Annals of Glaciology



Article

Cite this article: Rusin J, Lavergne T, Doulgeris AP, Scott KA (2024). Resolution enhanced sea ice concentration: a new algorithm applied to AMSR2 microwave radiometry data. *Annals of Glaciology* **65**, e15, 1–12. https://doi.org/10.1017/aog.2024.6

Received: 20 June 2023
Revised: 4 January 2024
Accepted: 21 January 2024

Keywords:
polar and subpolar oceans; remote sensing; sea ice

Corresponding author:
Jozef Rusin; Email: jozefjr@met.no

© The Author(s), 2024. Published by Cambridge University Press on behalf of International Glaciological Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Resolution enhanced sea ice concentration: a new algorithm applied to AMSR2 microwave radiometry data

Jozef Rusin[1,2], Thomas Lavergne[1], Anthony P. Doulgeris[2] and K. Andrea Scott[3]

[1]Research and Development Department, Norwegian Meteorological Institute, Oslo, Norway; [2]Department of Physics and Technology, UiT The Arctic University of Norway, Tromsø, Norway and [3]Department of Systems Design Engineering, University of Waterloo, Waterloo, Canada

Abstract

Passive-microwave sea ice concentration (SIC) algorithms employ different frequencies and polarisations in their operational implementations. Commonly, these algorithms utilise combinations such as 19/37 GHz, yielding reduced measurement uncertainties but at a coarse spatial resolution. Alternatively, these algorithms can solely use 89 GHz, producing a higher spatial resolution but with increased measurement uncertainties. This study evaluates the application of a resolution-enhancing SIC algorithm (reSICCI3LF), initially developed for the coarser Special Sensor Microwave Imager / Sounder, on the Advanced Microwave Scanning Radiometer. By applying reSICCI3LF, we aim to produce a 5 km SIC for 2013–2020 in the Fram Strait and the Barents and Kara Sea region that gains the benefits of both types of algorithms, high spatial resolution and low measurement uncertainty.

We present the algorithm tuning, spectral analysis of spatial resolutions, and validation against the Round Robin Data Package of 0% and 100% SIC points and SIC derived from Landsat-8. The findings demonstrate that the reSICCI3LF algorithm produces a SIC field with fine details, achieving a balance between high spatial resolution and lower measurement uncertainties compared to a 89 GHz based SIC. Consequently, this resolution-enhanced SIC technique can potentially initialise higher-resolution coupled ocean and sea ice forecasting systems through data assimilation.

1. Introduction

Sea ice concentration (SIC), the fraction of an ocean area covered with sea ice, is a crucial variable for understanding and monitoring the state of the Earth's climate system, with sea ice influencing the Earth's energy balance, ocean currents and atmospheric circulation (Stroeve and Notz, 2018). Additionally, SIC data are used to initialise sea ice forecast models (Zhang and others, 2021) and provide input and boundary conditions to coupled climate and atmospheric models (Tietsche and others, 2013; Browne and others, 2019).

With sea ice being a critical component of the climate system and a key indicator of climate change (Meier and Stroeve, 2022), the global climate observing system has identified sea ice as an essential climate variable (Bojinski and others, 2014). SIC is one of these key parameters since it can be used to succinctly capture the state of the sea ice environment (Lavergne and others, 2022a). Passive microwave (PMW) sensors have enabled frequent and large-scale measurement of polar sea ice concentration and its derived quantities (sea ice extent and area) since the late 1970s, thanks to their polar orbits and wide swaths. PMW sensors are particularly suited to polar applications due to the microwave emissions they observe being independent of solar illumination and having generally low sensitivity to atmospheric influences in the atmospheric window frequencies, except for in the presence of intense precipitation and wind-roughened open ocean (Meier and Stewart, 2020), with greater sensitivity to atmospheric water vapour occurring in the 89 GHz channel (Ivanova and others, 2015). This has meant that PMW sensors have produced a near continuous climate data record that has enabled the identification of changes to the ice melt timing and duration (Stroeve and others, 2014; Serreze and others, 2016) and reductions in Arctic summer sea ice extents (Stroeve and Notz, 2018).

An established PMW sensor is the Advanced Microwave Scanning Radiometer (AMSR2) onboard the Japan Aerospace Exploration Agency (JAXA) Global Change Observation Mission for Water (GCOM-W) satellite. Due to the larger antennae size, the AMSR2 sensor can provide higher spatial resolution observations than previous high-frequency PMW sensors prior to 2012. The spatial resolution capabilities and observed frequency channels for AMSR2 are shown in Table 1.

Operational SIC algorithms generally fall within two groups: 19 and 37 GHz algorithms or near-90 GHz algorithms. The near-90 GHz algorithms have a high spatial resolution of ~5 km when using 89 GHz AMSR2 data but are the most sensitive to atmospheric effects (Andersen and others, 2006; Ivanova and others, 2015). Algorithms based on 19 and 37 GHz show the smallest atmospheric effects (Andersen and others, 2006), thus producing lower measurement uncertainties (Ivanova and others, 2015), but will lead to SIC that is coarser in spatial resolution (Table 1). The latest GCOS implementation plan outlines the research needs for future SIC products. It states that SIC products need to target a higher spatial resolution of 1–5 km

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

Table 1. AMSR2 passive microwave radiometer sensor measured frequencies, their polarisations (vertical and horizontal), and their instantaneous field of view (IFOV) spatial resolutions

Central frequency (GHz)	Polarisations	IFOV (km)
6.925	V, H	35 × 62
10.65	V, H	24 × 42
18.7	V, H	14 × 22
36.5	V, H	7 × 12
89.0	V, H	3 × 5

for improved regional and coastal analysis as well as measurement uncertainties lower than 5% (Zemp and others, 2022). Additionally, operational models that assimilate SIC are increasing in grid resolution and consequentially have the potential to resolve higher-resolution sea ice features. Therefore, further SIC algorithms from PMW must be developed to produce a higher spatial resolution with a low measurement uncertainty to provide sufficient observations for assimilation. This operational need is the rationale behind the sea ice retrievals and data Assimilation in Norway (SIRANO) project (https://cryo.met.no/en/sirano), which aims to advance satellite observations of SIC and their assimilation into ocean/ice forecast models to improve sea ice forecasting capabilities.

Currently, the primary AMSR2 datasets assimilated into forecast models are the Ocean and Sea Ice Satellite Application Facilities (OSI SAF) OSI-408-a 19/37 GHz-based SIC algorithm, provided at a 10 km grid spacing, or the University of Bremen's ARTIST sea ice (ASI) dataset, derived from the 89 GHz channel and provided at a 6.25 km spacing. For this study, we focus on sea ice regions that are operationally modelled in Norwegian areas of interest; the Barents-2.5 km model domain, implemented by the Norwegian Meteorological Institute (Röhrs and others, 2023) and the Svalbard 4 km model domain, implemented at the Norwegian Polar Institute (Duarte and others, 2022).

Previous research has been undertaken to enhance the resolution of PMW sensors by adding higher resolution information from visible and thermal infrared sensors (Ludwig and others, 2019, 2020; Dworak and others, 2021), as well as synthetic aperture radar (SAR) (Scott and others, 2015; Wang and others, 2016; Karvonen, 2017). However, these methodologies have constraints when compared to using purely PMW data, such as restrictions due to the polar night (visible), cloud cover (visible & thermal infrared), reduced data coverage (visible, thermal infrared & SAR) when compared to the sub-daily imaging capabilities of AMSR2, and will additionally introduce new measurement uncertainties when compared to purely PMW data.

Consequentially this research focuses on improving SIC measurement using purely PMW data by utilising different frequencies and polarisations (19 V, 37 V, 37H, 89 V & 89H) to produce a SIC that has an improved spatial resolution at 5 km gridding that does not impact the measurement uncertainty. Previous work has been undertaken to develop algorithms that combine multiple PMW frequency channels to produce a higher resolution SIC such as the IceCREAM algorithm (Kilic and others, 2020) and the reSICCI3LF algorithm (Lavergne and others, 2021). The IceCREAM optimal estimation method was developed using AMSR2 channels but specifically for the upcoming Copernicus Imaging Microwave Radiometer (CIMR) mission and therefore does not utilise the high-resolution 89 GHz channel of AMSR2. The reSICCI3LF algorithm was defined in the European Space Agency (ESA) Climate Change Initiative (CCI) algorithm theoretical basis documentation (ATBD) by Lavergne and others (2021). Its methodology uses the near-90 GHz channel but has only been applied to the Special Sensor Microwave Imager / Sounder (SSMIS) data record to date.

This research will therefore migrate the application of the reSICCI3LF algorithm presented in the ESA CCI ATBD to AMSR2 data and fully evaluate its applicability in producing a higher-resolution SIC. The research aims are as follows:

1. Adapt the reSICCI3LF algorithm to the AMSR2 sensor.
2. Assess the spatial resolution capabilities of the AMSR2 reSICCI3LF SIC.
3. Validate the AMSR2 reSICCI3LF SIC to comprehend how its application impacts measurement accuracy and precision.

2. Data and methods

2.1 AMSR2 data processing

The PMW SIC fields produced for this analysis are derived using data from the AMSR2 sensor. For this analysis, eight complete years of data have been processed, starting from 2012 and ending in 2020. The frequencies and the instantaneous field of view (IFOV) footprint sizes of AMSR2 frequencies applied in this research are outlined in Table 1. Overall the processing chains used in this research are adapted from the ESA CCI SIC climate data record and OSI SAF methodologies (Lavergne and others, 2019, 2021) with the general overview of these chains being outlined in Fig. 1.

