EUMETSAT OSI SAF
OCEAN AND SEA ICE

Algorithm Theoretical Basis Document for Global Sea Ice Concentration Level 2 and Level 3

OSI-410-a, OSI-401-d, OSI-408-a

Version : 1.2

Date : 01/11/2022

Prepared by :

Fabrizio Baordo, Rasmus Tonboe and Eva Howe

The Danish Meteorological Institute

Document Change record

Document version	Software version	Date	Author	Change description
v0.1		26/08/2021	FAB	First version
V1.0		01/10/2021	FAB	Updated after PCR review. Corrected minor editorial things Section 1.5 : Added list of main developments Section 2.0 : Description updated Section 3.1 : Added justification for SSMIS Section 3.2 and 4.2.3: Removed description of Wentz and Meissner RTM Section 4.1 : Updated the description of the 30 days running mean Section 4.2.5 and 4.2.6 : Reformulation of text
V1.1		23/05/2022	FAB	Section 4.2.2: updated description of static tie-points Section 4.2.3: updated limits of NWP skin temperature screening test Section 4.2.6: updated description of screening tests to filter the sea ice concentration Added Section 4.2.7 to describe logic to handle the 'near coast' points Section 5: modified to improve description of L3 processing
V1.2		01/11/2022	FAB	Updated version following recommendations of the ORR Review. Correction of minor typographical errors

Table of contents

1. Introduction

1.1. The EUMETSAT Ocean and Sea Ice SAF

The Satellite Application Facilities (SAFs) are dedicated centres of excellence for processing satellite data – hosted by a National Meteorological Service – which utilise specialist expertise from institutes based in Member States. EUMETSAT created Satellite Application Facilities (SAFs) to complement its Central Facilities capability in Darmstadt. The Ocean and Sea Ice Satellite Application Facility (OSI SAF) is one of eight EUMETSAT SAFs, which provide users with operational data and software products. More on SAFs can be read at www.eumetsat.int.

The objective of the OSI SAF is the operational near real-time production and distribution of a coherent set of information, derived from earth observation satellites, and characterising the ocean surface and the energy fluxes through it: sea surface temperature, radiative fluxes, wind vector and sea ice characteristics. For some variables, the OSI SAF is also aiming at providing long term data records for climate applications, based on reprocessing activities.

The OSI SAF consortium is hosted by Météo-France. The sea ice processing is performed at the High Latitude processing facility (HL centre), operated jointly by the Norwegian and Danish Meteorological Institutes.

The sea ice products include sea ice concentration, the sea ice emissivity at 50 GHz, sea ice edge, sea ice type and sea ice drift and sea ice surface temperature (from mid 2014).

1.2. Disclaimer

Note: The comments that we get from our users is an important input when defining development activities and updates, and user feedback to the OSI SAF project team is highly valued.

Acknowledgement and citation

Use of the product(s) should be acknowledged with the following citations (product specific):

OSI-401-d: OSI SAF (2021): Global Sea Ice Concentration (SSMIS) Level 3, EUMETSAT SAF on Ocean and Sea Ice.

OSI-408-a: OSI SAF (2021): Global Sea Ice Concentration (AMSR-2) Level 3, EUMETSAT SAF on Ocean and Sea Ice.

OSI-410-a: OSI SAF (2021): Global Sea Ice Concentration (AMSR-2, SSMIS) Level 2, EUMETSAT SAF on Ocean and Sea Ice.

1.3. Scope of this document

This document describes the OSI SAF sea ice concentration products OSI-410-a, OSI-401-d and OSI-408-a.

The Level 2 global sea-ice concentration product (OSI-410-a) is derived from Level 1 passive microwave measurements from the SSMIS and AMSR2 instruments on-board the DMSP and JAXA's GCOM-W1 polar orbiting satellites respectively. The processing steps are identical for each of the instruments, but the final product is characterized by a different resolution because of the diverse frequencies and footprints of the sensors. The two Level 3 products (OSI-401-d and OSI-408-a) are daily gridded sea-ice concentration products, covering the Northern and Southern Hemispheres.

The Level 2 product is for users who require a temporal resolution of better than 24 hours and/or reduced timeliness compared to the Level 3 product. In addition, the Level 2 product is characterized by a lower uncertainty. This is because the L2 sea ice concentration products are not re-sampled and the original satellite projection is retained. This lower uncertainty is only achieved when the data is used together with a footprint operator so that the extent and weighting of the measurement is included.

The Level 3 sea ice concentration product is available 6 hours after midnight of the following day. Data used in the Level 3 product are collected for a full day from 00 to 24 hours. The requirements to the Level 2 product timeliness is that the products should be available to the user 150 minutes after satellite overpass (2 hours and 30 minutes).

1.4. Glossary

AMSR	Advanced Microwave Scanning Radiometer
ATBD	Algorithm Theoretical Basis Document
CC	Fraction of Cloud Cover (NWP parameter)
CDOP	Continuous Development and Operations Phase (OSI SAF project)
CIWC	Specific Cloud Ice Water Content (NWP parameter)
CLWC	Specific Cloud Liquid Water Content (NWP parameter)
CSWC	Specific Snow Water Content (NWP parameter)
DMI	Danish Meteorological Institute
DMSP	Defense Meteorological Satellite Program
ECMWF	European Centre for Medium range Weather Forecast
ESA SICCI	European Space Agency Sea Ice Climate Change Initiative
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
FG	First Guess
FoV	Field of View

FWHM	Full Width at Half Maximum
FYI	First year ice
HDF	Hierarchical Data Format
JAXA	Japan Aerospace Exploration Agency
MARS	Meteorological Archival and Retrieval System
MYI	Multi year ice
NSIDC	National Snow and Ice Data Center
NWP	Numerical Weather Prediction
NWP SAF	Numerical Weather Prediction Satellite Application Facility
OSI SAF	Ocean and Sea Ice Satellite Application Facility
RTM	Radiative Transfer Model
RTTOV	Radiative Transfer for TOVS
SAR	Synthetic Aperture Radar
SIC	Sea Ice Concentration
SSMIS	Special Sensor Microwave Imager Sounder
T_B	Brightness Temperature
TCWV	Total Column Water Vapour (NWP parameter)
TUD	Technical University of Denmark

1.5. Overview

This ATBD summarizes the upgrades which have been applied to the day-to-day sea ice concentration algorithm. The algorithm has been changed to unify the Level 2 and Level 3 processing chain in one unique framework which is capable of ingesting microwave observations from both AMSR2 and SSMIS (F-16, F-17, F-18). The sequential logic of the algorithm is unchanged although technical developments and new features were introduced to improve the accuracy of the Level 2 processing chain and to allow the dissemination of the Level 2 product. Currently, the operational algorithm implements radiative transfer calculations which are used to derive the magnitude of the atmospheric contribution that is applied to the brightness temperatures. In particular, the Wentz and Meissner radiative transfer model has always been employed in the operational processing chain. However, in this version of the algorithm, we have implemented the possibility of using the community NWP SAF radiative transfer model RTTOV. Apart from the improvements that can characterize the quality of the

simulated brightness temperatures, RTTOV surely ensures a much easier and quicker way to ingest microwave observations from new satellite sensors. As a general overview, the main developments which characterize the new version of the algorithm can be summarised as follows:

- Microwave observations from different satellites/sensors are processed by the same framework
- More robust diagnostic to possibly prevent the use of corrupted/missing microwave observations and anomalies in the algorithm processing chain
- Dissemination of the Level 2 product
- Update on the static Tie-Points used to identify the 'first guess' SIC
- RTTOV (version 13) as the new component to perform radiative transfer calculations (the python wrapper interface is used)
- L2/L3 products characterize by identical screening tests to 'filter' the sea ice concentration
- Changes in the interpolation of near coast points
- Python code upgraded to python 3.8 (previously python 2.7)

1.6. Reference and applicable documents

1.6.1. Applicable documents

[1] EUMETSAT OSI SAF
Product Requirements Document
SAF/OSI/CDOP3/MF/MGT/PL/2-001, version 1.9, 31/12/2021

1.6.2. Reference documents

[1] EUMETSAT OSI SAF
The EUMETSAT OSI SAF Sea Ice Concentration Algorithm, ATBD, OSI-401-b
SAF/OSI/CDOP/DMI/SCI/MA/189, version 1.5, 29/04/2016

[2] EUMETSAT OSI SAF
The EUMETSAT OSI SAF AMSR-2 Sea Ice Concentration Algorithm, ATBD, OSI-408
SAF/OSI/CDOP2/DMI/SCI/MA/248, version 1.0, 27/01/2016

[3] EUMETSAT OSI SAF
The EUMETSAT OSI SAF Level-2 Sea Ice Concentration Algorithm, ATBD, OSI-410
SAF/OSI/CDOP3/DMI/SCI/MA/341, version 1.3, 07/07/2020

2. Processing chain overview

The primary technical changes applied to the algorithm were developed to have a unique framework capable of ingesting and processing microwave observations from both AMSR2 and SSMIS and also to allow the generation and dissemination of the Level 2 (L2) sea ice concentration product (OSI-410-a). Figure 1 provides a schematic overview of the different operational implementation which now permits the L2 production. To process 1 single satellite swath file and be able to disseminate the corresponding L2 product in near real time, we needed to separate the task which provides the calculations of the

ocean and ice tie-points and statistics (standard deviations) relative to the sea ice concentration algorithm (in Figure 1 named as ‘calc tp & stats’). In the new implementation, if for instance 20210316 [YYYYMMDD] is the date of the L3 product, the ‘calc tp & stats’ task is applied to the swath files of the previous day (20210315) which have been through the L2 processing logic (shown in Figure 2). The ‘calc tp & stats’ outputs (stored in json and csv files) are then supplied to every single swath file to calculate the sea ice concentration and the errors which characterize the Level 2 sea ice concentration product. Finally, all the ‘L2 swath files’ will be used to generate the L3 product (20210316). This is the principal difference with respect to the previous operational processing logic where the Level 3 sea ice concentration product was the only output disseminated. In the old set up, we basically had ‘all in one process’ where the ‘calc tp & stats’ was executed on the same swath files which were used to calculate the L3 product (e.g. 20210316).

As shown in Figure 2, ‘calc tp & stats’ is directly related to the outputs of the L2 processing steps: to perform the calculation of the daily and 30-day average tie-points as well as the calculation of the sea ice concentration statistics, we need to consider the swath files of the previous day which have been successfully processed by the L2 logic. As a consequence the resultant ocean and ice tie-points and statistics depend on the total number of swath files which are available in 1 day. There are different possible causes which can impact on the result of the ‘calc tp & stats’ task. For instance, the Level 1 AMSR2/SSMIS dissemination might be delayed or some observations in the swath files might be corrupted. Additionally, during the summer season, the number of ice points (sea ice concentration greater than 95%) calculated by the static tie-points NASA team algorithm might be very small or even null and, consequently, we might have problem to identify the dataset used for the calculations of the ice tie-points (although this is very rarely: during summer 2020, only 2 times we identified less than 10 ice points in the Northern hemisphere). For these reasons, to have a more robust operational framework, in this new L2 processing logic, we introduced: a) a gross check on the Level 1 microwave observations (to prevent the use of corrupted or missing observations); b) a monitoring system (in terms of time series) for the daily and 30-day average tie-points and for the sea ice concentration statistics; c) an automatic warning system which can notify if the number of ocean and ice points is largely decreased respect to the mean value observed in the previous 15 days. As an additional safety precaution, if for any reason, ‘calc tp & stats’ have failed to generate tie-points and statistics for a specific date, the L2 processing will look for the first available tie-points and statistics files in the previous 7 days.

The general processing flow for the L2 and L3 production is highlighted in Figure 2 and the sequential logic can be summarized with the following steps:

1. The Level 1 microwave observations are corrected to account for land-spillover. This is a new feature implemented in this release of the algorithm and it is based on the same method as that described in Lavergne et al. (2019).
2. The Level 1 microwave observations corrected to account for land-spillover (output of step 1) are successively used to calculate the first guess sea ice concentration (FG SIC) following the NASA team algorithm which implements static tie-points.
3. Using the NWP data we compute radiative transfer calculations to derive the magnitude of the atmospheric contribution that is applied to the brightness temperatures.
4. Given the outputs of the ‘calc tp & stats’ task (tie-points and statistics) we calculate:
    a. the final OSI SAF sea ice concentration using the 30-day average tie-points together with the atmospheric corrected brightness temperatures of point 3;

b. the total error of the sea ice concentration at every point of the satellite swath.

Processing steps 1-4 are consistently applied to every single swath file in near real time to generate L2 sea ice concentration product (OSI-410-a) and successively all the L2 swath files are considered to generate the final L3 product (OSI-401-d or OSI-408-a). In section 3, we will provide more details on the input data and in section 4 and 5 we will thoroughly describe the algorithm logic which is schematically represented in Figure 2.

