

Article

SMAP Salinity Retrievals near the Sea-Ice Edge Using Multi-Channel AMSR2 Brightness Temperatures

Thomas Meissner * and Andrew Manaster

Remote Sensing Systems, 444 Tenth Street, Suite 200, Santa Rosa, CA 95401, USA; manaster@remss.com

* Correspondence: meissner@remss.com

Abstract: Sea-ice contamination in the antenna field of view constitutes a large error source in retrieving sea-surface salinity (SSS) with the spaceborne Soil Moisture Active Passive (SMAP) L-band radiometer. This is a major obstacle in the current NASA/Remote Sensing Systems (RSS) SMAP SSS retrieval algorithm in regards to obtaining accurate SSS measurements in the polar oceans. Our analysis finds a strong correlation between 8-day averaged SMAP L-band brightness temperature (TB) bias and TB measurements from the Advanced Microwave Scanning Radiometer (AMSR2) in the C-through Ka-band frequency range for sea-ice contaminated ocean scenes. We show how this correlation can be employed to develop: (1) a discriminant analysis that is able to reliably flag the SMAP observations for sea-ice contamination and (2) subsequently remove the sea-ice contamination from the SMAP observations, which results in significantly more accurate SMAP SSS retrievals near the sea-ice edge. We provide a case study that evaluates the performance of the proposed sea-ice flagging and correction algorithm. Our method is also able to detect drifting icebergs, which go often undetected in many available standard sea-ice products and thus result in spurious SMAP SSS retrievals.

Citation: Meissner, T.; Manaster, A. SMAP Salinity Retrievals Near the Sea-Ice Edge Using Multi-Channel AMSR2 Brightness Temperatures.

Remote Sens. **2021**, *13*, 5120.

<https://doi.org/10.3390/rs13245120>

Academic Editor:

Viviane V. Menezes

Received: 14 November 2021

Accepted: 14 December 2021

Published: 16 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: SMAP; AMSR2; sea surface salinity; sea-ice; cold water

1. Introduction

Three spaceborne L-band radiometers have been employed to measure sea-surface salinity (SSS) over the global ocean: (1) ESA's Soil Moisture and Ocean Salinity (SMOS) mission, which has been operating since 2009; (2) NASA's Aquarius mission, which operated between 2011 and 2015; and (3) NASA's Soil Moisture Active Passive (SMAP) mission, which has been operating since 2015. These instruments measure the electromagnetic passive microwave emission from the Earth's surface and atmosphere at a frequency of 1.41 GHz or wavelength of 21 cm. Retrieval algorithms for these sensors, which turn brightness temperature (TB) measurements into SSS, have been developed and successively refined [1]. Over open ocean scenes, it has been possible to achieve an average global accuracy of 0.2 psu or better for the retrieved SSS on monthly time scales and for 100-km spatial averages [2].

The retrieval accuracies of the L-band sensors strongly degrade in polar oceans. A major factor in this degradation is the decreased sensitivity of the L-band surface emission to SSS measurements at low sea-surface temperatures (SST). This results in a decreased signal-to-noise ratio in cold water (SST < 10 °C). To be specific, at colder SSTs, an error of +1 K in the TB measurement translates into an error of about −4 psu in the SSS retrieval. As such, improving SSS retrievals in or near polar regions has been identified as a major future science goal in ocean microwave remote sensing [3–6].

In this context, SSS retrievals close to the sea-ice edge have turned out to be particularly challenging. Even very small sea-ice fractions in the antenna field of view result in large spurious salinity biases. The degree of sea-ice contamination in the antenna field of

view can be characterized by the quantity g_{ice} , which is defined as the sea-ice concentration weighted by the antenna gain pattern [7]:

$$g_{ice} = \int dA G \cdot f_{ice} \quad (1)$$

here, G is the normalized antenna gain pattern projected on the Earth's surface. The dA denotes the differential surface area, and f_{ice} is the sea-ice concentration (SIC) at the Earth's surface. The integration in Equation (1) stretches over the whole Earth field of view. The measured emitted surface TB from the sea-ice contaminated Earth cell is given by:

$$T_{B,meas} = \int dA G \cdot [(1 - f_{ice}) \cdot T_{B,ocean} + f_{ice} \cdot T_{B,ice}] = \langle T_{B,ocean} \rangle + \Delta T_{B,c} \quad (2)$$

with

$$\Delta T_{B,c} \approx g_{ice} \cdot [\langle T_{B,ice} \rangle - \langle T_{B,ocean} \rangle] \quad (3)$$

in Equation (3), we have assumed constant average values for ocean and sea-ice TB within the field of view that contributes to the integral. The term $\Delta T_{B,c}$ in Equation (3) can be regarded as a measure of sea-ice contamination in the measured TB. The TB of typical thicker sea-ice and ocean scenes differ by about 125 K for vertical polarization (V-pol). For horizontal polarization (H-pol) the sea-ice–ocean TB difference is even larger. We realize that an antenna gain weighted sea-ice fraction g_{ice} of only 1% will result in a positive TB bias error of about +1.25 K, which, in turn, translates into a fresh bias error of about −5 psu in the SSS retrieval. At L-band frequencies, the sea-ice TB depends on ice-thickness, which can vary over location and season, in particular in the N hemisphere [8]. It introduces further variability into the problem, which can be an additional source of error. All of this demonstrates the challenge of measuring SSS if even small sea-ice contamination occurs within the field of view. We also note that sea-ice contamination can occur through the antenna sidelobes, and, thus, the value for g_{ice} in Equation (3) can be sizeable, even if the SIC f_{ice} at the Earth boresight or within the main beam is zero. For both the Aquarius and the SMAP sensors, a sizeable portion of the antenna gain pattern G falls onto the sidelobes.

This discussion demonstrates the necessity for (1) developing a reliable procedure to detect small sea-ice concentrations in the field of view of the L-band sensor antenna; and (2) developing a correction for it. The second task amounts to estimating the contamination $\Delta T_{B,c}$ in Equation (3) as accurately as possible and, in turn, removing it from the TB measurement in Equation (2). One major obstacle in this methodology is the fact that available ancillary SIC products are generally not accurate enough to be used for detecting and removing sea-ice contamination in the SMAP salinity observations. Most of these standard SIC products, for example [9] or [10], apply a weather filter in their retrievals to screen out weather systems (clouds, rain, high winds), which also removes most but not all of the observations whose SIC is below 15%. To avoid sea-ice contamination in the SSS retrievals, the current NASA/RSS Version 4 SMAP Salinity release [11] employs a conservative ice mask, which results in a notable loss of data in the polar oceans [12].

The purpose of our study is to develop, train, and test sea-ice flagging and sea-ice correction algorithms for the SMAP SSS retrievals that can be implemented in future version updates. Our methods will not use external ancillary SIC fields due to their aforementioned insufficient accuracy in regards to low SIC near the sea-ice edge. Instead, both the SMAP sea-ice flagging and the sea-ice correction algorithms will use TB measurements from the JAXA Advanced Microwave Scanning Radiometer (AMSR2), which takes measurements at frequencies between C-band and W-band and both vertical (V-pol) and horizontal (H-pol) polarizations. This idea was inspired by Figures 1 and 2, which compare the eight-day averaged TB biases of the SMAP vertical polarization (V-pol) channel and the AMSR2 C-band (6.93 GHz) near the Antarctic sea-ice edge for two different periods during winter and summer seasons. The TB biases are determined as the difference between the measured TB at the ocean surface and the TB value computed from a radiative

transfer model (RTM). We will give a more detailed account of the TB calculation in the next section. Visually, there is a strong correlation between these two TB biases near the ice edge. This suggests that there is the potential for the AMSR2 TBs to be directly integrated into the SMAP SSS retrieval algorithms for sea-ice flagging and correction. This is further supported by two key facts:

- (1) The spatial resolution of the SMAP footprint (≈ 40 km) and the lowest frequency AMSR2 frequency channel at C-band (≈ 50 km) are close. When multi-day aggregates are considered, in which several footprints are averaged together, then both sensors, SMAP and AMSR2, see approximately the same Earth scenes. Observations at the higher AMSR2 frequencies can be spatially resampled to the coarser C-band resolution.
- (2) Compared to the SMAP L-band channels, the AMSR2 channels are insensitive or only very weakly sensitive to SSS. The AMSR2 ocean measurements are affected by sea-ice contamination but not by signals such as freshwater river plumes or melt run-off. This enables us to separate the spurious sea-ice signal from the real freshwater signals.

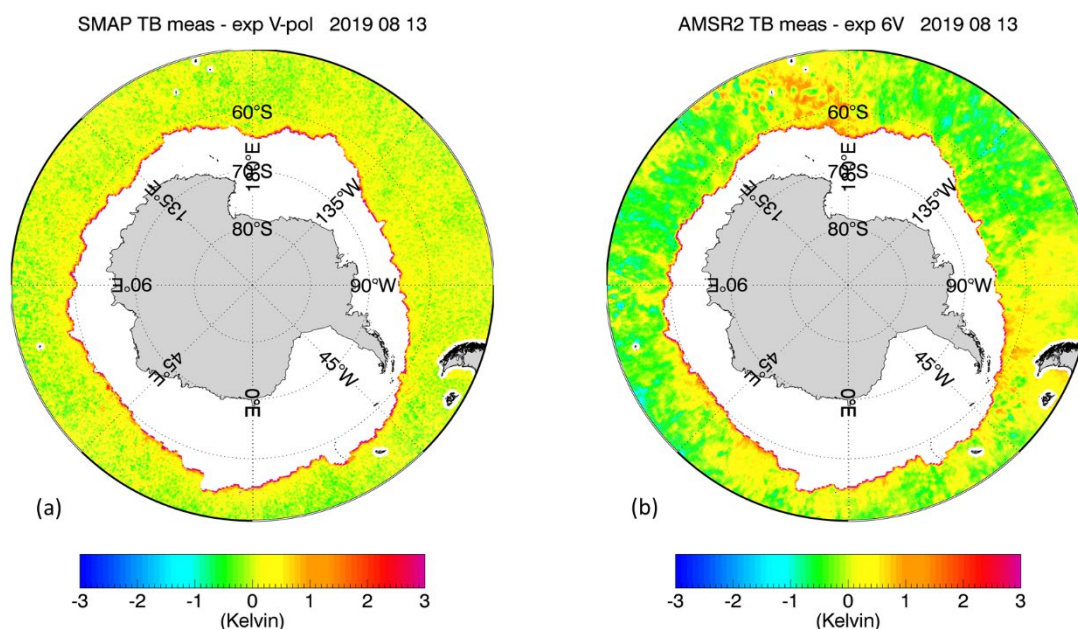


Figure 1. 8-day average TB biases near the Antarctic sea-ice edge for a scene during the Austral winter (13 August 2019) (a) SMAP V-pol; (b) AMSR2 6.93 GHz V-pol channel. The TB bias is defined as the difference between measured and expected (computed) specular surface TB. The sea-ice contamination results in a positive TB bias for both SMAP and AMSR2 near the ice edge, which results in a fresh (negative) bias in the retrieved SSS.