In summary, the Level 1 (L1) processing ingests and pre-processes the top of the atmosphere (TOA) L1 brightness temperature product (L1B) from JAXA observed by the satellite. The Level 2 (L2) chain then takes the L1 pre-processed (L1P) files and applies the SIC algorithms whilst maintaining the original swath coordinates. A key stage in the L2 chain is the correction loop in the second iteration of the processing chain (Fig. 1). During this step, corrections are made to the SIC fields to account for atmospheric contributions since it has been shown that correcting the temperature brightness (T_B's) observed by the PMW sensor improves SIC accuracy (Andersen and others, 2006; Ivanova and others, 2015; Lu and others, 2018, 2022). For this research, we applied the double-difference scheme outlined in Lavergne and others (2019) that advances the work of Tonboe and others (2016). An additional correction is also applied to the SIC field using the "sea ice curve" method outlined in Lavergne and others (2019), which enables the correction at 100% SIC regions to account for deviations that occur due to the different ice types, snow cover and seasonal variations. Finally, in the Level 3 (L3) processing chain, the individual L2 swath files are gridded onto a 5 km EASE2 polar grid. During the gridding, uncertainty estimates for the SIC algorithms are calculated using the methodology outlined in Lavergne and others (2019), where the total uncertainty is assumed to be composed of the inherent uncertainty from the algorithm (including sensor noise and temperature brightness variability around the mean open water and consolidated ice signatures) and the smearing uncertainty caused due to gridding of the swath data. In the final Level 3 (L3) product, open water filters (OWF) are also applied to the SIC fields. This step aims to remove erroneous sea ice that is present in open water regions caused by atmospheric influences such as wind, water vapour and cloud liquid water whilst minimising the removal of true ice at the ice edge (Lavergne and others, 2019). The OWF technique used in this analysis is the dynamic tuning approach used in Lavergne and others (2019, 2021, 2022b). It should be noted that despite the OWF being effective at omitting false ice in open water (OW) regions, their application will also remove true low SIC values present at the ice edge (Ivanova and others, 2015).

2.2 AMSR2 sea ice concentration

First developed during the ESA CCI sea ice project for the Special Sensor Microwave Imager / Sounder (SSMIS) sensor (Lavergne

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

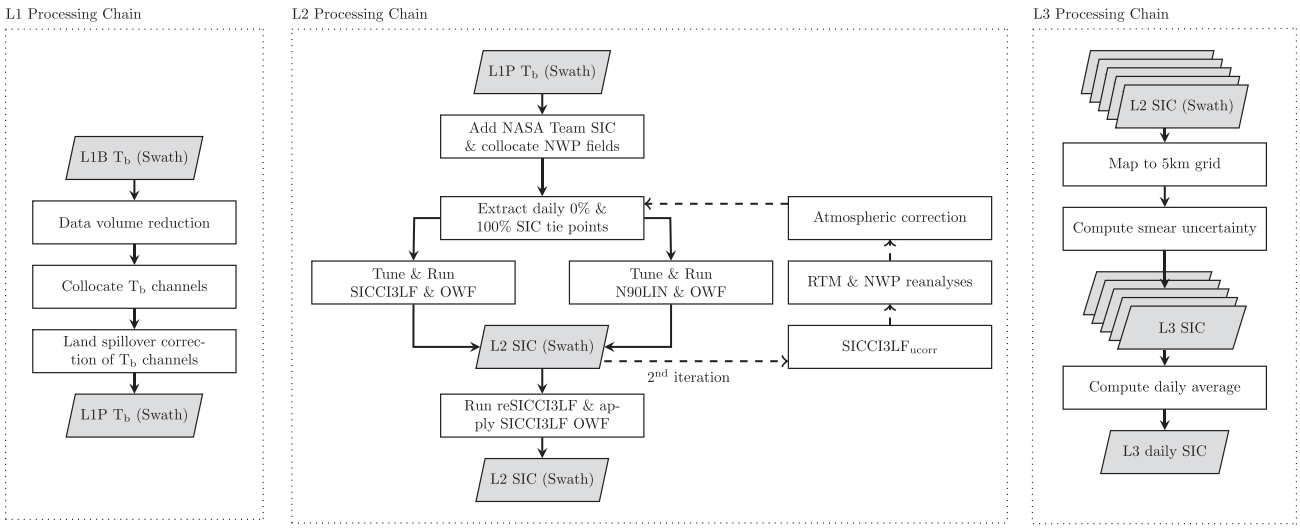


Figure 1. AMSR2 processing chains used to produce a 5 km reSICCI3LF sea ice concentration.

and others, 2022b), the resolution enhanced SICCI3LF (reSICCI3LF) algorithm (Eq. (1)) merges the SIC from a high-resolution SIC with a coarser SIC field, with the aim of adding higher resolution sea ice details into the coarser field. The rationale behind this is that reSICCI3LF will gain the benefits of both algorithms; the higher resolution capabilities, particularly at the ice edge, of using an 89GHz-based algorithm with the accuracy and low measurement uncertainty of a 19/37GHz-based algorithm within the ice pack.

For this research, we apply the reSICCI3LF algorithm to AMSR2 data for the first time. This means that we can produce a higher resolution product with a 5 km final gridding compared to the ESA CCI data produced at 12.5 km due to the coarser resolution restrictions of the SSMIS sensor. Additionally, we implemented the reSICCI3LF at an earlier stage of the production, and it is now produced during the L2 processing chain, whereas, for the ESA CCI project, this was done in the final stages of the L3 processing chain and on the daily average files. Implementing this at an earlier stage, rather than the daily averaging, enables the reSICCI3LF to be generated with an OWF more representative of the conditions at the time of observation. This means the OWF will therefore be more indicative of the dynamic atmospheric conditions that influence the OWF and will improve the filtering process. The reSICCI3LF algorithm is shown in Eq. (1):

$$SIC_{reSICCI3LF} = SIC_{SICCI3LF} + (SIC_{N90LIN} - SIC_{N90LIN\ blurred}), \quad (1)$$

where $SIC_{SICCI3LF}$ is the low-resolution SIC field and SIC_{N90LIN} is the high-resolution SIC field. In this research, the $SIC_{SICCI3LF}$ algorithm is the updated version outlined in (Lavergne and others, 2022b). The SICCI3LF and N90LIN are hybrid algorithms optimised for open water and consolidated ice pack conditions using the brightness temperatures from channels 19 V, 37 V and 37H for SICCI3LF and 89 V and 89H for N90LIN. The optimisation follows the methodology outlined in Lavergne and others (2019) with the final SIC being a weighted linear average between the optimisations at open water and consolidated ice for both SICCI3LF and N90LIN (Appendix).

The blurring of the N90LIN (SIC_{N90LIN} blurred) dictates where the added sea ice details from the N90LIN should be implemented. This is because the difference term in Eq. (1), referred to as delta hereafter, will adjust the SIC only where there are sharp SIC delta gradients, such as in the marginal ice zone (MIZ) or sea ice openings. This is because SIC_{N90LIN} blurred is blurred to approximately the same spatial resolution as $SIC_{SICCI3LF}$ and consequently means that only small scale features in SIC_{N90LIN} are added to $SIC_{SICCI3LF}$. Additionally, this should also mean that the delta is zero away from these sharp SIC delta gradients, such as within the ice pack, and consequently, the $SIC_{reSICCI3LF}$ will equal the $SIC_{SICCI3LF}$ field. A visual representation of the inputs for the reSICCI3LF algorithm is shown in Fig. 2.

As outlined in Lavergne and others (2021), for the correct amount of enhancement to occur in the SIC delta, the spatial resolution of SIC_{N90LIN} blurred should be equal to or be closely comparable to that of the $SIC_{SICCI3LF}$ field. In this research, all SIC fields are on a 5 km gridding, but their true spatial resolutions will vary depending on the SIC algorithm and the frequency channels applied (Table 1). Therefore to create SIC_{N90LIN} blurred a Gaussian kernel is applied to SIC_{N90LIN}. This results in a weighted averaging of SIC values when the Gaussian kernel passes over, with the rate of the weight decrease being controlled by the standard deviation or "sigma" value. For the first stage of the research, we want to confirm the correct smoothing scale. To test this, we incrementally increased this sigma value to gradually increase the amount of blurring in the SIC_{N90LIN} field. The blurred fields are then compared to the $SIC_{SICCI3LF}$ using spectral analysis (section 2.5) to evaluate which of the SIC_{N90LIN} blurred fields are most comparable to the $SIC_{SICCI3LF}$ field in spatial resolution. Section 3.1 describes this process in more detail.

