

OSI SAF Visiting Scientist Program

Final Report

Sea surface temperature metrics for evaluating OSI
SAF sea ice concentration products

OSI_AS19_04

Sandra Castro

October 2019

Abstract

The aim of this investigation is to do additional quality control on the EUMETSAT OSISAF and ESA CCI sea-ice concentration climate data records (CDRs), with emphasis on the 25 km OSI-450 and SICCI-25 km, gridded sea ice concentration (SIC) products for the Arctic region. This post-processing quality assessment is based on SIC comparisons relative to an independent SST product that has been vetted for quality. The reference SST product of choice is the level 4, 5-km OSTIA SST. To facilitate comparisons between the SIC and SST products, both OSI-450 and SICCI-25 have been re-gridded to the same 5 km-pixel grid of the SST product. SIC pixels with corresponding $SST \geq 3\text{ }^{\circ}\text{C}$, are flagged for additional quality screening in both products. This set, termed the “inconsistency set,” is used to identify spurious SIC, and their sources, and to design quality control checks to remove these sources of error from the SIC CDRs. It is found that the vast majority of spurious retrievals are located along the sea ice – open water edge, coastal boundaries and subpolar marginal seas. They result from uncorrected atmospheric effects impacting low concentration retrievals over water, land contaminated microwave retrievals along coastlines, and representation errors resulting from gridding/remapping data to different spatial resolutions. Interpolation procedures associated with changes in grid resolution can smear the ice edge over neighboring pixels resulting in proportionally larger areas of underestimated retrievals and/or erroneous retrievals spreading over wider areas away from the coasts and into the open waters. A new filter with multilevel thresholding is evaluated for its skill to eliminate SST-flagged spurious SICs from the inconsistency set. The filter chain is based on multiple thresholds that check for conservation of the SIC-SST dependence within mixed pixels (i.e., $SST = f(\text{SIC})$, maximum SST and maximum SIC), a minimum distance from land requirement to eliminate false/underestimated SIC retrievals due to land-spillovers, and a valid range of standard smearing uncertainties (SSE), as this variable was found to be the most sensitive to blurring edge effects introduced by interpolation incurred when remapping/regridding to finer spatial resolutions. Each SIC product reports an estimated SSE for each pixel. Evaluation of the SSE suggested that the OSISAF uncertainties are underestimated, requiring a more aggressive thresholding. A new “energy” conservation metric, measuring the continuity between SIC and SST within a pixel, is proposed to evaluate the impact of the noise corrections on subsequent truncations of the inconsistency set after each threshold eliminated a fraction of the data. The new metric appears to be highly sensitive to small changes resulting from masking out the SIC outliers. According to the metric, OSI-450 SIC retrievals are strongly impacted by corrections addressing noise resulting from land spillovers, followed by residual atmospheric effects. The reverse order is true for SICCI. Although the filtering methodology proposed here rejected 98% of the initial inconsistency set, the remaining de-noised set has important contributions to the variability of the ice edge, with standard deviations of $\sim 12 - 15\%$. Despite common objections for using an ancillary data set for independent quality assessments, a synergistic SIC – SST retrieval appears to be beneficial for improving the accuracy of both SIC and SST products. Since both products have non-complementary sources of error, this synergism can be exploited for mutually identifying/removing residual noise in the other product. Moreover, SSTs for the de-noised SIC appear to have standard deviations ≈ 0 , suggesting that they are not introducing trends of their own.

1. Introduction

It has been observed that, during the Arctic summer months, a fraction of sea ice concentration (SIC) retrievals have rather large sea surface temperature (SST) fluctuations (sometimes up to 25 °C) when compared to an independent SST product in the same grid resolution. It is safe to assume that some of these retrievals are, in fact, false SICs resulting from residual noise contamination. Since SIC and SST have non-complementary sources of error in the passive microwave (PMW) frequencies, they can be used to identify undetected erroneous retrievals in the other product as long as the sources of those errors are non-correlated with the other product's errors. The goal of this investigation is two-fold: 1. to explore the utility of SST as a diagnostics tool to identify cases when SIC retrievals with large SSTs are true ice and when they are noise, and 2. to propose corrections to eliminate spurious SIC retrievals that remain undetected by the noise corrections implemented as part of the processing chain of the EUMETSAT and ESA SIC products.

The retrieval algorithms of PMW sea ice concentrations are sensitive to emissivity and surface temperature of sea ice, atmospheric effects, melt ponds, and ice thickness among others. They are also prone to land contamination in the coastal areas due to the large microwave antenna patterns. Large uncertainty in SIC retrievals is found in the marginal ice zone, especially during the summer Arctic months, due to atmospheric and wind roughness influences of open water areas (see for instance Ivanova et al. (2014) for a comprehensive review of accuracy and precision of SIC algorithms in various sea ice and atmospheric conditions and inter-comparison of the SIC retrieval algorithms). Methods have been developed to minimize the impact of those error sources (Tonboe et al., 2016), from corrections applied to the brightness temperatures to account for the influence of water vapor in the atmosphere and open water roughness caused by wind to weather filters. These corrections are never all-encompassing, however, as there are atmospheric corrections that are not feasible to implement (e.g., clouds) or their impact is small and therefore neglected (e.g., geolocation errors); thus, some residual error is always present (Tonboe et al., 2016). In the end, residual error is included in the SIC uncertainty estimates. Uncertainty estimates are split into two components, algorithm error and smearing error. The former component includes error sources such as residual atmospheric noise, sensor noise, and ice surface emissivity variability, whereas the second component accounts for footprint mismatch errors and smearing errors resulting from mapping the SIC to grids with finer resolution than the sensor channels' resolution.

Routine monitoring of SIC products indicate that, although the atmospheric corrections and weather filters have reduced the weather effect considerably, a substantial number of suspect data remain in the SIC products. The radiometric signature of new ice, close to open waters, is difficult for the algorithm to interpret, as are the melting ponds during the melting season, and weather effects still give spurious ice data over the open sea. Because of the synergy between PMW SIC and SST retrievals, they have been used together in the past to perform quality control in the other product. Fernandez et al. (1998), for instance, used SSM/I SIC to provide an accurate description of the ice boundary in the SST product (the 55% SIC contour as opposed to the isotherm for the freezing temperature of sea water, i.e., -1.8 °C), while SSTs were used to quality-control the SICs. In this and other studies (Nomura, 1995), SICs with SST > 1.0 °C were considered false retrievals (i.e., open water pixels classified as ice), and rejected before SICs were assimilated into the ECMWF reanalysis project to improve the SST field. We will use the

same concept to do some post-processing quality control on the EUMETSAT and ESA SIC products.

Here, we consider the $SST = 3\text{ }^{\circ}\text{C}$ as the threshold for identifying the subset of SIC retrievals in potential need of extra quality assessment. We term the extracted subset, ($SIC > 0\%$ with $SST \geq 3\text{ }^{\circ}\text{C}$), the “inconsistency set,” and use it as the basis to derive filters, based on physically realistic thresholds, to mask out spurious sea ice detections. Maps of monthly SIC means from the inconsistency set (Fig. 1) indicate that inconsistencies tend occur along coastal areas, at the ice edge, and in marginal seas. Also, inconsistencies appear to be more numerous in June, July, and August. This is confirmed by the total counts in Tables 1(SIC) and 2 (SST) summarizing the monthly mean statistics for the inconsistencies shown in Fig. 1.

Table 1. Monthly Statistics for the 2014 SICCI Inconsistencies

month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
(cnt)e+4	3.10	1.14	2.77	2.76	5.34	32.61	64.56	20.93	4.87	1.14	0.85	1.94
mean	10.05	6.32	8.51	8.90	7.75	19.26	14.65	9.56	5.99	6.16	2.97	5.29
SD	13.17	8.34	11.31	11.29	10.49	21.58	16.82	11.47	7.67	8.29	4.03	6.36
min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25%	1.01	0.56	0.92	0.93	0.71	2.19	1.81	1.05	0.64	0.58	0.31	0.62
50%	4.59	2.87	3.94	4.27	3.51	10.18	8.07	4.66	2.77	2.71	1.32	3.00
75%	13.86	8.92	11.42	12.75	10.92	30.70	21.57	14.28	8.49	8.26	3.99	8.51
max	87.65	58.30	73.98	68.75	100.0	99.83	91.99	72.61	51.76	47.08	29.89	66.91

Table 2. Monthly Statistics for the OSTIA SSTs corresponding to the 2014 SICCI Inconsistencies

month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
(cnt)e+4	3.10	1.14	2.77	2.76	5.34	32.61	64.56	20.93	4.87	1.14	0.85	1.94
mean	4.11	3.63	3.93	3.97	4.56	4.14	5.98	4.41	3.90	4.80	4.82	4.73
SD	0.84	0.49	0.70	0.73	1.69	1.59	5.51	1.88	1.01	2.11	1.85	1.20
min	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
25%	3.38	3.23	3.34	3.36	3.39	3.26	3.29	3.25	3.26	3.33	3.23	3.71
50%	3.96	3.53	3.79	3.83	4.04	3.65	3.75	3.64	3.60	3.82	3.91	4.49
75%	4.67	3.96	4.40	4.44	5.12	4.33	4.80	4.64	4.14	5.60	6.26	5.63
max	6.67	5.82	6.24	7.19	14.37	16.35	24.12	17.77	10.32	11.89	10.33	7.79

Monthly maximum value composite (MVC) maps of total pixel counts for the 2014 SICCI inconsistencies are shown in Fig. 2 (The top panel shows the map in full color scale; the bottom panel shows the same map, but zooming in at smaller counts to better appreciate the structure of pixels reputedly flagged as containing an inconsistency). A MVC reduces the daily inconsistency SIC maps to a single map in which each pixel is assigned the total number of times during the year it was flagged as having an inconsistent SIC retrieval. If the location of the inconsistencies migrated with time, the colored regions should have spread over a wide latitude range, but Fig. 2 indicates that they do not appear to move much at this grid resolution (5-km spacing). This seems to be in contradiction with Lavergne et al. (2019) according to which the position of the daily sea ice edge can fluctuate substantially within a month, hence requiring that the climatological masks used to delineate the monthly maximum sea ice extent within the SIC

CDRs be expanded to a buffer zone of 150 km in the Northern Hemisphere when choosing training open water areas in the SIC algorithms. Of course, the values highlighted in Fig. 2 indicate recurrence of inconsistencies within a pixel-size area, so the contradiction stems from the fact that the inconsistencies in the ice edge do not migrate with the ice edge, but remain quasi-stationary. Maps of MVC of inconsistency counts suggest that the inconsistencies have preferred locations in terms of latitude, and are more abundant in the Hudson Bay, the south-east coast of Greenland, the Fram Straight, the Kara Sea and the Baltic Sea. This is an indication that the SIC inconsistencies are coherent in space and time, suggesting that they are associated with errors that are “systematic” in nature. In other words, they do not appear to be random at this point. The next step is then to identify the sources of the systematic errors causing some of the inconsistencies.

At the heart of this issue is what constitutes a valid joint SIC-SST retrieval and what is noise. For that we need to answer what is the maximum SST that is physically viable within a mixed pixel ($SIC > 0\%$). We will adopt a statistical approach in which marginal probability density functions of SIC related parameters, as well as a joint probability density function of SIC and SST within the marginal ice zone, are used to define thresholds for filtering residual noise-contaminated SIC retrievals. In Section 2 we describe the products used in the inconsistency analysis. In Section 3 we try to determine the primary sources of the errors by looking at individual examples of inconsistencies with specific spatial structures. In Section 4 we narrow in on the variables used to screen the inconsistency set for spurious retrievals and use pdfs to try to identify ranges of the specified variables that can be used in thresholds to eliminate the noise. In Section 5, we finalize the choice of the thresholds and evaluate the impact the different filters have on the inconsistency set. A summary of findings and recommendations is given in the conclusions.

