s⁻¹), V_{nwp} (total columnar water vapour, mm), L_{nwp} (total columnar liquid water content, mm), and T_S (2 m air temperature). θ_{instr} is the nominal incidence angle of the instrument series (varies upon satellites). The double-difference scheme is thus both a correction for the atmosphere influence on the T_B (as predicted by the NWP fields) and a correction to a nominal incidence angle. The typical values of δ_{TB} range from about 10 K over open water to few tenths of a kelvin over consolidated sea-ice. The liquid water content (L) fields from ERA-Interim (used in the CCI and OSI SAF “v2” CDRs) were found to not be accurate enough to be used in our atmospheric correction scheme (Lu et al., 2018). The TBs were thus not corrected for L ($L=0$ in both TB_{nwp} and TB_{ref}).

NEW in CCI+ (Year 2): The RTM-based correction of TBs was re-assessed during Year 2 of the CCI+ Sea Ice project, and we looked if using ERA5 (instead of ERA-Interim) L_{nwp} can have a positive impact. Preliminary results (to be consolidated in D2.3 Update to ADP) indicate that ERA-5 has some skills at L_{nwp} , and that including L_{nwp} in the RTM-correction step has a positive impact over Open Water for the algorithms using the high-frequency (85.5, 89.0 and 91.0 GHz) channels (Fig. 3).

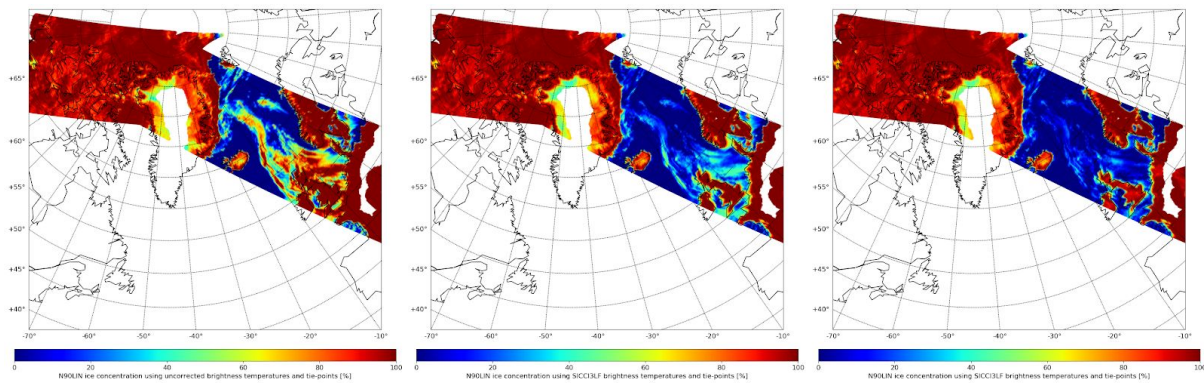


Figure 3: Example positive impact of using ERA5 reanalysis data to reduce noise in output sea-ice concentration using the “N90LIN_corrSICC13LF” algorithm (see Table 2). Left: no atmospheric correction, middle: atmospheric correction without L_{nwp} , right: with L_{nwp} . The impact of atmospheric correction is mostly over open water. In this case, including L_{nwp} is beneficial for noise reduction even along the sea-ice edge. SSMIS F18 data 2016-01-02.

In principle, any RTM will do for implementing F . Here we use the Remote Sensing Systems (RSSs) RTM, for which the tuning to different instruments is documented in Wentz (1983) for SMMR, Wentz (1997) for SSM/I and SSMIS, and Wentz and Meissner (2000) for AMSR-E and AMSR2.

5.4 The distance-along-the-line

Once the unit vector u is known (section 5.3.3 : u is the unit vector sustaining the 100% sea-ice line), a scalar quantity d can be computed for each vector T (e.g. triplet) of T_B :

$$d = u \cdot T$$

We call d the “distance along the line” (DAL). Since u points from multi-year ice to first-year sea ice (Fig. 2), older ice have lower DAL values than younger ice. DAL seems to relate to sea-ice type and sea-ice age (Fig. 5 in Laverigne et al. 2019).

Here, however, the DAL is computed as part of the SIC processing because it is the basis for two critical processing steps that are introduced next: the Open Water Filtering (OWF, section 5.5), and the closed ice correction (section 5.6).

5.5 Open-water filtering

The weather filters (WFs) of Cavalieri et al. (1992) have been used in basically all available SIC CDRs (except the earlier EUMETSAT OSI SAF data sets, Tonboe et al., 2016). WFs are algorithms that combine TB channels to detect when rather large SIC values (sometimes up to 50 % SIC) are in fact noise due to atmospheric influence (mainly wind, water vapour, cloud liquid water effects) and should be reported as open water (SIC = 0 %).