2.3 Round Robin Data Package

The Round Robin Data Package (RRDP) version 2 (Pedersen and others, 2019) is a validation data set produced for the ESA CCI that includes referenced sea ice concentrations at 0% and 100% SIC, covering the Arctic and Antarctic regions during different seasons and varying meteorological conditions. The 0% SIC RRDP validation points are defined as high latitude areas of no ice that have been identified to be ~100 km away from the ice edge using ice charts and are used to validate the ability of the algorithm to define open water correctly. The 100% SIC RRDP validation points are defined as regions of two consecutive days of converging ice drift that has been observed from synthetic aperture radar from the PolarView / MyOcean /CMEMS ice drift dataset (Pedersen and others, 2019), with the assumption that after two consecutive days of convergence, any remaining

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

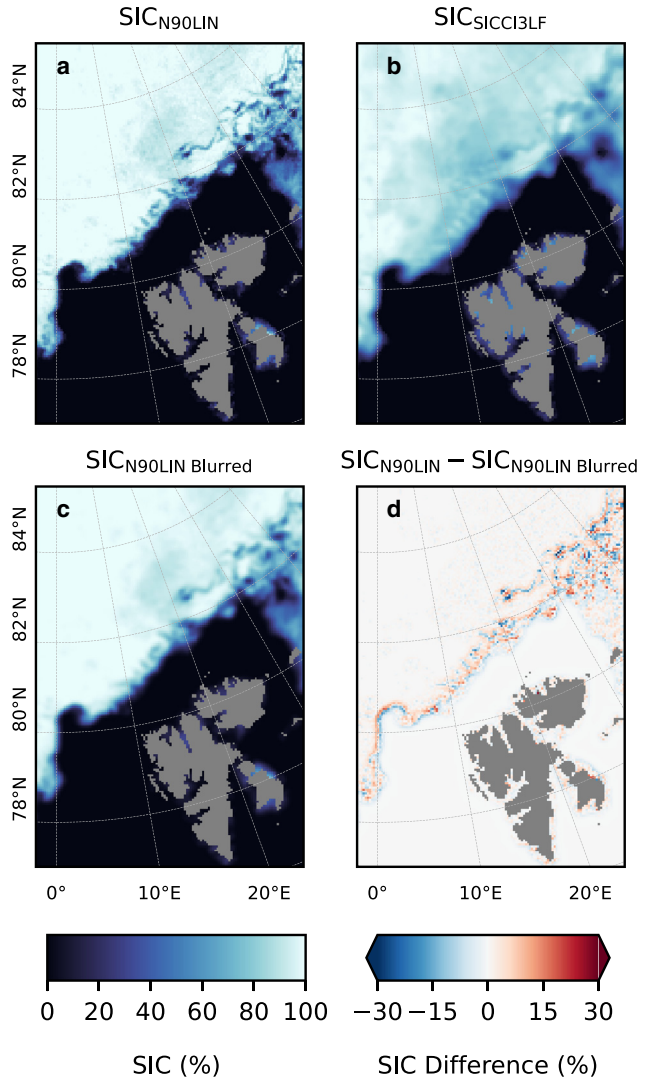


Figure 2. Inputs for the reSICCI3LF sea ice concentration (SIC) algorithm; (a) high-resolution SIC (N90LIN), (b) low-resolution SIC (SICCI3LF), (c) N90LIN blurred SIC, and (d) the difference of N90LIN - N90LIN$_{blurred}$. Svalbard SIC obtained from the 12th of December 2016.

openings will either be frozen over or closed due to ice motion. For the summer periods, the assumption that remaining openings will refreeze cannot be guaranteed; therefore, biases may be present. As a result, the study does not investigate the summer months for 100% SIC. The locations of the RRDP points used in the study are shown in Fig. 3.

2.4 Landsat validation data

For validating the PMW SIC algorithms, we produce a 5 km SIC using high-resolution multi-spectral Landsat-8 data. The data used in creating this SIC was derived from a 30 m ice/water supervised classification dataset produced by Kern (2021). For our research, a subset of 44 scenes of the original (Kern, 2021) total was used since only the Landsat-8 (L8) data from 2013 to 2015 intersected with our region of focus.

Overall, the classification is based on surface broadband shortwave albedo derived from channels 3 (533–590 nm), 4 (636–673 nm), and 5 (851–879 nm) of the L8 Operational Land Imager sensor. The surface broadband shortwave albedo images needed for the final classification can be produced with an assumption that the top of atmosphere (TOA) reflectances are equivalent to the TOA albedo (α_{TOA}), which can be assumed to be related to the surface albedo

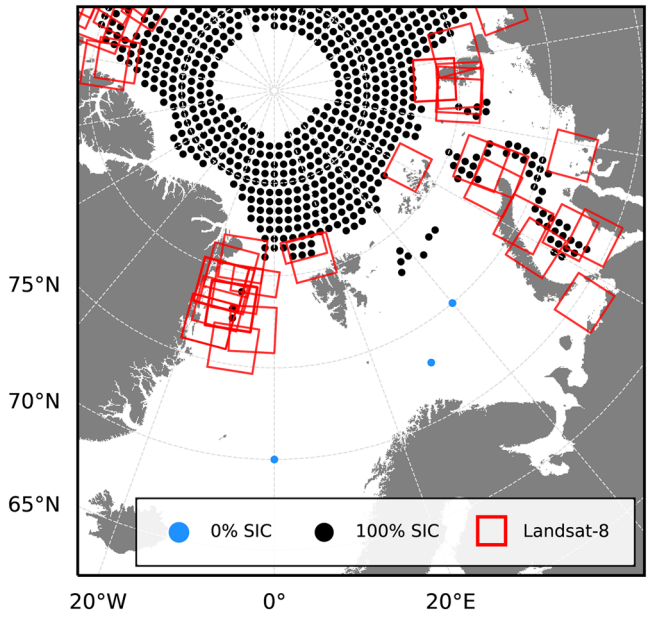


Figure 3. Round Robin Data Package v2 (RRDP) locations of 0% and 100% SIC validation points used in the study and the locations of the Landsat-8 classification swaths applied in the validation. A total of 20,105 RRDP 0% SIC points and a total of 23,295 100% SIC points were used in the validation, sampled from within the open water (blue) and consolidated ice (black) geographical locations from 2013 to 2019. A total of 44 Landsat-8 scenes were used in the analysis.

($\alpha_{surface}$) by using the following equation (Kern and others, 2022):

$$\alpha_{TOA} = a + b \times \alpha_{surface}, \quad (2)$$

where the coefficients a and b are related to the atmospheric conditions and the L8 wavelength respectively. For each channel, the $\alpha_{surface}$ is derived using Eq. (2) from a surface broadband shortwave albedo map made using channel bandwidths as weights. The surface broadband shortwave albedo images are then used in the supervised visual classification of open water and ice for the (Kern, 2021) dataset. An in-depth processing description can be found in (Kern and others, 2022).

The final L8 SIC at 5 km gridding used for the PMW validation is produced by re-projecting the 30 m L8 data to the same EASE2 polar projection as the PMW data. Then the 30 m ice/water classified cells intersecting with a 5 km PMW grid cell are counted. Consequently, for each 5 km grid cell, the intersecting water and ice L8 data ratio can be used to derive a 5 km L8 SIC that will align perfectly with the 5 km PMW SIC. Any 5 km grid cells that did not have more than 90% coverage, due to cloud cover, from the 30 m L8 source data were masked from the validation analysis. In total, this produces 34,345 pixels that can be used for validating the PMW algorithms. The locations of the 44 classified L8 scenes are shown in Fig. 3.

2.5 Variance spectra

Spectral analysis of the SIC in two dimensions enables us to examine different spatial scales in the sea ice cover. This allows us to identify dominant scales present in the data and estimate the spatial resolution of the three different SIC algorithms. For example, if a reduction in the spectral density at shorter wavelengths (higher frequencies) is observed, it would indicate that the SIC is composed of fewer smaller-scale features, hence indicating a coarser SIC. By applying this methodology, we can identify the smallest resolvable scales in the data.

Typically for spectral analysis, the discreet Fourier transform (DFT) is applied to decompose the field. For this paper, we

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

apply the alternative method outlined in Denis and others (2002), which uses a discrete cosine transform (DCT). The benefit of the DCT method over DFT is that we can calculate the spectra for non-periodic signals. The data must be periodic for DFT, achieved through spatial detrending (removal of non-periodic signals) or windowing (multiplying the SIC by a function to taper the data at the domain boundaries). However, the downside of spatial detrending is that it can result in the removal of meaningful spectral information and introduce artificial features at large scales that consequentially produce aliasing (Denis and others, 2002).

For the research, two regions were studied using this technique, one over the ice pack and one at the ice edge, as shown in Fig. 4a. For each day of the study period, 2013 to 2020, the SIC is transformed into two-dimensional spectral space, and the power variance spectrum is derived based on the method outlined in Denis and others (2002), using the updated weighted wavelength binning outlined in Ricard and others (2013). This same approach has also recently been applied to study the size of sea surface temperature features (Iversen and others, 2023).

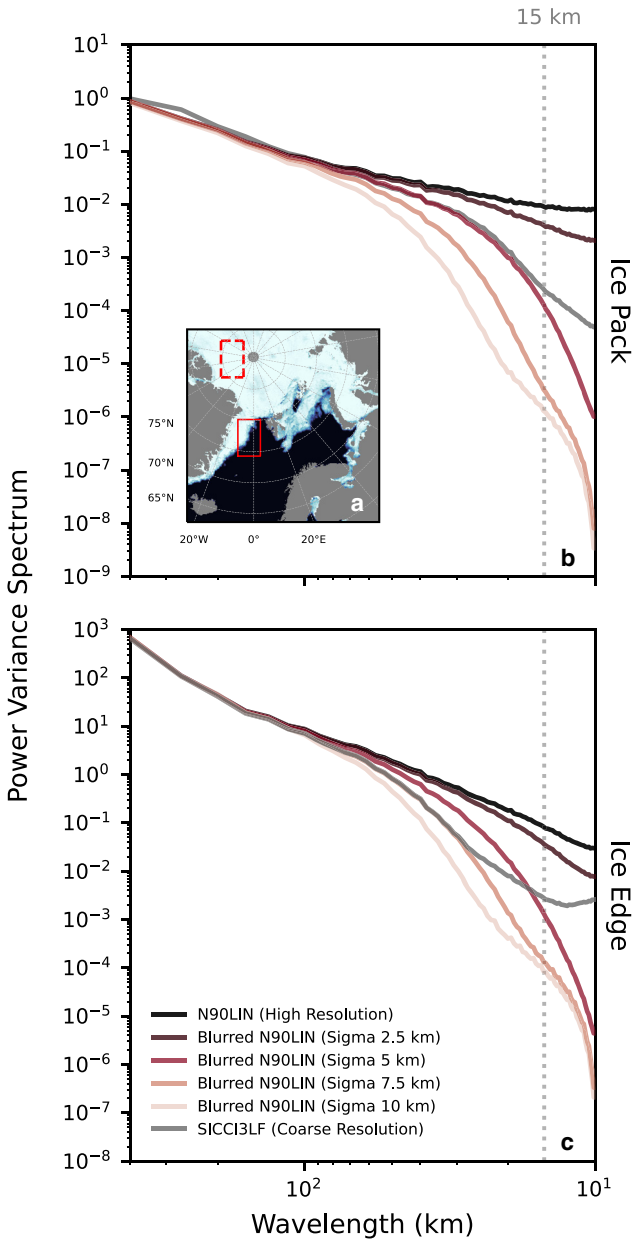


Figure 4. Median power spectra computed using daily averaged sea ice concentration (SIC) data from 2013 to 2020 for the two outlined regions shown in inset (a). The ice pack region, depicted by the dashed red region, and the ice edge region, represented by the solid red region, were studied. Plots (b–c) display the median power spectra for the SICCI3LF, N90LIN, and incrementally blurred N90LIN SIC fields for the ice pack and ice edge regions, respectively. The dotted grey line is located at a wavelength of 15 km to highlight when the sigma field of 5 km will approximately begin to have lower spectral values than the targeted SICCI3LF spectrum. Note that the y-axes ranges differ in plots (b-c).