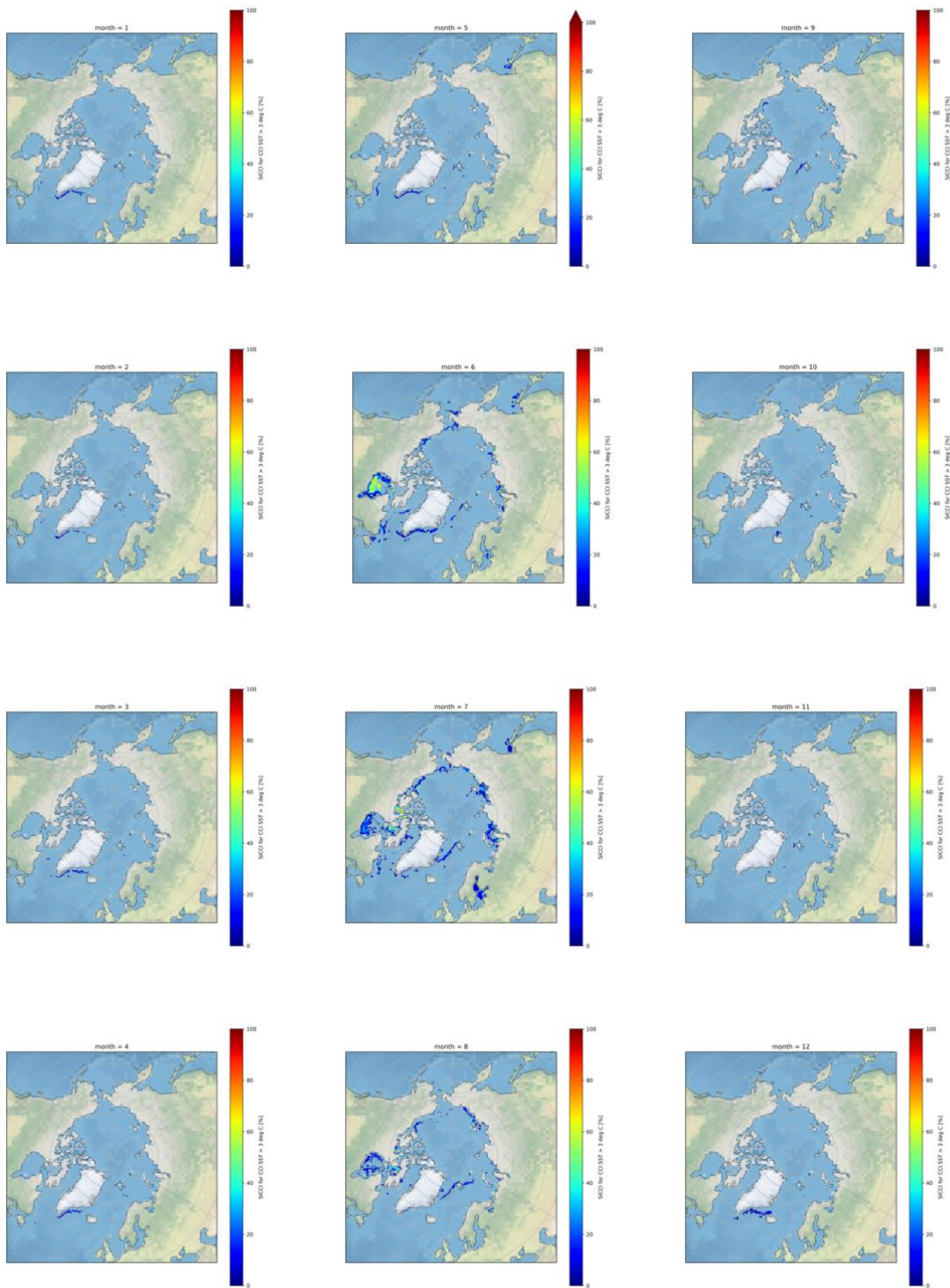


Figure 1. Monthly mean SICCI sea ice fraction inconsistencies for 2014

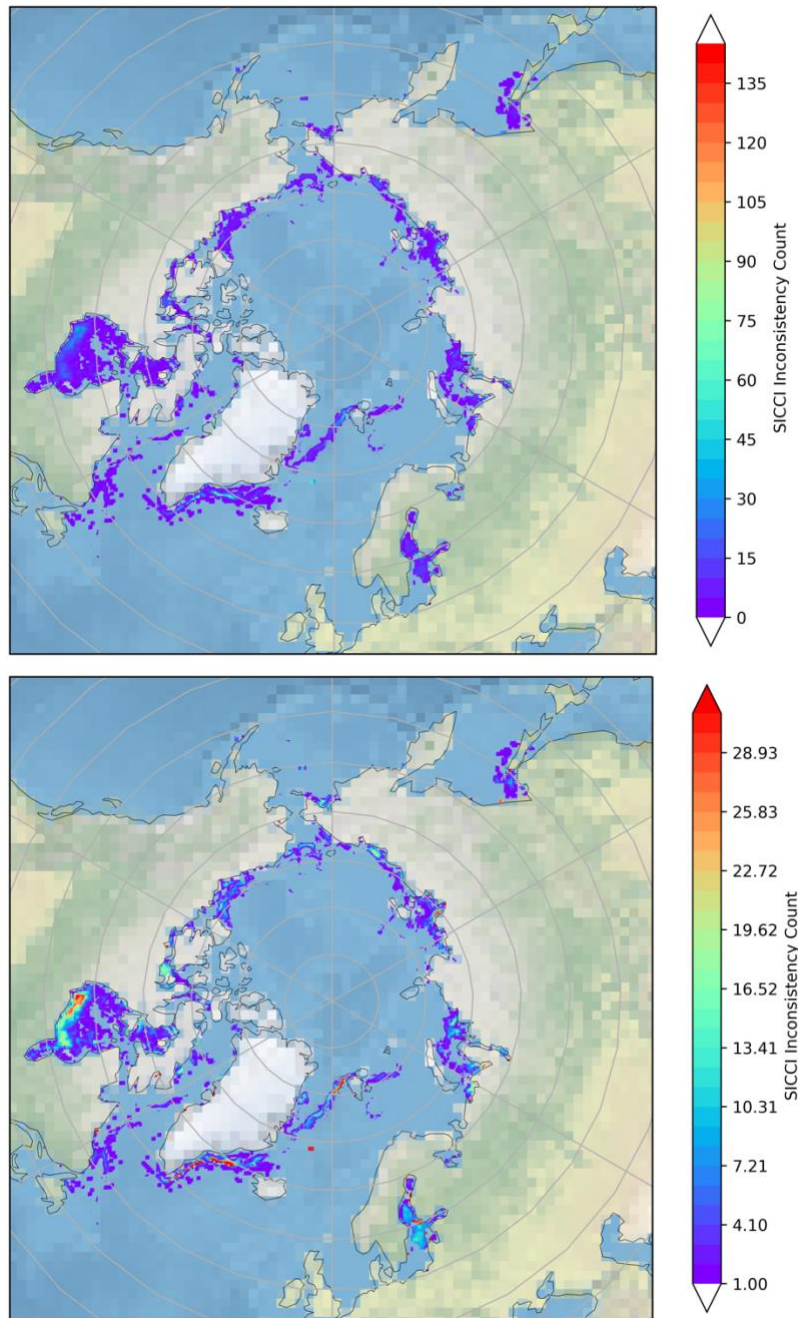


Figure 2. Map of SICCI inconsistency count for 2014. The color of the pixels indicates how many times in the year individual pixels were flagged as having a SIC inconsistent with an SST. Top map: Full range; bottom map: pixels with less than 30 counts. These maps indicate that the inconsistencies have preferred locations in terms of latitude, and are more abundant in the Hudson Bay, the south-east coast of Greenland, the Fram Strait between Greenland and Svalbard, the Kara Sea and the Baltic Sea.

2. Description of Data Sets

2.1 Sea ice concentration products

Two different SIC products have been used in this study: the European Space Agency Climate Change Initiative (ESA CCI) SIC product, SICCI and the EUMETSAT OSISAF SIC product, OSI-450 (Lavergne et al., 2019). While both SIC products use the same algorithms and processing chains, they differ in spatial resolution (footprint sizes of the sensors used in the algorithms). The OSISAF SIC product uses information from the coarse resolution (30 – 60 km) passive microwave (PM) instruments available since October 1978: the Satellite Multichannel Microwave Radiometer (SMMR), the Special Sensor Microwave / Imager (SSM/I), and the Special Sensor Microwave Imager and Sounder (SSMIS). The SICCI product is based on the medium resolution (15 – 25 km) PM sensors: the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E; 2002-2010 on board the US Aqua satellite) and Advanced Microwave Scanning Radiometer 2 (AMSR2; 2012-today on board the Japanese GCOM-W1 satellite). Here, we used data from the version 2, SICCI-25 km and OSI-450 products for 2014. Both of these products are released on Equal Area Scalable Earth 2 (EASE2) grids with 25×25 km resolution, but were resampled to $0.05^\circ \times 0.05^\circ$ regular longitude-latitude grids to match the resolution of the reference SST product. Fig. 3 shows 0.05° -resolution SIC maps above 50°N from SICCI (top-left) and OSI-450 (top-right) for 30 June 2014. Ice concentrations are expressed in percentages. Included in the products are maps of estimated uncertainties at each grid cell. These uncertainties give a quantification of the remaining noise from the two uncertainty components: the smearing standard error (SSE), which measures the increase in uncertainty due to mismatching spatial dimensions and the algorithm uncertainty; the total standard error (TSE), which is the sum of the previous two, is also provided with the products. The smearing uncertainty is largest, up to 40% SIC, at the ice edge and low concentrations, near 0% SIC, in areas where all contributing satellite footprints cover the same SIC (e.g., open water). It is the dominant term in the spatial domain of the inconsistencies, and thus it is used to help diagnose the validity of the SIC retrievals. Maps of SSE corresponding to SICCI and OSISAF for 30 June 2014 are shown in the center row of Fig. 3. A detailed description of the SIC products and the underlying algorithms can be found in Lavergne et al. (2019).

2.2. Sea surface temperature product

To be able to exploit the synergy between SIC and SST for additional quality control of the SIC products we need a high-quality SST reference. SST analyses (level 4 products) are spatially complete global maps obtained from blending data from multiple platforms (microwave and infrared sensors, buoys, etc.) on regular global grids and using some form of interpolation to fill in the gaps. In this study, we use the version 2, Operational Sea Surface Temperature and Ice Analysis (OSTIA) foundation SST (Donlon et al., 2012). This is a global, daily, optimally interpolated analysis produced by the UK Met Office. Although it is distributed in $0.05^\circ \times 0.05^\circ$ resolution grids, the smallest analysis feature resolution is based on the correlation length scale of 10 km. A foundation SST is an estimate of the surface temperature of the ocean devoid of diurnal warming variability. To minimize the influence of the diurnal cycle, the analysis merges nighttime data and daytime SST retrievals filtered for wind speeds > 6 m/s. For this study, the SIC products, originally at 25 km resolution, were re-gridded onto the same resolution of the OSTIA SST, i.e. from 25 km down to 0.05° (~ 5 km). This SST product is fully vetted as it is

used operationally as a boundary condition for all weather forecast models at the Met Office and at European Centre for Medium-range Weather Forecasting (ECMWF). The product is reported to have zero mean bias and an accuracy of ~ 0.57 K compared to in situ measurements. An example of this product is shown in Fig. 3, bottom-left panel for 30 June 2014.

2.3. Distance-from-land mask

Due to the coarse resolution of the PM sensors used, especially SMMR, SSM/I, and SSMIS, the brightness temperature data are influenced by land emissivity several tens of kilometers away from the coastline, and SIC retrievals tend to be overestimated near land. In other cases, open water areas are interpreted as ice. Even though the version 2 CDRs undergo extensive corrections for land effects both at level 2 (the brightness temperatures) and level 3, not all noise due to land contamination effects is successfully removed; hence, we use a distance-from-land mask to help sort out remaining SIC retrievals affected by land from the inconsistency set. The distances (Fig. 3, bottom right) are provided in kilometers and mapped to a 0.05° -resolution grid. The mask was produced by DMI.

2.4. SIC- and SST-derived products

In addition to the SIC and SST themselves, estimates of their temporal and spatial variability were also considered as diagnostic tools to help with the screening of the inconsistencies. The spatial variability of SIC and SST was estimated from the Sobel operator, the reasoning being that most of the inconsistencies are located at the ice edge, and the 3×3 convolution kernels of the Sobel operator respond maximally to edges, emphasizing high frequency (i.e., high variability) regions. It returns the absolute gradient magnitude at each pixel. Temporal variability, on the other hand, was estimated by two proxies: as the standard deviation (SD) of the parameters (SIC and SST) over a 3-day rolling window and, because a sample size of 3-point measurements might not give statistically significant SDs, as the statistical range over the same 3-day rolling window. This is the absolute value of the difference between the maximum and minimum value (i.e., $|\text{MAX} - \text{MIN}|$) of the variables, for each pixel, over a $[-1; +1]$ day rolling window. As the range depends solely on two-point measurements, the maximum and the minimum values, it is the de facto measurement of dispersion for small data sets. Interestingly enough, the spatial patterns of these two proxies for temporal variability were exactly the same, and they varied only in their magnitudes. Maps of the spatial and temporal gradients for SICCI, OSISAF, and OSTIA for 30 June 2014 are shown in Fig. 4.