The WF by Cavalieri et al. (1992) detects (and consequently forces SIC to 0 %) all observations with either $GR_{3719v} > 0.050$ and/or $GR_{2219v} > 0.045$ as open water. The GR notation stands for gradient ratio and this quantity is computed, e.g. as

$GR3719v = (TB37v - TB19v) / (TB37v + TB19v)$. $T3719v = 0.050$ and $T2219v > 0.045$ are two thresholds that must be tuned in order to reach consistent behaviour across satellites.

Indeed, while WFs are effective at removing false sea ice in open-water regions, they are also “greedy”: they will always falsely remove (detect as open water) some amount of low-concentration (and/or thin) sea ice, especially along the ice edge (Ivanova et al., 2015). The greediness of the filters is controlled by the thresholds $T3719v$ and $T2219v$.

In CCI Phase 2 and OSISAF, following Lu et al. (2018), we used a WF computed from T_B that has been corrected for atmospheric influence and features a test for $GR3719v$ only. In addition, we dynamically tuned $T3719v = 0.050$ on a daily basis based using the dynamic tie-points samples (section 5.2.4). See also 3.4.2 in Laverne et al. (2019).

The dynamic tuning of the Open Water Filter threshold in the (37V, 19V) space is effective at ensuring temporal consistency across satellite missions (e.g. Fig. 1 and Fig 2. in Kern et al. 2019) but we still observed variations of the effective greediness of the filter with seasons. These were attributed to the fact that the OWF was always applied in the (37V, 19V) space while the SIC is computed in other spaces (e.g. 19V, 37V, 37H for SICCI3LF or 6V, 37V, 37H). This mismatch of frequency channels and spatial resolution in CCI Phase 2 and OSI SAF triggered new developments in CCI+.

New in CCI+ (Year 2): In CCI+, the OWF is tuned and formulated using the same frequency channels as the SIC. In practice, the OWF is computed in the optimized projection plane (Fig. 2), along two axes: (x: d_{OWF} , y: SIC), with d_{OWF} a normalized version of the DAL parameter (section 5.4), such that d_{OWF} is 0 all along a line segment extending from the OW to the FYI tie-point in the (x: d_{OWF} , y: SIC) 2D space.

$$d_{OWF} = d - d_{ref}(SIC)$$

$$d_{OWF} = (u \cdot T) - ((1 - SIC) \times (u \cdot < OW >) + SIC \times (u \cdot < FYI >))$$

In this 2D space, the 100% SIC points will have SIC close to 1, and negative d_{OWF} (since they are “to the left” of the FYI tiepoint). The intermediate SICs will mostly be along or to the left of the (x=0) vertical segment, and most 0% SIC points will have SICs close to 0 and positive d_{OWF} (since the tail of TBs impacted by weather falls “to the right” of the OW tiepoint). In this 2D space, we define a “high-weather” point, d_{HW} at the 95% percentile of the d_{OWF} points and with SIC = 0.

The OWF in CCI+ is based on two binary tests.

- Test1 = (SIC ≤ 0.1);
- Test2 = (SIC ≤ (0.1 + 0.5 * d_{OWF} / d_{HW}));
- OWF = Test1 OR Test2.

T_B observations that fulfill at least one of the two conditions (Test1 or Test2) are flagged (OWF=True) as “probably open water”, and their SIC is set to 0% in the product. Noticeably, our Weather Filters are computed on swath projection (at L2).

We finally note that the name “weather filter” can be misleading as the non-expert could understand that it is meant for filtering out weather effects (false sea ice) from calm open-water and low-ice-concentration conditions. However, this is not how the $GR3719v$ filter works (Fig. 3), as it will remove true sea ice as well, even in calm weather conditions (OW samples below J). For these reasons, we often refer to such a filter as an “open-water filter” (OWF).

5.6 Reducing systematic errors at high-concentration range

By construction, most if not all SIC algorithms (Bootstrap, Bristol, our new dynamic algorithms) consider that the SIC is exactly 100 % when the input T_B falls on the consolidated ice line. The concept of an ice line has sustained the development of SIC algorithms for decades, since it allows algorithms to return SICs close to 100 % for all consolidated ice conditions, whatever the type of sea ice (multi-year ice, first-year ice, mixture of types).

In Lavergne et al. (2019) we introduced a change from the concept of a sea-ice line to that of a sea-ice *curve*, as explained below.