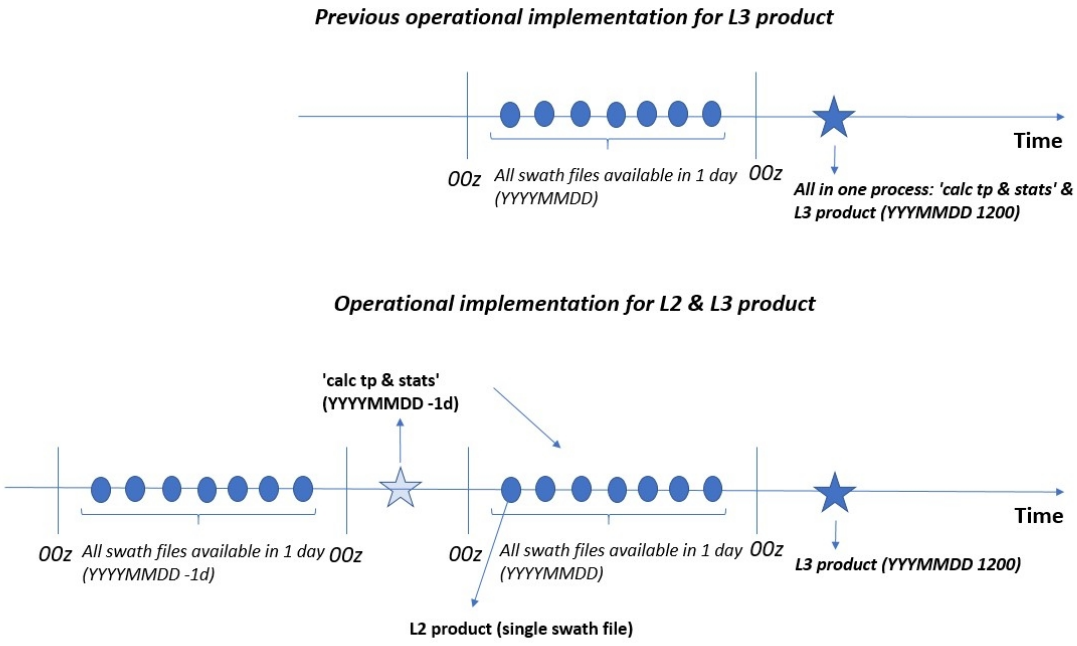

Figure 1: Schematic view of the operational Level 2 (L2) and Level 3 (L3) production.

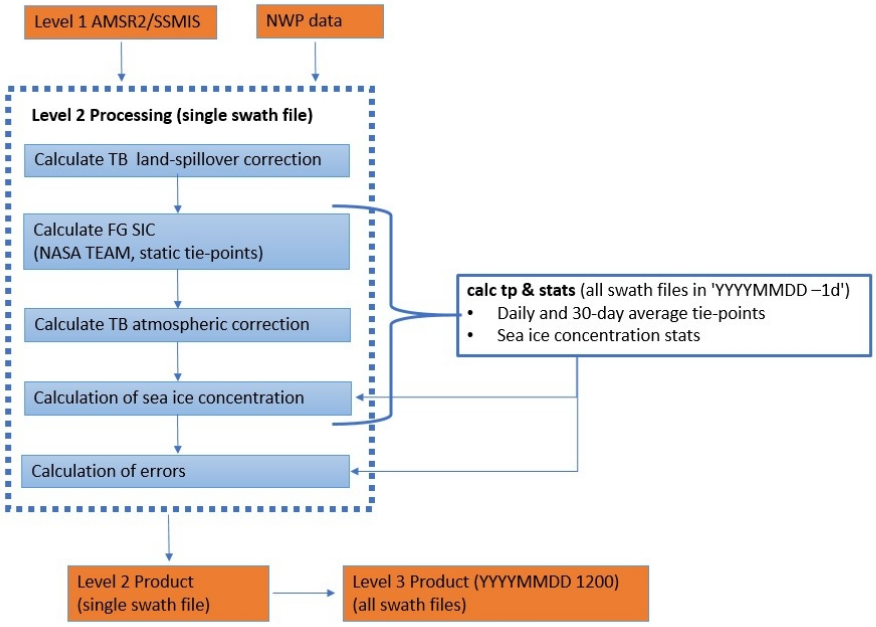

Figure 2: Flow chart for the Level 2 and Level 3 sea ice concentration processing chain.

3. Input Data

This section describes the necessary input data for the OSI SAF sea ice concentration algorithm which, as highlighted in Figure 2, requires near real time observations from microwave satellite sensors and data from Numerical Weather Prediction (NWP) systems. More details on the input data are provided below.

3.1. Observations

The primary input for the sea ice concentration algorithm is Level 1 geolocated microwave brightness temperatures. In the current operational configuration, the Level 2 processing chain ingests observations collected by the passive conically scanning microwave radiometers AMSR2 and SSMIS. As illustrated in the following 2 sections, AMSR2 and SSMIS have channels which are very close in frequency, but different in spatial resolution. Consequently, the sea ice products generated by AMSR2 observations are characterised by higher spatial resolution. However, the near real time sea ice concentration retrievals individually generated by AMSR2 and SSMIS are beneficial to provide a consistent temporal and spacial coverage of the Arctic and Antarctic regions. Since this new version of the algorithm is based on a more flexible processing framework which can handle observations from different microwave satellite sensors, it is now possible to easily introduce other operational microwave sensors (e.g. FY-3D MWRI) or sensors which will be on board of future satellite missions (e.g. MetOp-SG MWI).

3.1.1. AMSR2

AMSR2 is the successor to ADEOS-II and AMSR for EOS (AMSR-E). The AMSR2 instrument was launched in May 2012. The instrument has a 2 m diameter antenna, a constant incidence angle of 55° and a swath width of 1450 km. The instrument has a total of 14 channels covering microwave frequencies from 6.925 GHz to 89.0 GHz with both horizontal (H) and vertical (V) polarizations. As described in section 4, according to the sea ice concentration algorithm, we use observations collected by different channels. Particularly, for AMSR2, two algorithms are implemented: the OSI SAF hybrid and the TUD. The first algorithm uses 3 channels (18.7 GHz V and 36.5 GHz V & H), while the second employs 4 channels (18.7 and 36.5 GHz V, and 89.0 GHz V & H). Characteristics of the AMSR2 channels used in the sea ice concentration algorithms are given in Table 1. The original source of AMSR2 data (approximately 29 swaths per day) are Level 1B brightness temperature products ('L1SGBTBR') in HDF5 format which are disseminated by JAXA.

Channel Number	Central Frequency (GHz)	Polarization	Bandwidth (MHz)	Footprint (km x km)	Sampling Interval (km x km)
8	18.70	V	200	14 x 22	10 x 10
11	36.5	V	1000	7 x12	10 x 10
12	36.5	H	1000	7 x 12	10 x 10
13	89.0	H	3000	3 x 5	5 x 5
14	89.0	V	300	3 x 5	5 x 5

Table 1: Characteristics of the AMSR2 channels used in the sea ice concentration algorithms

3.1.2. SSMIS

SSMIS is the successor to the Special Sensor Microwave/Imager (SSM/I). SSMIS is flown on board the United States Air Force (DMSP) series of satellites, F-16, F-17, F-18 and F-19, which came into operation in November 2005, March 2008, March 2016 and May 2016, respectively (however, F-19 has not been operational since February 2016). The SSMIS instrument has a 0.8 m diameter antenna, a constant incidence angle of 53.1° and a swath width of about 1700 km. The instrument has a total of 24 channels covering microwave frequencies from 19 to 183 GHz. For SSMIS, only the OSI SAF hybrid sea ice concentration algorithm is implemented and, as for the AMSR2, 3 channels are used: 19.35 GHz V and 37.0 GHz V & H. Apart from the minor shift in the frequencies of the channels, the primary difference between the AMSR2 and SSMIS OSI SAF hybrid sea ice concentration algorithm arises from the higher spatial resolution of the AMSR2 channels. Consequently, the resultant L2 and L3 SSMIS products are characterized by lower resolution. A summary of the SSMIS channels which are used in the sea ice concentration algorithm is provided in Table 2. The SSMIS data (approximately 44 swaths per day considering F-16, F-17 and F-18) used in the operational chain are those which are disseminated via EUMETCast.

Channel Number	Central Frequency (GHz)	Polarization	Bandwidth (MHz)	Footprint (km x km)	Sampling Interval (km x km)
13	19.35	V	356	42.4 x 70.1	25 x 12.5
15	37.0	H	1580	27.5 x 44.2	25 x 12.5
16	37.0	V	1580	27.5 x 44.2	25 x 12.5

Table 2: Characteristics of the SSMIS (F-16, F-17, F-18) channels used in the sea ice concentration algorithm

3.2. NWP data

In the L2 processing chain, NWP data are used as an input for a radiative transfer model (RTM) which must simulate the microwave brightness temperatures that are observed by the satellite sensor at the frequencies selected for the sea ice concentration algorithms. Simulated observations are necessary to derive the magnitude of the atmospheric contribution that is applied to the brightness temperatures and this process is described in more detail in chapter 4. In this section, we only provide information on the NWP parameters which are mandatory as an input for the RTM.

In the day-to-day sea ice concentration algorithm, we use NWP data from the ECMWF high resolution atmospheric model (HRES). In the previous operational set up, the Wentz and Meissner RTM was used (Wentz 1997; Wentz and Meissner 2000) and it required the following NWP parameters: wind speed, 2 meter air temperature and total column water vapour. In the previous logic, the Wentz and Meissner RTM expected NWP data available for 5 forecast steps (12, 15, 18, 21 and 24) at the 00z and 12z analysis time which were gridded (through a MARS request) on a regular 0.5 by 0.5 latitude and longitude grid. In this new version of the algorithm, we have implemented the possibility of using the community NWP SAF radiative transfer model RTTOV (Saunders et al., 2018, Hocking et al., 2020). Apart from the technical developments which were needed to embed RTTOV in the processing logic, RTTOV also requires a larger number of NWP parameters as an input. RTTOV not only needs single level atmospheric parameters, but it also requires model level fields such as vertical profiles of temperature and humidity (or more parameters depending on the RTTOV configuration). As a consequence, we additionally implemented a new ‘MARS request’ to adapt the data stream from the ECMWF high resolution atmospheric model. In doing so, we also increased the spatial resolution of the regular latitude and longitude grid (increased to 0.25 by 0.25) and the temporal frequency of the forecast steps (3-hour steps from 0 to 24). This should provide better results in the NWP/swath collocation process (e.g. NWP data closer to the time of the observations and NWP resolution closer to the sampling interval of the sensor).

Before applying the RTM calculations, the NWP parameters must be collocated in time and space with the observations available in a satellite swath file (L2 processing of Figure 2). To do so, firstly we identify the central time of the swath file (centre of the time window calculated considering the first and the last satellite scan) and, in terms of date/time, we associate the NWP data file with the closest analysis/forecast time. After this, we use the python pyresample package to spatially interpolate the NWP data on the satellite swath by means of the nearest Gaussian weighting function (we use the pyresample ‘resample_gauss’ function). Finally, to the NWP interpolated data, we apply the RTM to

calculate the simulated brightness temperature at every pixel of the satellite swath. Obviously, the RTM calculations are performed for every sensor channel which is used by the sea ice concentration algorithm (e.g. Table 1 and Table 2).

Table 3 provides an overview of the NWP data which are needed according to the RTTOV configuration. In this new version of the OSI SAF sea ice concentration algorithm, we have the possibility of selecting RTTOV in two different configurations: RTTOV 'clear-sky' or simply RTTOV as well as RTTOV 'all-sky' or RTTOV-SCATT. RTTOV-SCATT is expected to be the RTM model which will be adopted in the new operational chain. More details on the different RTM are provided in section 4.2.3.

RTM	NWP Single Level	NWP Model Levels	NWP lat/lon regular grid	NWP temporal resolution (FC steps at 00z/12z AN)
RTTOV (OSI-401-d, OSI-408-a, OSI-410-a)	10m U/V Wind component 2m Temperature Surface Pressure Skin Temperature	Temperature Specific Humidity CLWC (optional)	0.25 x 0.25	0, 3, 6, 9, 12, 15, 18, 21, 24
RTTOV-SCATT (OSI-401-d, OSI-408-a, OSI-410-a)	10m U/V Wind component 2m Temperature Surface Pressure Skin Temperature	Temperature Specific Humidity CLWC CIWC CC CRWC CSWC	0.25 x 0.25	0, 3, 6, 9, 12, 15, 18, 21, 24

Table 3: Overview of the NWP data which are mandatory as an input to the different RTTOV configurations

4. Level 2 processing

The Level 2 processing chain implements the fundamental logic which, given geolocated Level 1 brightness temperatures, estimates the sea ice concentration at the same coordinates of the satellite swath. The steps described in this section are equally applied to any satellite swath independently of the microwave sensor. The processing flow of Figure 2 shows that we can distinguish 2 main blocks which respectively identify: a) the 'L2 logic' which is sequentially applied to every satellite swath; b) the 'calc tp & stats' task which, considering all the satellite swaths previously processed by the 'L2 logic', generates tie-points and statistics that are necessary for the calculation of the sea ice concentration and the corresponding uncertainties. Considering that the OSI SAF sea ice concentration algorithm strictly depends on the calculation of tie-points, the 'calc tp & stats' task is firstly described.