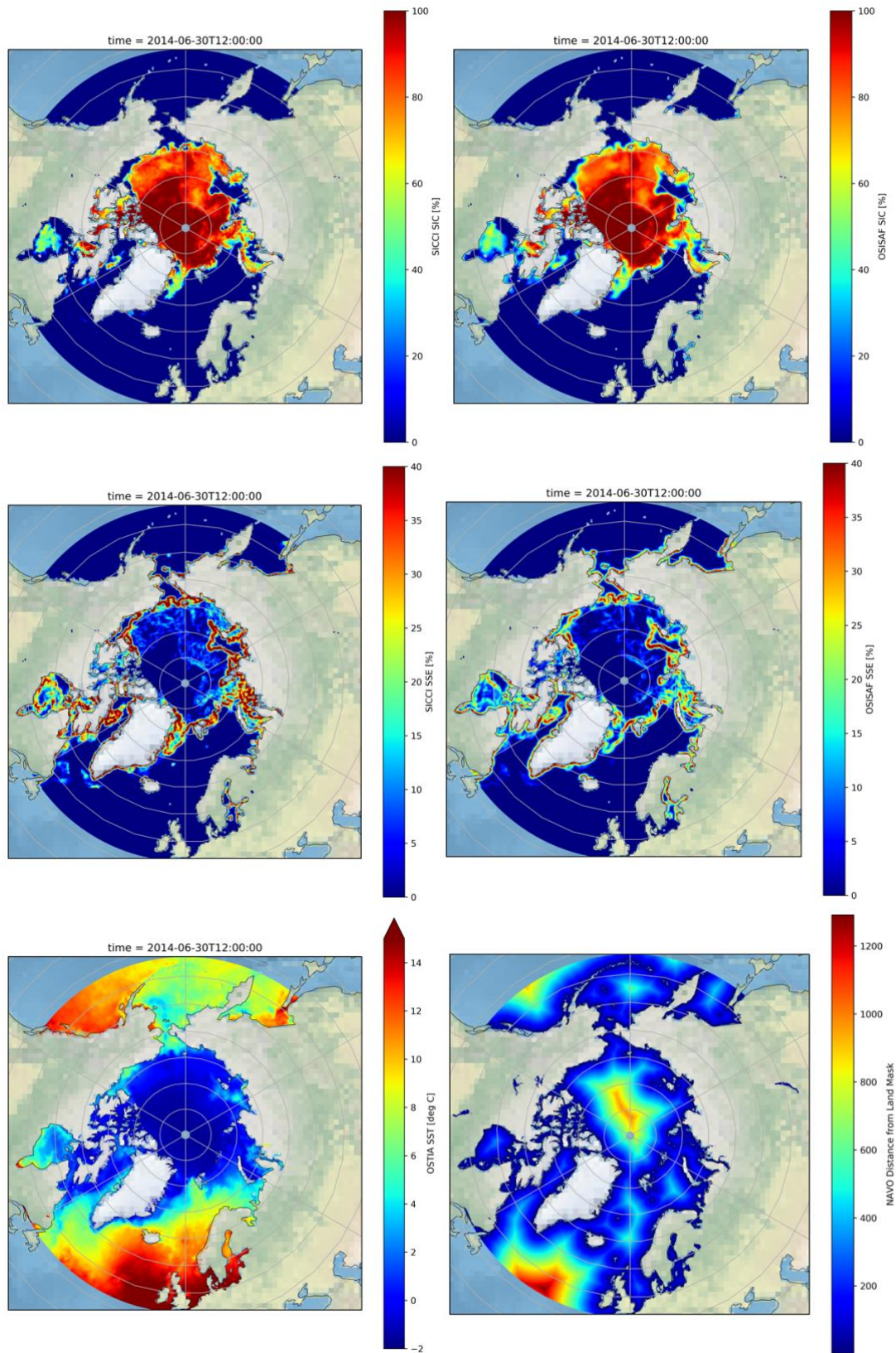


Figure 3. SIC products (top-left: SICCI; top-right: OSISAF) and related uncertainties (center row). Ancillary products: OSTIA SST (bottom-left) and Distance from land (bottom-right).

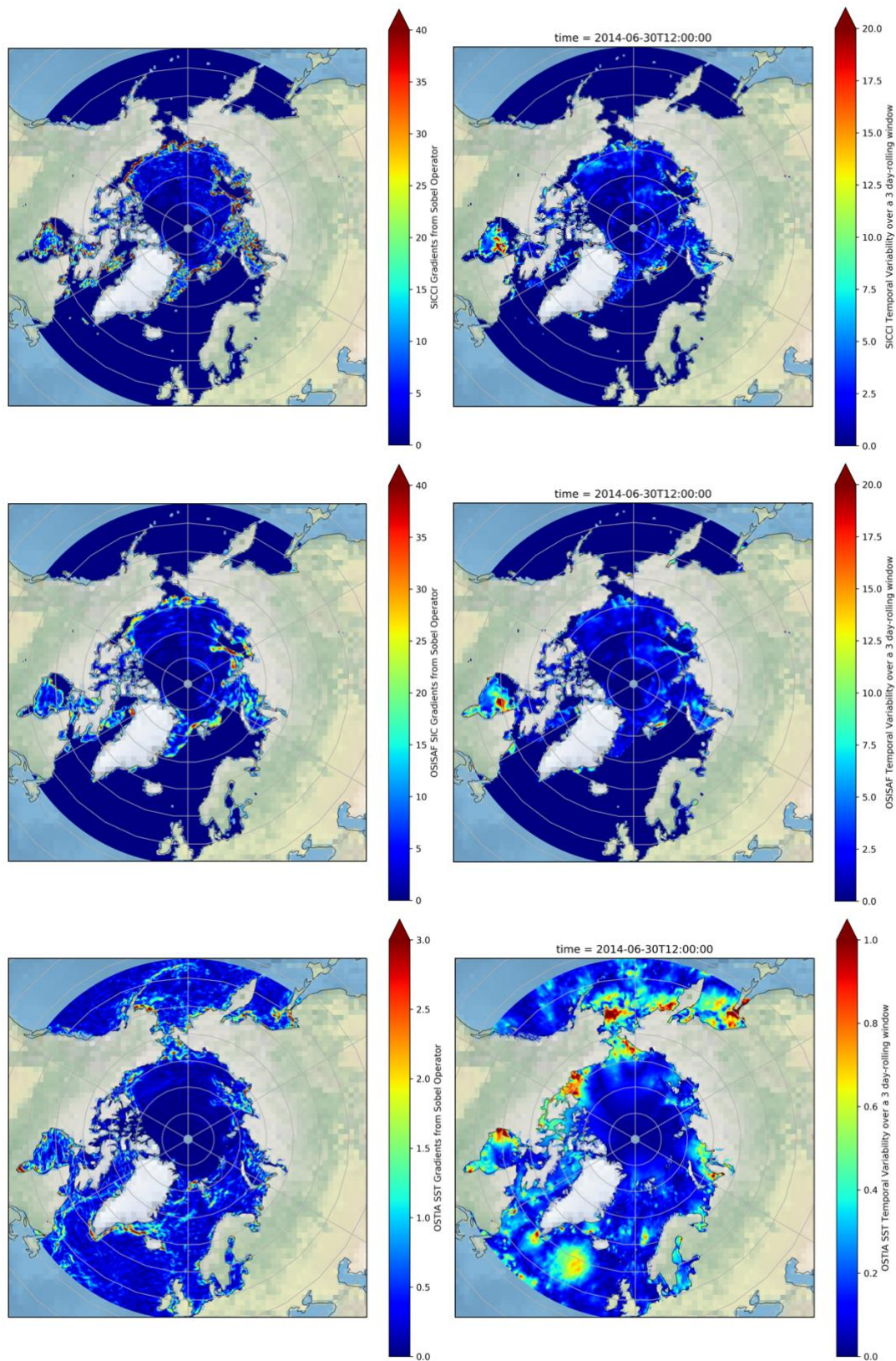


Figure 4. Spatial (left) and temporal variability (right) for: top: SICCI, middle: OSISAF, bottom: OSTIA SST

3. Characterization of Inconsistencies

The key questions here are: Are the inconsistencies real or are they flagging a spurious SIC retrieval? If the latter, how can they be filtered out? In order to propose corrections for residual noisy SIC retrievals we need to know first what is causing them. In order to answer these questions, we look at the summer inconsistencies in more detail as shown in Figs. 5 and 6 for June and July, respectively, and zoom in on specific features. Selected features were chosen taking into consideration geographical regions where inconsistencies seem to have preferential locations as indicated by the MVC of inconsistency occurrences for 2014 shown in Fig. 2.

3.1. Inconsistencies due to failure of the SIC algorithm's atmospheric correction: The Baltic Sea case

The maps of monthly MVC of SIC inconsistencies for July (Fig. 6, top row) and August (not shown) reveal two distinct regions for both SICCI and OSISAF where there should not be any ice for that time of the year, yet the SIC is above 20 – 30%. These are the Baltic Sea (Gulfs of Bothnia, Finland and Riga) and the area around the Amur river outlet into the Tatar Strait in the Sea of Okhotsk. Lending support to the claim that these inconsistencies are indeed spurious SIC retrievals is the fact that OSTIA SSTs for the same pixel grids (Fig. 6, mid row) show SSTs in excess of 20 °C, consistent with sub-polar, open water, summer temperatures (see maximum SST for July in Table 2). A close-up view of these inconsistencies for the Baltic Sea is shown in Fig. 7 for 28 July 2014. Not only are these SICs wrong (the correct value should have been SIC = 0% for open water), but there is nothing in the associated uncertainties to indicate that these retrievals are impacted by residual noise. In fact, the close-up views for the total errors in the Baltic Sea example (Fig. 7, bottom row) show that the largest errors associated with these retrievals, away from the coasts, are quite low (SSE < 10 – 15%) for both SICCI and OSISAF products. These false ice concentrations are most likely due to uncorrected atmospheric effects.

Typical corrections for atmospheric contributions to the SIC uncertainties are implemented at the brightness temperature (T_B) level and include the use of a radiative transfer model (RTM) forced with numerical weather prediction (NWP) reanalysis fields (Andersen et al., 2006; Ivanova et al., 2015; Tonboe et al., 2016; Lavergne et al., 2019). The two atmospheric sources with the greatest impact to the SIC retrievals are water vapor and liquid water content (LWC), since their microwave signature is similar to that of sea ice. Retrievals of LWC, however, are particularly challenging and current NWP LWC fields lack the accuracy required to be used in RTM-based microwave atmospheric corrections (e.g., Lu et al., 2018; Ivanova et al., 2015). Hence, most SIC products, including the ones considered here, go uncorrected for the influence of LWC and cloudy atmospheres. Confounding this issue is the fact that LWC can be retrieved in the PMW frequencies of 18 and 37 GHz where clouds are semi-transparent, but the OSI-450 and SICCI-25 km CDR also use the same MW frequency space (19V, 37V, 37H). Since the reflectivity of wet ice is similar to that of LWC, this can lead to erroneous classifications of open water for wet ice during the ice melting season if the cloud cover happens to be made of the “right type” of cloud, in this case dense, wet clouds. To ameliorate deficiencies in the atmospheric correction to the T_B , the RTM simulations are complemented with a weather filter based on the gradient ratio of two microwave channels selected based on their sensitive to some atmospheric influence. The SIC products considered here use an “open water” weather filter (OWF) based on the 19V and

37V GHz channels to remove false ice (due mainly to LWC, given the sensitivity of the channels used) in open water: erroneous SIC are set to $SIC = 0\%$ (i.e., they are reverted to open water).

PMW SSTs on the other hand are measured between 6 – 7 and 10 – 11 GHz. The spectral signature of the SST at these low frequencies is quite different from the surface roughness and the atmosphere, making the separation of the SST signal from the atmospheric contribution more effective. This is a clear example in which comparing a SIC product to a SST product is a good diagnostic tool for detecting spurious retrievals since their sources of error are independent and hence, they complement each other. Fig. 7 indicate that spurious SICs that can be traced back to failings of the SIC algorithm's atmospheric correction can be easily flagged/eliminated with a filter based on the maximum SST within the marginal ice zone (MIZ). The question then becomes, what is the range of physical SSTs in close proximity to ice? We aim to answer this question based on a joint SIC-SST distribution for the MIZ (Section 4).

3.2. Inconsistencies due to gridding/interpolation effects during the SIC CDR processing chain: The Fram Strait case

One of the most striking features of the MVC maps for June and July (Figs. 5 and 6) is the fact that many inconsistencies seem to closely follow the ice edge in long, narrow filaments. Zooming in on one of those areas for the Fram Strait on 18 July 2014 (Fig. 8), one can see that the SIC and the SST inconsistency criteria overlap in a narrow region of open water next to the ice edge, resulting in a filament-type structure that traces the $SIC = 0\%$ contour (Figs. 9 and 10, for SICCI and OSISAF respectively). The diagnostic variables mapped in Fig. 8 show the SIC related products extending all the way to the $SIC=0\%$ contour, whereas the SST related products have been cut off at the $3\text{ }^{\circ}\text{C}$ -isotherm, since this is the lower boundary that defines the inconsistency set. These two products overlap in the narrow filament shown in Figs. 9 and 10. Close inspection of this filament gives a good idea about its source. As can be appreciated in the SIC and SST inconsistency maps (top two panels in Figs. 9 and 10), the filament displays a pixelated or jagged edge on its open-water side and a smooth edge on its ice side. Moreover, in the SIC maps (top-left panels) the concentrations increase linearly, from 0% at the jagged edge to $\sim 40\%$ at the smooth edge, whereas the SSTs (top-right panels) increase linearly, directly opposite from the spots on the sides with higher concentration, from $3\text{ }^{\circ}\text{C}$ at the smooth edge to $\sim 5\text{ }^{\circ}\text{C}$ at the jagged edge. This is indicative of interpolation edge effects that result from remapping a coarse resolution pixel onto a fine resolution pixel with conventional linear interpolation techniques. Interpolation edge effects affect mixed pixels where there is a sharp edge/transition between two surfaces and can adopt different forms such as blurring/smearing and blocking. In this particular example, the SST edge (i.e., the $3\text{ }^{\circ}\text{C}$ -contour) is smooth because the level 4 (gridded and interpolated) OSTIA product is in its native 5 km -resolution grid. The level 4 SIC products, however, were generated at 25 km -resolution grids and then resampled to 5 km to be on the same grid as the SSTs. Resampling involves interpolation, and also possibly extrapolation, as non-zero ice concentration values are projected into regions with open water. Blocking is an effect of nearest-neighbor interpolation schemes in which many subpixels in the zoomed image are assigned the same value of the coarser pixel. Pixel replication at the ice edge gives it its blocky appearance, hence its name. The non-zero ice concentration values are extended throughout the effective coarser resolution pixel, while the higher resolution SST products can resolve open water values within that area. So, it appears that blocking in the

vicinity of the 0%-edge is most likely giving the resampled SIC its jagged appearance. This is visible not only in the open water edge of the inconsistency filament, but also in all the SIC related products shown in Fig. 8.