Indeed, careful analysis of the spread of consolidated ice samples along the ice line reveals that systematic deviations exist that are somewhat stable with time. These systematic deviations draw a sea-ice *curve*.

These deviations are computed in a coordinate system in which abscissae are computed as $u.T$ (the dot product of u -the unit vector sustaining the consolidated ice line- and T a T_B triplet) and the ordinate as $B_{CI}(T)$ (the result of the best-ice SIC algorithm for a given T_B triplet).

The proposed correction scheme moves the concept of an ice line to that of an ice curve, that more closely follows the $B_{CI}(T)$ samples along the u axis. A new ice curve L is tabulated for each day in the record by binning the $B_{CI}(T)$ values by their $u.T$ values. The correction algorithm for a given T_B observation triplet T goes then by:

- Compute $C = B_{CI}(T)$;
- Compute $d = u.T$
- Evaluate $L(d)$: value of curve L at index d (1D interpolation from tabulated values).
- Corrected SIC is $C / L(d)$
- Use the corrected SIC in the hybrid SIC formula (Sect. 5.2.1) in place of B_{CI} .

This consolidated ice curve defines the SIC 100 % isoline during the 2nd iteration of the L2 chain only, and has most effect on near-100% SIC regions. In the first iteration (before atmospheric correction), a straight ice line is used.

5.7 Uncertainties at L2

Uncertainty estimates are needed when the SIC data are compared to other data sets or when they are assimilated into numerical models. The mean accuracy of some of the more common algorithms, used to compute ice concentration from SSM/I data, such as NASA Team and Bootstrap are reported to be 4-6 % in winter (Andersen et al., 2007; Ivanova et al. 2015) but the actual value varies with instrument, region, ice condition, etc. and time-varying maps of uncertainties are needed.

We make the assumption that the total uncertainty can be written as

$$\sigma_{tot}^2 = \sigma_{algo}^2 + \sigma_{smear}^2$$

where σ_{algo} is the inherent uncertainty of the concentration algorithm, and σ_{smear} is the uncertainty due to resampling to a grid where the sensor footprint covers more than one pixel. In this section, we only cover σ_{algo} . The contribution σ_{smear} is indeed computed at L3, after gridding and daily compositing, and is documented in Sect. 6.2.

We first introduce how uncertainties are computed for a SIC algorithm, then how they are combined for resulting in the SIC uncertainty of an hybrid SIC algorithm.

5.7.1 Algorithm and tie-point uncertainty

Both the water surface and ice surface emissivity variabilities result in SIC uncertainties. Emission and scattering in the atmosphere also affects the T_b and the computed ice concentrations. Different algorithms have different sensitivities to these surface and atmospheric parameters (Andersen et al., 2007; Ivanova et al., 2015).

The algorithm uncertainty is the weighted average of the algorithm uncertainty obtained at 0% SIC and 100% SIC. Thus:

$$\sigma_{\text{algo}}^2 = (1 - C^2) \times \sigma_{CI}^2 + C^2 \times \sigma_{OW}^2$$

where σ_{OW} (resp σ_{CI}) is the uncertainty (1 standard deviation) obtained when applying the SIC algorithm on the 0% SIC (resp 100% SIC) training data samples.

The formula above is applied to both the B_{OW} and B_{CI} algorithms separately, and two uncertainty values are obtained: σ_{BOW} and σ_{BCI} . These two values are then combined to obtain the algorithm uncertainty of the hybrid algorithm, as described in the next section.

5.7.2 Uncertainty for the hybrid algorithm

The algorithm uncertainty of the hybrid algorithm is computed as a linear combination of the variances. Using linear combination of variances as the resulting variance is in line with the hypothesis that the uncertainties of both B_{OW} and B_{CI} algorithms are strongly correlated to each other ($\sim +1$). This approach prevents what would be an artificial reduction of uncertainties in the merging process. The same linear weights as used in the hybrid SIC algorithm are used for mixing σ_{BOW}^2 and σ_{BCI}^2 , and the resulting standard deviation σ_{algo} is stored in the product swath file.

6 GRIDDING AND DAILY COMPOSITING (L3)

6.1 Gridding and daily averaging

The gridding and daily averaging step loads all satellite observation within 24 hours, centered on 12:00 UTC, and grid these to the final output grids. There are two such grids, one for the northern and one for the southern hemispheres. Both have a grid spacing of 25 km.