3. Results

3.1 Tuning reSICCI3LF

As previously outlined, the optimal amount of blurring of the N90LIN field needs to occur to ensure that the correct amount of SIC is added and removed from the SICCI3LF field. Previously, the ESA CCI project (Lavergne and others, 2021) determined the sigma of the Gaussian kernel by analysing transects over the ice edge and comparing the SIC values to the SICCI3LF SIC. We advance this method by calculating each incrementally blurred N90LIN field's power spectrum and comparing the results to the SICCI3LF field, with a suitable blurring expected to bring the blurred N90LIN spectrum to align with the SICCI3LF spectrum (Fig. 4).

Figure 4b shows the median power spectra for the dashed red outlined region in Fig. 4a. The x-axis shows the wavelength of features in km, which can be interpreted as the size of sea ice features present, and the y-axis can be seen as the prevalence that these features occur. The largest wavelength is constrained to the size of the smallest length of the bounding box, whereas the smallest wavelength is limited to the size of two grid pixels, which is 10 km in this application.

Initially, all algorithms generally follow similar trends over the ice pack from the largest wavelength (140 km) until a wavelength of ~100 km. However, the SICCI3LF does show a small peak occurring at ~120 km, which is not present in the N90LIN or blurred N90LIN fields. At ~100 km onwards, the blurred fields using a 7.5 km and 10 km sigma decrease more rapidly than the SICCI3LF, N90LIN, sigma 5 km, and sigma 2.5 km fields. From ~100 km to 20 km, the SICCI3LF and sigma 5 km blurred field follow similar trends, indicating they observe similar distributions of ice features at these wavelengths. At 15 km, the 5 km sigma blurred field drops rapidly compared to SICCI3LF. This decrease is likely due to the Gaussian blurring removing small-scale heterogeneous sea ice features that are present. Compared to the other fields, the N90LIN spectrum decreases gradually and begins to flatten out at 15 km, with an overall higher power spectrum across the smaller wavelengths. Some of this signal will be due to the higher resolution capabilities of the N90LIN algorithm. However, this signal will also be affected by its sensitivity to atmospheric effects, which cause artificially low SIC features of variable size within the ice pack.

For the ice edge region (Fig. 4c), all algorithms have higher spectra when compared to the ice pack region (Fig. 4b) for the large wavelengths. This is because the open water section will be seen as one large continuous feature that will dominate at these wavelengths, even though ice is always present in the region (of varying spatial scale) throughout the study period. We see that the blurred field using a sigma of 10 km begins to deviate from SICCI3LF at ~100 km onwards the most, whereas a sigma of 7.5 km closely follows the SICCI3LF field for wavelengths up to 30 km but then deviates rapidly from the SICCI3LF spectrum. The 5 km sigma field has slightly larger values than the SICCI3LF spectra, but from ~20 km to 10 km, the 5 km sigma spectrum is closer to SICCI3LF than the 7.5 km sigma spectrum. Similar to the ice pack region, the blurring will remove small sea ice features at shorter wavelengths. This is more evident for the ice edge region since the blurring will also smooth the ice edge, i.e.

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

reducing the present sharp transitions and consequently causing the low spectra values at <15 km wavelengths for 5, 7.5 and 10 km sigma values. For both regions, the 2.5 km spectrum is closer to the N90LIN spectrum; this indicates that a sigma of 2.5 km does not suitably blur the SIC field to a resolution comparable to SICCI3LF. This means that if a 2.5 km sigma were used, fewer fine scale ice details would be present in the reSICCI3LF SIC since the delta term in Eq. (1) will be close to 0. Overall, based on Figs. 4b and 4c, a sigma of 5 km was selected to produce the blurred N90LIN field since it had comparable spectra to SICCI3LF for both the ice pack and ice edge regions.

3.2 Production of reSICCI3LF

A visual comparison of reSICCI3LF derived using this selected sigma value vs SICCI3LF and N90LIN is shown in Fig. 5. The top row comparison (Figs. 5a–5c) is an example for the Svalbard Archipelago and Zemlya Georga region from the 9th of March 2019. Visually we see that Fig. 5b, (reSICCI3LF), has finer sea ice details than Fig. 5a (SICCI3LF) and visually looks closer to Fig. 5c (N90LIN) in the size of spatial features present. This is also evident in Figs. 5d–5f, which shows an example of the ice edge in the Greenland Sea and Fram Strait. Again we see finer sea ice details visually in the reSICCI3LF field Fig. 5e when compared to SICCI3LF (Fig. 5d), which again appear close in spatial resolution to the N90LIN field visually (Fig. 5f). However, we do see that reSICCI3LF does not define all the openings as sharply compared to the N90LIN, with the reSICCI3LF appearing to retain more of the SICCI3LF field. Additionally, the reSICCI3LF SIC does not fully reproduce the apparent oversaturated 100% SIC areas in the N90LIN SIC.

Alongside the visual comparisons, the spectral analysis method introduced in section 2.5 can be used to evaluate further the spatial resolution of the final products (Fig. 6). Since the ice pack region is relatively homogeneous in its SIC values, it does not contain many variations from one pixel to another; this means it will not have

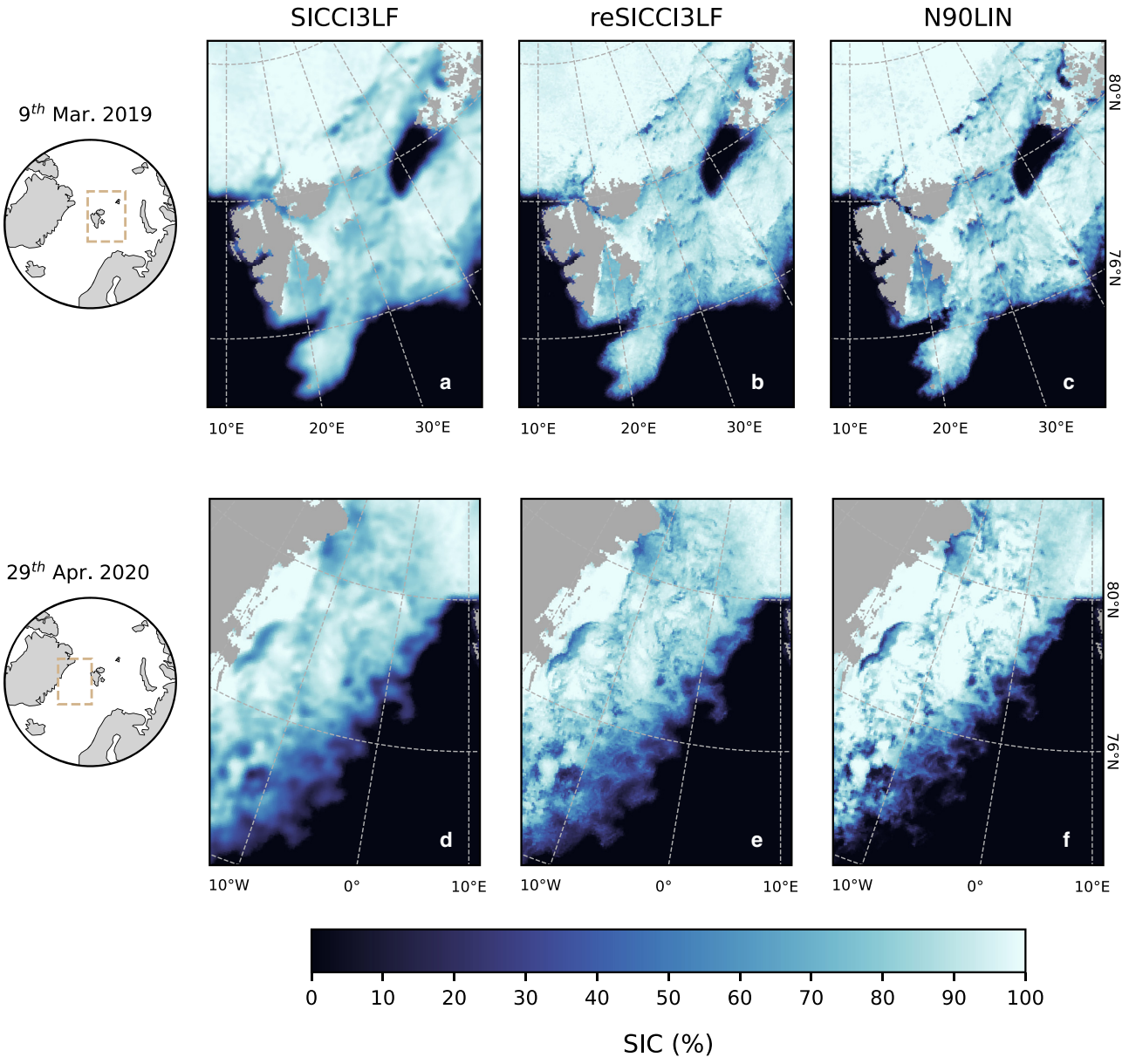


Figure 5. Sea ice concentration (SIC) algorithms, SICCI3LF, reSICCI3LF, and N90LIN, visualised with a regional focus on the Svalbard Archipelago and Zemlya Georga (a-c) and the Greenland Sea and Fram Strait (d-f). The SIC algorithms presented are the SICCI3LF (coarse resolution) in (a) and (d), the reSICCI3LF (resolution enhanced) in (b) and (e), and N90LIN (high resolution) in (c) and (f). The dates 9th of March 2019 and 29th of April 2020 were selected to highlight the varying capabilities of the SIC algorithms after applying open water filtering.