Clearly, inconsistencies along the ice edge are being affected by interpolation artifacts. What is not so clear is when in the processing chain they were introduced. In addition to the interpolation just mentioned to put the SIC and the SST products on a common grid to facilitate this analysis, the SIC products underwent other resampling/interpolation steps during the processing chain; that is, when transitioning from swath (level 2) data to gridded (level 3) data and when filling the gaps in the gridded data (level 4). In addition to these interpolation steps, the footprint sizes of the channels used in the mapping of the ice concentrations are uneven, ranging from ~50 – 70 km for the 19.3 GHz to ~30 km for the 37 GHz for the SMMR, SSM/I, and SSMIS sensors used in OSISAF, and from ~25 × 15 km (18.7 GHz) and ~15 × 10 km (37 GHz) for AMSR-E and AMSR2 used with SICCI. This aspect alone introduces footprint mismatch errors (Tonboe et al., 2016). Confounding this issue is the fact that the level 4 OSISAF products are delivered in a grid finer than the sensor's footprint resolutions (25 km vs. 30 – 70 km), which results in blurring/smearing in the gridded OSISAF SICs. This is evident in Fig. 8, where the OSISAF SICs (top-right) look smeared or blurred relative to the sharp SICCI concentrations (top-left). In other words, the small-scale features apparent in SICCI are lost in OSISAF.

Both SICCI and OSISAF products provide smearing uncertainties, which estimate the combined effect of the footprint mismatch error and the smearing error incurred when resampling to finer resolutions than the sensor's footprint. The smearing uncertainty is greatest where sharp spatial gradients occur, typically at the sea-ice edge. This appears to result in a different sort of inconsistency here, as OSISAF inconsistencies should inherently have increased smearing uncertainties than SICCI's, i.e. more representativeness error coming from sensors with a much larger footprint mismatch and a wider gap between sensor and the target grid resolutions. Yet, the bottom two plots of Figs. 5 and 6, for the June and July SSE MVCs, show that the OSISAF inconsistencies have the largest SSE uncertainties constrained to a narrow distance from the coast (bottom right panels), whereas the SICCI uncertainties have larger SSEs spreading over much wider distances from the edges (bottom left panels). In reality, the SICCI 25 km inconsistencies have sharper edges since they are not subject to the interpolation smearing inherent in OSISAF (see for instance the MVC SSE maps shown in Figs. 8 for SICCI and OSISAF). Therefore, since the smearing uncertainty is larger where spatial gradients occur, SICCI ends up with a larger smearing uncertainty than OSISAF's. This can be easily explained from the SSE function used in the current SIC CDRs, with $SSE = K \times (SIC_{max} - SIC_{min})_{3 \times 3}$, i.e., the scalar K, whose value depends on the diameter of the field of view of the instrument channels used for the SIC computation and the spatial spacing of the level 4 grid, is amplified by the difference between the maximum and the minimum SIC in a 3x3 grid box centered on each pixel. Whereas in theory $K_{OSISAF} > K_{SICCI}$, $K = 1$ for all EUMETSAT SIC products (Lavergne et al., 2018). SSE is then reduced to the 3x3 gradient filter, hence the SSE sensitivity to sharp gradients near the ice edge. Since $(SIC_{max} - SIC_{min})_{3 \times 3} \rightarrow 0$ for OSISAF in lieu of its smooth edges (not much SIC contrast), the result is $SSE_{OSISAF} < SSE_{SICCI}$. This value for OSISAF is likely underestimated on the fine grid. Confounding this issue is the additional re-projection and interpolation performed on the level 4, 25-km SIC products to the 5 km OSTIA SST grid. This

bilinear interpolation most likely blurred the SICCI uncertainty gradients across larger distances away from the boundaries as pixels with high intensity values relative to the surrounding neighborhood get smeared across the area by the interpolation. For OSISAF, the pixel intensity difference relative to its neighbors is small, and thus would not be affected as much at this latter interpolation stage.

Back to the filaments in Figs. 9 and 10, the discussion above explains why the OSISAF filament (Fig. 10, left panel) has lower SSE than SICCI's, except in a couple of spots closer to land as indicated by the panel to its right, and why the SSE is larger over a wider range of SICCI inconsistencies near the coasts, as shown in the bottom row panels in Figs. 5 and 6. It also explains why the diagnostic variable that best correlates with the SIC inconsistencies near the ice edge where interpolation artifacts are prominent is the smearing uncertainty (SSE) and, to a lesser degree for OSISAF at least, the distance from land (second row in Figs. 9 and 10). Interpolation artifacts at the ice edge do not necessarily mean that the SIC retrievals are incorrect at their native resolution. The error arises when estimating (guessing) values at a sub-pixel resolution (the higher resolution SST grid) at which no real microwave information is available. They are cosmetic glitches in a sense, that are largely unavoidable when one encounters abrupt boundaries. Thus, filters based on SSE won't entirely eliminate interpolation artifacts; they will, however, correct for blurring of the edges, especially with SICCI, resulting from converting all the products to the same resolution. Since the OSISAF SSE in the 5-km grid is not smeared too far away from the coast, the distance filter should eliminate most OSISAF SIC inconsistencies with high SSE adjacent to land. The full region influenced by misclassification can potentially be greater than that of the elevated SSE values.

3.3. Inconsistencies due to residual land contamination in the MW SIC retrievals

Further considering the region from the previous example, notice that in addition to the inconsistency feature due to gridding (the long filament in Figs. 9 and 10), the OSISAF inconsistency map in Fig. 10 reveals many more outliers around the coast of Svalbard that are not present in the SICCI map (Fig. 9). This is also made explicit by the MVC maps for June and July (Figs. 5 and 6) where the OSISAF composites show a significantly greater number of inconsistencies along the coasts. It is well known that in coastal areas, the emissivity of land is of the same order of magnitude as that of the sea ice, and both have much higher emissivity than the water. This, combined with the large beam width of the microwave sensor antenna pattern, results in land-to-ocean spillovers in which coastal signals can be mistaken for sea ice during the summer months. The severity of this impact depends on the footprint of the sensor and the satellite revisit time. The SIC products evaluated here use a statistical correction (Cavalieri et al., 1999) to reduce land-contaminated pixels in the vicinity of coastlines during the summer time, but the June and July inconsistency maps for OSISAF (Figs. 5 and 6) indicate that the land spillover correction being used is leaving behind a significant number of spurious SIC retrievals. Hence, the need for a post-processing filter to mask out spurious SIC within a minimum distance from the shore when the MIZ extends all the way to the coastline.

3.4. Inconsistencies with SIC > 80%: The Kara Sea case

The monthly statistics for the SICCI inconsistencies (Table 1) indicate the presence of a small fraction of SICCI inconsistencies with concentrations near the 100%-ice regime for the months

of May, June, and July (see maximum SIC values for those months). This suggests the presence of open water in pixels that appear to be simultaneously pack ice. In other words, this is a conflicting SIC vs. SST combination. The question is, are these pixels supposed to have SIC = 100% (all ice) or are the ice concentrations being overestimated? The latter alternative seems more likely as the former is not supported by the corresponding moderate SSTs. The MVC maps for the June 2014 inconsistencies (Fig. 5) show that the highest concentrations are located in the Gulf of Ob and coastal areas of the Kara Sea. As for the source of the errors, Tonboe et al. (2016) explains that land spillovers in coastal areas with intermediate ice concentrations will result in overestimation of the SIC retrieval, which “fits the bill” quite nicely here as pixels with overestimated SIC are all near land. Since the overestimated SIC retrievals are also the result of land contamination, one would think that a filter based on a minimum distance requirement as explained in Section 3.3 should take care of these inconsistencies. That should be the case if the overestimated SIC retrievals area contained within the minimum distance requirement stipulated in the filter, but if they extend passed it, another filter might be needed to constrain SIC retrievals within the MIZ by a physically meaningful maximum SIC. As with Section 3.1, the maximum concentration possible within the MIZ should come into focus with the aid of a joint SIC - SST distribution for the MIZ.

3.5. When the inconsistencies are not inconsistent after all: The Hudson Bay case

It is reasonable to expect that some SICs in the inconsistency set are valid; that is, we are open to the possibility that, given the scale of the footprints, a fraction of the retrievals for which the mean SIC > 0% and SST ≥ 3 °C corresponds to physically meaningful conditions observed in the MIZ. The Hudson Bay is an interesting case because a significant portion of the summer SIC retrievals inside the bay are flagged as “inconsistent” when compared to the OSTIA SST product. This can be appreciated in more detail in the SICCI and OSISAF inconsistency maps for the Hudson Bay shown in Figs. 11 and 12 corresponding to 30 June 2014. Data extractions, on the same day, for the whole set of variables used as diagnostics tools are also shown for the grid cells flagged in the SICCI (Fig. 11) and the OSISAF (Fig. 12) inconsistency sets. There is nothing to suggest that there is something wrong with the SIC retrievals flagged by the inconsistency criteria. The SIC inconsistencies from both products show an extensive marginal ice zone that expands from the center of the bay all the way to coast. Different patterns are seen in between products with SICCI showing a well-defined ice edge along the shoreline and SICs increasing linearly with distance from the coast. For the OSISAF product only the eastern shore of the bay shows signs of open water with intermediate to high concentrations in the rest of the MIZ that extend all the way to the coast, which is actually a more realistic representation of the typical spatial distribution of ice in the bay. SSTs also seem quite reasonable with most values less than ~5 °C, with the exception of some localized warm spots along the coast indicating river outlets. If we are correct in our assumption that these are valid SIC retrievals, then they should prevail after all the filtering is done.

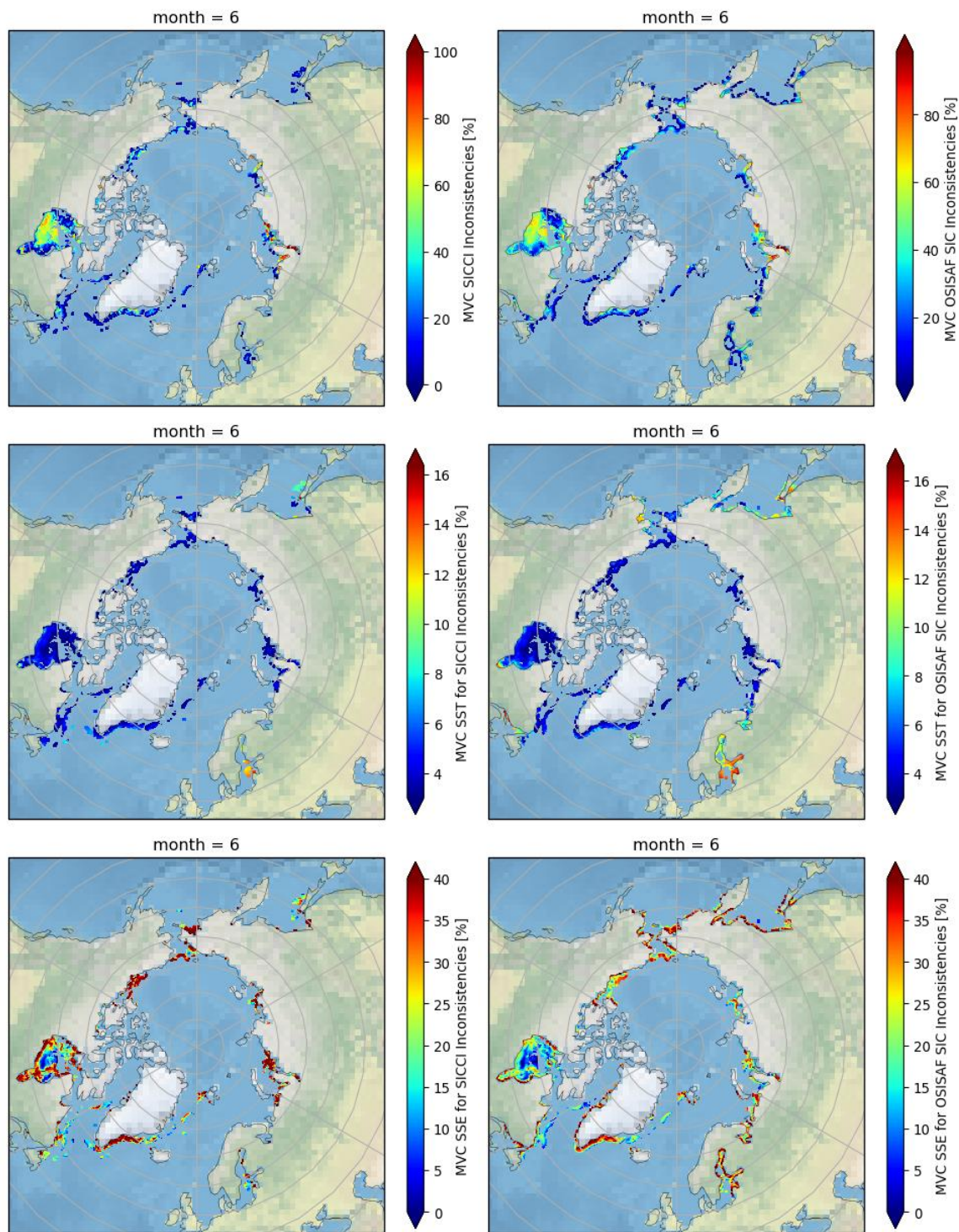


Figure 5. Monthly Maximum Value Composite (MVC) for SICCI (left column) and OSISAF (right column) inconsistencies for June 2014. Top row: SIC, Middle row: OSTIA SST, Bottom row: SSE.