The proposed sea-ice flagging algorithm is designed to detect sea-ice concentrations g_{ice} in the SMAP 40-km footprint of about 1.5% or more. This is a much lower detection threshold than the sea-ice edge, which is commonly defined as the location where the SIC is 15%. The low detection threshold is mandated by the aforementioned high sensitivity of the SSS retrieval to sea-ice contamination. On the other hand, it is also essential to keep the false alarm rate as low as possible. False alarms will either result in unnecessary data loss or overcorrection. It is quite evident in Figures 1 and 2 that the good correlation between SMAP and AMSR2 TB biases really only applies near the sea-ice edge and breaks down over the open ocean. Thus, applying a correction to observations that are not contaminated by sea-ice would introduce an unnecessary error in the SMAP SSS retrievals. The sea-ice detection algorithm is based on a discriminant analysis technique adapted from pattern recognition theory, which is designed to discriminate two different classes of data: uncontaminated open ocean scenes and sea-ice contaminated scenes. The discriminant analysis flag is then followed by additionally flagging the nearest neighbor (NN) to

a given observation, which will extend the area of flagged observations. The NN flagging is necessary to catch SMAP observations that contain sea-ice contamination in the antenna sidelobes. The sea-ice correction algorithm is a statistical regression designed to recover as many of the flagged SMAP SSS observations as possible. Both, the sea-ice flagging and the sea-ice correction algorithm input AMSR2 TB together with several external ancillary geophysical fields. The use of these ancillary fields will increase the performance of the algorithm.

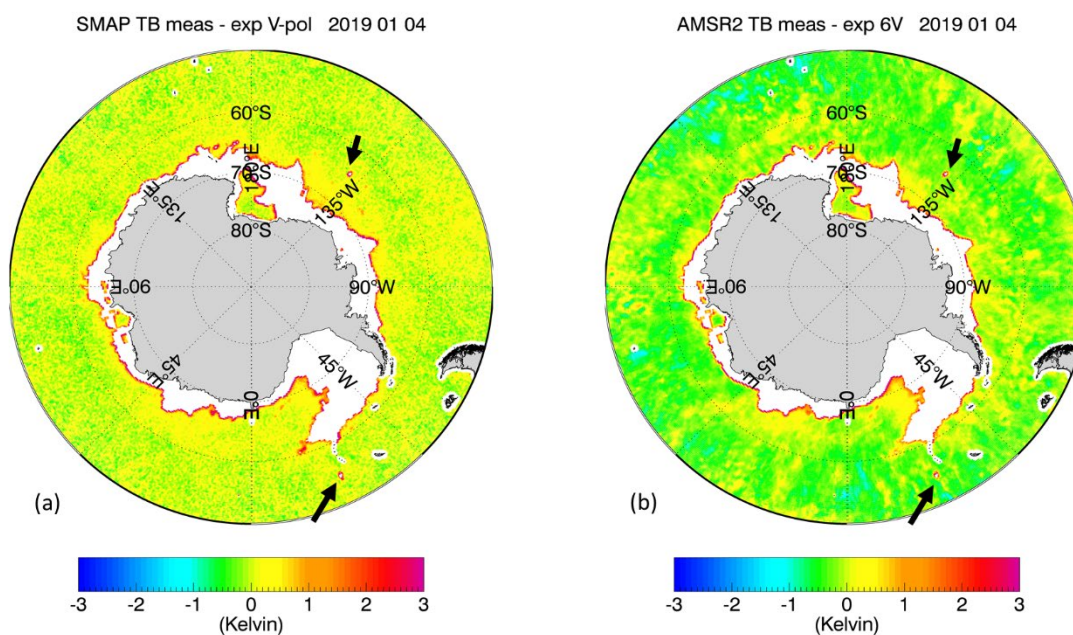


Figure 2. 8-day average TB biases near the Antarctic sea-ice edge for a scene during the Austral summer (4 January 2019) (a) SMAP V-pol; (b) AMSR2 6.93 GHz V-pol channel. The TB bias is defined as the difference between measured and expected (computed) specular surface TB. The sea-ice contamination results in a positive TB bias for both SMAP and AMSR2 near the ice edge, which results in a fresh (negative) bias in the retrieved SSS. The black arrows indicate possible candidates of very large floating icebergs (Section 3.5.).

The vast majority of the scientific data users of the NASA/RSS SMAP SSS data utilize the eight-day or monthly time-averaged product from the Level 3 files. The Level 2 data are too noisy for most science applications. We have therefore decided to optimize our sea-ice flagging and sea-ice correction algorithms to perform best for the eight-day time-averaged SSS products. Thus, the analysis presented herein will focus on eight-day aggregate SMAP and AMSR2 TB observations.

In our study, we will present results for the SMAP observations at the 40-km sensor resolution. The standard NASA/RSS SMAP SSS product has been smoothed to a spatial resolution of 70 km, which results in a noise reduction over open ocean scenes that makes the product more useful for most scientific applications [13]. In regards to the evaluation of biases due to sea-ice contamination and the bias mitigation after the sea-ice correction, there is, unfortunately, little error reduction gained when spatially smoothing the product from 40 km to 70 km. Thus, it is appropriate to perform the analysis at the 40-km spatial scale.

2. Materials and Methods

2.1. Study Data Sets

2.1.1. SMAP Brightness Temperatures

The source for the SMAP TB are the L2C files from the NASA/RSS SMAP SSS V4.0 Release [11], which have been sampled onto a 0.25° fixed Earth grid. For our study, we use the TB at the specular surface, labeled $T_{B0}(\text{SMAP})$. It is calculated from the SMAP antenna temperature (TA) measurement after removing (1) antenna effects (antenna emissivity and cross-polarization contamination); (2) celestial radiation coming from the cosmic microwave background, galaxy, sun, and moon; (3) Faraday rotation in the Earth's ionosphere; (4) atmospheric attenuation by oxygen, water vapor, and clouds; and (5) the wind induced rough surface emission [13–15]. We screen the observations for degrading conditions and discard data containing land contamination (land fraction > 1%), sun glint, large galactic radiation, rain, and very high wind speeds (>15 m/s). The L2C maps are then time-averaged into 8-day aggregate 0.25° SMAP T_{B0} maps.

2.1.2. AMSR2 Brightness Temperatures

The source for the AMSR2 TB are the L1A AMSR2 TA measurements. Our study uses observations from the 6.93, 10.65, 18.7, 23.8, and 36.5 GHz V-pol and H-pol AMSR2 channels. We do not use the 2nd C-band channels (7.3 GHz) in our study as they contain no additional information to the 6.93 GHz channels in regard to developing our sea-ice flagging and correction algorithms. We also do not use the 89 GHz channels, as they are heavily affected by clouds and rain, which could introduce errors in our analysis. The observations of all channels are spatially resampled onto the largest footprint, which is one of the C-band (6.93 GHz) channels (62 km × 35 km) [16]. The spatial resampling is based on the Backus-Gilbert optimum interpolation technique [17].

The AMSR2 TA are first transformed into the top of the atmosphere (TOA) TB, labeled $T_{B,TOA}(\text{AMSR2})$, using the calibration procedure outlined in [18]. We employ the RTM to calculate the atmospheric attenuation [19] and the wind induced rough surface emission [20] for the AMSR2 channels to compute the specular surface TB for each channel, labeled $T_{B0}(\text{AMSR2})$. In the calculation of the atmospheric attenuation, we use the simplified bulk expressions derived in [21] that depend on total columnar water vapor and total columnar cloud liquid water. The ancillary input fields to the RTM calculation will be discussed in Section 2.1.4. The specular surface emissivities $E_0(\text{AMSR2})$ of the AMSR2 channels are determined by:

$$T_{B0}(\text{AMSR2}) = E_0(\text{AMSR2}) \cdot T_s \quad (4)$$

where T_s is the SST in Kelvin.

The AMSR2 observations are screened for degraded conditions by removing measurements that are contaminated by land, sun-glint, known sources of radio frequency interference (RFI), and rain. When removing land contaminated measurements, we require that the land fraction within the AMSR2 6.93 GHz footprint is less than 1%.

Finally, the resampled AMSR2 TB are averaged in space and time into 8-day aggregate 0.25° maps, which comprise the same time periods as the SMAP TB map. When gridding the AMSR2 swath observations into regular 0.25° maps, we use a simple straight average (drop in the bucket). We will create and use separate sets of maps for the TOA $T_{B,TOA}(\text{AMSR2})$, and the specular surface $E_0(\text{AMSR2})$.

We note that the value of 1% for the land fraction threshold is the maximal land fraction that is allowed within a single AMSR2 C-band footprint. In each 0.25° cell several of these footprints are averaged together, so the actual land contamination is in general considerably less than 1% in the 0.25° cells, even at locations closest to the coast or near islands. The land fraction does not change over time, so by careful inspection of the cases used in this study we can check that land contamination is not a major error source in our

analysis. We refer to the Supplementary Material for maps that zoom into coastal areas and small islands.

2.1.3. Expected Specular Surface Brightness Temperatures and Emissivities

The expected specular surface TB, labeled $T_{B0,exp}$, is the result of the RTM calculation of the expected specular surface emissivity $E_{0,exp}$ multiplied by T_s . The expected surface emissivity $E_{0,exp}$ is obtained from the Meissner–Wentz model of the dielectric constant (permittivity) of sea-water [15,20,22]. Our study calculates and uses expected $T_{B0,exp}$ for SMAP, and expected $E_{0,exp}$ for AMSR2.

The SMAP $T_{B0,exp}$ are the basic metric for training and testing our proposed SMAP sea-ice flagging and sea-ice correction algorithms. At L-band frequencies, both $T_{B0,exp}$ and $E_{0,exp}$ depend on SST and SSS and, thus, their calculation requires an ancillary reference SSS field (see Section 2.1.4.). The SMAP $T_{B0,exp}$ are not an input for running the sea-ice flagging and sea-ice correction algorithms, but we use them only for developing the algorithms and evaluating their performances. Assuming that sea-ice contamination is the only uncertainty source in the SMAP SSS retrievals, then the difference between measured and expected SMAP specular surface TB is a proxy for the sea-ice contamination term $\Delta T_{B,c}$ in Equations (1) + (2):

$$\Delta T_{B,c} = \Delta T_{B,0}(\text{SMAP}) = T_{B0,meas}(\text{SMAP}) - T_{B0,exp}(\text{SMAP}) \quad (5)$$

we have performed our analysis for both SMAP V-pol and H-pol measured–expected TB. All of the results turned out to be very close for both polarizations. Therefore, we will only present results for the SMAP V-pol channel for the remainder of the manuscript.

The AMSR2 $E_{0,exp}$ will serve as actual input to the sea-ice flagging and sea-ice correction algorithms. In the C- through Ka-band frequency range of AMSR2, the $E_{0,exp}$ do not depend or only depend very weakly on SSS. Thus, no external SSS input is required for their computation, and we can set the SSS in the dielectric constant model to a typical value of 35 psu without penalty. That is a crucial point for our proposed method.

As we did for the measured SMAP and AMSR2 TB, we also create time averaged 8-day aggregate 0.25° maps for the expected SMAP $T_{B0,exp}$ and the AMSR2 $E_{0,exp}$.

2.1.4. External Ancillary Data

As noted earlier, our study requires various external ancillary fields. Specifically, we use the:

- (1) HYCOM SSS field [23] as the reference SSS. We are using the output from the operational ocean analysis that is run by the U.S. Navy [24]. In our study, this reference SSS is used in the computation of the SMAP $T_{B0,exp}$, which serve as the basic metric for algorithm development and evaluation. The reference SSS field is not an input to the proposed sea-ice flagging or sea-ice correction algorithms or to the NASA/RSS SMAP SSS retrieval algorithm.
- (2) SST from the Canadian Meteorological Center (CMC) [25] as ancillary SST in the RTM computations for both SMAP and AMSR2.
- (3) Cross Calibrated Multi-Platform (CCMP) V2.0 wind speeds and directions [26–28] for the calculation of the wind induced rough surface emission for both SMAP and AMSR2. The CCMP winds are created by a variational analysis method that assimilates various microwave satellite wind measurements together with a background field from a numerical weather prediction model.
- (4) Atmospheric profiles for pressure, temperature, relative humidity, and cloud water density as well as total columnar water vapor and total columnar cloud water from the General Forecast System (GFS) of the National Centers for Environmental Prediction (NCEP) [29]. These fields are used for computing the atmospheric attenuation for both SMAP and AMSR2.