The gridding uses a KD-Tree search based on the distance between the cartesian coordinates (Xg, Yg, Zg) of the centers of the target grid cells, and the cartesian coordinates (Xs, Ys, Zs)

of the center of the satellite's FoVs. The search for N neighbours is constrained to a radius of influence of 25 km around each grid cell. Once the N closest FoV neighbours are known for each grid cell, they are combined in a daily averaged value, all with equal weight. A large enough value is chosen for N that allows all available SSM/I and SSMIS swaths to be combined for each day.

The gridding is done for all areas with data coverage, including the coastal zone and land grid cells in the direct vicinity of ocean grid cells. A gridded field is made for all the variables that might be of interest: the ice concentration estimates (both based on corrected and uncorrected T_B), the algorithm uncertainties (they are gridded as variances, not standard deviations), open water filters, etc.

When gridding open water filters (a binary field), the value “probably open water” is assigned a grid cell when more than half FoVs that map to the grid cell (nearest neighbour) showed “probably open water”.

6.2 Gridding and smearing uncertainty

The smearing uncertainty is the error due to the sensor footprint covering more than one pixel in the L3 product grid. Footprint sizes for the channels used for ice concentration mapping range from over 50 km, for the 19 GHz channels, to about 30 km, for the 37 GHz channels (SSMIS values). These footprints, of uneven size, are combined in the algorithms when computing the ice concentration and this leads to an additional smearing effect. We call this the footprint mismatch error. In addition, the ice concentration data are gridded and represented on predefined grids that have finer resolution (in our case 25 km) than the footprints. All these effects lead to a “smearing uncertainty” term, that is parametrized and computed at L3. See Lavergne et al. (2019) and Tonboe et al. (2016) for a description of the parametrization.

In the processing chain, the smearing uncertainty σ_{smear} is parametrized with a proxy that measures the local variability of the SIC field. We found σ_{smear} to be proportional to the 3 x 3 pixel max - min sea ice concentration difference. This includes both the smearing and the foot-print mismatch and it is thereby the total smearing uncertainty. To avoid computing the smearing below the sea ice concentration noise floor and not to exceed the range of values which were computed with the smearing simulator, the smearing error is:

$$\sigma_{\text{smear}} = 3 \times 3 \text{ max - min SIC difference}$$

$$\text{if } \sigma_{\text{smear}} < \sigma_{\text{algo}} : \sigma_{\text{smear}} = 0 ;$$

$$\text{if } \sigma_{\text{smear}} \geq 40\% : \sigma_{\text{smear}} = 40\% ;$$

7 FINAL ANALYSIS AND FILTERING (L4)

7.1 Apply masks and filters

7.1.1 Masking for land regions

The binary land masks (Sect. 2.3) are used to mask the final product: all gridded SIC values falling over land are replaced by NaN.

7.1.2 Climatological maximum extent masking

The expanded monthly sea-ice climatology (Sect. 2.4) is used to mask the final product: all gridded SIC values falling outside the climatology are set to 0% SIC.

7.1.3 Prepare cleaned SIC maps

The daily gridded SIC in Level 3 files range between (approximately) -20% and +120% SIC, and still show impact by weather effects over open water. At Level 4, two “cleaned” SIC variables are prepared for distribution to the users:

- `ice_conc` is the main SIC variable. It is obtained from the L3 SIC via two steps. Set all grid cells where the daily gridded OWF indicates that this is “probably open water” (see Sect. 5.3) to 0% SIC. Set all grid cells where the L3 SIC is > 100% SIC to exactly 100% SIC.
- `raw_ice_conc_values` is the secondary SIC variable. It contains the original (“raw”) L3 SIC values at the locations (within the climatological maximum) where the two steps described above (and 7.1.4) were applied.

These two SIC variables are an important feature of the SIC CDRs as they front the users with a “ready to use” SIC `ice_conc` variable that has minimum noise from weather effects and is restrained to the physical range [0;100% SIC] while at the same time preserving the full (symmetric) error distribution at both 0% and 100% SIC (by combining `ice_conc` and `raw_ice_conc_values`).

7.1.4 Additional land spill-over correction (coastal regions)

An additional grid-based correction step similar to that of Cavalieri et al. (1999) is implemented at Level 4. This correction step has less impact than the swath-based correction step described in section 4.4.

7.2 Gap filling by interpolation

For easing the use of the climate data record, some level of spatial interpolation is performed for reducing the occurrence of data gaps. Only missing data are interpolated. Interpolated

data points are clearly marked in the product file, so that users can choose to discard them and only ingest retrievals that rely on satellite signal.