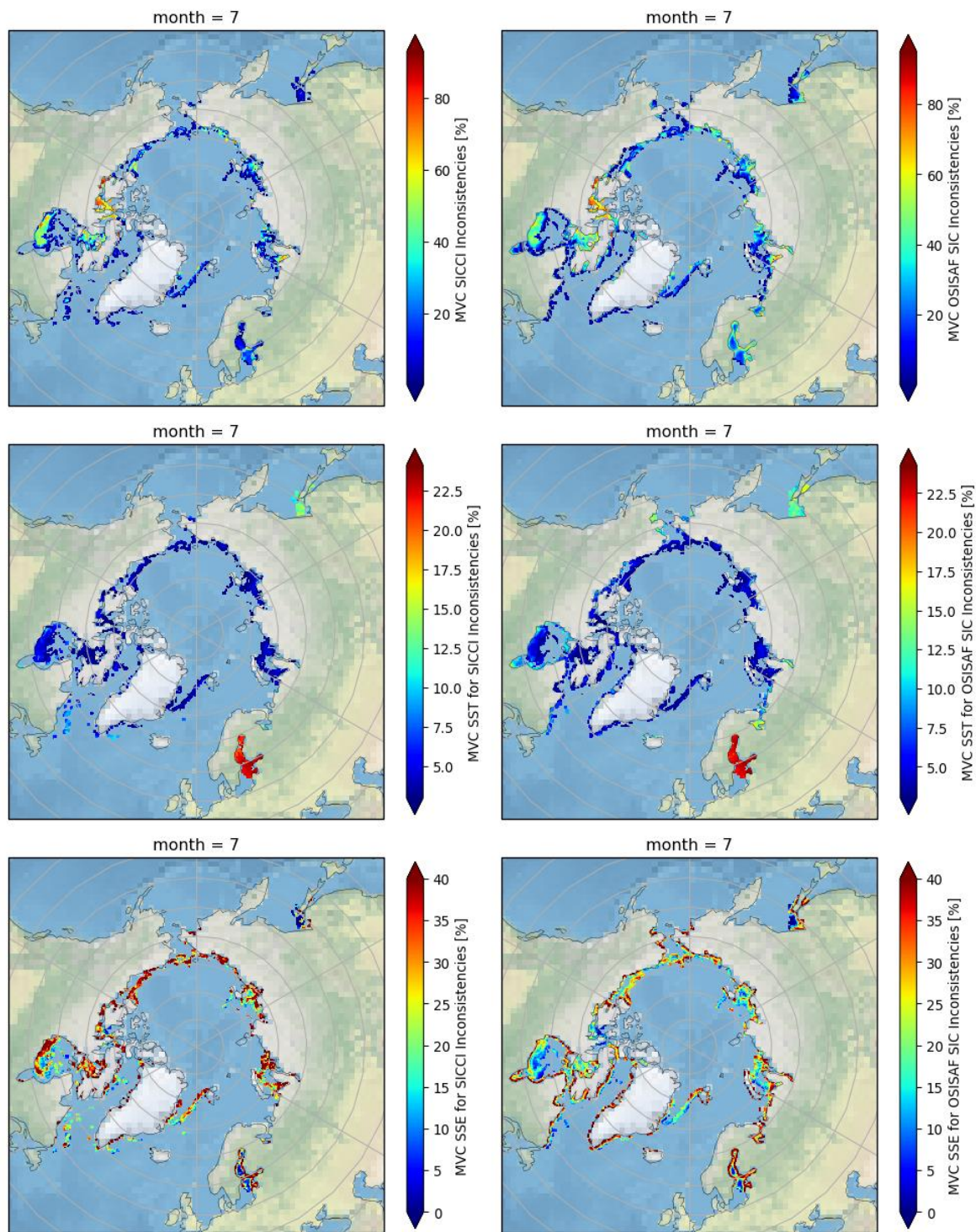


Figure 6. Monthly Maximum Value (MVC) Composite for SICCI (left column) and OSISAF (right column) inconsistencies for July 2014. Top row: SIC, Middle row: OSTIA SST, Bottom row: SSE

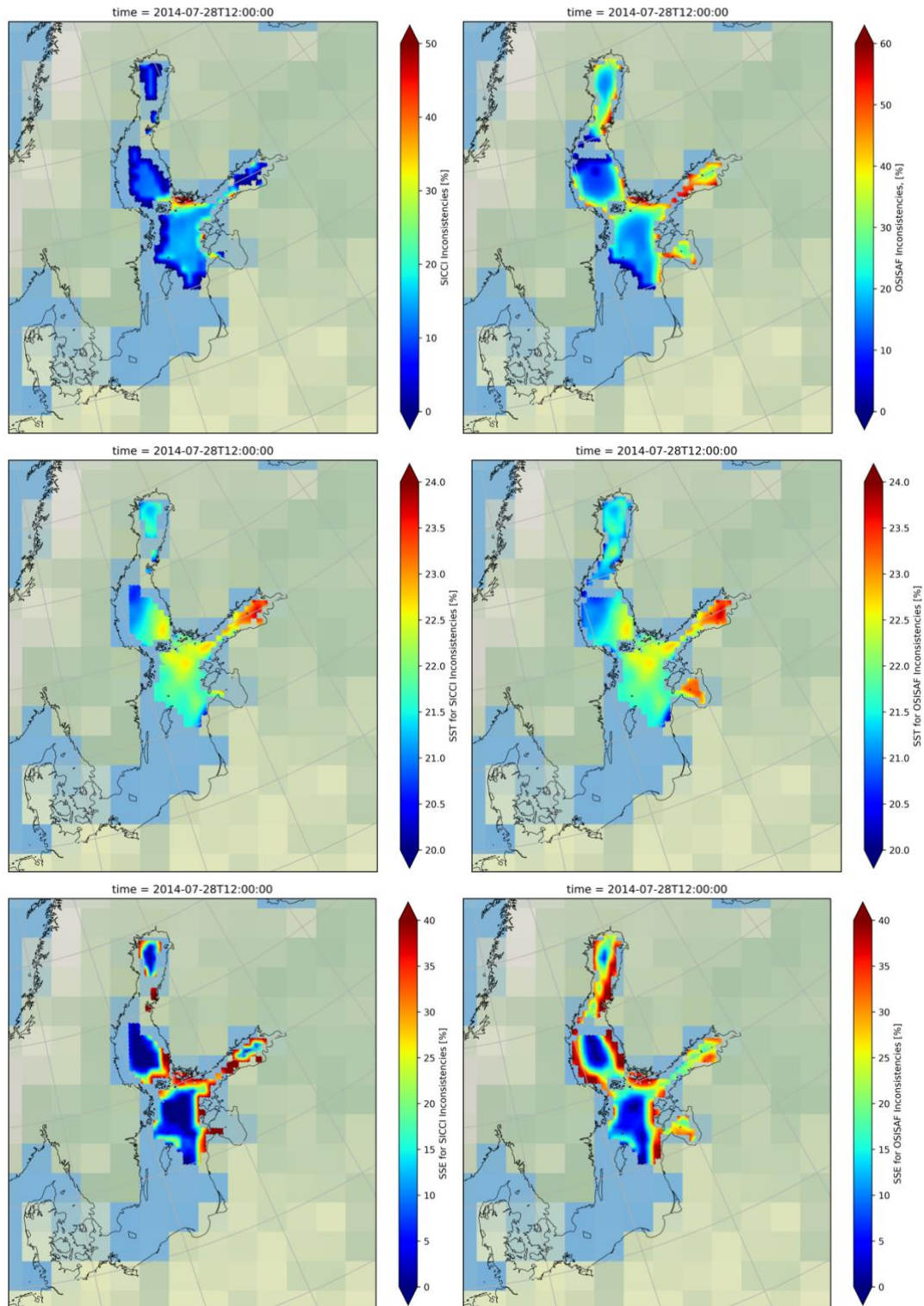
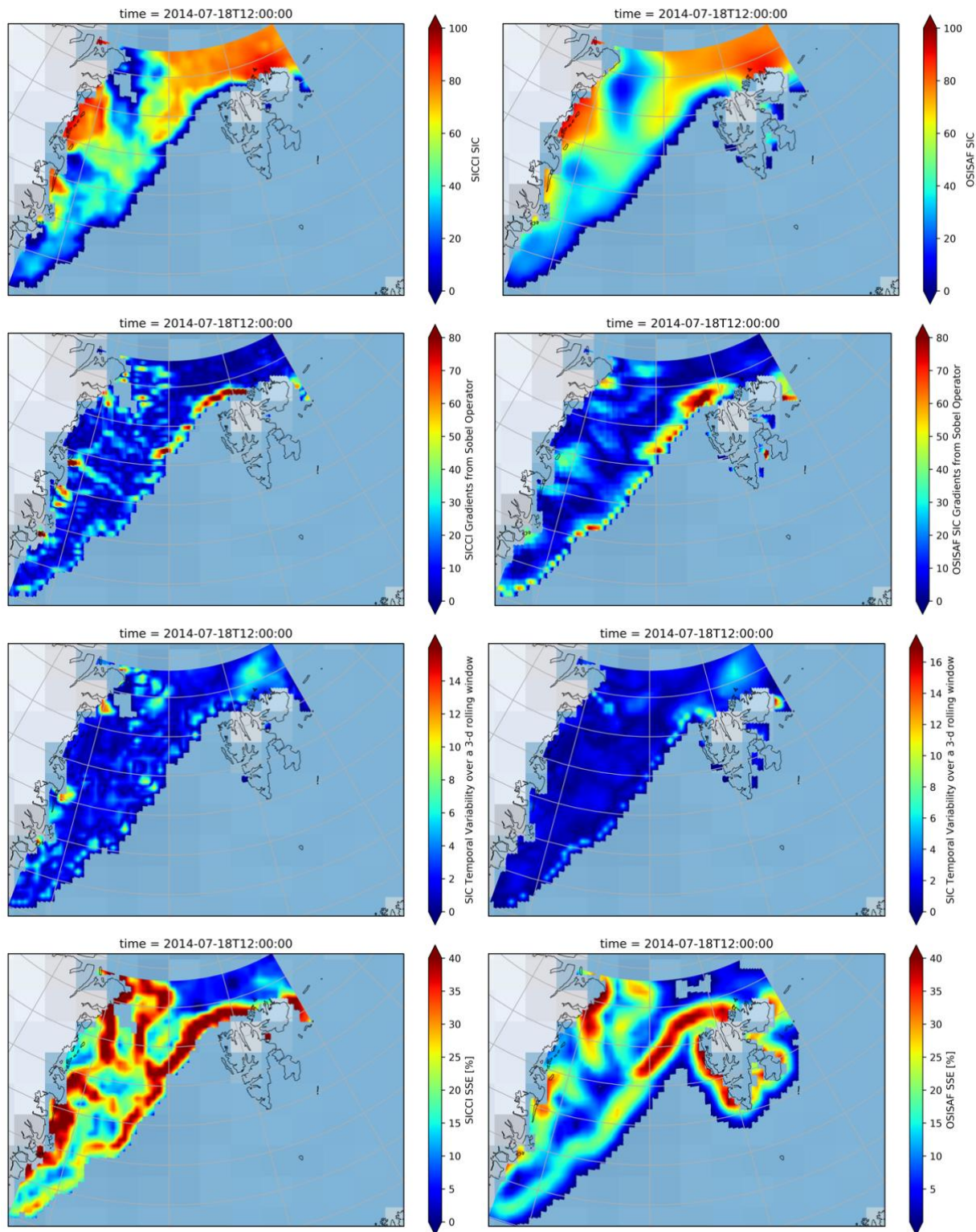


Figure 7. SIC Inconsistencies on the Baltic Sea for July 28, 2014 for the SICCI (left column) and OSISAF (right column) products. Top row: SIC inconsistencies, middle row: OSTIA SSTs, bottom row: corresponding SSE.