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

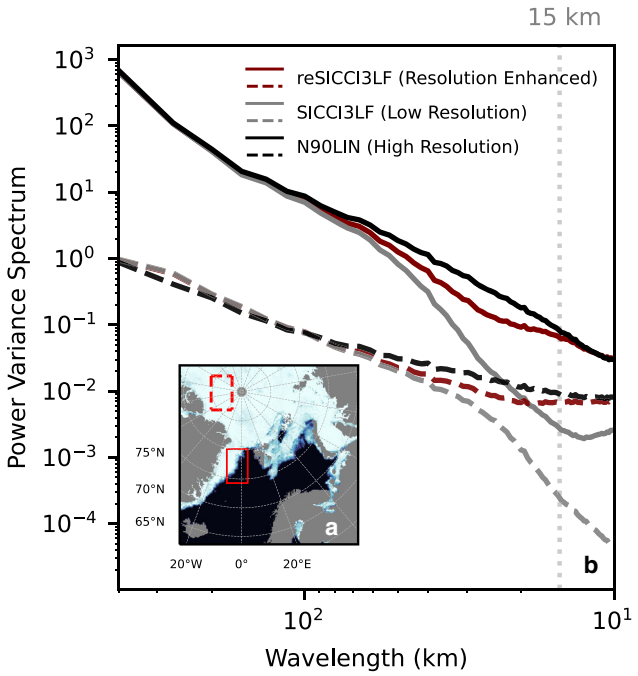


Figure 6. Median power spectra computed using daily averaged sea ice concentration (SIC) data from 2013 to 2020 of the reSICCI3LF, SICCI3LF and N90LIN sea ice concentration (SIC) fields. Inset (a) shows the location of the two regions studied, with the ice pack region represented by the dashed red region and the ice edge region indicated by the solid red region. For (b), the dashed lines represent the ice pack region, and the solid lines represent the ice edge region. The dotted grey line is located at a wavelength of 15 km to highlight when the reSICCI3LF and N90LIN algorithms approximately begin to diverge.

much high-frequency information, and consequently, the spectrum with a lower amplitude has less power at the high spatial frequencies. When comparing the different algorithms for both regions, the spectra show that the reSICCI3LF spectrum is brought closer to the N90LIN spectrum (when starting from the SICCI3LF spectrum). This confirms the visual assessment that the sea ice details in reSICCI3LF are "enhanced" to a similar spatial resolution to what is observed in N90LIN, particularly for wavelengths 10 km to ~15 km. However, the reSICCI3LF diverges from the N90LIN spectra between ~15 and 70 km. This means the reSICCI3LF does not observe the same amount of sea ice details at these scales compared to N90LIN. However, we only want to match the N90LIN partially due to the more significant contributions of atmospheric effects on the SIC and its power spectra. From 100 km up to 140 km for the ice pack region, the reSICCI3LF follows exactly the SICCI3LF spectra. This would also confirm that a suitable sigma was selected since the reSICCI3LF and SICCI3LF spectra are identical at large scales. We saw this in Fig. 4b, where the 5 km blurring follows mostly directly on the SICCI3LF spectrum. Confirming that the reSICCI3LF algorithm is performing as expected since the algorithm should only be using data from the SICCI3LF SIC field for large-scale features.

3.3 Validation against the Round Robin Data Package

In this section, we evaluate the reSICCI3LF vs N90LIN and SICCI3LF products at 0% and 100% SIC using the RRDP. Comparing the algorithm outputs at these known concentrations enables us to assess the algorithm's accuracy (mean values) and precision (distribution around the mean) in defining OW and closed ice (CI) regions. For this validation, we follow the approach of Kern and others (2019); we validate the SIC before the OWF application and before thresholding at 100% SIC. This unfiltered product provides more insight into the algorithmic variability and uncertainty of the SIC at 0% and 100%. If we had studied the SIC after the OWF application, the SIC at these tie-points would be set to exactly 100% SIC if >100% and 0% if <0%, biasing the "true" distribution of the algorithmic uncertainties present. The algorithmic distributions around the RRDP validation points are shown in Fig. 7.

We see that N90LIN (Fig. 7a) has the broadest distribution for CI regions when compared to the SICCI3LF and reSICCI3LF, indicating that N90LIN has the least precision. In contrast reSICCI3LF and SICCI3LF have similar distributions over the CI region, with the main differences being the locations of the outliers. For the OW validation (Fig. 7b), reSICCI3LF and SICCI3LF again appear to have comparable distributions, with reSICCI3LF having a slightly larger spread but in general both algorithms have means, medians and upper and lower quarterlies closely falling around 0, again indicating similar precision between reSICCI3LF and SICCI3LF. Fig. 7b shows the limitations of using purely N90LIN due to its broad OW distributions.

Summary results from the RRDP validation are shown in Table 2 for CI & Table 3 for OW validation points. A mean value closer to 100% (Table 2) or 0% (Table 3) is used to indicate the algorithm accuracy in correctly defining the 0% and 100% SIC regions. The root mean squared error (RMSE) is calculated using the RRDP validation points as the "truth", where a lower RMSE value indicates more minor deviations of the PMW SIC from the validation data. As expected for the CI region (Table 2), reSICCI3LF and SICCI3LF data have similar mean and RMSE values since the reSICCI3LF algorithm is designed to use the SICCI3LF SIC values over the ice pack. Additionally, we see that the N90LIN has overall the highest RMSE value and the

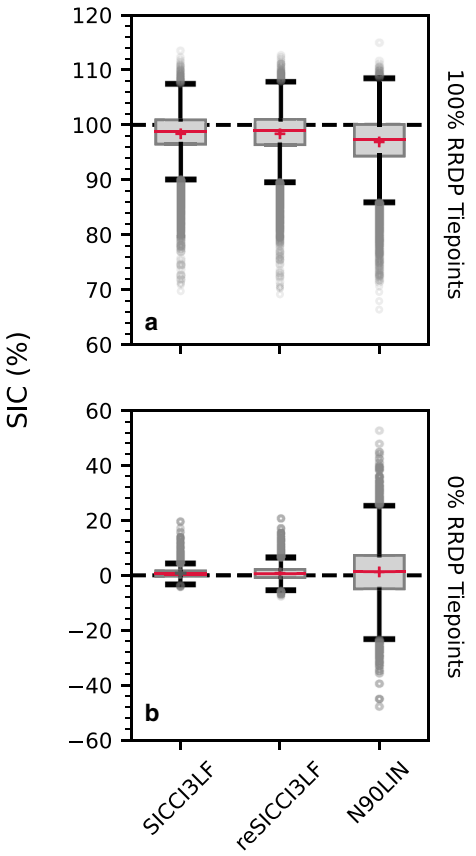


Figure 7. Validation results of algorithms (prior to open water filtering) compared to 100% (a) and 0% (b) sea ice concentration (SIC) validation points from the Round Robin Data Package Phase 2 dataset. The Red line represents the median, and the red cross represents the mean values of the distributions.

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

Table 2. Round Robin Data Package v2 100% sea ice concentration (SIC) validation results vs passive microwave SIC results derived from SICCI3LF, reSICCI3LF and N90LIN prior to the open water filtering being applied

	SICCI3LF		reSICCI3LF		N90LIN	
Period	Mean	RMSE	Mean	RMSE	Mean	RMSE
January	99.0	3.8	99.0	3.9	97.1	5.4
February	98.7	4.2	98.7	4.4	96.4	6.4
March	98.9	3.7	98.9	3.8	97.2	5.5
April	99.2	2.6	99.2	2.8	97.9	4.8
November	96.5	6.1	96.5	6.2	96.5	6.3
December	97.6	5.1	97.6	5.2	96.3	6.1
Overall	98.4	4.4	98.4	4.5	96.9	5.7

Table 3. Round Robin Data Package v2 0% sea ice concentration (SIC) validation results vs passive microwave SIC results derived from SICCI3LF, reSICCI3LF and N90LIN prior to the open water filtering being applied

	SICCI3LF		reSICCI3LF		N90LIN	
Period	Mean	RMSE	Mean	RMSE	Mean	RMSE
January	0.9	1.8	0.9	3.0	2.0	10.4
February	0.9	1.7	0.7	2.6	1.4	11.3
March	1.0	1.7	0.9	2.6	1.8	10.0
April	0.5	1.1	0.5	2.2	2.3	10.8
May	0.2	0.9	0.3	1.9	1.3	12.3
June	−0.1	1.5	0.0	2.1	−0.6	9.9
July	0.8	2.8	0.7	3.1	2.2	10.0
August	0.7	2.1	0.6	2.5	1.4	10.0
September	1.3	2.3	1.2	2.7	−0.2	9.2
October	1.3	1.9	1.5	3.1	1.3	10.1
November	1.0	1.8	1.0	2.6	0.9	9.8
December	1.0	1.8	1.0	2.7	2.1	9.8
Overall	0.8	1.9	0.8	2.6	1.2	10.3

lowest mean value indicating lower accuracy and precision than SICCI3LF and reSICCI3LF.