- (5) Precipitation rates from the Integrated Multi-satellitE Retrievals for the Global Precipitation Mission (IMERG) Version 6 [30] for rain-flagging of both SMAP and AMSR2 observations.

We note that these ancillary fields are the same that are already being used in the V4.0 NASA/RSS SMAP salinity retrieval algorithm for the purpose of turning measured SMAP TOA TB into surface emissivity. Extensive analysis has been conducted and shown that they are well suited for that purpose [14,31].

2.1.5. Training and Test Sets

Our study is based on several training and test sets of SMAP and AMSR2 data, which are summarized in Table 1. For each of the two different years, 2018 and 2019, we have chosen one set during each of the 4 seasons (winter, spring, summer, fall) to accommodate a wide range of possible geophysical conditions near the ocean-sea-ice interface. Sets 1–4 from the year 2018 are used for training, and Sets 5–8 from the year 2019 are used for testing. That guarantees that training and testing data are independent. We use Data Set 9 from early 2021 for testing our algorithms for observations near floating icebergs.

Table 1. Date and orbit ranges for the 8-day SMAP and AMSR2 TB data sets that are used for training and testing the sea-ice flagging and sea-ice correction algorithms.

Set	Date Range	SMAP Orbit Range	AMSR2 Orbit Range	Pole	Used for
1	21–28 January 2018	15,875–15,991	30,209–30,325	N	training discriminant flagging
				S	training regressions for sea-ice correction
2	14–21 April 2018	17,089–17,205	31,418–31,533	N	training discriminant flagging
				S	training regressions for sea-ice correction
3	16–23 July 2018	18,448–18,565	32,772–32,888	N	training discriminant flagging
				S	training regressions for sea-ice correction
4	9–16 October 2018	19,692–19,808	34,010–34,125	N	training discriminant flagging
				S	training regressions for sea-ice correction
5	1–08 January 2019	20,920–21,037	35,233–35,349	S	testing discriminant flagging testing regressions for sea-ice correction
6	24–31 March 2019	22,120–22,236	36,427–36,543	S	
7	10–17 August 2019	24,153–24,269	38,452–38,567	S	
8	21–28 September 2019	24,767–24,883	39,063–39,179	S	
8	21–28 September 2019	24,767–24,883	39,063–39,179	N	Arctic example (Section 3.4.)
				S	iceberg A-68 (Section 3.5.)
9	27 February–6 March 2021	32,445–32,561	46,708–46,824	S	

As discussed in Section 2.1.4., we use the HYCOM SSS from the U.S. Navy operational ocean analysis as a reference field. The reference SSS is generally more reliable near the Antarctic sea-ice edge than near the Arctic sea-ice edge. In the Southern polar oceans, there exist reliable in-situ measurements from ARGO drifters, which serve as an important input to the operational ocean analysis data assimilation, whereas in the Northern polar oceans those in-situ measurements are sparse. Moreover, there are many freshwater river outflows into the Northern polar oceans, particularly around Northern Russia. In those instances, the reference SSS is often a poor representation of the true SSS field, which can result in a significant mismatch between measured and expected TB that is not caused by any sea-ice contamination. Last but not least, the L-band TB variability due to varying sea-ice thickness is much less prevalent in the Antarctic oceans than in the Arctic oceans, where large areas of thin sea-ice form near the sea-ice edge during fall. To avoid tainting the performance statistics, we will use only scenes in the Southern hemisphere for algorithm testing. We will show results for a case in the Arctic Ocean (Section 3.4.) to

demonstrate that the algorithm works there as well, but the observations from this case will not be included in the performance statistics.

The training of the algorithms turns out to be very robust and using scenes from the Northern hemisphere is not harmful. For training the sea-ice detection and flagging algorithm, we have found that it is actually slightly beneficial to use Arctic scenes, but only if the observations in known areas of freshwater river outflows are discarded.

2.1.6. A-Priori Sea-Ice Flagging

We impose a climatological monthly sea-ice mask as a-priori condition for the sea-ice detection and correction algorithms. This climatological sea-ice mask was compiled from many years of historical SIC data before the launch of SMAP and AMSR2. Its sole purpose is to provide a simple indicator for locations in the ocean where sea-ice can form. The mask also includes areas where drifting icebergs have been detected. Figure 3 shows the extent of the sea-ice mask for the example scenes from Figures 1 and 2. In addition, we also require the value of the ancillary SST to be lower than 10 °C. If either one of these two a-priori conditions are not met, the sea-ice detection algorithm is not run, the observation is not flagged as sea-ice, and no sea-ice correction is performed. The reason for applying these a-priori constraints is to avoid over-flagging and over-correcting observations that are very unlikely contaminated by sea-ice.

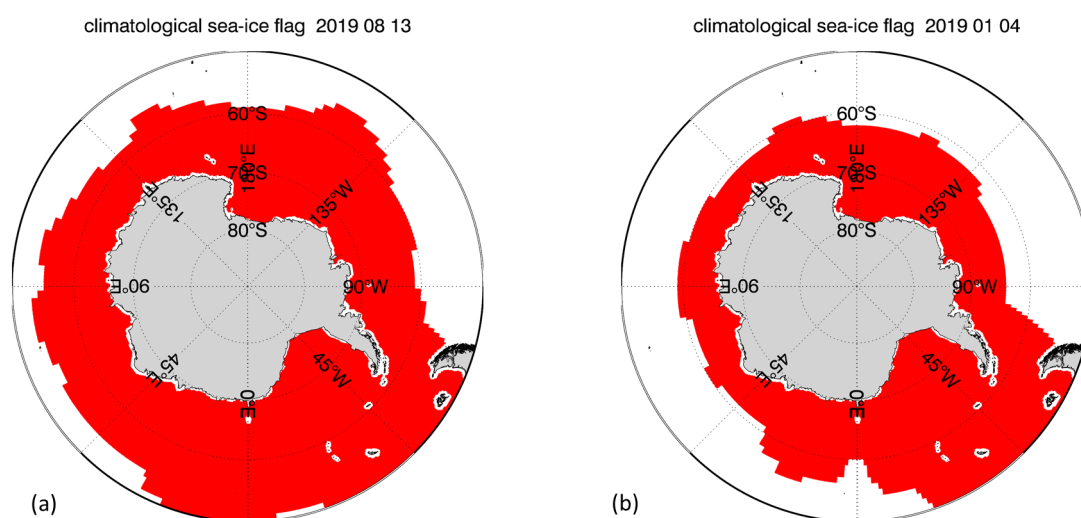


Figure 3. Climatological sea-ice masks near Antarctica for two sample days: (a) 13 August 2019 (Set 7-S); (b) 4 January 2019 (Set 5-S).

2.2. Sea-Ice Detection and Flagging

2.2.1. Discriminant Analysis

The first and biggest step of the sea-ice detection and flagging algorithm employs the Fisher Discriminant Analysis, which is a well-established method in pattern classification for distinguishing two different classes of data points in a higher-dimensional space. Appendix A gives a summary and outline of this methodology.

When training the discriminant analysis, we draw data from two classes both of which are subsets of the 4 training sets listed in Table 1. These classes are defined as:

$$\begin{aligned} \Delta T_{B0}(\text{SMAP})_{V\text{-pol}} < e_1 &\Rightarrow \text{Class 1} \\ e_2 < \Delta T_{B0}(\text{SMAP})_{V\text{-pol}} < e_3 &\Rightarrow \text{Class 2} \end{aligned} \quad (6)$$

we have threshold values of $e_1 = 0.4$ K, $e_2 = 2.0$ K, $e_3 = 4.5$ K. The RMS difference for 8-day aggregate $\Delta T_{B0}(\text{SMAP})_{V\text{-pol}}$ over open ocean scenes is found to be about 0.2 K, so any observation that falls within Class 2 is very likely contaminated by sea-ice. The sea-ice detection threshold e_2 of 2.0 K corresponds to an approximate sea-ice concentration g_{ice} of 1.5%. If the discriminant analysis determines that an observation falls into this class, then it is flagged as sea-ice. Observations that fall within Class 1 might still contain some sea-ice contamination, and we address them in the second step of the sea-ice flagging algorithm. It is necessary to leave a large enough separation between the upper threshold e_1 of Class 1 and the lower threshold e_2 of Class 2. If the values for these thresholds are too close together, the discriminant analysis is less effectively able to distinguish between the two classes, which would result in an increased number of missed detections, false alarms, or both. The reason for imposing an upper threshold e_3 in the definition of Class 2 will be discussed shortly.

The inputs to the discriminant analysis are the measurements \mathbf{X} of the 10 AMSR2 TB channels, each of which represents a point in a 10-dimensional data space. We consider two cases for the AMSR2 data points \mathbf{X} : Case 1 uses the TOA AMSR2 TB. Case 2 uses the measured–expected AMSR2 surface emissivities:

$$\begin{aligned} \text{Case 1: } \mathbf{X}_k &= T_{B,\text{TOA}}(\text{AMSR2})_k \\ \text{Case 2: } \mathbf{X}_k &= \left[E_{0,\text{meas}}(\text{AMSR2}) - E_{0,\text{exp}}(\text{AMSR2}) \right]_k \cdot T_{\text{eff}} \end{aligned} \quad (7)$$

where the typical SST of $T_{\text{eff}} = 273.15$ K serves to convert the dimensionless emissivity values into quantities with dimension K. The index k in Equation (6) runs over the 10 AMSR2 channels.

The Fisher Linear Discriminant Analysis determines an optimal direction \mathbf{W} within this 10-dimensional space, which results in the best separation between the two classes after the data are projected on it. The elements of the projection vector \mathbf{W} can be regarded as optimal weighting coefficients of the 10 AMSR2 channels. Table 2 lists the results for the components of \mathbf{W} that were determined from the training Sets 1–4 from Table 1 for both Case 1 and Case 2. We notice very low weights in the high-frequency channels, 23.8 and 36.5 GHz for Case 2, in which the atmospheric attenuation is explicitly removed. These higher frequency channels are more sensitive to the atmosphere and thus more impacted by uncertainties in the ancillary fields which are used for removing the atmospheric effects, in particular cloud liquid water and water vapor. The emissivity difference of the lower frequency AMSR2 channels, in particular C-band, correlates better with the observed SMAP ΔT_{B0} . The discriminant is given by the projection of the data vector \mathbf{X} onto the projection direction \mathbf{W} , which is the scalar product $\mathbf{W}^T \cdot \mathbf{X}$ of the two vectors. For the determination of the criterion to distinguish between the two classes, we plot the probability density functions (PDFs) or histograms of the discriminant $\mathbf{W}^T \cdot \mathbf{X}$ for each class (Figure 4). The intersection point of the two PDFs defines the value d of the decision boundary and the decision criterion (A10) in Appendix A. The values for d for the two cases are also listed in Table 2. The values of \mathbf{W} and d are determined from the training data sets and these values are subsequently used in the test runs.

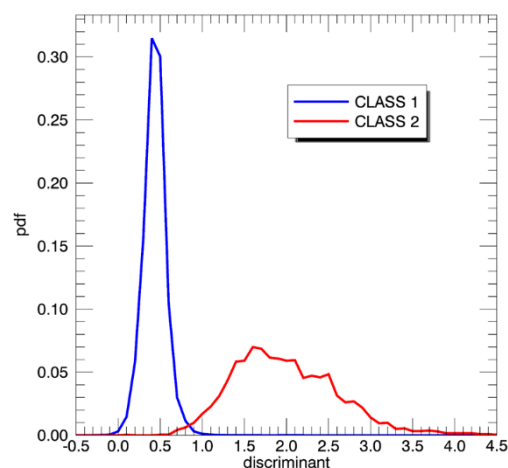


Figure 4. Probability density functions (pdf) of the two classes defined in Equation (5). The two pdf curves were computed for the training data sets (Table 1) and refer to case 2, which uses the AMSR2 measured minus expected specular surface emissivities as input. The x -axis values are the discriminant $\mathbf{W}^T \cdot \mathbf{X}$. The decision boundary d is at the intersection between the two pdf, which in this example is $d = 0.85$.