Data gaps can occur in several forms, such as missing scan lines, missing orbits and the polar observation hole (NH only). While simple spatial interpolation might be efficient in filling small gaps (e.g. one or two missing scan lines), it blurs the sea ice concentration features. This effect becomes overwhelming when large areas are missing. To overcome this issue, yet implementing a general approach for all cases, the ice concentration estimates from the previous and next daily products are used in the interpolation as well. In the case of SSM/I, SSMIS, AMSR-E, and AMSR2, it means that interpolation on a given date D uses pixels from 3 data files: $D-1$, D and $D+1$. We use $D-2$, D , and $D+2$ for SMMR.

Gap-filling by interpolation is implemented in two steps: first a temporal interpolation, then a spatial interpolation.

7.2.1 Temporal interpolation

All gaps (non-land grid cells with missing SIC data) at day D are identified. For these gaps, the average of the SIC from $D-1$ and $D+1$:

$$SIC_{i,j,D} = 0.5 \times (SIC_{i,j,D-1} + SIC_{i,j,D+1})$$

In cases where only one of the $D-1$ or $D+1$ maps have data at i,j coordinate, $SIC_{i,j,D}$ is set to this value. Many gaps will be filled by the temporal interpolation step, but some will remain, for example the polar observation hole in the NH.

7.2.2 Spatial interpolation

If there are still gaps in the SIC map at day D after the temporal interpolation step, these are filled by spatial interpolation using only the data from day D . This is implemented by a Gaussian weighting function of the distance. As in Tonboe et al. (2016), we use $R = |\text{lat}_{i,j}|$ (the absolute value of the latitude in degrees at grid cell i,j) as one standard deviation of the weighting function. This radius allows longer interpolation lengths at high latitudes (where the polar observation hole is), than at lower latitude (where the data gaps are smaller, and often filled by the temporal interpolation step). The weight is additionally localized into a $[-3R; +3R] \times [-3R; +3R]$ neighbourhood of grid cell (i,j) .

7.3 Compute total uncertainty

The total uncertainty (variance) is computed at L4 as the sum of the variances of the (gridded and daily averaged) algorithm uncertainty (Sect. 5.5), and of the smearing uncertainty (Sect. 6.2):

$$\sigma_{tot}^2 = \sigma_{algo}^2 + \sigma_{smear}^2$$

7.4 Possible melting or high T2m flag

A quality flag using the ERA5 T2m (air temperature at 2 meters) field is added in the processing. The T2m ERA5 values that have been interpolated in time and space to each FoV are compared to +5C. The binary results of the flags (1 if T2m > 5C and 0 otherwise) are gridded and daily averaged with the same procedure as SIC (see Sect. 6.1). The resulting map shows the frequency of “high T2m value” during the day, and is used for triggering flags in the product files. Such a flag can however only be used for warning the users of possible melting events (or false sea ice), not correcting the SIC values. This is because of the uncertainty of ERA5 T2m at high latitudes, and the possible trends it could carry into the final SIC climate data record.

8 TEMPORAL AGGREGATION

This ATBD so far described how L1B T_b swath data is turned into L2 SIC (still on swath), and finally composited to daily maps (L3 and L4). Some users also require SIC products aggregated on longer periods, e.g. weekly, monthly, seasonally, yearly. The present section describes how such temporally aggregated SIC products are prepared.

As introduced in Sect. 7.1.3, the L4 daily files have two daily files: `ice_conc` and `raw_ice_conc_values`. The first holds the filtered SIC values, limited to the range [0%-100%], while the second holds SIC values that result from the same algorithms, but that are either outside the [0%-100%] range, or are detected as “probably open water” by the Open Water Filter.

When averaging SIC over longer periods (e.g. a month), it is important to take into account the off-range values, so that not to introduce additional biases. Indeed, and considering the case of 100% SIC, uncertainties in the T_b data and the retrieval algorithms lead to an error distribution in SIC that is almost Gaussian in shape and ideally centered on 100% (see e.g. Fig. 11 in Laverne et al. 2019). Since `ice_conc` is truncated at 100%, an average of these values will tend to be biased low wrt 100% SIC. To average the values from the full SIC distribution is more correct, and allows to keep the mean value closer to 100% SIC in the regions of full sea-ice cover. The same yields at 0% SIC.

The temporal aggregation algorithm implemented is:

1. For each day in the month, read `ice_conc` and `raw_ice_conc_values`;
2. For each day, combine `ice_conc` and `raw_ice_conc_values` to obtain a daily field with all “raw” values: below 0%, above 0% but detected by the OWF, and above 100%.
3. Compute the monthly mean (and standard deviation);
4. Split the monthly mean variable in two variables: `ice_conc` that holds only mean SICs in range [0-100%] and `raw_ice_conc_values` that holds mean SICs above 100%, and below a threshold e.g. 10%.
5. Write `ice_conc`, `raw_ice_conc_values`, and `ice_conc_variability` (the standard deviation of the daily values).