4.1. Level 2 calculation of tie-points and statistics ('calc tp & stats')

The sea-ice concentration algorithm is based on the principle that it is possible to distinguish sea ice and water with a passive microwave instrument, due to the difference in their emissivity signatures. At microwave frequencies, in clear sky conditions, the ocean surface emissivity is generally lower than ice and consequently the brightness temperatures over ocean (which is seen as a 'cold background') are characterized by lower values than those observed over sea ice areas. Based on this physical assumption, the tie points represent typical signatures of ice and water which are derived by selecting brightness temperature from regions of known open water and ice. This is the primary goal of the calculation of the tie points: we want to cluster brightness temperature over sea ice concentration areas of 100% (ice) and 0% (ocean) and then use these values as a reference in the sea ice concentration

retrieval algorithm. Usually these tie-points are static in time and space, but they can be adjusted to follow the seasonally changing signatures of sea-ice and open water. Static tie-points are prone to be affected by sensor drift, inter sensor calibration differences and climatic trends in surface and atmospheric emission. The data must therefore be carefully calibrated before computing the sea ice concentrations. Here, we use dynamic tie-points, a method that minimizes these unwanted effects, with or without prior calibration. During winter, in the consolidated pack ice area well within the sea-ice edge, the sea-ice concentration is very near 100% (Andersen et al., 2007). This has been established using high resolution SAR data, ship observations and by comparing the estimates from different sea-ice concentration algorithms. The apparent fluctuations in the derived sea-ice concentration in the near 100% sea-ice regime are primarily attributed to snow/ice surface emissivity variability around the tie-point signature and only secondarily to actual sea-ice concentration fluctuations. In the marginal sea-ice zone, the atmospheric emission may be significant. The fluctuations due to atmospheric and surface emission are systematic. In fact, different algorithms with different sensitivity to atmospheric and surface emission compute very different trends in sea-ice area on seasonal and decadal time scales (Andersen et al., 2007). This means that not only the sea-ice area has a climatic trend, but the atmospheric and surface constituents affecting the microwave emission are also changing (for example, different wind patterns, water vapour and liquid water concentrations in the atmosphere, snow depth over sea-ice, ice thickness). In an attempt to compensate for the influence of these unwanted trends, in the OSI SAF sea ice algorithm the tie-points are derived dynamically: we compute tie-points on a daily basis and these outputs are combined into a 30 day running mean tie-points which are those used to compute the final OSI SAF Level 2 sea ice concentration. The strategy of using dynamical tie-points based on a temporal window (30 day running mean in our algorithm) has been discussed in previous studies such as Lavergne et al. (2019) and Tonboe et al. (2016). This dynamical tuning helps the sea ice concentration algorithm to adapt to inter seasonal variations of the sea-ice and open-water emissivity, mitigate sensor drift and compensate for trends potentially arising from the NWP data used within the atmospheric correction scheme.

For each swath grid point, the following approach is implemented:

1. Using static tie-points based on the NASA Team algorithm (Comiso et al. 1997), we calculate an estimate of the sea ice concentration (we name this initial estimate as ‘first guess sea ice concentration’, FG SIC). More details on this step are provided in section 4.2.2.
2. As described in paragraph 4.2.3, the FG SIC together with the NWP data is passed to the RTM and we derive the magnitude of the atmospheric contribution that is applied to the brightness temperatures.
3. The corrected brightness temperatures are then divided in two separate clusters, ocean and ice, which are necessary for the calculation of the tie-points:
    a. We identify the open water area as close as possible to the sea-ice edge. To do that the tie points are geographically selected along two belts in the Northern and Southern hemisphere defined by the NSIDC maximum sea-ice extent climatology plus an additional 100 km. In addition, data points below/above the latitude of 50/-50 degrees are rejected.
    b. On the contrary, the area for the ice tie-point is defined where the FG SIC is larger than 95% and within the limits of the NSIDC maximum sea ice extent climatology. Additional masks ensure that samples are taken away from the coastal regions. The NASA Team algorithm is a standard relatively simple algorithm with little sensitivity to ice temperature

variations (Cavalieri et al., 1986) and, for this reason, it was chosen to provide an initial estimate for the sea ice concentration.

The daily ocean and ice clusters are characterized by a number of brightness temperatures which have been selected by means of step 1, 2 and 3 considering all the satellite observations collected in one day at latitudes greater/less than 40/-40 degrees and excluding measurements over land (a land mask is applied). The two daily dataset are finally used to compute the tie-points respectively for the Bristol algorithm, the Bootstrap algorithm and the algorithm 89 GHz linear. In particular, the two sets for the first two algorithms are successively used by the OSI SAF hybrid algorithm, and the sets for the last two algorithms are used by the TUD hybrid algorithm. Table 4 provides a schematic summary of the OSI SAF sea ice concentration algorithms.

As an example of daily tie-points, Figure 3 shows the ocean (SIC = 0%) and ice (SIC ~ 100%) tie-points which have been derived by the Bootstrap algorithm considering all the AMSR2 observations (channel 8 and 11) over the Southern hemisphere for 1 August 2020. In this example, the 1 August daily tie-points are firstly used to update the 30 day running mean and, successively, the 30 day running mean tie-points are provided as an input to every swath file of the next day (2 August 2020).

In Figure 3, the black cross indicates the open water tie-point (x = T_B at 19 GHz vertical polarization; y = T_B at 37 GHz vertical polarization) which is computed as the average of all corrected brightness temperatures that are selected over ocean (blue points). On the contrary, the ice tie-point is given by the slope and offset of the ice line (shown in FIgure 3) which are computed using the principal component analysis (PCA) when applied to all corrected brightness temperatures that are selected over ice (red points). However, to check the consistency of the ice cluster, we also monitor the average of all corrected brightness temperatures that are selected over ice (black star in Figure 3). The latter value is simply used to monitor and compare the variability of the ocean and ice tie points.

The seasonal variability of ocean and ice tie-points derived from AMSR2 observations (channel 8 and 11) is explored from Figure 4 to Figure 7 which show time series plots (from 1 August 2020 to 30 June 2021) of daily and 30 day running mean tie-points. The tie-points are computed by means of the Bootstrap algorithm (19 and 37 Ghz Vertical polarization) and are separated for the Northern and Southern hemisphere. The monitoring of ocean and ice tie-points is a new feature of the algorithm and time series plots (as those of Figures 4-7) are automatically generated by the operational run.

Algorithm	Frequencies used for tie-points calculation
Bristol	19V, 37V, 37H
Bootstrap	19V, 37V
89 GHz linear	89H, 89V
OSI SAF sea ice concentration algorithm Hybrid (Bootstrap + Bristol) TUD (Bootstrap + 89 GHz linear)	

Table 4: Summary of OSI SAF sea ice concentration algorithms

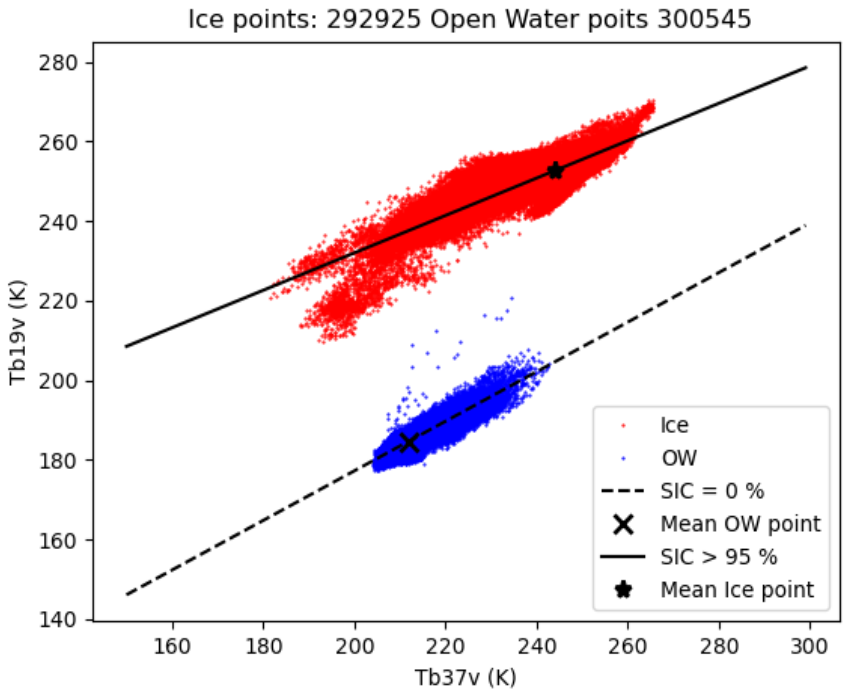

Figure 3: Open water (black cross) and ice (black star) tie-points for the Bootstrap algorithm in the 37 and 19 GHz (vertical polarization) brightness temperatures space. The ocean (blue points) and ice (red points) clusters are derived considering the AMSR2 observations (channel 8 and 11) over the Southern hemisphere for 1 August 2020.

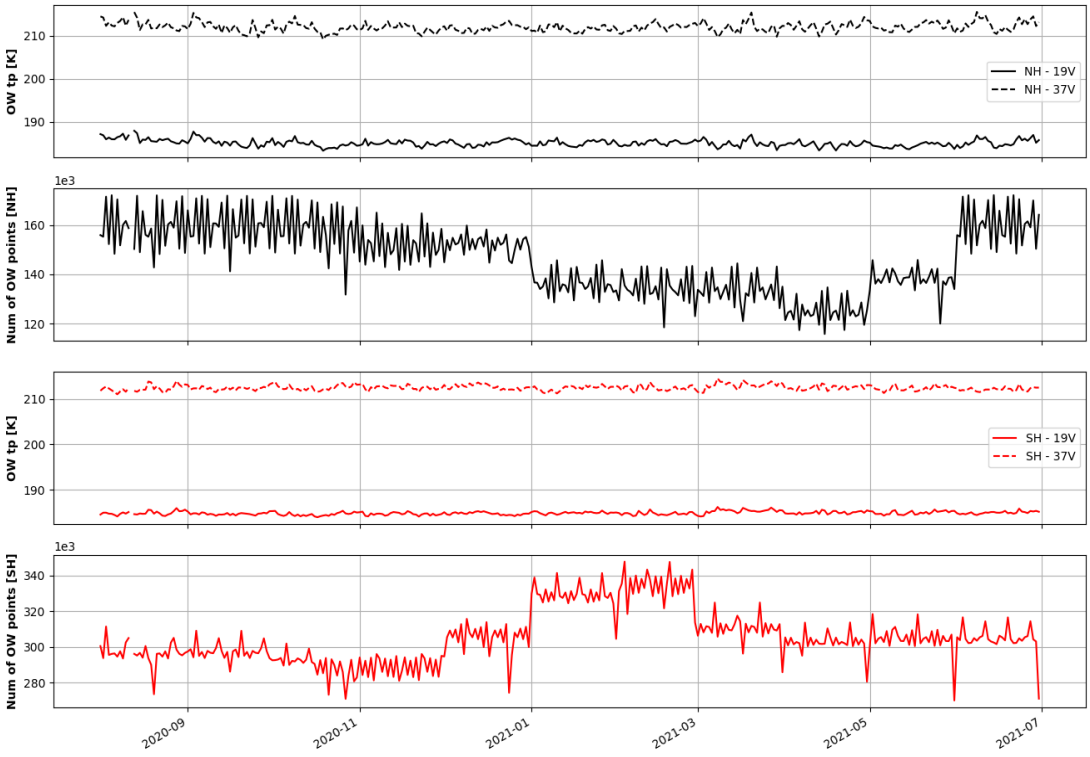

Figure 4: Time series (from 1 August 2020 to 30 June 2021) of daily Open Water (OW) tie-points for the Bootstrap algorithm (19 and 37 Ghz Vertical polarization) derived from AMSR2 observations (channel 8 and 11). Tie-points are separately shown for the Northern (NH) Southern (SH) hemisphere.