Table 2. Values of the components of the Fisher Linear Discriminant Analysis projection vector \mathbf{W} (Appendix A) within the 10-dimensional space spanned by the AMSR2 TB channels in the training data set (Table 1). The length of the vector \mathbf{W} has been normalized to 1. Case 1 refers to the algorithm using the AMSR2 TOA as input. Case 2 refers to the algorithm using the AMSR2 measured minus expected specular surface emissivities as input. The last column contains the values for the decision boundary d for each case.

Channel i	6.93 V	6.93 H	10.65 V	10.65 H	18.7 V	18.7 H	23.8 V	23.8 H	36.5 V	36.5 H	d
\mathbf{W}_i (case 1)	0.140082	-0.46514	0.254423	-0.08172	-0.62169	0.486014	0.168304	-0.12771	-0.15391	0.03985	52.05
\mathbf{W}_i (case 2)	0.01366	-0.50493	0.43747	-0.10526	-0.70372	0.20662	-0.00025	0.06365	-0.00406	0.02058	0.85

Including an upper threshold value e_3 in the condition for Class 2 in Equation (5) might look surprising at first. The reason for including this upper threshold is merely technical. For the numerical calculation of the optimal projection vector \mathbf{W} from Equation (A9), it is desirable that the scatter matrices \mathbf{S}_i in (A7) for both classes $i = 1, 2$ have the same order of magnitude. This can be achieved by introducing an upper threshold in the definition of Class 2, as we have done. One strength of the Fisher Linear Discriminant Analysis is that it is robust when adding additional data points to Class 2. The detection algorithm has no difficulty placing any event for which $\Delta T_{B0}(\text{SMAP})$ exceeds the value e_3 into Class 2, thus classifying it correctly as a sea-ice contaminated observation.

Table 3 lists the basic characteristics of this first step in the sea-ice detection algorithm, i.e., the missed detection and the false alarm rate. When calculating these values, we have only included observations that fall within the a-priori sea-ice mask (Section 2.1.6.). That means that a rate of $x\%$ refers to $x\%$ from all observations that have been a-priori masked. We see that Case 2, which inputs AMSR2 specular surface emissivity, has better skill than Case 1, which inputs AMSR2 TB TOA, particularly in regard to false alarms. The calculation of the AMSR2 specular emissivity from the TOA TB requires external ancillary input fields for atmosphere, wind speed, wind direction, and SST followed by an RTM calculation of the atmospheric attenuation and the wind induced rough surface emission, which are then both removed from the TOA TB. The ancillary fields contain valuable information and removing the atmospheric and roughness effects helps to improve the skill of the detection algorithm. The benefit of removing atmospheric and wind roughness effects on sea-ice detection and SIC retrievals has been demonstrated in several studies [32–34].

Table 3. Missed detection and false alarm rate for the discriminant sea-ice flag in the test data set (Table 1). Only observations within the a-priori sea-ice flagging (Section 2.1.6.) are considered for calculating the rates. Case 1 refers to the algorithm using the AMSR2 TOA as input. Case 2 refers to the algorithm using the AMSR2 measured minus expected specular surface emissivities as input. A missed detection is defined as an observation, which the detection algorithm places into Class 1 but for which $\Delta T_{B,0}(\text{SMAP}) > e_2 = 2.0$ K. A false alarm is defined as an observation, which the detection algorithm places into Class 2 but for which $\Delta T_{B,0}(\text{SMAP}) < e_1 = 0.4$ K.

	Case 1	Case 2
missed detection rate	0.15%	0.12%
false alarm rate	0.39%	0.06%

2.2.2. Nearest Neighbor Flagging

The discriminant analysis used in the 1st step of the sea-ice detection algorithm is designed to safely detect sea-ice contaminations of 1.5% or more as part of Class 2. However, observations that the discriminant analysis does not put into Class 2 can still be contaminated by sea-ice and thus might need to be flagged as well. Lowering the threshold value e_2 in the class definition (5) would achieve that, but it would also result in an increased number of false alarms, which is not desirable. In order to catch the majority of these low SIC cases, we add a NN flagging as a 2nd step of the sea-ice detection algorithm. The basic assumption is that most of the sea-ice contaminated observations that went undetected in the 1st step are in close spatial proximity to at least one of the events that have already been flagged. Our sea-ice detection flags both the nearest neighbors and the next to nearest neighbors of any cell that the discriminant analysis has put into Class 2. This amounts to flagging the 24 surrounding cells of the 0.25° grid together with any cell from Class 2.

2.3. Sea-Ice Zones

The combined discriminant and NN sea-ice flagging scheme allow us to define sea-ice zones, which classify the severity of sea-ice contamination for each observation. This will become helpful for developing and evaluating the sea-ice correction algorithm. Figure 5 shows the schematic flow and decision tree for sea-ice zone classification. An observation is placed into Zone 3 if it is flagged by the discriminant detection algorithm and at least one of its NN is not flagged. An observation that is flagged by the discriminant detection algorithm and at least one of its NN has been classified to be within Zone 3 is placed into Zone 4. The same decision is repeated for Zone 5 classification. An observation that has not been flagged by the discriminant algorithm but has at least one flagged NN, is placed within Zone 2. That decision is again repeated twice for classifying observations within Zone 1 and Zone 0. Observations that lie within Zone 0 are regarded as open ocean scenes. Figure 6a illustrates the five sea-ice zones for a sample scene near the Antarctic. It is evident that the distance to the denser sea-ice pack and thus the severity of sea-ice contamination increases when going from Zone 0 to Zone 5. For comparison, we have plotted the antenna gain weighted sea-ice fraction from the OSI-SAF SIC product [10] in Figure 6b. The sea-ice zones found by our algorithm match visually well with locations where g_{ice} (OSI-SAF) exceeds about 1%.

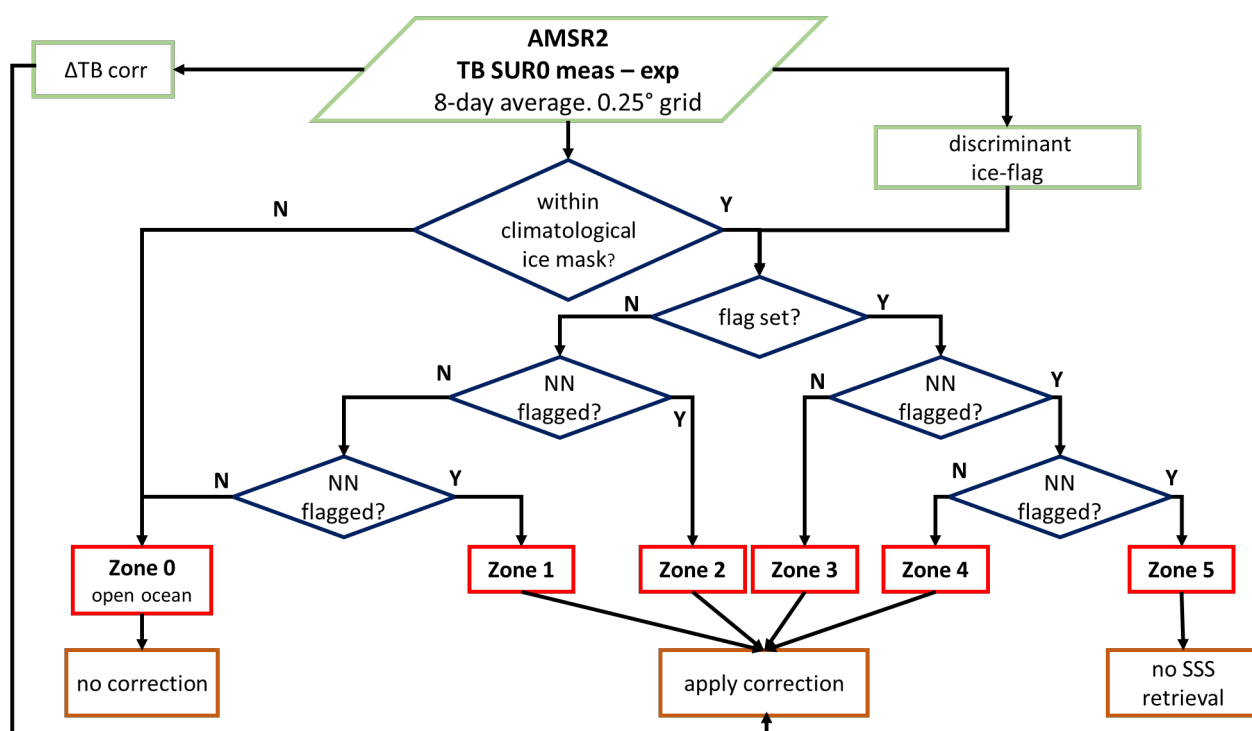


Figure 5. Schematic flow diagram of the sea-ice detection and flagging algorithm and definition of the sea-ice zones 0–5.

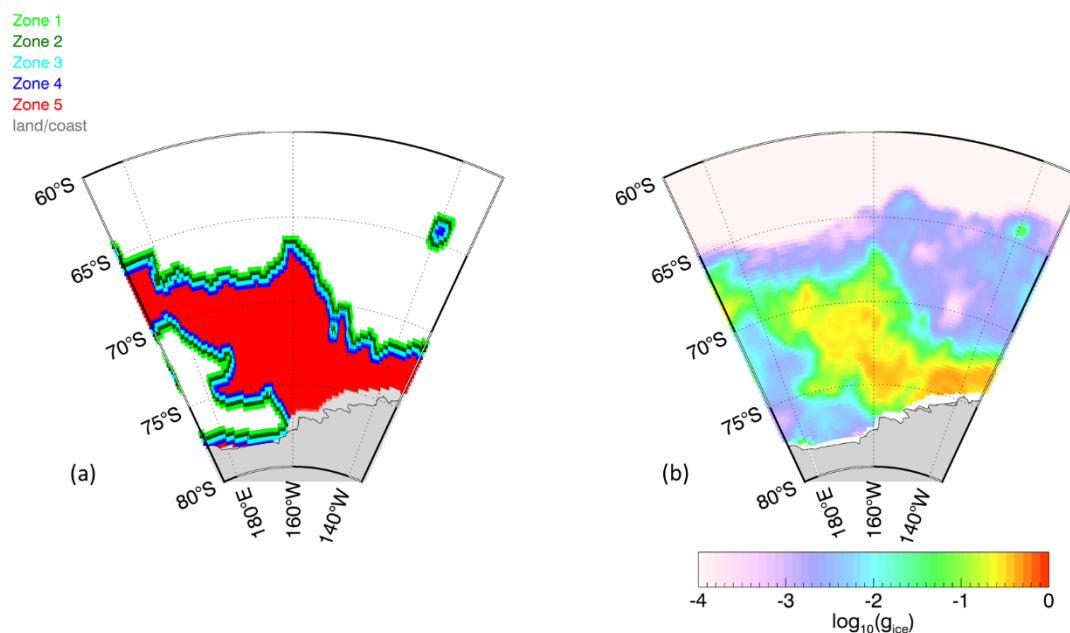


Figure 6. (a) Illustration of the sea-ice zones 0–5 for an area near the Antarctic ice shelf. Zone 0 (open ocean) observations are plotted white. The area near (65S, 140W) depicts a very large iceberg. (b) Antenna gain weighted sea-ice fraction g_{ice} from the OSI-SAF SIC product [10].