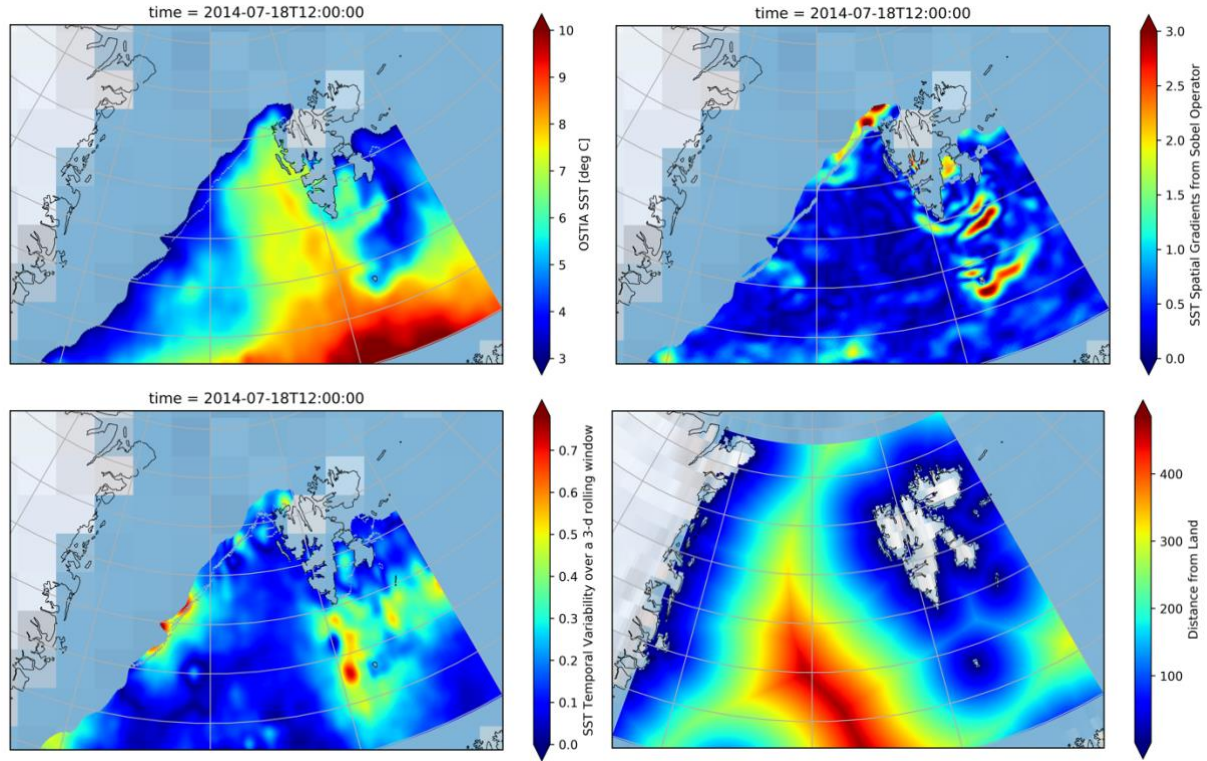


Figure 8. SIC products and related variables between Greenland and Svalbard for July 18, 2014. The SIC-related products are shown where the SIC > 0% and the SST-related products are shown where the OSTIA SST ≥ 3 °C. The inconsistencies are defined at the intersection of these two domains along the ice edge.

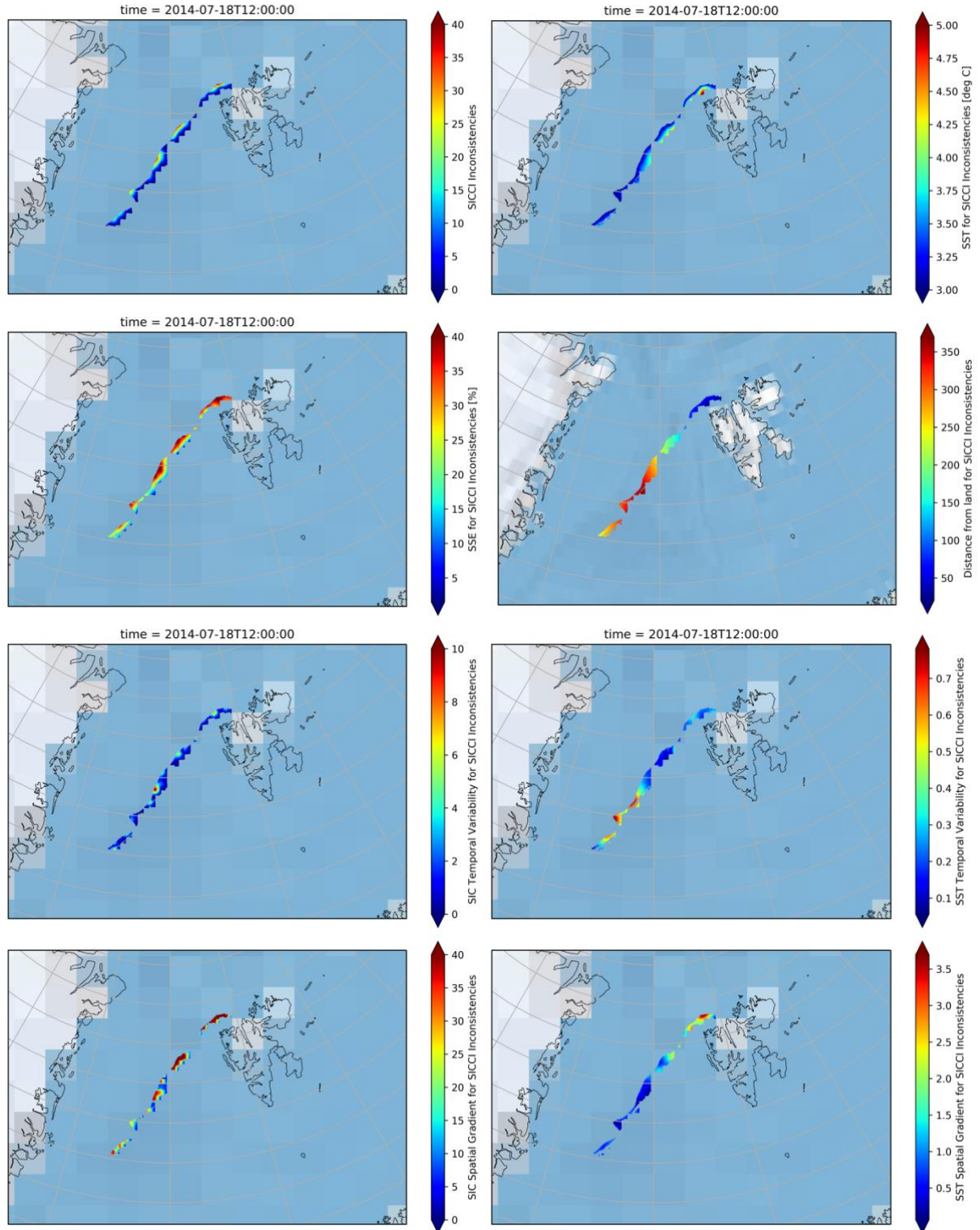


Figure 9. SICCI inconsistencies and corresponding parameters for the Svalbard case