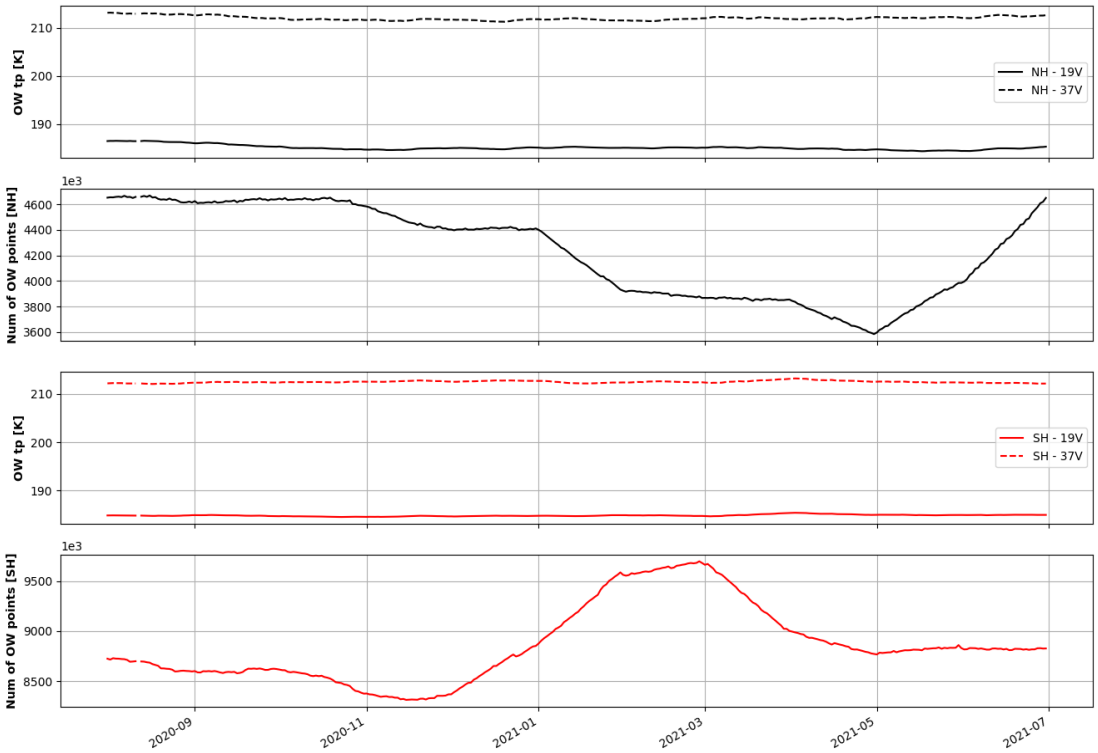

Figure 5: Time series of 30 day running mean Open Water (OW) tie-points derived from daily tie-points shown in Figure 4.

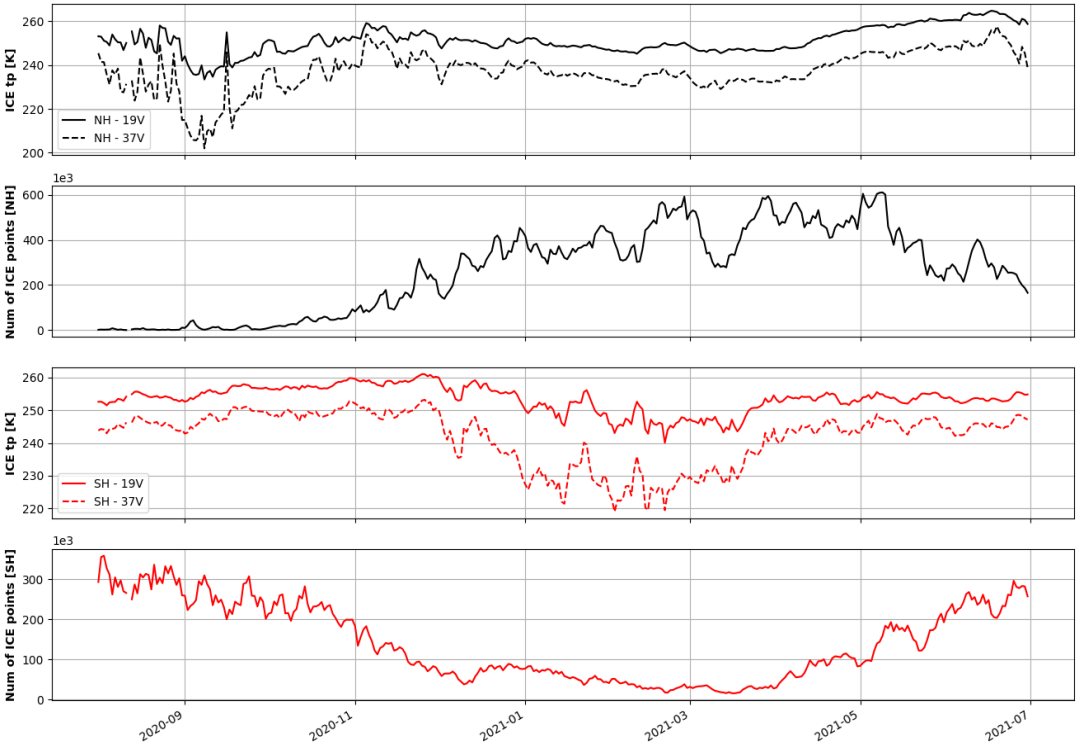

Figure 6: As Figure 4, but showing the time series of daily ice tie-points.

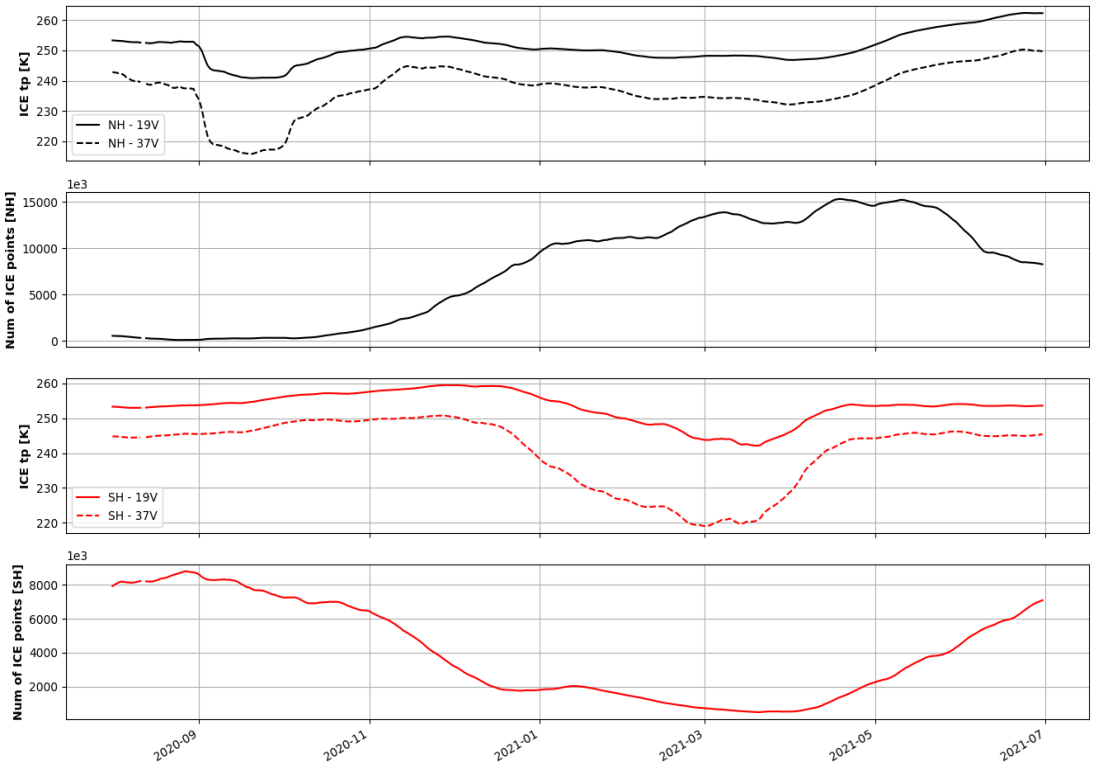

Figure 7: As Figure 5, but showing 30 day running mean ice tie-points.

The ‘calc tp & stats’ task additionally handles the computation of the statistics which are required for deriving the sea ice concentration errors (section 4.2.5). As for the daily tie-points, to execute the calculation of statistics, we need that the satellite swaths have been processed by the ‘L2 logic’. In particular, it is expected that the final OSI SAF sea ice concentration (section 4.2.4) has been computed for each swath grid point and this means that the Hybrid or the TUD OSI SAF algorithm has been executed taking as an input the 30 day running mean tie-points together with the atmospheric corrected brightness temperatures. The statistics (mean, standard deviation and counts) are calculated selecting the OSI SAF sea ice concentrations over the same geographical locations which are previously identified and used for the calculation of the daily ocean and ice tie-points (e.g. the location of the brightness temperatures which are represented by the two clusters of points in Figure 3). Similarly to the tie-points, we have daily statistics separated for surface type (open water or ice) and hemisphere (Northern or Southern). The resultant daily standard deviations over open water and ice (σ_{OW} and σ_{Ice}), which for instance are computed for the 1 August 2020, are successively provided as an input to every swath file of the next day (2 August 2020) so that the calculation of the total sea ice concentration error (section 4.2.5) can be executed. The open water and ice standard deviations quantify the tie-point uncertainty which includes residual atmospheric noise, sensor noise and ice surface emissivity variability.

The daily open water and ice statistics for the OSI SAF sea ice concentrations (Hybrid or TUD) are monitored by means of time series plots which, as a new feature of the algorithm, are automatically generated by the operational run. Time series (from 1 August 2020 to 30 June 2021) of daily open water and ice statistics for the Hybrid OSI SAF sea ice concentrations are respectively shown in Figure 8 and 9.

Together with the time series monitoring plots, the new operational logic of the algorithm also implements safety checks to control the daily number of ocean and ice points which are selected for the calculation of the ocean and ice tie-points and statistics. In particular, the mean value of the ocean and ice points accumulated in the last 15 days is computed and compared against the daily number (e.g. the daily value which is shown in the time series plots). Two automatic warnings are triggered if the daily number is 40% or 80% below the mean of the last 15 days. This automatic warning system is particularly helpful because a severe (>80%) or consistent (>40%) reduction of data over the ocean and ice most likely indicates a problem with the dissemination of the Level 1 satellite data. If we lose a large number of satellite segments, the L3 product will be degraded and we will observe gaps in the sea ice concentration over different geographical areas. Generally, a consistent or severe reduction of points over ice can be observed during the summer season when the NASA Team FG SIC can struggle to identify regions where the sea ice concentration is greater than 95%. Operationally, when a data reduction is captured by the warning system, an email is automatically sent to notify the event. However, the ‘calc tp & stats’ task continues to operate unless the ocean or ice daily number goes below 50 points. This threshold has been arbitrarily chosen to stop the execution of the ‘calc tp & stats’ task. In this case, daily tie-points and statistics are not computed and the ‘L2 logic’ looks back up to 7 days to search for the last tie-points and statistics available. As a reference, in 1 year, we only found two cases in the Northern hemisphere when the total number of points identified as closed-ice (FG SIC > 95%) went below 50 (28 July and 12 August 2020).

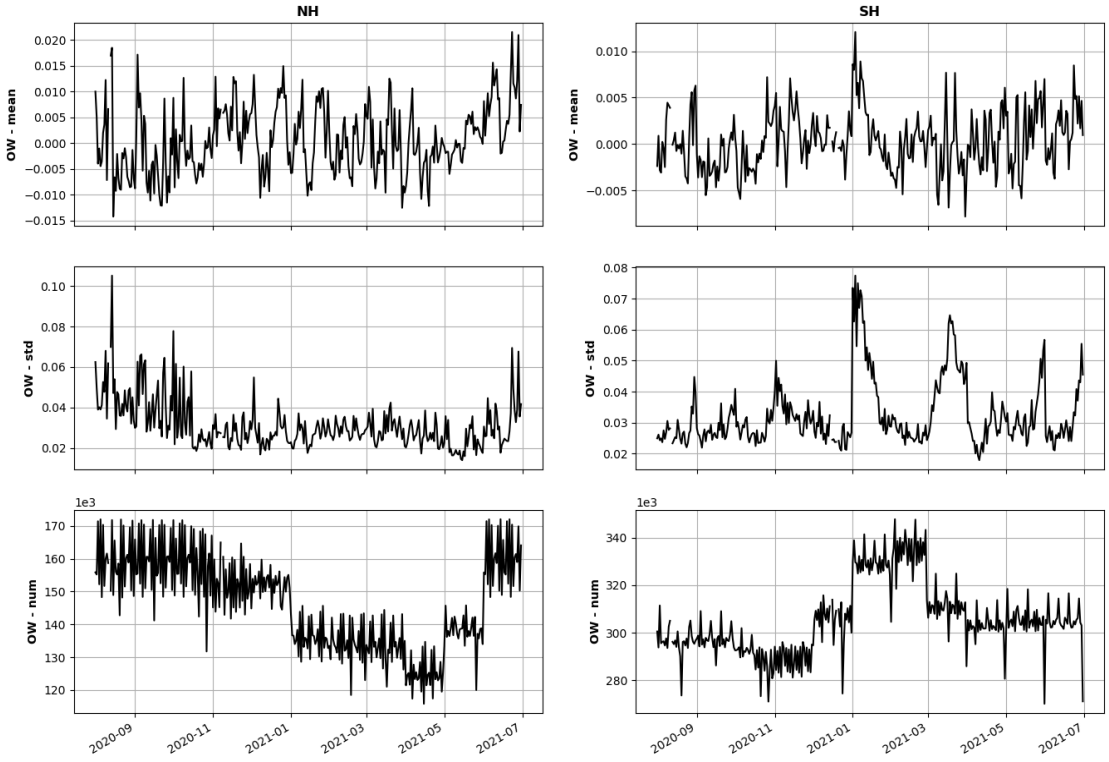

Figure 8: Time series (from 1 August 2020 to 30 June 2021) of daily open water (OW) statistics (mean, standard deviation and counts) for Hybrid OSI SAF sea ice concentrations. Statistics are separately computed for the Northern (NH) and the Southern (SH) hemisphere.