2.4. Sea-Ice Correction

The purpose of the sea-ice correction algorithm is to estimate the contamination term $\Delta T_{B,c}$ for the sea-ice flagged observations (given by Equations (2), (3) and (5)) and then subtract it from the measured SMAP $T_{B,0}$ before the SSS is retrieved. As we did for the discriminant flagging algorithm, we again consider the two cases defined in Equation (6). Case 1 uses the TOA AMSR2 TB as input, and Case 2 uses the measured–expected AMSR2

surface emissivities as input. The corrections are determined as statistical linear regressions, which have the forms:

$$\begin{aligned}\Delta T_{B,\text{corr}1} &= \Delta T_{B,0,1}(\text{SMAP}) = \alpha_0 + \sum_{k=1}^{10} \alpha_k \cdot T_{B,\text{TOA}}(\text{AMSR2})_k \\ \Delta T_{B,\text{corr}2} &= \Delta T_{B,0,2}(\text{SMAP}) = \sum_{k=1}^{10} \beta_k \cdot T_S \cdot [E_{0,\text{meas}}(\text{AMSR2}) - E_{0,\text{exp}}(\text{AMSR2})]_k\end{aligned}\quad (8)$$

The indices 1 and 2 distinguish the two cases 1 and 2.

The regression coefficients α_k and β_k in Equation (7) are derived for the training set and then evaluated for the test set listed in Table 1. We derive separate regressions for each of the Zones 1–4. No sea-ice correction is performed for observations in Zone 0 (i.e., open ocean), in which sea-ice contamination is non-existent or small. Performing a correction, in this case, would likely introduce unnecessary errors into the SMAP SSS retrievals. Observations within Zone 5 are considered non-salvageable for the purpose of retrieving SSS.

We train and apply separate corrections for the SMAP V-pol and H-pol channels. Note that the regression for Case 2 in Equation (7) does not contain a constant term. The assumption is that the correction term should go to zero if the measured and the expected AMSR2 surface emissivities are the same, as in this case the g_{ice} goes to zero as well. We also note that sea-ice contamination always results in a positive TB bias. Because of noise in the observations, the calculated values of $\Delta T_{B,c}$ in Equation (7) are sometimes negative. In these instances, we set $\Delta T_{B,\text{corr}} = 0$, and we do not apply a sea-ice correction.

3. Results

3.1. Correlation between SMAP and AMSR2 Brightness Temperatures

A pivotal point for the performance of the proposed flagging and correction method is the very good correlation between the SMAP and the AMSR2 measured–expected TB for the 8-day aggregate observations near the sea-ice edge. The $\Delta T_{B,0}$ of the two lowest AMSR2 frequency channels, 6.93 GHz and 10.65 GHz, correlate best with the SMAP $\Delta T_{B,0}$. This is mostly due to the fact that the higher AMSR2 frequencies are more strongly affected by atmospheric effects (i.e., water vapor and cloud water absorption), which act as noise sources in the detection and correction algorithms. Figure 7 shows a 2-dimensional logarithmic joint histogram of the SMAP V-pol specular surface TB bias $\Delta T_{B,0}(\text{SMAP V})$ and the scaled AMSR2 6.93 GHz V-pol specular TB bias $\lambda \cdot \Delta T_{B,0}(\text{AMSR2 6V})$ for the test set scenes (Table 1) within sea-ice zones 1–4. The scaling factor λ can be determined from a linear regression and is found to be $\lambda = 1.15$. The correlation coefficient is 0.96 (Table 4). Similar values are found (Table 4) for the correlation coefficients between the $\Delta T_{B,0}(\text{SMAP V})$ and the sea-ice correction terms $\Delta T_{B,\text{corr},1}$ and $\Delta T_{B,\text{corr},2}$ that were derived in Section 2.4. It is because of these high correlations, and the fact that the AMSR2 TB are insensitive to SSS, that we can employ them for sea-ice flagging and sea-ice correction in the SMAP SSS retrieval algorithm.

Table 4. Pearson correlation coefficients between the SMAP V-pol specular surface TB bias $\Delta T_{B,0}(\text{SMAP V})$ for the test set scenes (Table 1) within sea-ice zones 1–4. The $\Delta T_{B,0}(\text{6V})$ is the AMSR2 6.93 GHz V-pol specular TB bias. The $\Delta T_{B,\text{corr}1}$ and $\Delta T_{B,\text{corr}2}$ are the values for the sea-ice corrections that were derived in Section 2.4.

Product	$\Delta T_{B,0}(\text{6V})$	$\Delta T_{B,\text{corr}1}$	$\Delta T_{B,\text{corr}2}$
correlation	0.96	0.96	0.96

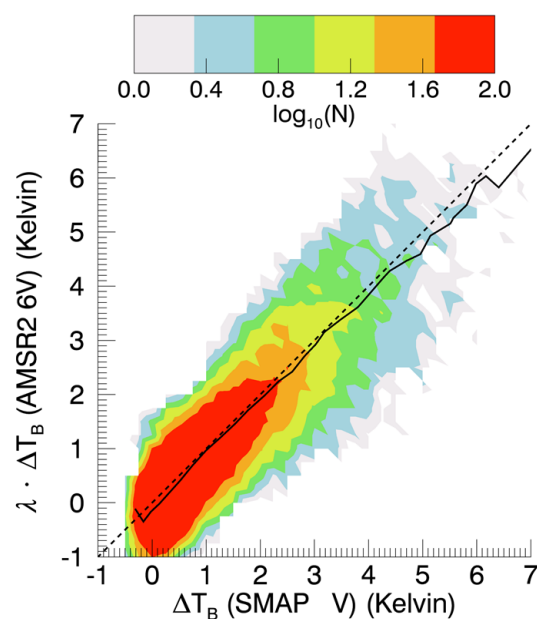


Figure 7. 2-dimensional logarithmic joint histogram of the SMAP V-pol specular surface TB bias $\Delta T_{B,0}(\text{SMAP V})$ and the scaled AMSR2 6.93 GHz V-pol specular TB bias $\lambda \cdot \Delta T_{B,0}(\text{AMSR2 6V})$ for the test set (Table 1) within sea-ice zones 1–4. The bin size is 0.25 K. The value of the scaling factor is $\lambda = 1.15$. The dashed line indicates the 1:1 line (ideal case). The full line indicates the binned average between y -axis versus x -axis and x -axis versus y -axis. The Pearson correlation coefficient is 0.96.

We note that the lowest frequency AMSR2 channels (C-band and X-bands) correlate best with SMAP, though the higher AMSR2 channels (Ku–Ka-bands) still contain valuable information when it comes to sea-ice contamination, and it is, therefore, beneficial to include them in our method.

3.2. Performance Evaluation of the Sea-Ice Correction Algorithm

Table 5 summarizes the performance evaluation statistics for the two cases of the sea-ice correction algorithm applied to the test cases within Zones 1–4 and compares them to the case without applying a correction. We have also included estimated values for the sea-ice fraction g_{ice} in each zone.

Table 5. Performance statistics of the sea-ice correction algorithm (Section 2.4.) for the test set (Table 1) within sea-ice zones 0–4. The table lists biases, standard deviations and RMS (all in K) for the performance metric, which is the SMAP V-pol specular surface TB bias $\Delta T_{B,0}(\text{SMAP V})$. Case 1 refers to the algorithm using the AMSR2 TOA as input. Case 2 refers to the algorithm using the AMSR2 measured minus expected specular surface emissivities as input. The estimated average g_{ice} values in the 2nd column are obtained by dividing the values for the TB biases without correction in each zone by 125 K, which is a typical value for the L-band TB difference between sea-ice and ocean scenes.

Zone	estimated g_{ice}	No Correction			Case 1			Case 2		
		Bias	Std.Dev	RMS	Using AMSR2 TB TOA	Using AMSR2 E0 Meas-Exp	Bias	Std.Dev	RMS	
0	<0.05%	0.07	0.20	0.21						
1	0.3%	0.33	0.28	0.43	−0.03	0.27	0.27	0.02	0.28	0.28
2	0.6%	0.70	0.56	0.89	0.06	0.52	0.52	0.02	0.50	0.50
3	2.2%	2.76	2.02	3.42	0.19	1.32	1.33	−0.10	1.33	1.33
4	9.6%	11.62	8.07	14.14	−0.01	3.12	3.12	−0.51	3.35	3.39

The correction algorithm was designed and trained to debias the SMAP $\Delta T_{B,0}$ in each of the zones. The values show that this goal has been largely achieved. It is also

encouraging to see reduced values for the standard deviations. The relative improvement compared to the uncorrected cases is strong for Zone 3 observations and even stronger for Zone 4 observations. We should note, however, that the remaining errors after the correction for Zone 3 and Zone 4 will still result in large errors in the SSS retrievals. A TB error of 1 K translates into an SSS error of about 4 psu in cold water (Section 1), and, thus, the expected SSS errors are about 5 psu in Zone 3 and about 12 psu in Zone 4. SSS errors of these magnitudes might still be too large for many scientific or operational oceanographic applications. Based on the results shown in Table 5, our analysis is able to give a reliable estimate of the SSS error in each of the zones. We expect that sea-ice corrected observations within Zone 1 and Zone 2 will be useful for the majority of scientific studies. The following sections will show several examples. Ultimately, it depends on the size of the signal of interest, and it will be up to the data users to decide if they want to discard the data for Zones 3 and 4.

Table 5 indicates that the Case 2 algorithm, which employs external ancillary fields and the RTM performs slightly better in Zones 1 and 2, whereas the Case 1 algorithm, which does not use external ancillary fields, performs slightly better in Zone 4. This is likely due to the fact that the quality of the external ancillary fields degrades with increasing sea-ice concentration. In particular, the ancillary SST and wind fields depend on satellite measurements, which become less accurate or are unavailable for observations within Zone 4. In Case 1, the higher frequency channels are indirectly employed to remove the atmospheric, wind, and SST effects from the TB TOA. In Case 2, this role is taken by the ancillary fields and the RTM calculation of atmospheric attenuation and wind induced rough surface emission. For Case 1, it is conceivable to develop a more elaborate machine learning technique, for instance, a neural network. It will need to be seen if this results in better performance than the statistical regression algorithm.

3.3. Antarctic Scene Examples

Figure 8 shows the results of the sea-ice flagging algorithm when applied to the sample test scenes from Figures 1 and 2. The figure plots only observations that fall within Zone 0. This means that all of the sea-ice contaminated data have been removed resulting in almost bias-free scenes. Figure 9 shows the result within Zones 0–2 after the sea-ice correction algorithm has been applied. Only few cells are visible that are still contaminated by sea-ice. The vast majority of grid cells are free of large sea-ice contamination, and the SSS retrievals are of similar quality as over the open ocean. We will discuss the residual errors in $\Delta T_{B,0}(\text{SMAP})$ that remain after the sea-correction is applied in more detail in Section 4.