For the uncertainty fields, the compositing is implemented as:

1. Compute monthly average of `smearing_standard_error` and `algorithm_standard_error` after the daily maps are turned into variances (i.e. squared).
2. Compute `total_standard_error` (variance) as the sum of the two monthly averaged standard error fields from the earlier step.
3. Write all 3 monthly averaged uncertainties in the product file as standard deviations (square root).

9 ALGORITHMS FOR ESMR (1972 - 1977)

The Electrically Scanning Microwave Radiometer (ESMR) instrument on board the NIMBUS 5 satellite was a one channel 19.35 GHz horizontally polarised microwave radiometer operating from Dec. 1972 to May. 1977. The data have recently been made available online by NASA. Even though ESMR was a predecessor of modern multi frequency radiometers there are still parts of our SICCI processing methodology which can be applied to the data to derive the sea ice extent globally and potentially sea ice type and sea ice concentration. In fact both the dynamical tie-points and the atmospheric noise reduction of the Tb's which are used in the OSISAF/CCI processing chain for modern radiometer data (SMMR, SSMI, AMSR...) can reduce the noise over ice and open water consistently with ESMR. The ESMR data will extend sea ice climate record with an important reference from the 1970ies.

The ESMR instrument was a cross-track scanner measuring at 78 scan positions from nadir to 50 degrees perpendicularly to both sides of the flight track. The orbit height was about 1100 km with an inclination of 81 degrees. The phased array antenna size was about 85 x 83 cm and the spatial resolution about 25 km at nadir increasing to about 160 x 45 km at the edges of the swath. The full swath was about 3000 km which is much more than modern conically scanning radiometers. The data from one swath (Fig. 4a) and the combination of data from one day (Fig. 4b) is shown below. ESMR has complete daily coverage globally and there is very good daily coverage in polar regions. The excellent coverage is exploited in the processing of data and derivation of sea ice parameters.

The complete ESMR Level 1 data record contains calibrated radiances expressed in units of brightness temperature is available from NASA where it is archived in the original IBM binary proprietary format, also referred to as a binary TAP file. These TAP files have been converted to NetCDF for further processing.

Algorithms for processing sea ice edge maps from ESMR data are under development, and will be documented when they are more mature.

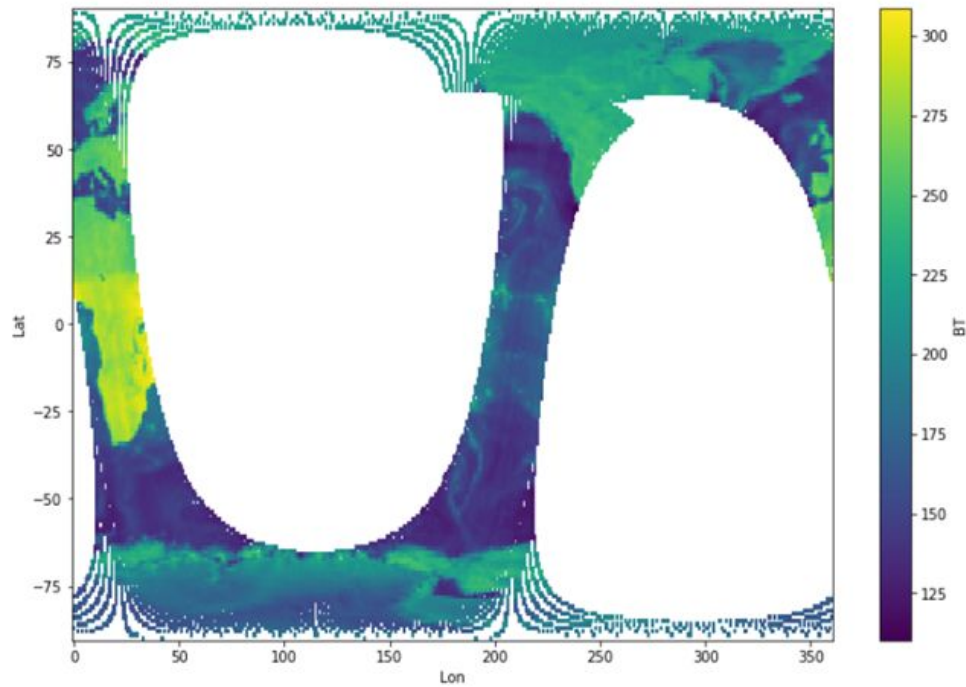


Figure 4a. The 19 GHz horizontally polarised brightness temperatures in Kelvin for approximately one orbit, on 14. Dec.1972 about 9 UTC. The swath width is about 3000 km.