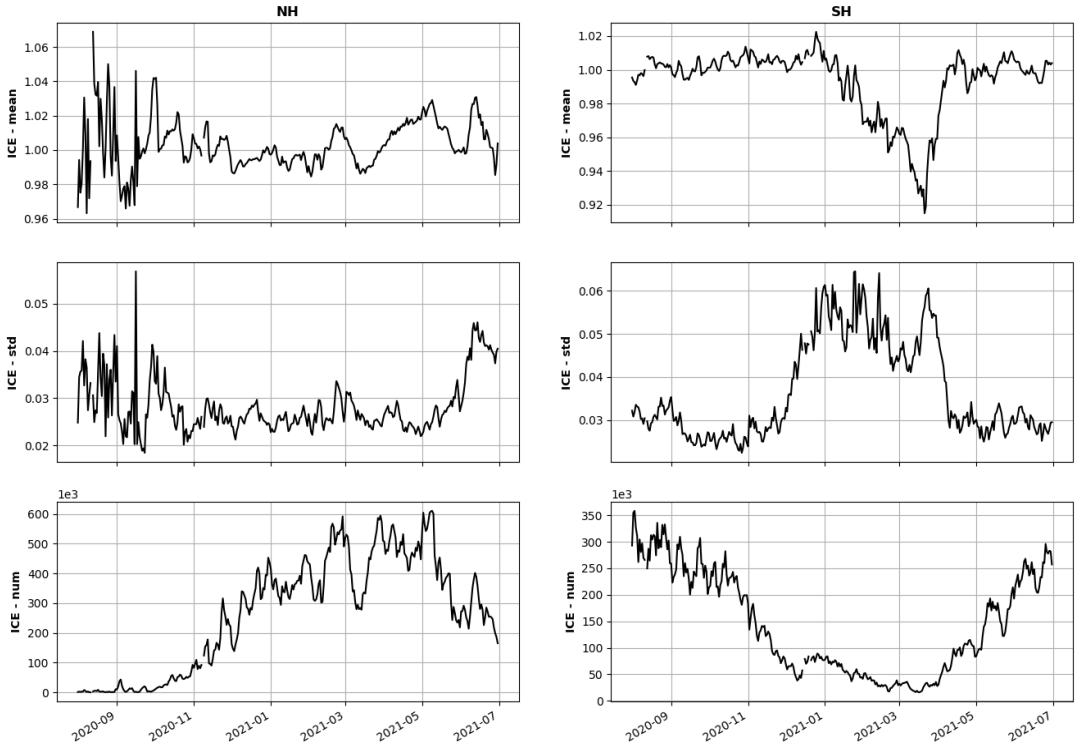

Figure 9: As Figure 8, but showing the time series of the daily ice statistics.

4.2. Level 2 swath file processing

In this section we describe the sequential logic ('L2 logic') which is applied to every satellite swath to generate the L2 sea ice concentration product.

4.2.1. Land-spillover correction

Due to the coarse resolution of the sensors, the T_B are influenced by the land emissivity several kilometres away from the coastline (land-spillover). At the microwave frequencies used by the algorithm (Table 1 and Table 2), the emissivity of land is much higher than the ocean and it is comparable to sea ice emissivity and, consequently, the sea ice concentration might be overestimated in coastal regions. For this reason, the very first step of the processing logic is a correction scheme which is applied to the Level 1 microwave observations so that they are corrected to account for land-spillover. This swath-based correction step is a new feature implemented in this release of the algorithm and it is based on the same method described in Lavergne et al. (2019). The basic principle is that a fine-resolution land mask is used together with the antenna viewing geometry to estimate (and correct for) the simulated contribution of land emissivity to the observed T_B. This correction scheme was adopted from Maass and Kaleschke (2010) with some modifications 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); b) the approximation of the antenna pattern functions as Gaussian 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 of FoV that overlap land and ocean are corrected for the contribution by land and they are then used in the following steps of the processing chain.

4.2.2. Calculation of sea ice concentration from static tie-points

The corrected T_B for land-spillover are provided as an input to the NASA Team algorithm (Comiso 1986; Comiso et al. 1997) to calculate an initial estimate of the sea ice concentration ('first guess sea ice concentration', FG SIC). The NASA Team algorithm uses static tie points for the 19 GHz Vertical and Horizontal polarization and the 37 GHz Vertical polarization. As discussed in Lavergne et al. 2019, the NASA Team algorithm is an acceptable choice for the purpose of selecting closed-ice samples and consequently the FG SIC is used to identify the ice cluster (concentrations greater than 95%) for the dynamical tie-points calculation. The FG SIC is also used as an input to the RTM which is used for correcting atmosphere influence on the brightness temperature. Table 5 provides a detailed summary of the static tie-points which are utilized for the AMSR2 and SSMIS. Previously, the static tie-points were identical for the two sensors and based on the values provided by Comiso et al. 1997. In this version of the algorithm, we found beneficial upgrading the tie-points as follows: for AMSR2, we use tie-points values as those provided by Ivanova et al. 2015; for SSMIS, we updated the tie-points from Comiso et al. 1997 to Ivanova et al. 2015 only for the Southern Hemisphere.

NH	**AMSR2**			**SSMIS**		
	OW	**FYI**	**MYI**	**OW**	**FYI**	**MYI**
19V	183.72	252.15	226.26	177.10	258.20	223.20
19H	108.46	237.54	207.78	100.80	242.80	203.90
37V	209.81	247.13	196.91	201.70	252.80	186.30
SH	**AMSR2**			**SSMIS**		
	OW	**FYI**	**MYI**	**OW**	**FYI**	**MYI**
19V	185.34	258.58	246.10	185.02	259.92	246.27
19H	110.83	242.80	217.65	118.0	244.57	221.95
37V	212.45	253.84	226.51	209.81	254.39	226.46

Table 5: Static tie-points used in the NASA Team algorithm to derive the 'first guess' sea ice concentration in the OSI SAF algorithm. Static tie-points are diversified according to the hemisphere (NH and SH) and the surface type: open water, (OW), first year ice (FYI) and multi year ice (MYI).

4.2.3. Calculation of atmospheric correction

The sea ice concentration retrieval is characterized by random or systematic errors which can lead to departure from 0% and 100% sea ice concentration or generate unrealistic sea ice concentrations greater than 0% over open water ('false ice'). For instance, over ocean, the sea ice algorithm can be prone to errors due to the high variability of local atmospheric conditions such as humidity, cloud and precipitation as well as to the wind roughening of the sea surface. These effects generally bring to an overestimation of the sea ice concentration in the ice-edge area or to false ice over ocean. On the contrary, over a drier atmosphere above the consolidated ice pack, the impact of weather-related effects is negligible and the main source of error for the algorithm comes from the large variability of the ice surface emissivity which is due to different causes such as the type and thickness of the ice, the temperature of the emission layer and the snow depth. As described in literature (Andersen et al 2006; Ivanova et al 2015; Tonboe et al 2016; Lavergne et al 2019), to improve the accuracy of the sea ice concentration retrieval and minimize the impact of the different error sources, the observed brightness temperatures are corrected for atmospheric contribution by using a radiative transfer model. Successively, the corrected brightness temperatures are used within the calculation of the ocean and ice tie-points.

As previously described in section 3.2, the RTM ingests a number of NWP parameters (Table 3) and computes the brightness temperatures to simulate those that are observed by the differ channels of the satellite sensor. Independently of the RTM, the logic of the correction scheme is as follows:

- NWP data are firstly selected considering the analysis/forecast date/time closest to the central time of the satellite swath and, successively, by means of the python 'pyresample' package, are spatially interpolate on the satellite swath.
- The FG SIC, also necessary as an input for the RTM, is clipped to the interval [0,1] (values smaller than 0 become 0, and values larger than 1 become 1). To improve the RTM calculation and remove possible false ice over ocean, in the new version of OSI SAF algorithm, we added a screening based on the NWP skin temperature: below/above the latitude of 66/-66 degree, if the skin temperature is greater than 276 K, the FG SIC is forced to be zero. We found this screening effective in removing false ice over the ocean without interfering with the ice-edge and closed-ice areas.
- The magnitude of the atmospheric correction (δT_B) is finally computed. If the simulated T_B is given by T_B = RTM(NWP), δT_B is calculated as follows:

  $\delta T_B = T_B(NWP) - T_B(NWP_{ref})$,

  where $T_B(NWP)$ is the output of the RTM when the calculations are done with the expected NWP data as described in Table 3, while $T_B(NWP_{ref})$ is computed considering a reference atmospheric state with the same air and surface temperature of $T_B(NWP)$, but forcing the other NWP parameters to zero.

The atmospheric correction is calculated for each swath grid point at latitudes greater/less than 40/-40 degrees and screening out land (a land mask is applied). δT_B is algebraically summed with the observed brightness temperature and the resultant value is used in the calculations of the tie-points.

In this version of the algorithm, we have implemented the possibility of using the community NWP SAF radiative transfer model RTTOV (Saunders et al., 2018). The RTTOV version 13 (Hocking et al., 2020) is the new operational RTM adopted for the Level 2 processing chain. RTTOV is a fast RTM which can simulate observations collected by a large number of passive visible, infrared and microwave satellite sensors. For the purpose of the OSI SAF algorithm, RTTOV is configured to compute the microwave brightness temperatures in each of the channels of the sensor which is used for the sea ice concentration retrieval. The core of RTTOV is a fast parameterisation of layer optical depths due to gas absorption. Depending on the NWP which are supplied as an input (Table 3), RTTOV can be configured in two different ways:

- 'clear-sky' (simply named RTTOV): given an atmospheric profile of temperature (T) and water vapour (Q) together with surface parameters and a viewing geometry, RTTOV computes the T_B at different frequencies of the sensor being simulated. We call this simulation 'clear-sky' because RTTOV only considers gas absorption from water vapour and there is no attempt to model scattering from cloud and precipitation. In the clear-sky configuration, RTTOV can additionally receive as an input the profile of cloud liquid water content (CLWC) which is still considered as a purely absorbing medium. As a result, this simulation (T+Q+CLWC) can be affected by the presence of clouds, but there is no modelling of scattering.
- 'all-sky' or RTTOV-SCATT: this is the interface which is capable of simulating microwave radiances affected by cloud and precipitation. In this configuration, RTTOV-SCATT implements a two-independent column approximation and the resultant brightness temperature is given by a weighted average of the brightness temperatures calculated from a 'clear-sky' column (only T+Q) and a 'cloud-precipitating' column which adds the scattering effects from water and ice. The simulation of 'all-sky' brightness temperatures is possible only if 5 NWP hydrometer profiles are supplied as an additional input to the RTM: the cloud liquid and ice water content (CLWC and CIWC), the rain and snow water content (CRWC and CSWC) and the fraction of cloud cover (CC). The RTTOV-SCATT calculations are executed by means of lookup tables for hydrometeor scattering properties ('hydrotables'). Further details on RTTOV-SCATT and the generation of the bulk optical properties in the hydrotables can be found in Geer et al., 2021.

In the algorithm, we left the flexibility of selecting both RTTOV and RTTOV-SCATT which can be valuable for conducting assessment studies. RTTOV-SCATT is surely the best performing model for simulating microwave observations in all-sky conditions (Geer et al., 2021) and it is used in the operational ECMWF data assimilation system. We expect to use RTTOV-SCATT in the operational L2 processing chain. Apart from the benefit we might gain in the accuracy of the sea ice concentration algorithm, RTTOV will surely simplify the possibility of ingesting microwave observations from existing operative sensors or those sensors which will be part of future satellite missions.

4.2.4. Calculation of sea ice concentration

This section describes the two algorithms that are used to compute the sea-ice concentrations: the OSI SAF hybrid and the TUD. The OSI SAF hybrid algorithm is applied to generate sea-ice concentration to both the AMSR2 and SSMIS product; it uses the 19H, 19V and 37H channels. The TUD algorithm is only applied to generate the AMSR2 product and not the SSMIS. It uses the 19V, 37V, 89V and 89H channels, and is capable of higher resolution than the OSI SAF hybrid algorithm, but is more susceptible to noise due to water vapour, cloud and precipitation.