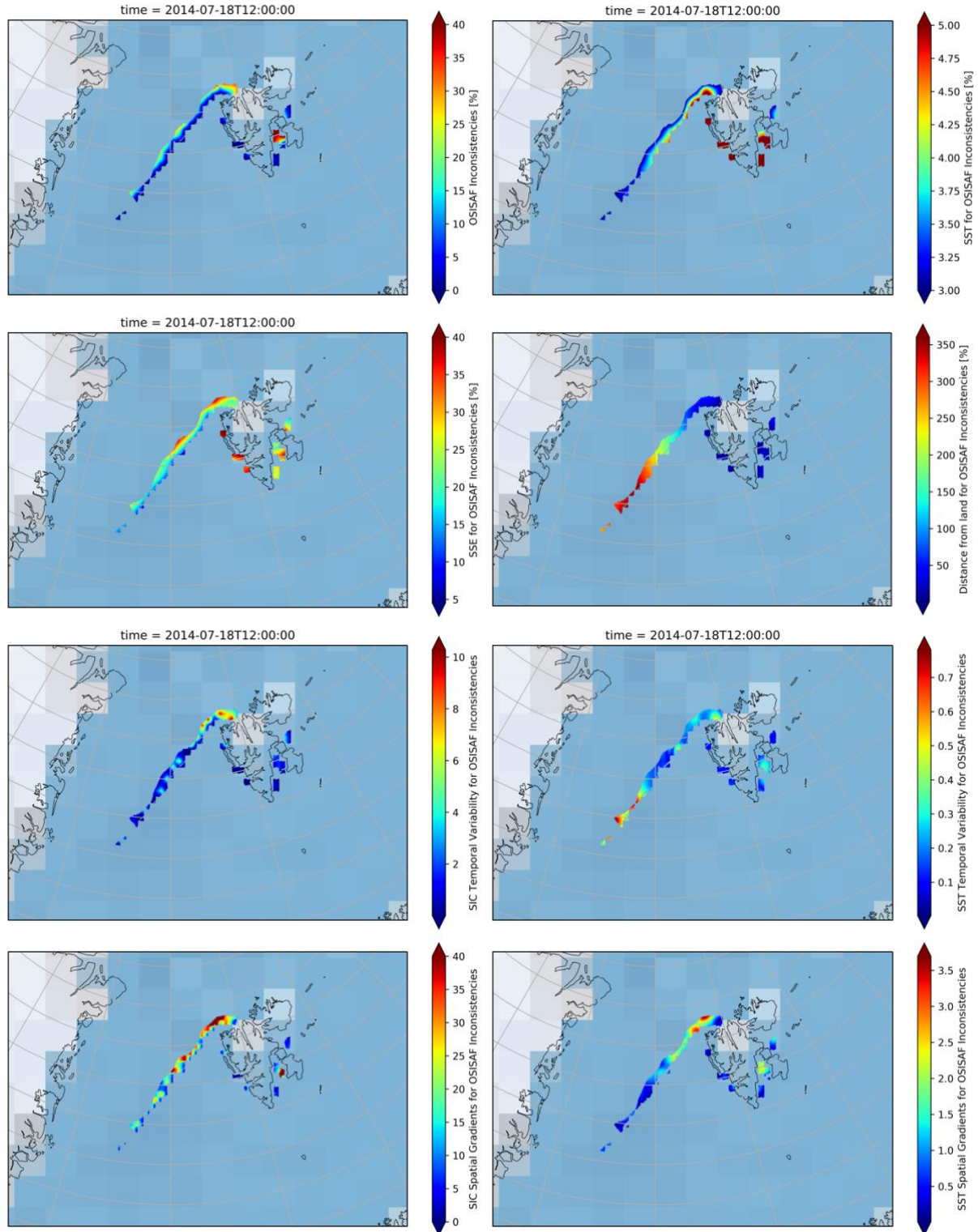


Figure 10. OSISAF inconsistencies and related parameters for the Svalbard case

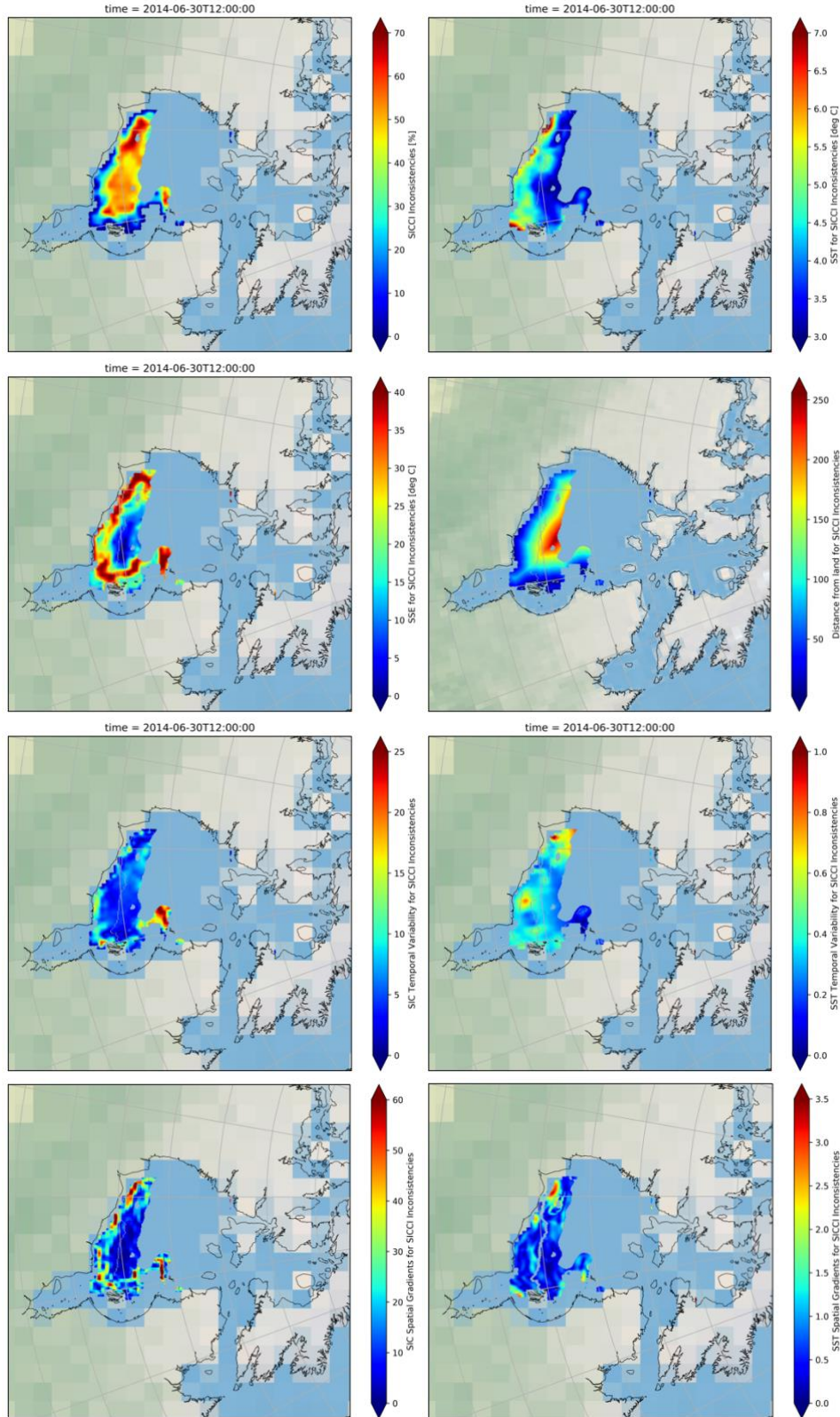


Figure 11. SICCI inconsistencies and related parameters for the Hudson Bay case

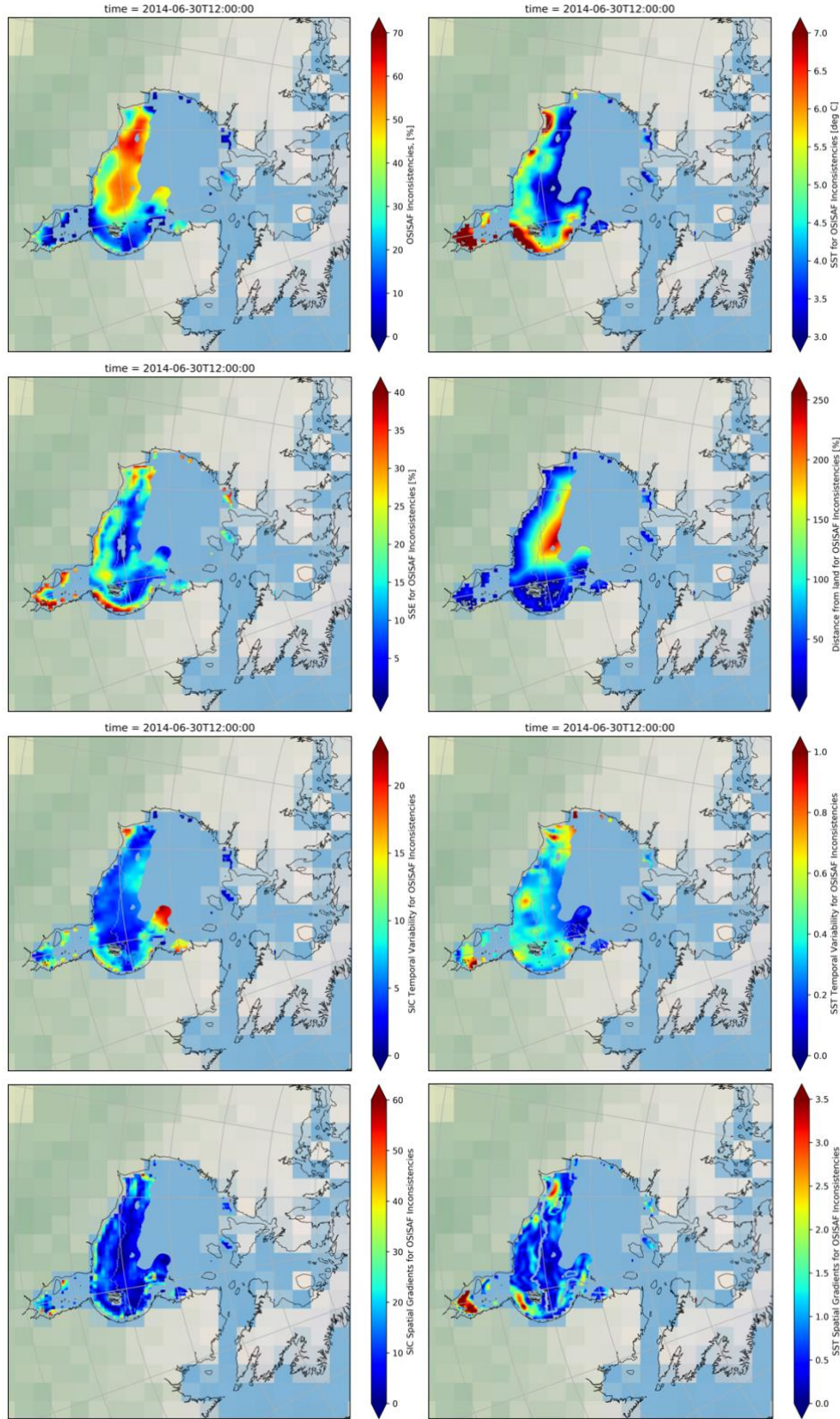


Figure 12. OSISAF inconsistencies and related parameters for the Hudson Bay case

4. Filter Determination

4.1 Filters to reduce spurious SIC inconsistencies resulting from land spillovers and interpolation/gridding effects

Visual inspection of the maps of inconsistencies illustrated in the examples of Section 3 suggests that, in addition to the SST, the other two parameters that might be able to help filter out spurious SIC retrievals in both SICCI and OSISAF products are the SSE and the distance from land. This is further supported by their probability density functions (pdf). Evaluation of the pdfs was done using kernel density estimation (KDE) and results are shown in Fig. 13 for all the variables being considered so far to help discriminate which SIC inconsistencies are false and which are real.

The pdfs for the SIC uncertainties and their distance from land show spikes that are not present in the probability distributions of the other variables being tested for their diagnostic potential. It is hypothesized that these spikes are the result of systematic effects in the inconsistency data set, originating from a common noise source. A first step in designing filters for noise reduction is to try to identify the common sources of noise responsible for the spikes in the pdfs (Section 4). We then identify thresholds exceeding the candidate spike segments, and eliminate the spike segments for the candidate thresholds in a systematic manner (Section 5). The final filter is based on the threshold that produces the trimmed data subset with the lowest standard deviation, since the standard deviation shrinks as the largest outliers are removed.

For the SSE, there appears to be spikes at both ends of the distribution, with a small spike close to $SSE = 0\%$ and a sharper, more prominent spike for high SSE. The smearing uncertainty (Lavergne et al., 2019) is largest at the ice edge (SSE up to 40% ; the atmosphere is more variable over the ice edge) and lowest near open water ($SSE = 0\%$ for $SIC = 0\%$; the atmosphere is more transparent over open water yielding better accuracies). The small spike at low SSE suggests the presence of false ice in open water (low SSE that should have been 0% SIC). In other words, the spikes at low SSE indicate that there are false positives in the smearing uncertainty determination, since a low SSE is understood to be an indication of confidence that the retrievals should be accurate, but the noise spike at the low end of the distribution suggests otherwise. The large spikes for high SSE suggest shortfalls in the atmospheric correction at low ice concentrations (i.e., the ice edge). The impact of the systematic effects on inconsistencies with high SSE (i.e., the ice edge and low concentration areas ($SIC < 15\%$)) is particularly prominent in the SICCI distribution, showing a significantly larger and wider spike than in the OSISAF SSE pdf. Note also that the low concentration spike, clearly evident in the SICCI pdfs for the two uncertainty terms, is still visible in the SSE pdf associated with the OSISAF inconsistencies, but it is completely absent from the pdf for the total standard errors (TSE). We took this as supporting evidence that the SSE was more sensitive than the TSE to the presence of false ice, and thus a better discriminator of the inconsistencies. Additionally, since the smearing uncertainty is the dominant source of error in the SIC retrievals we use SSE, instead of TSE, as the diagnostic parameter of choice from hereafter.

In the same spirit of this observation, we can say that the smearing uncertainty, computed as the difference between the highest and lowest SIC values in a 3×3 neighborhood grid around each pixel in the grid (Lavergne et al., 2019), is another measure of spatial variability, very similar to the 3×3 Sobel operator considered here as proxy for the local variability of the SIC

inconsistencies. Note, however, that the pdfs for the Sobel-derived spatial gradients shown in Fig. 13, do not exhibit any spikes in the distributions like the ones in the SSE pdfs. We take this as an indication that our Sobel estimate of the SIC spatial variability is less sensitive to noise in the inconsistency set, and as such we remove it from the diagnostic tools from this point forward. Hence, the two spikes in the SSE distribution not only suggest residual noise due to atmospheric influence, undetected by the atmospheric correction and the OWF, but potential residual noise from the SIC spatial variability in the ice edge, probably due to the enhanced atmospheric variability in the same ice-edge domain. Finally, open water pixels near the ice edge can reach higher SSTs and thus can have a significant impact on the inconsistencies, especially when SIC retrievals are gridded or when mixed pixels are binned/interpolated to finer resolution grids. This emphasizes the importance of this variable as discriminator for inconsistencies associated with interpolation/gridding effects.

The pdfs for the inconsistencies' distance from land also show spikes for SIC in close proximity to land in both products, but whereas the spike completely dominates the distribution of the OSISAF inconsistencies, the SICCI distribution appears to be less impacted with a narrower spike and a quick recovery in terms of pdf smoothness for SICs not too far off from land. The dominant impacts of SSE with SICCI and distance with OSISAF revealed in the pdfs give support to our unsubstantiated claims, made in the previous section based, solely on the visual inspection of the inconsistency examples shown in Figs. 9 and 10.

The last variable considered in the preliminary set of potential discriminators of true and false inconsistencies is the temporal variability of the SIC. Temporal variability could be useful to identify errors more random in time, but the inconsistencies proved to be more systematic in nature. This is supported by the smoothness of the corresponding pdfs shown in the last row of Fig. 13. Since there are no visible spikes in the density distributions of the temporal variability of the SIC inconsistencies in either SICCI or OSISAF, we eliminated this variable from the diagnostic tools. This leaves SSE and minimum distance-from-land (DIS_{min}) as main variables to help discriminate true from false inconsistencies.

The sensible thing to do in the presence of these systematic effects is to vary the range of data used to extract the result or to implement data cut offs to ensure the quality of the data. Thus, we propose filters for the correction of systematic errors in the SIC inconsistencies based on thresholds for a SSE range, (SSE_{min} , SSE_{max}), and a DIS_{min} requirement, aimed at reducing the spikes in the associated pdfs. There is no unequivocal indication of where these thresholds reside just from the look of the probability distributions. Visual inspection of the SICCI and OSISAF SSE pdf suggests that SSE_{min} is around 2 – 5%. The SSE_{max} , on the other hand, is not so easy to pin down especially for SICCI since the high smearing spike decreases gradually from ~38% to ~32%. In an effort to narrow down on a potential range for SICCI SSE_{max} , we also looked at the joint probability density function of SIC and SSE shown in Fig. 14. The joint pdf shows that most the SICCI inconsistencies have smearing uncertainties > 35%, suggesting a perhaps SSE_{max} should be ~32 – 35%. For OSISAF the high SSE spike is better constrained suggesting a preliminary range ~36 – 38%. As for DIS_{min} , it is a well-known fact that, due to the coarse resolution of microwave radiometers, data may be influenced by land up to 70 km from the coastline (Cavalieri et al., 1999; Tonboe et al., 2016). The pdfs for distance indicate that the land spillover is affecting SICs retrieved within ~40 km from land. Hence, we proposed an

exploratory range for the DIS_{min} threshold between 30 and 70 km. Since this effect appears to be less prominent in the SICCI inconsistencies, it suggests that perhaps the DIS_{min} requirement can be relaxed with SICCI.

4.2 Statistical characterization of the SIC vs. SST functional dependence for filter applications

Throughout this work we define an inconsistency as an ice concentration pixel ($SIC > 0\%$) that, when compared with an independent SST product on the same grid resolution, has a corresponding SST pixel with a temperature $\geq 3\text{ }^{\circ}\text{C}$. We termed the pair ($SIC > 0\%$, $SST \geq 3\text{ }^{\circ}\text{C}$) to be “inconsistent” because, intuitively, we question the validity of having a significant fraction of ice coexisting with what appears to be a well-above freezing SST at polar latitudes. The SIC and SST products being compared in this study are retrieved independently and individually. In reality, however, a pixel with, say, $SIC = 30\%$ indicates that the other 70% of the pixel is open water. So, in a sense, this is a “mixed pixel problem” with the pixel representing the average of two spectral classes emitted by two different surfaces on the ground, one ice and the other water, but each surface being the focus of a separate retrieval. If they were retrieved simultaneously, one would have to consider how the two surfaces interact along the common edge. In other words, a joint retrieval would require explicit characterization of the relationship between SIC and SST in the MIZ, (the surface domain of the mixed pixel). In reality, we are working with independently-derived SIC and SST retrievals.

Scatter plots of SIC vs SST for the 2014 inconsistency set for both the SICCI and the OSISAF products, are shown in the top two panels of Fig. 15. At first sight, the scatter plots for both products seem to have very little in common, but there is substantial overlapping obscuring what is really going on, especially for OSISAF. A hexbin plot is an alternate representation of the relationship between two variables that can be particularly useful when there are a lot of data points, case in point, the scatter plot of SIC vs. SST for OSISAF shown in Fig. 15. The hexbin plot partitions the spatial domain of the (SIC, SST)-pairs in hexbins and gives a visual representation of the total counts per hexbin, as indicated by the color of the bins. This is shown in the bottom row panels of Fig. 15. In the selected color scale, black corresponds to bins with very few counts. If we ignore these for the time being, then the colored bins for SICCI and OSISAF, although not entirely equal, show some similarities in that both products have the greatest concentration of points near the origin (yellow and blue colors).