For the OW validation (Table 3), SICCI3LF and reSICCI3LF have the lowest overall means with values closest to 0 at 0.8%, followed then by N90LIN at 1.24%. The RMSE values for reSICCI3LF do increase for all months compared to SICCI3LF for the OW regions, indicating lower precision caused by incorrect SIC being added through the delta term in Eq. (1). Using purely N90LIN results in an overall RMSE 4.0 times larger compared to reSICCI3LF.

3.4 Validation against landsat imagery

The RRDP analysis can only assess the performance of the SIC algorithms at the 0% and 100% ends of the SIC range. We use L8 5 km analysis to validate the algorithms capabilities at these SIC values, but also across a range of intermediate values between 0% and 100%. This is particularly interesting due to the desire to improve the measuring capabilities in the marginal ice zone (MIZ) where these intermediate values are observed. However, the L8 validation will be constrained to clear sky conditions (no cloud cover) and coastal regions where L8 collects sea ice data. Consequently, the validation will only represent low atmospheric conditions for the PMW validation with a regional focus on the East Greenland coast (Fig. 3) since this is where most of the validation data is obtained. In total 34,345 match-up points with the L8 data across the study region have been used to produce the validation plots shown in Fig. 8.

We again study the pre-OWF SIC from the PMW for this validation. All three algorithms agree positively with the L8 validation, as seen with the line of best fit calculated using regression analysis. Additionally, the error bars, representing one standard deviation of the mean, overlap the 1:1 line, indicating a good correlation between the variables.

From Figs. 8a and 8b, we see that both the reSICCI3LF and the SICCI3LF overestimate SIC values between 5% and 35% and underestimate SIC values <80%. By applying the reSICCI3LF, we see that the overestimation present in the SICCI3LF is reduced with the mean and median values falling closer towards the 1:1 line with a smaller y-intercept for the line of best fit. Specifically, looking at Fig. 8a for these concentrations, we see that for these values, the distributions are positively skewed (mean > median) for the 5–10% and the 10–15% bins of the SICCI3LF algorithm whereas in Fig. 8b for reSICCI3LF the data is less positively skewed for these bins with mean values for bins 5–10%, 10–15%, 15–20%, 25–30% and 30–35% falling closer towards the 1:1 line. This indicates that the reSICCI3LF adjusts the concentrations at 5 to 35% SIC, which causes the larger y-intercept in the SICCI3LF data and produces a more accurate SIC at low values when compared to SICCI3LF. This SICCI3LF inaccuracy is likely due to the coarser algorithm producing higher SIC values in heterogeneous and ice transition zones, such as the ice edge and MIZ, or open water in consolidated ice regions, where the low-resolution algorithms will be more smeared.

The N90LIN results, Fig. 8c, show that the algorithm tends to underestimate the SIC across the full range of values. This underestimation increases as the SIC value increases and is most evident for SIC >30% SIC. This means that the N90LIN tends to produce lower SIC than the truth, which we also previously observed when validating against the RRDP data at 100% (Table 2). However, this negative bias present in N90LIN will not detrimentally impact the reSICCI3LF SIC since we only use the N90LIN delta at sharp SIC gradients (Fig. 2d), not its absolute value.

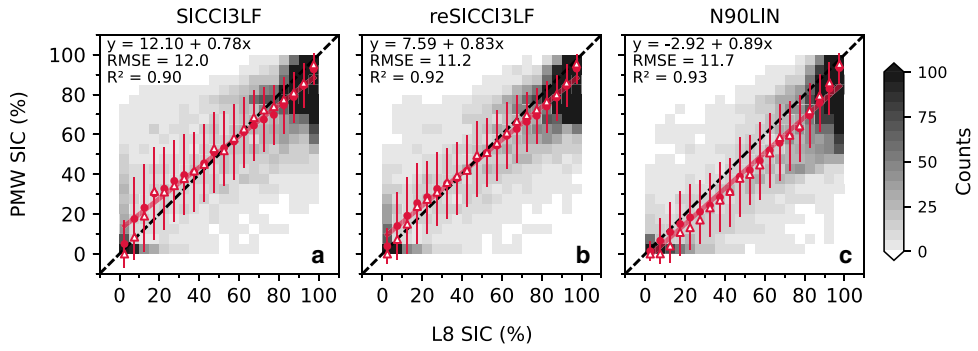


Figure 8. Validation of (a) SICCI3LF, (b) reSICCI3LF and (c) N90LIN against a 5 km SIC derived from Landsat-8 (L8) data obtained from 2013–2015. A total of 34,345 samples were used in the validation of each algorithm, with red circle markers representing the mean passive microwave (PMW) SIC binned at 5% Landsat SIC intervals (e.g. 0–5%, 5–10%, 10–15%), with the 2D histogram counts representing the number of PMW SIC samples binned at 5% intervals. The red outline triangles represent the median values, and the error bars represent one standard deviation around the mean.

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

Table 4. Sea ice concentration (SIC) results of the passive microwave (PMW) validated against a 5 km Landsat-8 (L8) SIC. PMW SIC results are derived from the SICCI3LF, reSICCI3LF and N90LIN algorithms before open water filters are applied. The differences (Diff.) and their subsequent statistics are derived by subtracting the L8 values from the PMW SIC.

Statistic	SICCI3LF	reSICCI3LF	N90LIN
Mean Diff.	−5.29	−5.27	−6.48
Median Diff.	−3.77	−3.60	−4.90
Std. Diff.	10.81	9.89	9.98
Gradient	0.78	0.83	0.89
Intercept	12.10	7.59	−2.92
R^2	0.90	0.92	0.93
RMSE	11.99	11.17	11.74

Overall Fig. 8 shows that SICCI3LF appears to have the weakest performance due to the overestimation at the low SIC, with both reSICCI3LF and N90LIN showing stronger positive correlations with the L8 data. The lower RMSE of the reSICCI3LF over the N90LIN would indicate that the reSICCI3LF SIC is better than the N90LIN in correctly retrieving the SIC. This is supported when looking at the statistics in Table 4. We see that reSICCI3LF also has the lowest mean, median, and standard deviation of the differences (PMW SIC - L8 SIC), further supporting that the reSICCI3LF is an improvement on using SICCI3LF and N90LIN individually.

4. Discussion and summary

This study presents the first application and evaluation of the reSICCI3LF algorithm on AMSR2 data to produce a 5 km SIC. Overall, the results show that reSICCI3LF can be applied to AMSR2 to produce a resolution-enhanced SIC with limited impact on the measurement uncertainty. We document that when applying the reSICCI3LF algorithm to AMSR2 data, the high-resolution SIC field (N90LIN) requires a Gaussian kernel using a sigma value of 5 km (Fig. 4) to suitably match the low-resolution SIC field (SICCI3LF). This ensures that only the high-resolution details from N90LIN are preserved when combining the SIC fields and not the large-scale features. However, this was not tested at different grid spacing, so the results may differ if applied to alternative SIC products and sensors. This is a crucial prior step to determine for reSICCI3LF implementation since the blurring of N90LIN SIC will dictate where the finer sea ice details from the N90LIN will be added into the SICCI3LF field.

Once the suitable sigma value was defined, we showed that the reSICCI3LF method successfully produced realistic SIC fields that are visually comparable to the high-resolution N90LIN field (Fig. 5). In the examples provided, the reSICCI3LF field has a higher spatial resolution than SICCI3LF, with the ice edges and ice openings appearing sharper and better defined, whereas, in the original SICCI3LF field, the ice edge appears smeared. Further spectral analysis confirmed the higher resolution capabilities of reSICCI3LF (Fig. 6).

After visually confirming that reSICCI3LF has improved spatial resolution capabilities, the SIC is validated against the RRDP and L8 SIC values to see how the uncertainties are affected. When comparing against the RRDP 0% and 100% SIC points, we find that the reSICCI3LF RMSE values are slightly larger than the low values of SICCI3LF. However, the reSICCI3LF algorithm has improved accuracy at low SIC values (5–30%) compared to SICCI3LF, attributed to its higher spatial resolution improving SIC measurement in heterogeneous sea ice environments such as the MIZ. As a result, by applying the reSICCI3LF algorithm to AMSR2 data, we can obtain the benefits of both the SICCI3LF and N90LIN algorithms to produce a higher-resolution SIC with low measurement uncertainty. The ability to better represent transitional ice zones, such as the ice edge, is crucial for producing accurate synoptic forecasts (Meier and Stewart, 2020) with higher resolution SIC shown to improve the accuracy of ice edge forecasting (Posey and others, 2015). The assimilation of this SIC into the Barents-2.5 km ocean/ice forecasting model and its impacts are presented in Durán Moro and others (2022).

Observable differences did occur between reSICCI3LF and N90LIN when comparing openings in the sea ice, with the N90LIN having sharper transitions to open water (Fig. 5f) and with more extensive open water areas being present in the ice openings (Fig. 5c). These differences are attributed to how the different frequency channels used in the SIC algorithms interact with different ice types and the atmosphere. For example, all SIC algorithms have been shown to underestimate SIC for sea ice thickness <35 cm, with 89 GHz-based algorithms shown to have the least sensitivity to thickness, attributed to its shorter wavelength (Ivanova and others, 2015). However, using 89 GHz is more prone to weather influences (Andersen and others, 2006), particularly for open water and near the ice edge (Ivanova and others, 2015), with atmospheric water vapour and wind-roughened water scattering causing erroneous sea ice over open water and both underestimation and overestimation within the ice pack (Lu and others, 2022). OWF are applied to SIC algorithms to filter pixels contaminated by strong atmospheric influences and set them to 0% SIC. Therefore, whilst these openings appear sharper and higher resolution in the N90LIN field, it does not mean that the values correctly represent the SIC conditions in those particular regions because open water filters will always remove true ice in these openings and at ice edges (Andersen and others, 2006; Ivanova and others, 2015).