- OSI SAF hybrid algorithm

  A total of 30 algorithms retrieving Arctic sea-ice concentration from satellite passive microwave data are described and compared in detail in the SICCI ATBD (ESA SICCI project consortium, 2013). The analysis of atmospheric sensitivity in Andersen et al. (2006) showed that the Bootstrap frequency mode algorithm (Comiso, 1986) had the lowest sensitivity to atmospheric noise over open water. Furthermore, the comparison to high resolution SAR imagery in Andersen et al. (2007) revealed that among the algorithms using the low frequency channels (19 and 37 GHz), the Bristol algorithm (Smith, 1996) had the lowest sensitivity to sea-ice surface emissivity variability. In addition, this algorithm had a low sensitivity to atmospheric emission in particular at high ice concentrations. Consequently, a hybrid algorithm has been established as a linear combination of two of the tested algorithms, the Bristol algorithm and the Bootstrap frequency mode algorithm. To ensure an optimum performance over both marginal and consolidated sea-ice, and to retain the virtues of each algorithm, the Bristol algorithm is given little weight at low concentrations, while the opposite is the case over high sea-ice concentrations.

  The Bootstrap algorithm (Comiso, 1986) is based on the observation of linear clustering of sea-ice T_B's in scatter plots of 19V vs 37V whereas open water T_B's cluster around a single point. It assumes only two surface types: sea-ice and open water, taking into account the variability of both to optimize the detection of small sea-ice concentrations. The total sea-ice concentration ($C_{Bootstrap}$), corresponding to an observed brightness temperature T_B (e.g. for a single channel), is calculated as follows:

  $$C_{Bootstrap} = (T_B - T_B^{OW}) / (T_B^{Ice} - T_B^{OW}) \quad (1)$$

  where T_B^{OW} and T_B^{Ice} are respectively the open water and the ice tie points which are derived in the 37V and 19V brightness temperatures space.

  The Bristol algorithm (Smith, 1996) is conceptually similar to the Bootstrap algorithm. In a three-dimensional scatter plot spanned by 19V, 37V and 37H the sea-ice T_B's tend to lie in a plane. The only difference to the Bootstrap algorithm is that instead of viewing the data in the 19V, 37V space, the Bristol algorithm views the data perpendicular to the plane in which the data lies, i.e. in a transformed coordinate system:

  $$Bristol_x = T_{37V} + 1.045\,T_{37H} + 0.525\,T_{19H}, \quad (2)$$

  $$Bristol_y = 0.9164\,T_{19V} - T_{37V} + 0.4965\,T_{37H}. \quad (3)$$

  The remaining analysis is identical to the Bootstrap algorithm. The Bootstrap algorithm is used over open water and the Bristol algorithm is used over sea-ice. At concentrations of up to 40% ('threshold' in equation 5) the sea-ice concentration is an average weighted linearly between the two algorithms. This hybrid OSI SAF sea-ice concentration is formulated as follows:

  $$SIC_{OSI\ SAF\ hybrid} = (1 - weight)\,C_{Bristol} + weight\,C_{Bootstrap}, \quad (4)$$

  where the weight is given by:

  $$weight = (|threshold - C_{Bootstrap}| + threshold - C_{Bootstrap}) / 2\,threshold \quad (5)$$

- OSI SAF TUD algorithm

  The TUD sea-ice concentration algorithm is using the part of the Bootstrap algorithm which is normally used over open water (Comiso, 1986; Comiso et al., 1997) together with the scaled H and V polarisation difference of the near 90 GHz channels (Pedersen, 1998). The Bootstrap algorithm in frequency mode has relatively coarse resolution but it is almost independent of different weather conditions.

  The TUD hybrid algorithm is configured so that the Bootstrap algorithm (equation 1), which is applied to the brightness temperatures in the 37V and 19V space, is equally used for the higher frequency and resolution channels 89V and 89H. The final TUD sea ice concentration is given by a combination of the coarse resolution sea-ice concentration (C_f), estimated from the lower frequency 19V and 37V channels, and the high resolution sea-ice concentration (C_{89}) which is derived by the 89 GHz channels (V & H). A comparison between C_f and C_{89} shows that there is larger scatter, in particular at the low sea-ice concentrations derived at 89 GHz compared to the lower frequency channels (Pedersen, 1998). In order to reduce the noise in C_{89}, the TUD OSI SAF sea-ice concentration is calculated as follows:

  if $C_{89} > 0$ and $C_f > 10$:

  $$SIC_{OSI\ SAF\ TUD} = \sqrt{C_f C_{89}} \quad (6)$$

  otherwise:

  $$SIC_{OSI\ SAF\ TUD} = C_f. \quad (7)$$

4.2.5. *Calculation of sea ice concentration errors*

Uncertainty estimates are needed when the sea-ice concentration data are compared to other data sets or when the sea-ice concentrations are assimilated into numerical models. The mean accuracy of some of the more common algorithms, used to compute sea-ice concentration from SSM/I data, such as the NASA Team and Bootstrap algorithms are reported to be 1%-6% in winter (Andersen et al., 2006B). This is also achieved with the OSI SAF algorithm (Ivanova et al., 2015).

The polar atmosphere is generally transparent for microwave radiation in between the sounding channels called the atmospheric windows near 19, 37, 91, and 150 GHz. For typical polar atmospheric states the down-welling emission at the surface is about 5-15 K at 18 GHz, 20-40 K at 36 GHz, 30-100 K at 90 GHz. For comparison, the sea-ice surface emission is typically 150-260 K. When computing the sea-ice concentration using the atmospheric window channels, the atmospheric emission and scattering is an error source. The tie-points are typical sea-ice and water signatures representative on a hemispheric-scale. Deviations from the typical surface emission signatures result in sea-ice concentration uncertainties. The microwave instruments have relatively large foot-prints on the ground, and the algorithms with the lowest sensitivity to both atmospheric and surface emissivity variability use T_B's at different frequencies with different footprint size. Representing these large footprints on a finer, predefined grid results in a representativeness error. In addition there is the geo-location error, sensor noise, drift, and sea-ice variability over the sampling period.

In the OSI SAF sea ice concentration algorithm, we assume two independent source of errors: a) the tie-point uncertainty ($\sigma_{tie\text{-}point}$) which includes residual atmospheric noise, sensor noise and ice surface

emissivity variability; b) the representativeness error (σ_{smear}) that is the uncertainty due to the footprint mismatch to a grid where the sensor footprint covers more than one pixel. In addition to these two error components, there is the geolocation error which is due to uncertainties in the orientation of the satellite. For both SSMIS and AMSR2, the geolocation error is approximately 5 km for the 6.9 GHz channel (~10% of the beam FWHM footprint dimensions), reducing to 1 km for the 89 GHz channel (~20% of the beam FWHM footprint dimensions). Considering that the magnitude of the geolocation error is small in comparison with the large footprints of the channels used in the OSI SAF algorithm, the contribution of the geolocation uncertainty is neglected. As a result, the total uncertainty of the OSI SAF algorithm is estimated as the squared sum of the tie-point and representativeness errors:

$$\sigma_{total}^2 = \sigma_{tie-point}^2 + \sigma_{smear}^2.$$

- Tie-point uncertainty

  Both the water surface and sea-ice surface emissivity variabilities result in sea-ice concentration uncertainties. Emission and scattering in the atmosphere also affects the T_B's and the computed sea-ice concentrations. Although we use the atmospheric correction scheme to minimize these uncertainties, the OSI SAF sea ice concentration algorithm provides ice concentrations which can be greater than 100% and less than 0% respectively over the ice (SIC ~ 100%) and open water (SIC = 0%) tie-points. To quantify this remaining noise of the algorithm, we use the daily standard deviations, which, as described in section 4.1, are separately calculated for both the hemispheres selecting the OSI SAF sea ice concentrations over the same geographical locations that are identified for the ocean and ice tie-points. As an example, Figure 8 and 9 show respectively the time series of the mean and the standard deviation (σ_{OW} and σ_{Ice}) of the OSI SAF sea ice concentration over open water and ice. These time series plots are helpful to systematically monitor the behaviour of the algorithm over the open water and ice tie-points in terms of the fluctuation around the expected SIC value of 0 % (open water) and 100 % (ice) and the corresponding spread given by σ_{OW} and $\sigma_{Ice.}$ To conclude, the tie-point uncertainty ($\sigma_{tie\text{-}point}$) at each grid point of the swath file is computed as:

  $$\sigma_{tie-point} = \sqrt{(1-SIC)^2\sigma_{OW}^2 + SIC^2\sigma_{Ice}^2}. \quad (8)$$

  In equation 8, SIC is the sea ice concentration calculated by the OSI SAF algorithm and, to provide the proper weight $(1\text{-SIC})^2$ or SIC^2, we simply consider SIC = 0 and SIC = 1 respectively if SIC < 0 and SIC > 1.

- Representativeness uncertainty

  Footprint sizes for the channels used for ice concentration mapping range from about 50 km for the 19 GHz channels to about 30 km for the 37 GHz channels. The ice concentration data are normally represented on a finer grid (for the L3 product is typically 10 km) than the sensor footprint sizes (30–70 km). This effect is called smearing. The combination of footprints of uneven size in the ice concentration algorithm results in an additional smearing effect. This we call the footprint mismatch error. The smearing and the footprint mismatch error cannot be estimated separately. However, the combined error can be estimated if all other error sources and the ice cover reference are known a priori. It can also be simulated using high-resolution

ice concentration reference data and a model for the satellite measurement footprint patterns. In the OSI SAF algorithm, the smearing and the footprint representativeness uncertainty (in the following test, it is simply named as smearing) is empirically derived using a microwave radiometer imaging simulator as described in Tonboe et al., 2016.

At the end of the L2 processing chain, the OSI SAF algorithm provides the uncertainty values (tie-point, smearing and total) at each swath grid point. As an example, Figure 10 shows the tie-point, smearing and total uncertainty as a function of the OSI SAF hybrid sea ice concentration which has been derived from AMSR2 observations corresponding to the satellite overpass on 1 August 2020 from 1236z to 1327z. The uncertainties are separately shown for the Northern (NH) and Southern (SH) hemisphere and diversified according to the L2 and L3 uncertainty. The smearing uncertainty is zero for open water and for 100% ice and at these two points on the curve the total uncertainty equals the tie-point uncertainty (including sensor and residual atmospheric noise) for open water and ice respectively. The smearing uncertainty reaches a maximum at intermediate concentrations between '0% + $\sigma_{tie\text{-}poin}$' and '100% − $\sigma_{tie\text{-}poin}$''. For the resolution of the L3 product, because the sea ice concentration is provided on a relatively fine grid of about 10 km compared to the actual resolution of the sensor, the L3 smearing uncertainty is the component which is dominating the total L3 uncertainty for most of the sea ice concentration range. On the contrary, for the L2 product, where the grid resolution is equivalent to the footprint size of the sensor, the L2 smearing uncertainty is at the minimum and the L2 total error is comparable in magnitude to L2 the tie-point uncertainty. The L3 uncertainty, which is calculated at each swath grid point by the L2 processing chain, is successively used in the gridding and daily average to generate the final uncertainty for the Level 3 product.

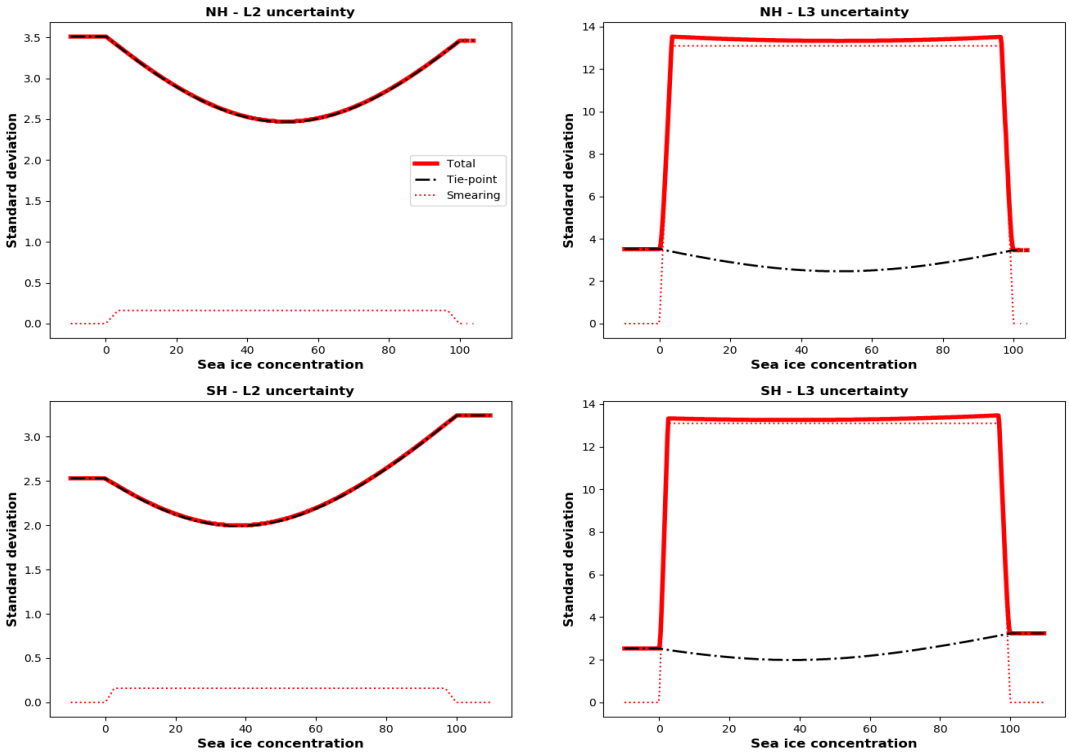

Figure 10: Estimate of the OSI SAF algorithm uncertainties as a function of the OSI SAF hybrid sea ice concentration: total error, red line, tie-point error, dash-dotted black line, smearing error, dotted red line. In this example, the OSI SAF hybrid sea ice concentration has been derived from AMSR2 observations corresponding to the satellite overpass on 1 August 2020 from 1236z to 1327z. Algorithm uncertainties are separately shown for the Northern (NH) and Southern (SH) hemisphere and diversified according to the L2 and L3 uncertainty.