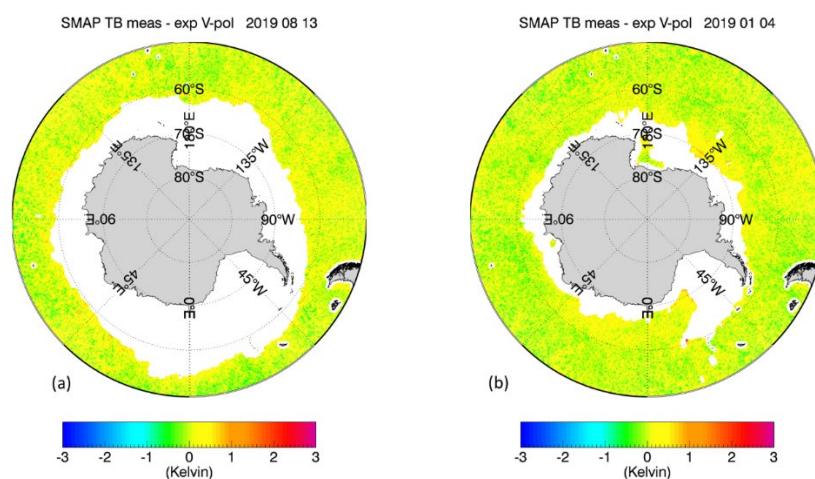


Figure 8. SMAP V-pol TB biases for the test cases from Figure 1 and Figure 2 after discriminant and NN flagging. The resulting observations lie all within sea-ice Zone 0. (a) 13 August 2019; (b) 4 January 2019.

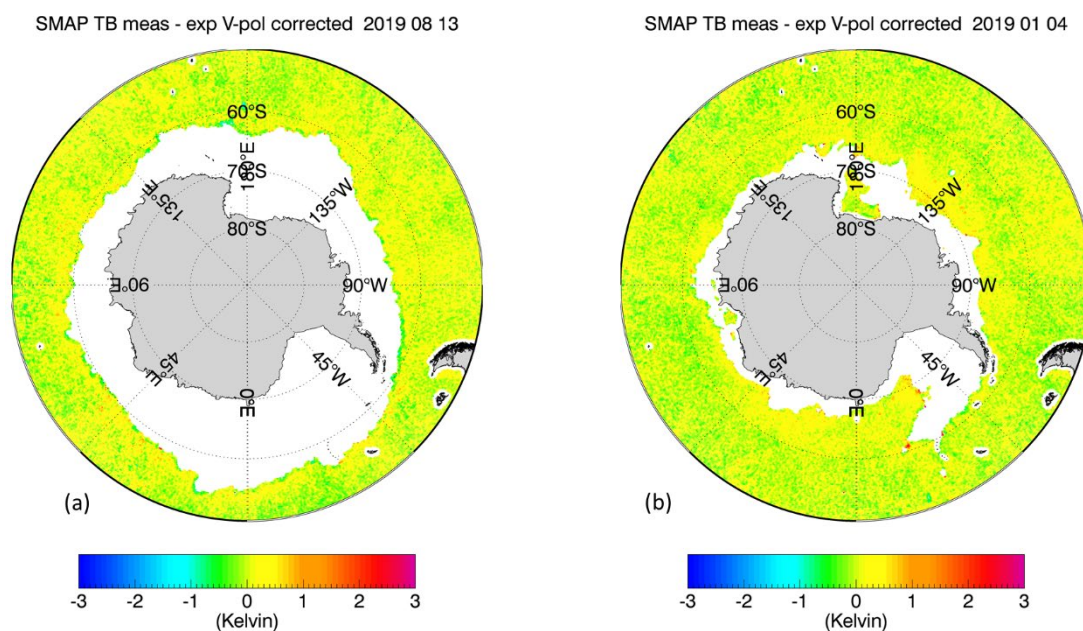


Figure 9. SMAP V-pol TB biases for the test cases from Figure 1 and Figure 2 after flagging and sea-ice correction within sea-ice Zones 0–2. (a) 13 August 2019; (b) 4 January 2019.

3.4. Arctic Scene Example

Figure 10 shows an example of how the algorithms perform in the Arctic. The large red areas are freshwater outflows from Siberian rivers. This positive SMAP TB bias is most likely caused by the HYCOM reference SSS underestimating the freshening caused by these river outflows whereas SMAP picks it up. AMSR2 is not sensitive to the freshening, and, thus, these areas remain unflagged, and no sea-ice correction is done. This is the desired outcome i.e., areas with no sea-ice should not be flagged as such. On the other hand, it is evident from Figure 10 that our algorithm is capable of detecting, flagging, and correcting the sea-ice contamination near the sea-ice edge, where spurious SMAP TB biases are observed.

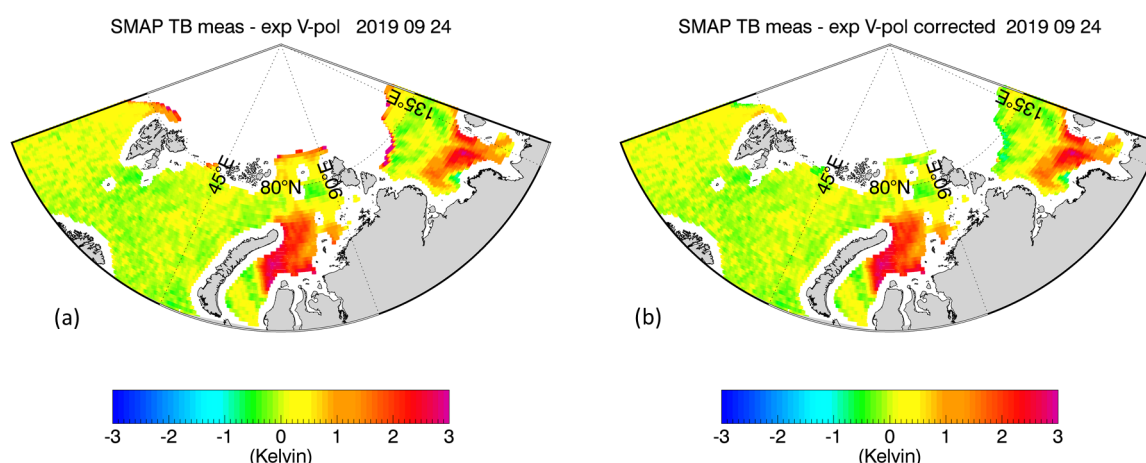


Figure 10. SMAP V-pol TB biases for an Arctic test scene on 24 September 2019 (Set 8 from Table 1). (a) before flagging and correction in zones 0–3; (b) after flagging and correction showing zones 0–2. The large red areas indicate freshwater river outflows. These areas are not affected by the sea-ice flagging and correction algorithms.

3.5. Icebergs

Very large drifting icebergs are a potential source of serious errors in the SMAP SSS retrievals. Even though icebergs are not sea-ice, the electromagnetic properties of sea-ice

and icebergs result in a similar degree of contamination in the SSS measurement. The a-priori climatological ice mask (Section 2.1.6.) deliberately includes areas that are prone to contain drifting icebergs. That is essential to enable the sea-ice flagging algorithm to also detect icebergs. A closer look at Figure 2 shows some very large icebergs near Antarctica, which are both visible in the SMAP TB and in the AMSR2 6V TB. Checking the results of the sea-ice flagging algorithm (Figure 8b) and the sea-ice correction algorithm (Figure 9b) for the same scene indicates that our method is able to detect these icebergs and correct the contamination to a good extent.

A good example of this detection and correction is provided by the very large iceberg A-68 [35], which was drifting near the South Georgia Islands and several hundred kilometers away from the Antarctic sea-ice edge in the early months of 2021. It was the cause of large spurious fresh SMAP SSS values in that location, whose fresh biases exceeded 10 psu. This resulted in a surface area of about 250,000 km² of undetected contaminated SMAP SSS data over the course of several weeks. Figure 11 shows (a) the iceberg contaminated SMAP data and (b) the result after flagging and correcting the contaminated observations. Our method appears to be very effective in detecting and mitigating the problem.

We note that neither of the products [9,10] shows this iceberg in their SIC data.

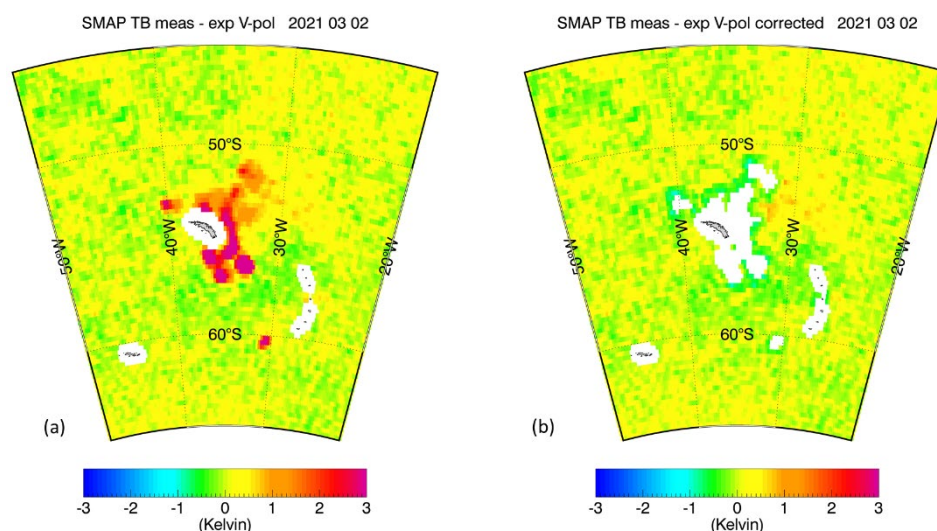


Figure 11. SMAP V-pol TB biases for an area around the South Georgia Islands near the Antarctica on 02 March 2021 (Set 9 from Table 1) that was affected by the very large drifting iceberg A-68 [35]. (a) before flagging and correction; (b) after flagging and correction showing zones 0–2.

4. Discussion

This section discusses the uncertainty and error sources in our methodology and its limitations.

- (1) Radiometer noise in the AMSR2 TB measurements is a source of error. However, the noise in the individual measurements is strongly reduced in the eight-day averages. Thus, the impact of AMSR2 radiometer noise on our detection and correction algorithms is small.
- (2) A larger uncertainty source are errors in the external ancillary fields, which might not get suppressed in the eight-day averaging. This impacts the algorithms from Case 2, which use the ancillary fields to remove atmospheric and wind roughness effects from the AMSR2 TOA TB. We expect the errors in the ancillary fields for SST, wind speed, and direction to increase with increasing sea-ice fraction g_{ice} and decreasing distance from the sea-ice edge. The ancillary data sets for these parameters are Level 4 products that rely on microwave satellite observations as input. When approaching the ice edge, the microwave satellite observations are flagged, and, thus, fewer

satellite input data become available. This can result in lower accuracy of the ancillary Level 4 product [26].

- (3) One of the major assumptions when using the AMSR2 TB to flag and correct the sea-ice contaminated SMAP TB was that the spatial resolutions of the SMAP and the AMSR2 TB measurements are similar. When averaged over eight days, the different orientations of the antenna footprints at a certain location have less of an impact. This means that SMAP and AMSR2 see approximately the same Earth scene. That assumption is of course only approximately fulfilled, and in reality, there is some spatial mismatch between the SMAP and the AMSR2 observations. It is possible to reduce this uncertainty source by resampling the SMAP and all the AMSR2 channels to the same target location and the same spatial resolution by employing the Backus-Gilbert optimum interpolation technique [17]. This resampling would use a polar stereographic projection for selecting the target locations, which is better suited than a regular 0.25° grid in polar regions. This is a project that is currently being considered, but it will require considerable effort.
- (4) There are scenarios that can result in a degradation of the good correlation between the SMAP and the AMSR2 TB. The most important one is thin sea-ice, which impacts the higher AMSR2 frequencies more than the lower ones or the SMAP L-band. We note that, at L-band, the sea-ice penetration depth is approximately 0.5 m, and the penetration depth grows with the electromagnetic wavelength. Large areas of thin ice tend to form in the Arctic oceans during fall. In the S hemisphere, this problem is less prevalent. Further studies are necessary to examine if and how much this poses a problem for our flagging and correction method.
- (5) Undetected RFI in the AMSR2 TB also constitutes an error source.
- (6) A problematic case is sea-ice close to the coast, which can result in simultaneous land and sea-ice contamination of the SMAP antenna field of view. Our proposed sea-ice detection and correction algorithms were not designed to handle this case. Likewise, we expect that the sidelobe correction for land contamination that is currently applied in the NASA/RSS SMAP SSS retrievals [13,14] is not able to cope with this scenario either.
- (7) The flagging method is not able to detect observations of very small sea-ice fractions that are significantly lower than the detection threshold of the discriminant analysis (1.5%) and that are not NN or next to NN of an observation that has been flagged by the discriminant analysis.
- (8) The statistical linear regressions that are used in the sea-ice correction algorithm are designed to de-bias the $\Delta T_{B,0}(\text{SMAP})$ in each sea-ice zone, which they do (Table 5). The residual errors after applying the correction approximately add to zero when summing over all test scenes with standard deviation values listed in Table 5. However, we note that the spatial correlation scales of these residual errors can be quite large. For instance, in Figures 9a, 10b, and 11b green bands are visible in the corrected observations, which indicate larger areas where the algorithm slightly overcorrects. In Figure 9b, a few yellow–orange areas are visible, which indicate that the algorithm undercorrects in these locations. That means that we should not expect that this residual error can be significantly reduced by spatial smoothing. Consequently, we anticipate that the remaining errors in the 70-km NASA/RSS SMAP SSS product will have a similar magnitude than in the 40-km product near the sea-ice edge.