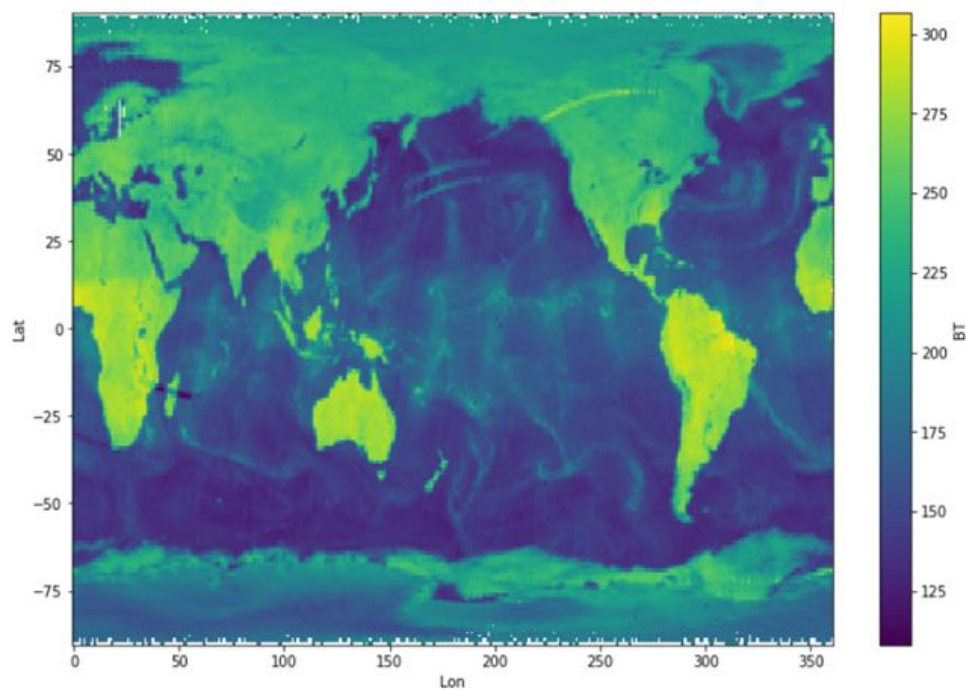


Figure 4b. The combined brightness temperatures (T_b) in Kelvin for one day on the 14. Dec.1972 (showing the mean T_b value at points with multiple observations).

10 REFERENCES

- Andersen, S., Tonboe, R., Kern, S., and Schyberg, H.: Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using Numerical Weather Prediction model fields: An intercomparison of nine algorithms, *Remote Sens. Environ.*, 104, 374–392, 2006.
- Andersen, S., Toudal Pedersen, L., Heygster, G., Tonboe, R., and Kaleschke, L.: Intercomparison of passive microwave sea ice concentration retrievals over the high concentration Arctic sea ice, *J. Geophys. Res.*, 112, C08004, <https://doi.org/10.1029/2006JC003543>, 2007.
- Cavalieri, D. J., Gloersen, P., and Campbell, W. J.: Determination of Sea Ice Parameters With the NIMBUS 7 SMMR, *J. Geophys. Res.*, 89, 5355–5369, 1984.
- Cavalieri, D. J., Crawford, J., Drinkwater, M., Emery, W. J., Eppler, D. T., Farmer, L. D., Goodberlet, M., Jentz, R., Milman, A., Morris, C., Onstott, R., Schweiger, A., Shuchman, R., Steffen, K., Swift, C. T., Wackerman, C., and Weaver, R. L.: NASA sea ice validation program for the DMSP SSM/I: final report, NASA Technical Memorandum 104559, National Aeronautics and Space Administration, Washington, D.C., 126 pp., 1992.
- Cavalieri, D. J., Parkinson, C. L., Gloersen, P., Comiso, J. C., and Zwally, H. J.: Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets, *J. Geophys. Res.*, 104, 15803–15814, <https://doi.org/10.1029/1999JC900081>, 1999.
- Comiso, J. C.: Characteristics of arctic winter sea ice from satellite multispectral microwave observations, *J. Geophys. Res.*, 91, 975–994, 1986.
- Donlon, C. J., Martin, M., Stark, J. D., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice analysis (OSTIA), *Remote Sens. Environ.*, 116, 140–158, <https://doi.org/10.1016/j.rse.2010.10.017>, 2012.
- Gloersen, P., and F. T. Barath. A scanning multichannel microwave radiometer for Nimbus-G and SeaSat-A. *IEEE Journal of Oceanic Engineering* OE-2(2), 172-178, 1977.
- Gloersen, P., W. J. Campbell, D. J. Cavalieri, J. C. Comiso, C. L. Parkinson, H. J. Zwally. Arctic and Antarctic sea ice, 1978-1987: satellite passive-microwave observations and analysis. NASA SP-511, Washington D. C., 1992.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N. (2020), The ERA5 Global Reanalysis. *Q J R Meteorol Soc.* Accepted Author Manuscript. doi:10.1002/qj.3803
- Hobbs, W., Massom, R., Stammerjohn, S., Reid, P., Williams, G., and Meier, W.: A review of recent changes in Southern Ocean sea ice, their drivers and forcings, *Global Planet. Change*, 143, 228–250, <https://doi.org/10.1016/j.gloplacha.2016.06.008>, 2016.
- Ivanova, N., Pedersen, L. T., Tonboe, R. T., Kern, S., Heygster, G., Laverigne, T., Sørensen, A., Saldo, R., Dybkjær, G., Brucker, L., and Shokr, M.: Inter-comparison and evaluation of

sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations, *The Cryosphere*, 9, 1797–1817, <https://doi.org/10.5194/tc-9-1797-2015>, 2015.

Kern, S., Lavergne, T., Notz, D., Pedersen, L. T., Tonboe, R. T., Saldo, R., and Sørensen, A. M.: Satellite passive microwave sea-ice concentration data set intercomparison: closed ice and ship-based observations, *The Cryosphere*, 13, 3261–3307, <https://doi.org/10.5194/tc-13-3261-2019>, 2019.

Kwok, R.: Sea ice concentration estimates from satellite passive microwave radiometry and openings from SAR ice motion, *Geophys. Res. Lett.*, 29, 1311, <https://doi.org/10.1029/2002GL014787>, 2002.

Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjær, G., Eastwood, S., Gabarro, C., Heygster, G., Killie, M. A., Brandt Kreiner, M., Lavelle, J., Saldo, R., Sandven, S., and Pedersen, L. T.: Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records, *The Cryosphere*, 13, 49–78, <https://doi.org/10.5194/tc-13-49-2019>, 2019.

Lu, J., Heygster, G., and Spreen, G.: Atmospheric Correction of Sea Ice Concentration Retrieval for 89 GHz AMSR-E Observations, *IEEE J-STARS.*, 11, 1442–1457, <https://doi.org/10.1109/JSTARS.2018.2805193>, 2018.

Maass, N. and Kaleschke, L. Improving passive microwave sea ice concentration algorithms for coastal areas: applications to the Baltic Sea, *Tellus A*, 62, 393–410, <https://doi.org/10.1111/j.1600-0870.2010.00452.x>, 2010.

Meier, W., Fetterer, F., Savoie, M., Mallory, S., Duerr, R., and Stroeve, J.: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 3, NSIDC: National Snow and Ice Data Center, Boulder, Colorado, USA, <https://doi.org/10.7265/N59P2ZTG>, 2017.

Smith, D. M.: Extraction of winter total sea ice concentration in the Greenland and Barents Seas from SSM/I data, *Int. J. Remote Sens.*, 17, 2625–2646, 1996.

Smith, D. M. and Barrett, E. C.: Satellite mapping and monitoring of sea ice, CB/RAE/9/2/4/2034/113/ARE, RSU, University of Bristol, Bristol, UK, 1994.

Tonboe, R. T., Eastwood, S., Lavergne, T., Sørensen, A. M., Rathmann, N., Dybkjær, G., Pedersen, L. T., Høyer, J. L., and Kern, S.: The EUMETSAT sea ice concentration climate data record, *The Cryosphere*, 10, 2275–2290, <https://doi.org/10.5194/tc-10-2275-2016>, 2016.

Wentz, F. J.: A model function for ocean microwave brightness temperatures, *J. Geophys. Res.*, 88, 1892–1908, <https://doi.org/10.1029/JC088iC03p01892>, 1983.

Wentz, F. J.: A well-calibrated ocean algorithm for SSM/I, *J. Geophys. Res.*, 102, 8703–8718, <https://doi.org/10.1029/96JC01751>, 1997.

Wentz, F. J. and Meissner, T.: AMSR ocean algorithm version 2, RSS Tech, Proposal 121599A-1, Remote Sensing Systems, Santa Rosa, California, 2000.