4.2.6. Filtered sea ice concentration

Ideally, a perfect algorithm should predict SIC = 0 % over ocean surfaces and values in the range 0 % < SIC <= 100 % in those sea-ice regions which extend from the edge to the inner areas with more consolidated ice (e.g. MYI). As already discussed in the previous sections, this is not the case for the retrieved OSI SAF sea ice concentration and we need to filter out those retrievals which are 'unphysical'. The 'unphysical' retrievals can be categorized as follows:

- $SIC_{OSI\ SAF} > 100\ \%$
- $SIC_{OSI\ SAF} < 0\ \%$
- physical magnitude, $0\ \% < SIC_{OSI\ SAF} < 100\ \%$, but most likely not ice (open water or close the ice edge).

In this section we describe the logic that we adopt to possibly capture and remove these retrievals and to generate the so called 'filtered' sea ice concentration. The 'filtered' sea ice concentration is complementary to the OSI SAF sea ice concentration: at each grid point of the swath file we associate

the sea ice concentration value calculated by the OSI SAF algorithm and the corresponding 'filtered' value. Firstly, $SIC_{OSI\ SAF}$ > 100 % is observed in sea-ice regions where the surrounding OSI SAF retrievals consistently have high values of concentrations (e.g. > 90 %) indicating that we are presumably observing consolidated ice areas. In this case, as filtered SIC, we simply consider SIC = 100 %. Swath grid points where $SIC_{OSI\ SAF}$ < 0 represent the opposite case: we constantly found these retrievals over open water regions. As filtered SIC, we could simply use SIC = 0, but we do not do that, because the logic to handle the third case is also capable to adjust these retrievals. The retrievals which have physical magnitude (0 % < $SIC_{OSI\ SAF}$ < 100 %), but most likely are not representing 'true' ice, must be analysed by means of a screening logic. These SIC retrievals are observed in areas of open ocean or close to the ice edge which, in different degree, are the result of the residual noise due to atmospheric contamination and surface emissivity not handle by the atmospheric correction scheme. Three screening tests are applied at each grid point of the swath file. All the tests are executed so that more than 1 test can be verified at the same time. When at least 1 test results 'true' the filtered SIC is set to 0. The description of the screening tests is summarised as follows:

1. the so called 'Open-water-filter' (OWF) screening which is implemented in the same way as described in Lavergne et al. 2019. This test is based on the dynamical computation of the 19V and 37V gradient ratio (GR3719V), which in the 19V and 37V observations space, can help to identify regions of typical open-water and low-concentration ice. The dynamic calculation of the GR3719V threshold (T) in the 19V and 37V space is implemented so that the 10% SIC can be intercepted. To conclude, the filtered SIC is set to 0 if GR3719V >= T or $SIC_{OSI\ SAF}$ <= 10%. For more details on the OWF, the reader can refer to Lavergne et al. 2019.

2. A screening based on the use of the NWP skin temperature. As done for the FG SIC to improve the radiative transfer calculation (see section 4.2.3), we check the skin temperature below/above the latitude of 66/-66 degree and, if the temperature is greater than 276 K, the filtered SIC is set to 0.

3. A screening based on the so called 'symmetric' 37 GHz polarization (P37) difference which is used as a cloud predictor in the logic implemented for the all-sky microwave data assimilation over ocean (Geer and Baur, 2011). The calculation of the 'symmetric' P37 is possible when RTTOV-SCATT is used and it is given by the mean of the P37 which is separately calculated from the observations and the simulated brightness temperatures. The 'symmetric' P37 varies between 0, which represents a profile with opaque cloud, and 1, which identifies clear-sky. We can use the 'symmetric' P37 to identify those areas of intense scattering over ocean which are the most complicated to handle by the sea ice algorithm. To do this, in our algorithm, we control where the 'symmetric' P37 is less than 0.7, but only below/above the latitude of 60/-60 degree. For more details on the 'symmetric' P37, the reader can refer to Geer and Baur (2011).

4. As a final test, we control if a retrieved SIC > 0% is located outside the expected NSIDC maximum sea ice climatology extent (more details can be found in the PUM). If this is the case, the sea ice concentration is set to zero.

The result of the screening tests is stored in a 'bit mask' variable ('status_flag') and they are provided at each grid point of the swath file. In the PUM, the user can find a detailed description of the status_flag for the Level 2 and Level 3 products. To conclude, at each grid point of the swath file, the user can find the $SIC_{OSI\ SAF}$, the $SIC_{filtered}$ and the 'status_flag' so that it is easy to identify which tests are applied to filter the $SIC_{OSI\ SAF}$.

4.2.7. *Near coast points*

For the L2 product, we do not apply a land/coast mask, but very simply, we check the position (latitude and longitude) of the satellite observations with respect to the land/coast mask which is used for generating the L3 product (Figure 11). Using the nearest neighbour approach, we classify the satellite points over land or 'near coast' (shore, near-shore, off-shore) according to the L3 mask which is shown in Figure 11. These observation points are very complicated to handle because, due to the low spatial resolution of the microwave sensors, land and coastal areas contaminate the satellite footprint. Due to the higher resolution of the AMSR2 footprint, the 'near coast' classification is done only considering the observation points close to 'shore' and 'near-shore', while in the classification for the SSMIS, we also include the 'off-shore'. Although the Level 1 observations are corrected to account for land-spillover, land and coastal contamination is also difficult for the radiative transfer calculations. Particularly, the forward operator (RTTOV) requires as an input the surface emissivity which, near to land and coastal regions, is very tricky to handle at the microwave frequencies used by the sea ice algorithm. As a conservative approach, in this version of the algorithm, we decided to reject the points classified as land and near coast so that we do not attempt a Level 2 SIC retrieval. Future developments of the algorithm will attempt to improve the atmospheric correction scheme in these complicated mixed surface areas (land/coast/ocean/ice) and extend the L2 SIC retrieval closer to the coast. However, as described in section 5, by means of an interpolation method, for the L3 product we attempt to provide an estimate of the sea ice concentration for all the 'near coast' points.

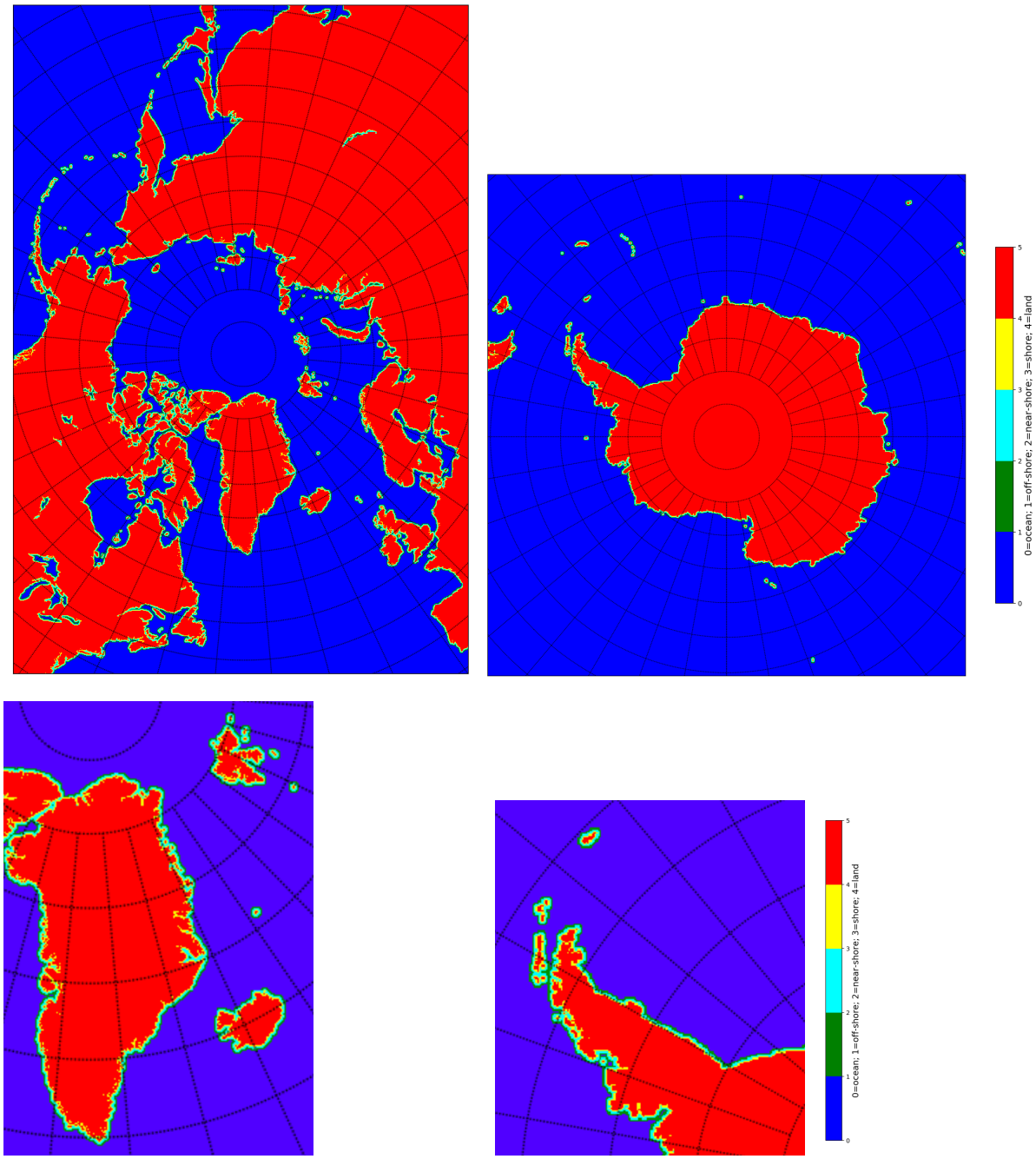

Figure 11: Polar Stereographic (10 km grid spacing) ocean/lakes/coast/land mask for the Northern (left) and the Southern (right) hemisphere. To better display the 'near coast' points the plots in the second row show a zoomed area of the Northern and Southern hemisphere. The coastline has been derived from the World Vector Shoreline. The 'near coast' points are classified as shore (yellow), near-shore (cyan), off-shore (green) which extend from the land (red) over sea (blue) areas. The mask is derived so that the off-shore points are about 12.5 km distant from the coastline. In the Northern hemisphere, all the lakes have been masked out with the exception of those displayed in the plot.

5. Level 3 processing

After all the satellite swath files have been received and processed in one day (24 hours from 00z), the Level 3 processing is executed. The L3 processing can be seen as a task that combines all the Level 2 products to generate the daily sea ice concentration estimate which covers the entire Northern and Southern hemisphere. Generally, the name of the Level 3 product file contains the string <YYYYMMDD1200> to easily identify the central date of the sea ice concentration analysis. The target grid for the L3 product, respectively for the Northern and Southern hemisphere, is provided in figure 11 which shows a polar stereographic projection with a grid spacing of 10 km.

The generation of the L3 product can be briefly summarised as follows: firstly, all the separate swath files are concatenated together so that all the variables, which have been generated during the L2 processing, are grouped together in a single array. Secondly, the array of every L2 variable is resampled according to the specification of the L3 grid resolution. The resultant value of the variable in the target grid cell is influenced by different points of the satellite swath. In particular, the final L3 value in one grid cell is given by the average of the L2 swath grid points which is calculated by a Gaussian weighting function. The Gaussian weighting, provided by the 'resample gauss' method of the 'pyresample' package, is given by the function $\exp(-R^2/\sigma^2)$, where R is the radius of influence (which is referred to the centre of the L3 grid cell) and σ is calculated from 3 dB FoV levels. R and σ are chosen according to the satellite sensor: 36 km and 18 km for AMSR2 and 75 km and 56 km for SSMIS. This logic is applied to all the L3 grid points apart from the near coast points. Figure 11 provides the image of the L3 target grid as well as the mask which is used to classify the 'near coast points'. As discussed in section 4.2.7, the sea ice retrieval is not attempted over the shore, near-shore and off-shore points for SSMIS and over the shore and near-shore points for AMSR2. However, for the L3 product, we search for a sea ice concentration interpolated value which can be assigned to the 'near coast points'. The interpolation is done by the 'resample custom' method of the 'pyresample' package and we calculate the sea ice concentration as the weighted average of the L2 SIC retrievals where the weighting function is given by 1 - r/R, where r is the distance between the observation and the centre of the L3 grid cell and R is the radius of influence (the weight assigned is zero when the distance of the observation is equal to the radius of influence). R is chosen independently of the satellite sensor and equal to 100 km to possibly give an estimation of sea ice concentration in those complex inner coastal regions. If we do not find at least 8 L2 swath grid points, we assign a missing value. More generally, missing values in the L3 products can be also the results of missing data in some part of the satellite swath. Finally, consistent with the L2 product, also for the L3, we implement the equivalent screening tests to filter the sea ice concentration. The screening tests can be checked looking at the L3 'status_flag' variable which can be also used to identify the land mask, the 'near coast points' and the missing values. A detailed description of the L3 status_flag can be found in the PUM.