As well as OWF, the tie point selection of the T_B's at 0% and 100% SIC is a crucial step required in the SIC algorithm application. However, the T_B's of the sea ice surface and open water will vary depending on specific physical properties. For example, wind-roughened open water, snow surface characteristics and ice thickness will produce tie points with a variable range of T_B values. Consequently, selecting the 0% and 100% tie points can cause a negative or positive SIC bias (Ivanova and others, 2015). These negative biases are present in all three SIC fields when comparing to the 100% RRDP validation points (Fig. 7), with N90LIN having the most notable bias for both the 100% and 0% regions. Therefore, the biases present are likely attributed to the tie point selection.

For this study, the tie points are dynamically selected for each day using the same method used by OSISAF (Lavergne and others, 2022b), which has been shown to reduce systematic SIC biases (Ivanova and others, 2015). Despite this, SIC biases are still evident compared to the RRDP validation points. This is likely caused by using tie points derived from the entire Arctic region rather than producing tie points derived from our specific regional application. This would mean that the 0% and 100% tie points may not fully represent the ice and water regimes in the focus region, consequently influencing the SIC values and impacting the biases. Therefore, in future work, it is recommended to use tie points derived from within the region of focus to mitigate this issue and produce improved SIC values, whilst paying attention to the specific challenges of using regional tie points (spatial discontinuities). In regards to reSICCI3LF specifically, the larger negative bias observed in N90LIN (Fig. 7a) should not impact the SIC values in the reSICCI3LF field. This is because the N90LIN bias does not impact the delta term in Eq. (1) since it is the difference value between N90LIN and N90LIN blurred. Overall, we have demonstrated that we can now produce a higher resolution SIC that better incorporates finer sea ice details without the

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

downside of the high measurement uncertainty that occurs when using a purely 89 GHz-based SIC field.

5. Outlook

Since the resolution-enhancing methodology presented in this paper is applied to SIC fields rather than brightness temperature, the technique offers flexibility for future development. SIC derived from different sensors (with different spatial resolution capabilities) can easily be introduced as the high-resolution field used to enhance a coarser SIC, for example, SAR. This would enable adding much finer sea ice details but will have less overlapping data for merging than SIC derived purely from PMW sensors. Additionally, the method presented in this paper uses two datasets to merge the SIC at different spatial scales. This could be further developed so that more than two datasets utilising a range of spatial scales could be applied to produce an enhanced SIC. For example, multiple SIC fields could be decomposed into different spatial scales (i.e. small, medium and large features). Then for each spatial scale, the SIC can be derived from the multiple SIC inputs, weighted using a coefficient to account for the confidence in the ability of that particular input to resolve a specific spatial scale. A similar concept was applied in Buehner and Shlyaeva (2015), where scale-dependent localisations produced an improved ensemble-based estimate for varying covariances.

A similar method would require additional SIC inputs, such as SIC from a 6 GHz-based algorithm. A 6 GHz algorithm such as the single channel 6H (Ivanova and others, 2015) or using the same methodology as N90LIN (Appendix), but with the 6 GHz channels, could potentially bring important information into a merged SIC. This is because microwave emissions at 6 GHz have a low sensitivity to water vapour and cloud liquid water over open water (Ivanova and others, 2015; Tonboe and others, 2021) as well as little emissivity difference between first-year ice and multiyear ice (Lee and others, 2017). A 6 GHz algorithm in this multi-scale multi-input merging could provide additional coarse information for large-scale features since its spatial resolution is low (Table 1). This 6 GHz SIC would be weighted with that from a 19/37 GHz algorithm and an 89 GHz algorithm to provide medium and fine-scale features respectively. The aim is that weighting each SIC input at multiple spatial scales has the potential to minimise the disadvantages of using a singular SIC input, such as the limitations of SICCI3LF and N90LIN previously discussed or the low-resolution and sensitivity to melt ponds and sea ice thickness of the 6 GHz channels (Ivanova and others, 2015; Kern and others, 2016). As mentioned already this multi-input method is also not limited to using purely PMW data, with SIC derived from different sensors also potentially being incorporated.

This multi-frequency methodology proposed above and presented in this paper for enhancing SIC is also relevant for the upcoming Copernicus imaging microwave radiometer (CIMR) mission proposed to launch at the earliest in 2028. Since CIMR is a multi-frequency mission that will measure 36.5 and 18.7 GHz at ~5 km spatial resolution, 10.65 and 6.9 GHz at ~15 km and 1.4 GHz at 55 km (Kilic and others, 2018), a similar methodology to that presented above can be adapted to exploit the high-resolution multi-frequency capabilities of CIMR, similar to the method presented in Kilic and others (2020). Additionally, the methodology presented in this paper, if applied to the entirety of the AMSR-E and AMSR2 dataset, would produce a nearly complete 5 km SIC time series from 2002 to the current date. This further work would complement the ~5 km spatial resolution CIMR SIC, greatly extending the period of future high-resolution SIC climate data records. Work is currently being undertaken to produce such a dataset for the Sustainable Development of the Arctic Ocean project (https://framsenteret.no/forskning/sudarco/) by applying the methodology presented in this paper to produce a regional 5 km SIC using AMSR-E and AMSR2 data. The expected publication of this data will be in 2024.

6. Conclusion

The research presented is a step closer to the current SIC goals set by GCOS to produce a SIC at 1–5 km resolution with a measurement uncertainty <5% (Zemp and others, 2022). To achieve this the paper outlined three research aims in section 1:

1. Adapt the reSICCI3LF algorithm to the AMSR2 sensor; we found that the reSICCI3LF can successfully be adapted to the AMSR2 sensor. For tuning the reSICCI3LF, we found that a sigma of 5 km was suitable to bring the high-resolution N90LIN SIC to a resolution comparable to the coarse-resolution SICCI3LF SIC (Fig. 4). This is a crucial step to resolve as it dictates where the high-resolution details are added into the reSICCI3LF SIC through the delta term in Eq. (1).
2. Assess the spatial resolution capabilities of the AMSR2 reSICCI3LF SIC; we found that the reSICCI3LF algorithm produces a SIC that contains sharper details such as at the ice edges and openings (Fig. 5) when compared to the coarse-resolution SICC3LF algorithm and has a closer spatial resolution to the high-resolution N90LIN SIC (Fig. 6).
3. Validate the AMSR2 reSICCI3LF SIC to comprehend how its application impacts measurement accuracy and precision; we demonstrated that the application of the reSICCI3LF algorithm does not deteriorate the SIC accuracy and precision with the reSICCI3LF producing comparable results to the coarse but low measurement uncertainty algorithm SICCI3LF (Fig. 7). Additionally, reSICCI3LF was the best performing algorithm showing the most agreement with a Landsat derived SIC (Fig. 8).

Acknowledgements. We acknowledge funding by the Research Council of Norway through project 302917 (SIRANO) and by the Norwegian FRAM Flagship program project SUDARCO (No. 2551323). We acknowledge the R&D contribution of ESA CCI Sea Ice and EUMETSAT OSI SAF projects to prepare the Level-2 sea-ice concentration data products used in this study.

References

Andersen S, Tonboe R, Kern S and Schyberg H (2006) 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 Environment* **104**(4), 374–392. doi: 10.1016/j.rse.2006.05.013

Bojinski S and 5 others (2014) The concept of essential climate variables in support of climate research, applications, and policy. *Bulletin of the American Meteorological Society* **95**(9), 1431–1443. doi: 10.1175/BAMS-D-13-00047.1

Browne PA, de Rosnay P, Zuo H, Bennett A and Dawson A (2019) Weakly coupled ocean–atmosphere data assimilation in the ECMWF NWP system. *Remote Sensing* **11**(3), 234. doi: 10.3390/rs11030234

Buehner M and Shlyaeva A (2015) Scale-dependent background-error covariance localisation. *Tellus A: Dynamic Meteorology and Oceanography* **67**(1), 28027. doi: 10.3402/tellusa.v67.28027

Cavalieri DJ, Gloersen P and Campbell WJ (1984) Determination of sea ice parameters with the NIMBUS 7 SMMR. *Journal of Geophysical Research: Atmospheres* **89**(D4), 5355–5369. doi: 10.1029/JD089iD04p05355

Denis B, Côté J and Laprise R (2002) Spectral decomposition of two-dimensional atmospheric fields on limited-area domains using the discrete cosine transform (DCT). *Monthly Weather Review* **130**(7), 1812–1829. doi: 10.1175/1520-0493(2002)130<1812:SDOTDA>2.0.CO;2

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

Duarte P and 10 others (2022) Implementation and evaluation of open boundary conditions for sea ice in a regional coupled ocean (ROMS) and sea ice (CICE) modeling system. *Geoscientific Model Development* **15**(11), 4373–4392. doi: 10.5194/gmd-15-4373-2022

Durán Moro M and 6 others (2023) Assimilation of satellite swaths vs daily means of sea ice concentration in a regional coupled ocean-sea ice model. *The Cryosphere Discussions* **2023**, 1–37.

Dworak R, Liu Y, Key J and Meier WN (2021) A blended sea ice concentration product from AMSR2 and VIIRS. *Remote Sensing* **13**(15), 2982. doi: 10.3390/rs13152982

Ivanova N and 10 others (2015) Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations. *The Cryosphere* **9**(5), 1797–1817. doi: 10.5194/tc-9-1797-2015

Iversen SC, Sperrevik AK and Goux O (2023) Improving sea surface temperature in a regional ocean model through refined sea surface temperature assimilation. *Ocean Science* **19**(3), 729–744. doi: 10.5194/os-19-729-2023

Karvonen J (2017) Baltic sea ice concentration estimation using sentinel-1 SAR and AMSR2 microwave radiometer data. *IEEE Transactions on Geoscience and Remote Sensing* **55**(5), 2871–2883. doi: 10.1109/TGRS.2017.2655567

Kern S (2021) Landsat surface type over water from supervised classification of surface broadband albedo estimates (Version 2021_fv0.01) [Dataset] (doi: 10.25592/uhhfdm.9181).