5. Conclusions

The AMSR2 TB measurements between C-band and Ka-band contain valuable information that can be employed in flagging and correcting sea-ice contaminated SMAP observations. This will result in better SMAP SSS measurements in polar oceans that currently suffer from sea-ice contamination, which is otherwise difficult to detect and remove. We plan to implement both the sea-ice flagging and the sea-ice detection algorithm

into the upcoming Version 5 release of the NASA/RSS SMAP SSS product. The calculated correction $\Delta T_{B,corr}$ (Equation (7)) allows us to provide an approximate value for the gain weighted sea-ice fraction g_{ice} in the SMAP antenna field of view and thus of the severity of the sea-ice contamination. The performance results from Table 5 can be used to estimate the residual RMS error in the retrieved SSS after the sea-ice correction is performed.

More extensive validation of the sea-ice corrected SSS retrieval is warranted and will be carried out. It will be particularly important to conduct a comparison with in-situ ground-truth measurements in the polar oceans, e.g., saildrones [36].

Possible future improvements or refinements include spatial resampling of the SMAP and AMSR2 observations to the same location and spatial resolution, as mentioned in Section 4. It is also desirable to see if the temporal resolution can be improved. We have designed the SMAP sea-ice detection and correction algorithms for eight-day time-averaged observations as our main goal was to improve the Level 3 eight-day SMAP SSS products, which are used in most scientific applications. However, sea-ice concentration and contamination can change within eight days. In particular, near the ice edge, sea-ice motion can exceed 100 km within this eight-day time window both along and across the ice edge. Our method is based on the time-averaged SIC at a certain location. This time-averaged SIC is approximately the same for the AMSR2 and the SMAP eight-day aggregates, as long as there are no temporal gaps in data acquisition for either of the two sensors. A shorter time window for the AMSR2 aggregates is warranted when dealing with Level 2 retrievals, which are of interest for most operational applications. That said, a smaller time window will increase the noise in the sea-ice detection and correction algorithms. We plan to explore if the time averaging window could be narrowed without increasing the missed detection and false alarm rates too much.

We expect that the proposed methodology could be applied to Aquarius SSS retrievals as well. This would require that the AMSR2 observations being spatially resampled to the lower Aquarius resolution of 100–150 km. Adapting this methodology to SMOS, which is a synthetic aperture radiometer, is more difficult and it is not clear if it is feasible. It is safe to assume that the method can be continued with AMSR3, which is planned to be launched in the 2023 time frame, if SMAP is still operating by then. A very exciting future mission that can apply our methodology is the European Copernicus Imaging Microwave Radiometer (CIMR) [37], which is planned to start operating by the end of this decade. CIMR will simultaneously observe the same Earth scene at a wide frequency range between L-band and Ka-band, and, thus, provide most of the necessary input for our methodology with one single instantaneous measurement.

Supplementary Materials: The following are available online at https://data.remss.com/smap/SSS/support_data/remote_sensing_special_issue/: Regression coefficients: Table S1: regression coefficients ΔT_{B0} (SMAP) versus $T_{B,TOA}$ (AMSR2), Table S2: regression coefficients ΔT_{B0} (SMAP) versus $E_{0,meas}$ (AMSR2)– $E_{0,exp}$ (AMSR2). For details see Equation (7). Support figures: SMAP and AMSR2 measured–expected TB zooming into coastal areas and near islands. The figures are provided to demonstrate that there is no significant land contamination in either the SMAP or the AMSR2 TB.

Author Contributions: Conceptualization, T.M.; methodology, T.M.; software, T.M. and A.M.; formal analysis, T.M. and A.M.; investigation, T.M. and A.M.; resources, T.M. and A.M.; data curation, T.M. and A.M.; writing—original draft preparation, T.M.; writing—review and editing, T.M. and A.M.; visualization, T.M. and A.M.; supervision, T.M.; project administration, T.M.; funding acquisition, T.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded under the NASA Salinity Continuation Processing System under JPL/RSS subcontract # 1664013 (Production System for NASA’s Ocean Salinity Science Team) and under NASA grant 80NSSC21K0545 (Development of Ocean Retrieval Algorithms for the Copernicus Imaging Microwave Radiometer).

Data Availability Statement: The NASA/RSS SMAP Version 4.0 salinity are publicly available at <https://www.remss.com/missions/smap/> (accesses on 30 October 2021) and on

https://podaac.jpl.nasa.gov/dataset/SMAP_RSS_L2_SSS_V4 (accessed on 30 October 2021) Future releases will be available at the same locations. The aggregated TB maps for both SMAP and AMSR2, as well as the ancillary fields for SST, wind speed, columnar water vapor and columnar cloud water that we have used for our analysis, are made available on https://data.remss.com/smap/SSS/support_data/ (accessed on 30 October 2021). Sample read routines for the map files are also available at this location.

Acknowledgments: We would like to thank the members of the NASA Ocean Salinity Science Team OSST for providing valuable input and suggestions over the course of this work. In particular, we would like to thank Semyon Grodsky (University of Maryland) for bringing to our attention the drifting iceberg A-68 near South Georgia Islands in early 2021 and the problem it caused on the SMAP salinity retrievals.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Fisher Linear Discriminant Analysis

Our sea-ice detection algorithm is based on the Fisher Linear Discriminant Analysis, which is a well-established statistical technique in pattern classification for distinguishing two classes of data [38,39]. This appendix gives a brief summary of the method.

Consider two data sets Ω_i , $i = 1, 2$, each of which contains N_i data points. Each data point itself is a measurement in a d -dimensional space and thus labelled as a d -dimensional vector \mathbf{X} . In our case, the measurement vector \mathbf{X} consists of the 10 AMSR2 TOA TB or the measured–expected specular surface emissivities (Table 2). The sample means \mathbf{M}_i of each set Ω_i are given by:

$$\mathbf{M}_i = \frac{1}{N_i} \cdot \sum_{\mathbf{X} \in \Omega_i} \mathbf{X} \quad (\text{A1})$$

which are both again d -dimensional vectors. The Fisher Discriminant Analysis considers the projection of the d -dimensional data vectors onto a fixed d -dimensional vector \mathbf{W} . The projected means are thus:

$$\tilde{\mathbf{M}}_i = \frac{1}{N_i} \cdot \sum_{\mathbf{X} \in \Omega_i} \mathbf{W}^T \cdot \mathbf{X} = \mathbf{W}^T \cdot \mathbf{M}_i \quad (\text{A2})$$

where \mathbf{W}^T denotes the transposed vector and thus $\mathbf{W}^T \cdot \mathbf{X}$ is the scalar product between the two vectors \mathbf{W} and \mathbf{X} . The distance between the projected means is:

$$\left| \tilde{\mathbf{M}}_1 - \tilde{\mathbf{M}}_2 \right| = \left| \mathbf{W}^T \cdot (\mathbf{M}_1 - \mathbf{M}_2) \right|. \quad (\text{A3})$$

The variances and scatters of each projected data set are:

$$\tilde{\sigma}_i^2 = \frac{1}{N_i} \cdot \sum_{\mathbf{X} \in \Omega_i} \left(\mathbf{W}^T \cdot \mathbf{X} - \tilde{\mathbf{M}}_i \right)^2 =: \frac{1}{N_i} \cdot \tilde{s}_i^2 \quad (\text{A4})$$

and the variance of the combined (pooled) data is:

$$\tilde{\sigma}^2 = \frac{1}{N_1 + N_2} \cdot (\tilde{s}_1^2 + \tilde{s}_2^2) \quad (\text{A5})$$

The goal is to determine the projection vector \mathbf{W} that leads to the maximum separation between the two data sets. This is achieved by maximizing the objective function $J(\mathbf{W})$, which maximizes the distance between the projected data points by simultaneously minimizing the pooled projected variance:

$$J(\mathbf{W}) = \frac{|\tilde{\mathbf{M}}_1 - \tilde{\mathbf{M}}_2|}{\tilde{s}_1^2 + \tilde{s}_2^2} \quad (\text{A6})$$

The optimization of $J(\mathbf{W})$ employs the covariance and scatter matrices of the individual sets Ω_i :

$$\text{Cov}(\Omega_i) = \frac{1}{N_i} \cdot \sum_{\mathbf{x} \in \Omega_i} (\mathbf{x} - \mathbf{M}_i) \cdot (\mathbf{x} - \mathbf{M}_i)^T =: \frac{1}{N_i} \cdot \mathbf{S}_i \quad (\text{A7})$$

and the within-scatter matrix \mathbf{S} :

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2. \quad (\text{A8})$$

The optimized projection vector \mathbf{W} is determined by the expression:

$$\mathbf{W} = \mathbf{S}^{-1} \cdot (\mathbf{M}_1 - \mathbf{M}_2) \quad (\text{A9})$$

and the value of the discriminant is given by $\mathbf{W}^T \cdot \mathbf{X}$. The normalization of the projection vector is arbitrary. In our study we have normalized it to have unit length, i.e., $\mathbf{W}^T \cdot \mathbf{W} = \mathbf{1}$. In order to obtain the decision criterion between the two classes, one creates plots of the probability density functions pdf_i ($i = 1, 2$) of the discriminant values in each class (Figure 4). The standard way of deciding if an event falls into class 1 (class 2) is the check if the discriminant pdf of class 1 (class 2) is larger than the discriminant pdf for class 2 (class 1). That puts the value d for the decision boundary between the two classes at the abscissa of the intersection point between the two pdf. In the context of Figure 4:

$$\begin{aligned} (\mathbf{W}^T \cdot \mathbf{X}) > d &\Rightarrow \text{pdf}_2 > \text{pdf}_1 \Rightarrow \mathbf{X} \in \text{Class 2} \\ (\mathbf{W}^T \cdot \mathbf{X}) < d &\Rightarrow \text{pdf}_1 > \text{pdf}_2 \Rightarrow \mathbf{X} \in \text{Class 1} \end{aligned} \quad (\text{A10})$$