Figure 12 and Figure 13, respectively for the Northern and Southern hemisphere, compare maps of the L3 sea ice concentration which is generated from the previous operational system and those obtained from the upgraded algorithm described in this ATBD. The maps show how the interpolation method consistently fills up the SIC missing values which previously categorised the two products in a different way. Overall, the unified processing logic in the upgraded algorithm is capable of providing a more uniform sea ice concentration analysis in the two hemispheres. However, as mentioned in section 4.2.7, in the L2 SSMIS processing, we also classify the 'off-shore' points as 'near coast' point. As a consequence, as displayed in Figure 12 (right panels), the L3 SSMIS product shows some regions in

the NH of missing values which instead are populated with an estimate of sea ice concentration in the AMSR2 product.

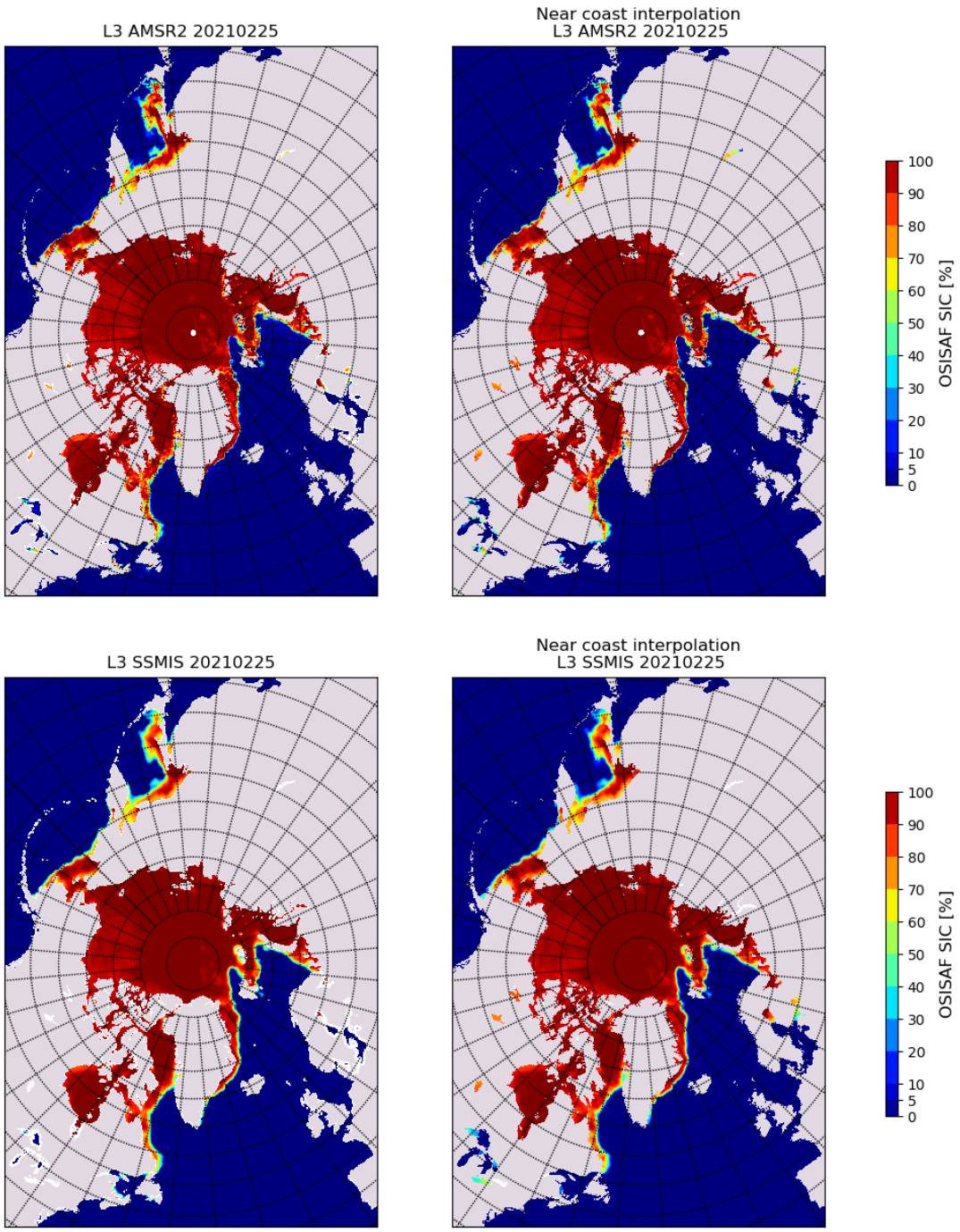

Figure 12: Maps which compare the L3 AMSR2 (top row) and L3 SSMIS (bottom row) sea ice concentration obtained from the previous operational system (left column) with those generated from the upgraded algorithm. White areas indicate missing values. The sea ice concentration analysis is for 25 February 2021. For the L3 AMSR2 product, we have missing data in the North Pole which is not covered by the satellite orbit.

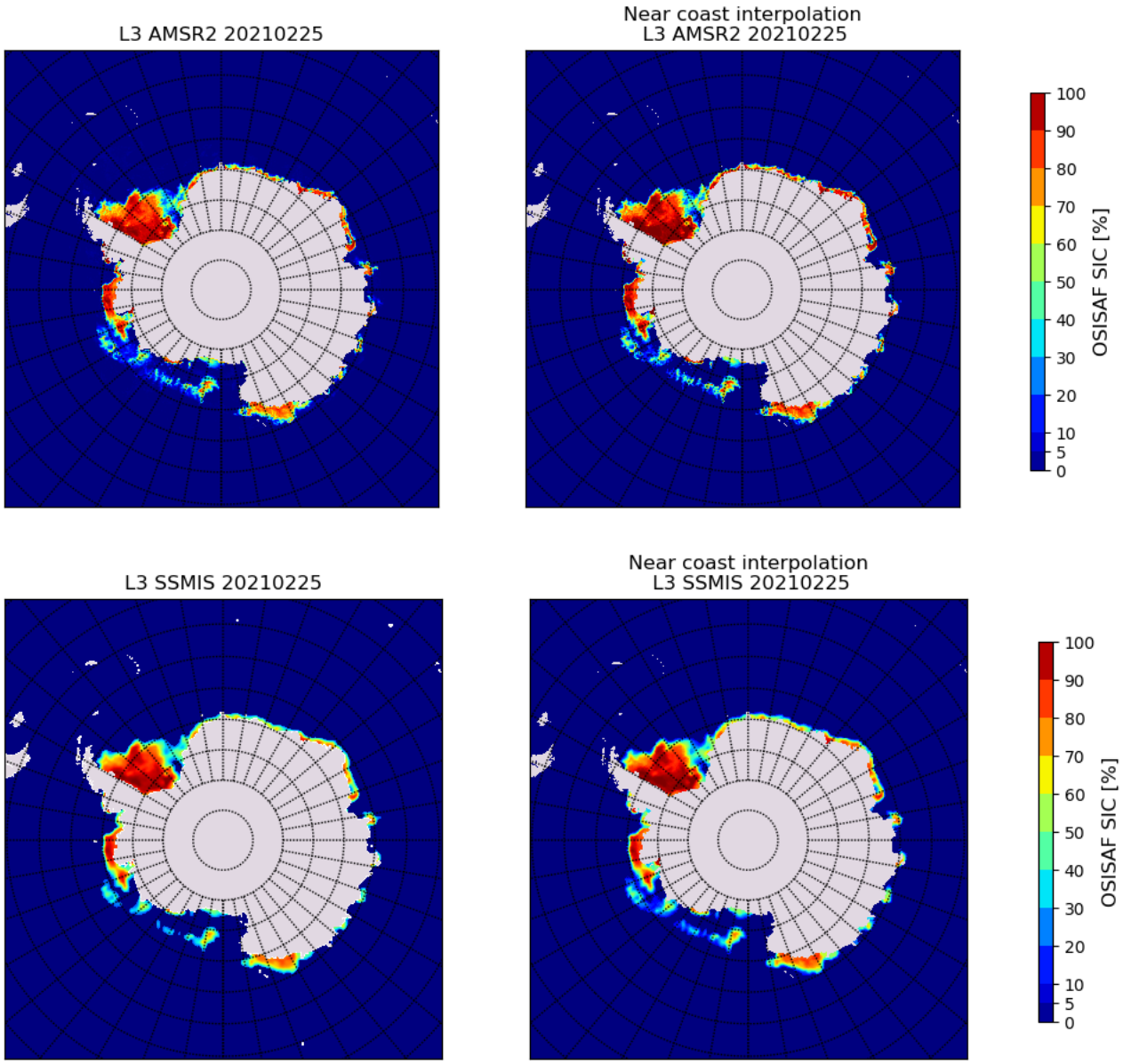

Figure 13: Same as Figure 12, but the L3 products are for the Southern hemisphere.

6. References

Andersen, S., R. Tonboe, S. Kern, and H. Schyberg. Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using Numerical Weather Prediction model fields: An intercomparison of nine algorithms. Remote Sensing of Environment104, 374-392, 2006.

Andersen, S., R. T. Tonboe and L. Kaleschke. Satellite thermal microwave sea ice concentration algorithm comparison. Arctic Sea Ice Thickness: Past, Present and Future, edited by Wadhams and Amanatidis. Climate Change and Natural Hazards Series 10, EUR 22416, 2006B.

Andersen, S., L. Toudal Pedersen, G. Heygster, R. Tonboe, and L Kaleschke, Intercomparison of passive microwave sea ice concentration retrievals over the high concentration Arctic sea ice. Journal of Geophysical Research 112, C08004, doi10.1029/2006JC003543, 2007.

Comiso J.C. Characteristics of arctic winter sea ice from satellite multispectral microwave observations. Journal of Geophysical Research 91(C1), 975-994, 1986.

Comiso J.C, D.J. Cavalieri, C.L. Parkinson, and P. Gloersen. Passive microwave algorithms for sea ice concentration: A comparison of two techniques. Remote Sensing of Environment 60, 357-384, 1997.

ESA SICCI project consortium. D2.6: Algorithm Theoretical Basis Document (ATBDv1), ESA Sea Ice Climate Initiative Phase 1 Report SICCI-ATBDv1-04-13, version 1.1, 2013.

Geer, A and Baur P.: Observation errors in all-sky data assimilation, Q. J. R. Meteorol. Soc., 137 2024–2037, 2011.

Geer, A, Baur P., Lonitz K., Barlakas V., Eriksson P., Mendrok, J., Doherty A., Hocking J., and Chambon P.: Bulk hydrometeor optical properties for microwave and sub-mm radiative transfer in RTTOV-SCATT v13.0, GMD, 2021. DOI: 10.5194/gmd-2021-73.

Hocking, J., Saunders, R., Geer, A. and Vidot, J 2020: RTTOV v13 User Guide, 2020. Available from https://nwp-saf.eumetsat.int/site/software/rttov/documentation.

Ivanova, N., L. T. Pedersen, R. T. Tonboe, S. Kern, G. Heygster, T. Lavergne, A. Sørensen, R. Saldo,G. Dybkjær, L. Brucker, and M. Shokr. 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, 2015.

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. (2019): 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.

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.

Markus T, Cavalieri D. The AMSR-E NT2 Sea Ice Concentration Algorithm: its Basis and Implementation. Journal of The Remote Sensing Society of Japan. 2009; 29 (1):216-225.

Pedersen, L.T. Chapter 6.2 in Sandven et al. IMSI report no. 8. Development of new satellite ice data products (Chapter 6.2). Bergen, Norway: NERSC Technical Report no.145, Nansen Environmental and Remote Sensing Center, Bergen, Norway, 1998.

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C., 2018: An update on the RTTOV fast radiative transfer model (currently at version 12), Geosci. Model Dev., 11, 2717-2737, https://doi.org/10.5194/gmd-11-2717-2018.

Smith, D. M. Extraction of winter total sea ice concentration in the Greenland and Barents Seas from SSM/I data. International Journal of Remote Sensing 17(13), 2625-2646, 1996.

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 well-calibrated ocean algorithm for SSM/I. Journal of Geophysical Research 102(C4), 8703-8718, 1997.

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