Kern S and 5 others (2016) The impact of melt ponds on summertime microwave brightness temperatures and sea-ice concentrations. *The Cryosphere* **10**(5), 2217–2239. doi: 10.5194/tc-10-2217-2016

Kern S and 6 others (2019) Satellite passive microwave sea-ice concentration data set intercomparison: closed ice and ship-based observations. *The Cryosphere* **13**(12), 3261–3307. doi: 10.5194/tc-13-3261-2019

Kern S and 6 others (2022) Satellite passive microwave sea–ice concentration data set intercomparison using landsat data. *The Cryosphere* **16**(1), 349–378. doi: 10.5194/tc-16-349-2022

Kilic L and 8 others (2018) Expected performances of the copernicus imaging microwave radiometer (CIMR) for an all-weather and high spatial resolution estimation of ocean and sea ice parameters. *Journal of Geophysical Research: Oceans* **123**(10), 7564–7580. doi: 10.1029/2018JC014408

Kilic L and 5 others (2020) Ice concentration retrieval from the analysis of microwaves: a new methodology designed for the copernicus imaging microwave radiometer. *Remote Sensing* **12**(7), 1060. doi: 10.3390/rs12071060

Lavergne T and 16 others (2019) Version 2 of the EUMETSAT OSI SAF and ESA CCI sea–ice concentration climate data records. *The Cryosphere* **13**(1), 49–78. doi: 10.5194/tc-13-49-2019

Lavergne T, Sørensen AM, Tonboe R and Pedersen LT (2021) CCI+ sea ice ECV sea ice concentration algorithm theoretical basis document. Technical report, European Space Agency, Available at: https://climate.esa.int/media/documents/SeaIce_CCI_P1_ATBD-SIC_D2.1_Issue_3.1_signed.pdf [Accessed: 22nd May 2023].

Lavergne T and 28 others (2022a) A new structure for the sea ice essential climate variables of the global climate observing system. *Bulletin of the American Meteorological Society* **103**(6), E1502–E1521. doi: 10.1175/BAMS-D-21-0227.1

Lavergne T and 8 others (2022b) OSI SAF: global sea ice concentration climate data records, algorithm theoretical basis document. Technical report, Ocean & Sea Ice Satellite Applications Facilities, Available at: https://osisaf-hl.met.no/sites/osisaf-hl/files/baseline_document/osisa f_cdop3_ss2_atbd_sea-ice-conc-climate-data-record_v3p0.pdf [Accessed: 22nd May 2023].

Lee SM, Sohn BJ and Kim SJ (2017) Differentiating between first-year and multiyear sea ice in the Arctic using microwave-retrieved ice emissivities. *Journal of Geophysical Research: Atmospheres* **122**(10), 5097–5112. doi: 10.1002/2016JD026275

Lu J, Heygster G and Spreen G (2018) Atmospheric correction of sea ice concentration retrieval for 89 ghz amsr-e observations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **11**(5), 1442–1457.

Lu J, Scarlat R, Heygster G and Spreen G (2022) Reducing weather influences on an 89 GHz sea ice concentration algorithm in the arctic using retrievals from an optimal estimation method. *Journal of Geophysical Research: Oceans* **127**(9), e2019JC015912. doi: 10.1029/2019JC015912

Ludwig V and 5 others (2019) The 2018 north greenland polynya observed by a newly introduced merged optical and passive microwave sea-ice concentration dataset. *The Cryosphere* **13**(7), 2051–2073. doi: 10.5194/tc-13-2051-2019

Ludwig V, Spreen G and Pedersen LT (2020) Evaluation of a new merged sea-ice concentration dataset at 1 km resolution from thermal infrared and passive microwave satellite data in the arctic. *Remote Sensing* **12**(19), 3183. doi: 10.3390/rs12193183

Meier WN and Stewart JS (2020) Assessment of the stability of passive microwave brightness temperatures for nasa team sea ice concentration retrievals. *Remote Sensing* **12**(14), 2197. doi: 10.3390/rs12142197

Meier WN and Stroeve J (2022) An updated assessment of the changing arctic sea ice cover. *Oceanography* **35**(3/4), 10–19.

Pedersen LT and 9 others (2021) Reference dataset for sea ice concentration [Dataset].

Posey PG and 10 others (2015) Improving arctic sea ice edge forecasts by assimilating high horizontal resolution sea ice concentration data into the US Navy's ice forecast systems. *The Cryosphere* **9**(4), 1735–1745. doi: 10.5194/tc-9-1735-2015

Ricard D, Lac C, Riette S, Legrand R and Mary A (2013) Kinetic energy spectra characteristics of two convection-permitting limited-area models AROME and Meso-NH. *Quarterly Journal of the Royal Meteorological Society* **139**(674), 1327–1341. doi: 10.1002/qj.2025

Röhrs J and 12 others (2023) Barents-2.5 km v2.0: an operational data-assimilative coupled ocean and sea ice ensemble prediction model for the barents sea and svalbard. *Geoscientific Model Development Discussions* **2023**, 1–31. doi: 10.5194/gmd-2023-20

Scott KA, Ashouri Z, Buehner M, Pogson L and Carrieres T (2015) Assimilation of ice and water observations from SAR imagery to improve estimates of sea ice concentration. *Tellus A: Dynamic Meteorology and Oceanography* **67**(1), 27218. doi: 10.3402/tellusa.v67.27218

Serreze MC, Crawford AD, Stroeve JC, Barrett AP and Woodgate RA (2016) Variability, trends, and predictability of seasonal sea ice retreat and advance in the chukchi sea. *Journal of Geophysical Research: Oceans* **121**(10), 7308–7325. doi: 10.1002/2016JC011977

Stroeve J and Notz D (2018) Changing state of arctic sea ice across all seasons. *Environmental Research Letters* **13**(10), 103001. doi: 10.1088/1748-9326/aade56

Stroeve JC, Markus T, Boisvert L, Miller J and Barrett A (2014) Changes in arctic melt season and implications for sea ice loss. *Geophysical Research Letters* **41**(4), 1216–1225. doi: 10.1002/2013GL058951

Tietsche S, Notz D, Jungclaus JH and Marotzke J (2013) Assimilation of sea-ice concentration in a global climate model–physical and statistical aspects. *Ocean Science* **9**(1), 19–36. doi: 10.5194/os-9-19-2013

Tonboe RT and 8 others (2016) The EUMETSAT sea ice concentration climate data record. *The Cryosphere* **10**(5), 2275–2290. doi: 10.5194/tc-10-2275-2016

Tonboe RT and 9 others (2022) Simulated geophysical noise in sea ice concentration estimates of open water and snow-covered sea ice. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **15**, 1309–1326. doi: 10.1109/JSTARS.2021.3134021

Wang L, Scott KA and Clausi DA (2016) Improved sea ice concentration estimation through fusing classified SAR imagery and AMSR-E data. *Canadian Journal of Remote Sensing* **42**(1), 41–52. doi: 10.1080/07038992.2016.1152547

Zemp M and 8 others (2022) The 2022 GCOS ECVs Requirements (GCOS 245). Technical Report 245, World Meteorological Organization, Available at: https://library.wmo.int/doc_num.php?explnum_id=11318 [Accessed: 22nd May 2023].

Zhang YF and 6 others (2021) Assimilation of satellite-retrieved sea ice concentration and prospects for september predictions of arctic sea ice. *Journal of Climate* **34**(6), 2107–2126. doi: 10.1175/JCLI-D-20-0469.1

Appendix A. SICCI3LF & N90LIN Algorithms

The SICCI3LF algorithm is the latest SIC algorithm used by the OSI SAF and ESA CCI climate data records. The algorithm is a hybrid approach optimised for both open water (OW) / low SIC conditions (B_{OW}) and consolidated ice pack conditions (B_{CI}) using the brightness temperature channels 19 V, 37 V and 37H from PMW missions such as SSMIS and AMSR2. The optimisation in (19 V, 37 V & 37H) T_b space for B_{OW} & B_{CI} follows the methodology outlined in Lavergne and others (2019). Once optimised for B_{OW} and B_{CI} the SICCI3LF algorithm is simply a linear weighted average between B_{OW} and B_{CI}. The weighting of the two algorithms is dependent on the OW weight

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press

(W_{OW}) and the SIC fraction, represented as values between 0.0 and 1.0 (Eq. (A.1)):

$$SIC_{SICCI3LF} = W_{OW}B_{OW} + (1 - W_{OW})B_{CI} \ .$$

Where,

$$W_{OW} = \begin{cases} B_{OW} < 0.7 & = 1 \\ B_{OW} > 0.9 & = 0 \\ B_{OW} \in [0.7, 0.9] & = \left(1 - \frac{B_{OW}-0.7}{0.9-0.7}\right) \ . \end{cases} \quad (A.1)$$

The 100% and 0% training data used to tune the algorithm are derived using the NASA Team (NT) SIC algorithm (Cavalieri and others, 1984) where the OW points are extracted in a 150 km wide strip 150 km from the ice edge (defined as 15% SIC in the NT algorithm) and the CI points are those for which SIC NT >95 %.

The near-90 GHz linear (N90LIN) algorithm is based on the same methodology as the SICCI3LF algorithm but is undertaken in the 2-dimensional data space using the T_b channels of 89 V and 89H. Since the algorithm is applied in 2-dimensional data space, the optimisation part of SICCI3LF and the merging of B_{CI} and W_{OW} does not occur. This is because the "optimal" T_b space will be the data projection through the OW and CI training data samples.

https://doi.org/10.1017/aog.2024.6 Published online by Cambridge University Press