References

1. Reul, N.; Grodsky, S.A.; Arias, M.; Boutin, J.; Catany, R.; Chapron, B.; D'Amico, F.; Dinnat, E.; Donlon, C.; Fore, A.; et al. Sea surface salinity estimates from spaceborne L-band radiometers: An overview of the first decade of observation (2010–2019). *Remote Sens. Environ.* **2020**, *242*, 111769, doi:10.1016/j.rse.2020.111769.
2. Kao, H.-Y.; Lagerloef, G.; Lee, T.; Melnichenko, O.; Meissner, T.; Hacker, P. Assessment of Aquarius Sea Surface Salinity. *Remote Sens.* **2018**, *10*, 1341, doi:10.3390/rs10091341.
3. Tang, W.; Yueh, S.; Yang, D.; Fore, A.; Hayashi, A.; Lee, T.; Fournier, S.; Holt, B. The Potential and Challenges of Using Soil Moisture Active Passive (SMAP) Sea Surface Salinity to Monitor Arctic Ocean Freshwater Changes. *Remote Sens.* **2018**, *10*, 869, doi:10.3390/rs10060869.
4. Olmedo, E.; Gabarró, C.; González-Gambau, V.; Martínez, J.; Ballabrera-Poy, J.; Turiel, A.; Portabella, M.; Fournier, S.; Lee, T. Seven Years of SMOS Sea Surface Salinity at High Latitudes: Variability in Arctic and Sub-Arctic Regions. *Remote Sens.* **2018**, *10*, 1772, doi:10.3390/rs10111772.
5. Supply, A.; Boutin, J.; Vergely, J.-L.; Kolodziejczyk, N.; Reverdin, G.; Reul, N.; Tarasenko, A. New insights into SMOS sea surface salinity retrievals in the Arctic Ocean. *Remote Sens. Environ.* **2020**, *249*, 112027, doi:10.1016/j.rse.2020.112027.
6. Vinogradova, N.; Lee, T.; Boutin, J.; Drushka, K.; Fournier, S.; Sabia, R.; Stammer, D.; Bayler, E.; Reul, N.; Gordon, A.; et al. Satellite Salinity Observing System: Recent Discoveries and the Way Forward. *Front. Mar. Sci.* **2019**, *6*, 243, doi:10.3389/fmars.2019.00243.
7. Meissner, T.; Wentz, F.J.; Le Vine, D.M. Aquarius Salinity Retrieval Algorithm Theoretical Basis Document (ATBD). End of Mission Version. RSS Technical Report 120117. 2017. Available online: images.remss.com/papers/tech_reports/2017/Meissner_AQ_ATBD_EOM_final.pdf (accessed on 30 October 2021).
8. Heygster, G.; Huntemann, M.; Ivanova, N.; Saldo, R.; Pedersen, L.T. Response of passive microwave sea ice concentration algorithms to thin ice. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 3618–3621.
9. Meier, W.N.; Fetterer, F.; Savoie, M.; Mallory, S.; Duerr, R.; Stroeve, J. NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 3. 2017. Available online: <https://nsidc.org/data/g02202/versions/3> (accessed on 30 October 2021). doi:10.7265/N59P2ZTG.

10. Danish Meteorological Institute Global Sea Ice Concentration (AMSR-2) | OSI SAF. Processed by the OSI-SAF Hybrid Dynamic Algorithm. Available online: <https://osi-saf.eumetsat.int/products/osi-408> (accessed on 30 October 2021).
11. Meissner, T.; Wentz, F.J.; Manaster, A.; Lindsley, R. NASA/RSS SMAP Level 2C Sea Surface Salinity V4.0 Validated Dataset. 2019. Available online: https://podaac.jpl.nasa.gov/dataset/SMAP_RSS_L2_SSS_V4 (accessed on 30 October 2021). doi:10.5067/SMP40-2SOCS.
12. Fournier, S.; Lee, T.; Tang, W.; Steele, M.; Olmedo, E. Evaluation and Intercomparison of SMOS, Aquarius, and SMAP Sea Surface Salinity Products in the Arctic Ocean. *Remote Sens.* **2019**, *11*, 3043, doi:10.3390/rs11243043.
13. Meissner, T.; Wentz, F.J.; Manaster, A.; Lindsley, R. NASA/RSS SMAP Salinity: Version 4.0 Validated Release: Release Notes, Algorithm Theoretical Basis Document (ATBD), Validation, Data Format Specification. RSS Technical Report 082219. 2019. Available online: https://data.remss.com/smap/SSS/Release_V4.0.pdf (accessed on 30 October 2021).
14. Meissner, T.; Wentz, F.J.; Vine, D.M.L. The Salinity Retrieval Algorithms for the NASA Aquarius Version 5 and SMAP Version 3 Releases. *Remote Sens.* **2018**, *10*, 1121, doi:10.3390/rs10071121.
15. Meissner, T.; Wentz, F.J.; Ricciardulli, L. The emission and scattering of L-band microwave radiation from rough ocean surfaces and wind speed measurements from the Aquarius sensor. *J. Geophys. Res. Oceans* **2014**, *119*, 6499–6522, doi:10.1002/2014JC009837.
16. Ashcroft, P.; Wentz, F.J. AMSR Level 2A Algorithm Theoretical Basis Document (ATBD). RSS Technical Report 121599B-1. 2000. Available online: https://images.remss.com/papers/rsstech/2000_121599B-1_Wentz_AMSR_Level2A_Algorithm_ATBD.pdf (accessed on 30 October 2021).
17. Poe, G.A. Optimum interpolation of imaging microwave radiometer data. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 800–810, doi:10.1109/36.58966.
18. Wentz, F.J. AMSR-2 Air-Sea Essential Climate Variables RSS Version 8.2. RSS Technical Report 05112021. 2021. Available online: https://images.remss.com/papers/tech_reports/2021/AMSR-2_Air-Sea_Essential_Climate_Variables_RSS_Version_8.2.pdf (accessed on 30 October 2021).
19. Wentz, F.J.; Meissner, T. Atmospheric absorption model for dry air and water vapor at microwave frequencies below 100 GHz derived from spaceborne radiometer observations: Atmospheric Absorption Model. *Radio Sci.* **2016**, *51*, 381–391, doi:10.1002/2015RS005858.
20. Meissner, T.; Wentz, F.J. The Emissivity of the Ocean Surface Between 6 and 90 GHz Over a Large Range of Wind Speeds and Earth Incidence Angles. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3004–3026, doi:10.1109/TGRS.2011.2179662.
21. Wentz, F.J.; Meissner, T. AMSR Ocean Algorithm Theoretical Basis Document (ATBD). RSS Technical Report 121599A-1. 2000. Available online: images.remss.com/papers/rsstech/2000_121599A-1_Wentz_AMSR_Ocean_Algorithm_ATBD_Version2.pdf (accessed on 30 October 2021).
22. Meissner, T.; Wentz, F.J. The complex dielectric constant of pure and sea water from microwave satellite observations. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1836–1849, doi:10.1109/TGRS.2004.831888.
23. Chassignet, E.; Hurlburt, H.; Metzger, E.J.; Smedstad, O.; Cummings, J.; Halliwell, G.; Bleck, R.; Baraille, R.; Wallcraft, A.; Lozano, C.; et al. US GODAE: Global Ocean Prediction with the HYbrid Coordinate Ocean Model (HYCOM). *Oceanography* **2009**, *22*, 64–75, doi:10.5670/oceanog.2009.39.
24. Global Ocean Forecasting System GOFS 3.1: 41-layer HYCOM + NCODA Global 1/12° Analysis. Available online: <https://www.hycom.org/data/glby0pt08/expt-93pt0> (accessed on 20 October 2021).
25. Canada Meteorological Center GHRSSST Level 4 CMC0.1deg Global Foundation Sea Surface Temperature Analysis (GDS version 2). 2017. Available online: <https://podaac.jpl.nasa.gov/dataset/CMC0.1deg-CMC-L4-GLOB-v3.0> (accessed on 20 October 2021). doi:10.5067/GHCMC-4FM03.
26. Mears, C.A.; Scott, J.; Wentz, F.J.; Ricciardulli, L.; Leidner, S.M.; Hoffman, R.; Atlas, R. A Near-Real-Time Version of the Cross-Calibrated Multiplatform (CCMP) Ocean Surface Wind Velocity Data Set. *J. Geophys. Res. Oceans* **2019**, *124*, 6997–7010, doi:10.1029/2019JC015367.
27. Wentz, F.J.; Scott, J.; Hoffman, R.; Leidner, M.; Atlas, R.; Ardizzone, J. Remote Sensing Systems Cross-Calibrated Multi-Platform (CCMP) 6-hourly ocean vector wind analysis product on 0.25 deg grid, Version 2.0. 2015. Available online: www.remss.com/measurements/ccmp (accessed on 30 October 2021).
28. Atlas, R.; Hoffman, R.N.; Ardizzone, J.; Leidner, S.M.; Jusem, J.C.; Smith, D.K.; Gombos, D. A Cross-calibrated, Multiplatform Ocean Surface Wind Velocity Product for Meteorological and Oceanographic Applications. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 157–174, doi:10.1175/2010BAMS2946.1.
29. National Centers for Environmental Prediction NCEP Data Products GFS and GDAS. Available online: <https://www.nco.ncep.noaa.gov/pmb/products/gfs/> (accessed on 30 October 2021).
30. Precipitation Processing System At NASA GSFC GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06. 2019. Available online: https://disc.gsfc.nasa.gov/datacollection/GPM_3IMERGHH_06.html (accessed on 30 October 2021). doi:10.5067/GPM/IMERG/3B-HH/06.
31. Meissner, T.; Wentz, F.J.; Scott, J.; Vazquez-Cuervo, J. Sensitivity of Ocean Surface Salinity Measurements From Spaceborne L-Band Radiometers to Ancillary Sea Surface Temperature. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 7105–7111, doi:10.1109/TGRS.2016.2596100.
32. Andersen, S.; Tonboe, R.; Kern, S.; Schyberg, H. Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using numerical weather prediction model fields: An intercomparison of nine algorithms. *Remote Sens. Environ.* **2006**, *104*, 374–392, doi:10.1016/j.rse.2006.05.013.

33. Ivanova, N.; Pedersen, L.T.; Tonboe, R.T.; Kern, S.; Heygster, G.; Lavergne, T.; Sørensen, A.; Saldo, R.; Dybkjær, G.; Brucker, L.; et al. Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations. *Cryosphere* **2015**, *9*, 1797–1817, doi:10.5194/tc-9-1797-2015.
34. Tian, T.; Tonboe, R.; Lavelle, J. The EUMETSAT OSI SAF AMSR-2 Sea Ice Concentration Algorithm Theoretical Basis Document (ATBD). Product OSI-408. Version 1.0. February 2015. Available online: https://osisaf-hl.met.no/sites/osisaf-hl.met.no/files/baseline_document/osisaf_cdop2_ss2_atbd_amsr2-sea-ice-conc_v1p1.pdf (accessed on 8 December 2021).
35. U.S. National Ice Center Antarctic Iceberg Tracking. Available online: <https://usicecenter.gov/Resources/AntarcticIcebergs> (accessed on 30 October 2021).
36. Vazquez-Cuervo, J.; Gentemann, C.; Tang, W.; Carroll, D.; Zhang, H.; Menemenlis, D.; Gomez-Valdes, J.; Bouali, M.; Steele, M. Using Saldrones to Validate Arctic Sea-Surface Salinity from the SMAP Satellite and from Ocean Models. *Remote Sens.* **2021**, *13*, 831, doi:10.3390/rs13050831.
37. Donlon, C. Copernicus Imaging Microwave Radiometer (CIMR) Mission Requirement Document. 2020. Available online: https://esamultimedia.esa.int/docs/EarthObservation/CIMR-MRD-v3.0-20190930_Issued.pdf (accessed on 30 October 2021).
38. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification*, 2nd ed.; Wiley: New York, NY, USA, 2001; ISBN 978-0-471-05669-0.
39. Bishop, C.M. *Pattern Recognition and Machine Learning*; Information Science and Statistics; Springer: New York, NY, USA, 2006; ISBN 978-0-387-31073-2.