The red bins, separating the very dense from the very sparse bins, are particularly interesting especially for SICCI because they cluster in a curve that resembles the coexistence curve in the phase diagram for water (i.e., the solid-liquid phase boundary). The curve of red bins has a negative slope (i.e., the tangent) consistent with the slope of the coexistence curve for ice-water boundary, defined by the Clausius-Clapeyron equation for melting/freezing. The coexistence curve denotes the (thermodynamically favorable) conditions at which two distinct phases may coexist. Following this train of thought, we will use the red curve apparent in the SICCI hexbin plot to try to define a sort of coexistence curve that characterizes the functional dependence between SIC and SST in mixed pixels. The color bins in the hexbin plot suggests that the (SIC, SST) inconsistency pairs seem to fall in one of two areas of their joint space domain, whereby for medium to high concentrations ($SIC > \sim 10\%$), the SST decreases gradually with increasing

concentration from about 10 °C to 3 °C (our lowest SST by definition of what constitutes an inconsistency), but for low ice concentrations ($SIC < 10\%$) the open water portion of the pixel can achieve significantly higher SSTs. It seems to make sense to use this functional dependence to filter out the inconsistencies outside the boundary established by the red bins (i.e., black bins). The main problem is, though we have a good idea of the shape of the SIC – SST functional dependence from the pattern revealed by the red bins in Fig. 15, we do not know its exact form.

Fig. 16, which indicates what inconsistencies in the scatter plots of SIC vs SST come from inside or outside the Arctic circle (points above 66.5°N are represented in red and below in blue), suggests that retrievals from outside the Arctic circle are solely responsible for the tail end of the low concentrations with high SST departures described above (i.e., they come from the Baltic Sea, the Hudson Bay and the Sea of Okhotsk). It seems reasonable, to a point, that mixed pixels from subpolar latitudes are able to reach warmer temperatures, but how high can the SSTs really get inside the ice edge ($SIC < 15\%$)? Similarly, it does not appear to be reasonable to have closed ice pixels ($SIC > 80\%$) with $SST = 3$ °C, as seen at the tail end of the high concentrations with low SSTs. Obviously, there should be a maximum in SST and in SIC, for the two phases to coexist without breaking the laws of thermodynamics.

A parameterization for $SST = f(SIC)$, based on the monthly scatter plots of the SICCI SIC vs. OSTIA SST (Fig. 17), was derived by trial and error as:

$$SST_{lim} = 6.24 (5.5/SIC)^{0.025} - 25.8, \quad (1)$$

and a series of SIC corrections are proposed based on this functional dependence. In one, for inconsistencies with $SIC > 10\%$, only those with matching $SST < SST_{lim}$ are kept. For $SIC \leq 10\%$, all SIC values are retained. Note that the precise functional form could depend on the resolution of the products, since it represents a potentially realistic balance within a given pixel. This filter is illustrated by the black curve in Fig. 17 for the monthly scatter plots of SICCI SIC vs SST. As can be seen in the monthly plots, the filter eliminates a significant number of outliers for the months of June, July and August, but leaves inconsistencies for other months mostly untouched, thus confirming that the parametric coexistence curve correction works as expected. We also consider an alternate variation of this correction in which all SIC with $SST < SST_{lim}$ are passed by the filter. These filters need to be used in combination with boundary conditions that limit the range in which Eq. (1) applies, i.e., with filters that restrict $f(SIC, SST)$ to a domain specified by a maximum allowable SST and maximum SIC for physically meaningful SIC-SST pair combinations. That is, in addition to the filters just described, we also eliminate SIC inconsistencies with $SST > SST_{max}$ and $SIC > SIC_{max}$. These two filters, exploiting the SIC vs. SST functional dependence, constitute a significant departure from the current DMI approach that eliminates all the inconsistencies for which $SIC > 15\%$ and $SST > 3$ °C.

Once again, we have no clear indication of the values of SST_{max} and SIC_{max} just by looking at the scatter plots of SIC vs. SST. An exploratory analysis in which the SST_{max} threshold was systematically lowered, starting from $SST = 20$ °C, followed by a visual assessment of the impact of the threshold on the monthly MVC maps of remaining SIC for July and August, indicates that SST_{max} should be ~10 – 12 °C to eliminate false SICs in the Baltic Sea and the Sea of Okhotsk in the summer. This range worked for both SICCI and OSISAF. As far as the value

of SIC_{max} , the threshold should establish the maximum SIC at which a fraction of the pixel can still be open water, and that is the MIZ. Thus, SIC_{max} should be around 70 – 80%, which is the concentration threshold used by different authors for the start of the pack/closed ice (i.e., Tonboe et al., 2016; Peng et al., 2018), i.e., the upper SIC threshold that defines the MIZ boundary.

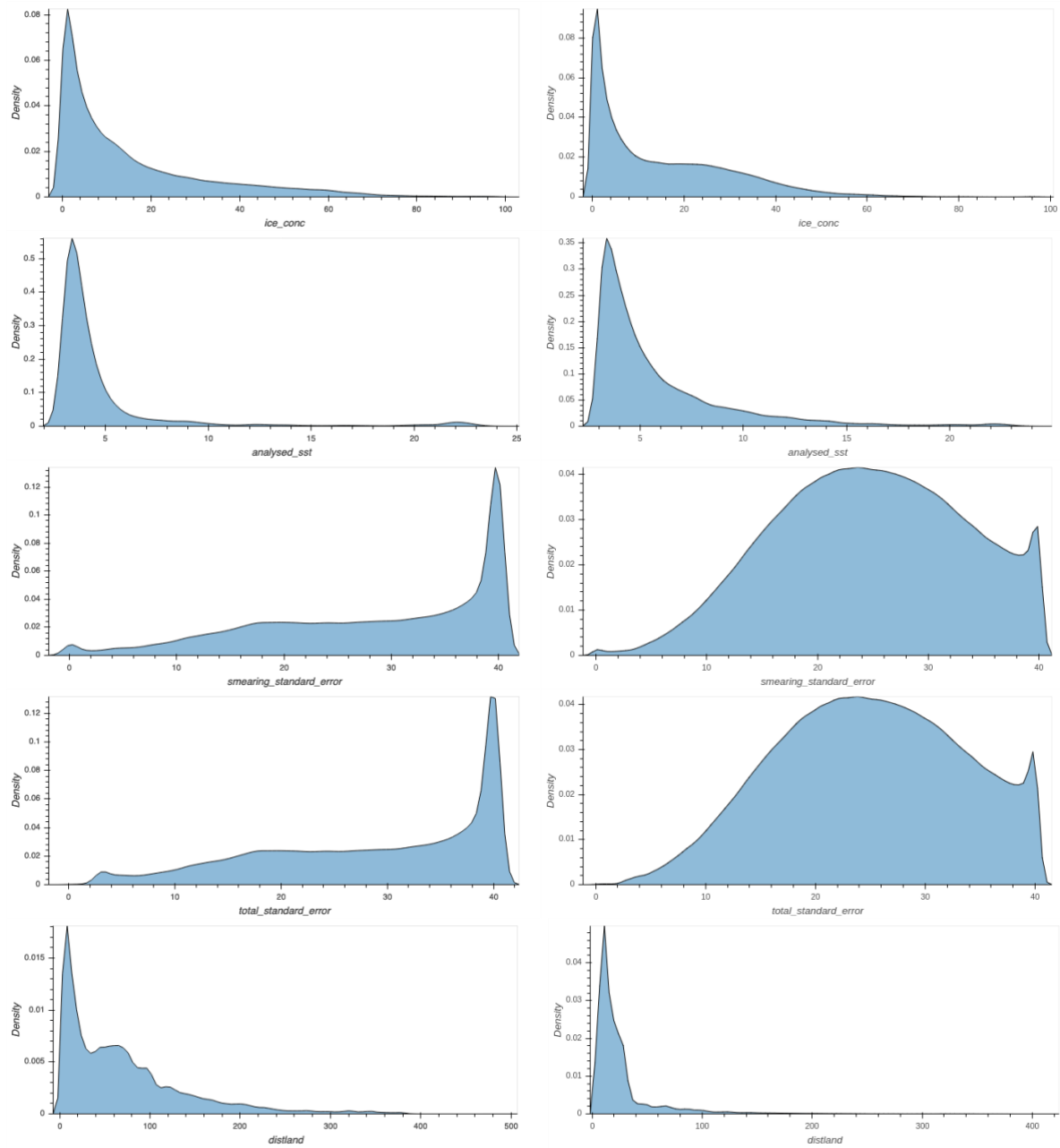
4.3 Summary of proposed filters

Inconsistencies are filtered out if:

1. Filter 1 (F1): $DIS < DIS_{min}$: Minimum distance from land (km)
2. Filter 2 (F2): $SIC > SIC_{max}$: Maximum SIC for physically meaningful SST (%)
3. Filter 3 (F3): $SST > SST_{max}$: Maximum SST for a physically meaningful SIC ($^{\circ}C$)
4. Filter 4 (F4): $SSE < SSE_{min}$ or $SSE > SSE_{max}$: Minimum and Maximum SSE to eliminate the spikes at both ends of the pdf for the SSE corresponding to the SIC inconsistencies.
5. Filter 5 (F5): $SST > SST_{lim} = f(SIC)$ for $SIC > 10\%$
6. Filter 6 (F6): $SST > SST_{lim} = f(SIC)$ for $SIC > 0\%$

and everything outside the threshold will be untouched. We suggest to use these filters in combination, not in isolation. Note that the F1, F2, and F3 filters (DIS_{min} , SIC_{max} , and SST_{max}) help eliminate residual noise that escaped detection during processing of the CDR. That is, they result from deficiencies in the noise corrections implemented with the SIC algorithm. While F1 and F2 address wrong retrievals resulting from proximity-to-land artifacts, F3 targets residual noise due to atmospheric influence. In other words, these three filters remove erroneous SIC values that should have been reported as open water (i.e., $SIC = 0\%$), but were somehow (a cloud, a narrow inlet or piece of land) misclassified as ice ($SIC > 0\%$). This opens the possibility of reclassifying inconsistencies eliminated by these filters as $SIC = 0\%$. In that sense, they would resemble traditional weather filters reported in the literature (Ivanova et al., 2015, Lavergne et al., 2019) and implemented with traditional SIC algorithms elsewhere, but with the main difference that they are being applied directly to the SIC retrievals (at level 3 processing), and not at the brightness temperature level (level 2). The F4, F5, and F6 filters (i.e., SSE and SST filters based on the functional curve of SIC vs. SST), on the other hand, reduce noise by eliminating inconsistencies where the SIC and the SSTs are not in balance according to the coexistence SIC-SST curve (e.g., excessive ice for the coincident temperature or vice versa). Thus, while the former three filters reduce noise resulting from misclassification issues, the latter three reduce noise resulting from misrepresentation issues.

In our approach here, we applied the misclassification corrections first, followed by the misrepresentation corrections. While both F5 and F6 use the same parameterized curve for SST_{lim} (Eq. 1), they differ in the way the function is implemented at the ice edge ($SIC < 10\%$); hence, they are considered separately. This gives rise to two alternate filtering schemes that, for lack of a better name, are termed F5- and F6-chains for the final filter used in the sequence.



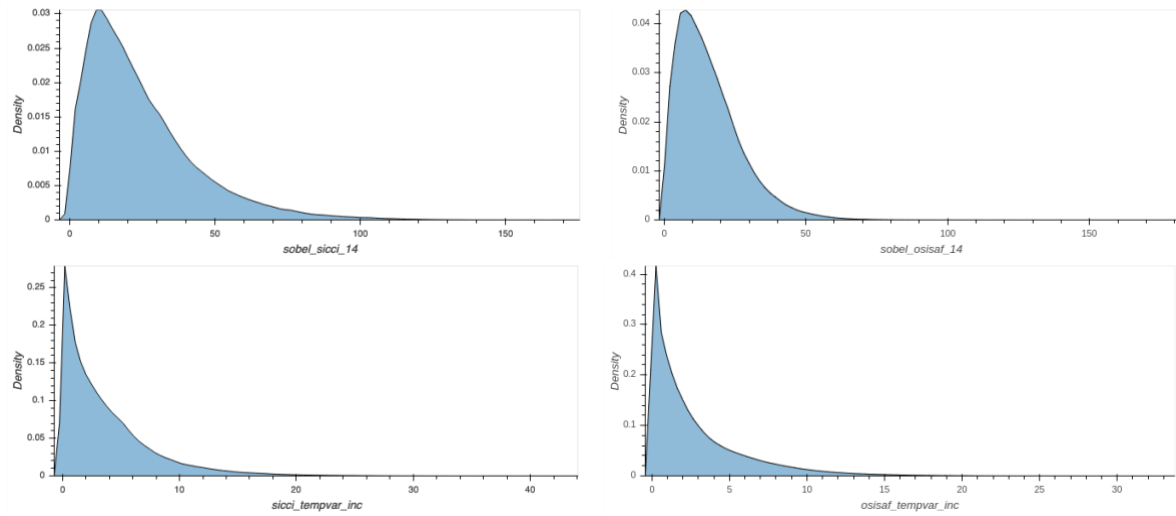


Figure 13. Probability density functions of individual variables (SIC, SST, SSE, TSE, SIC distance from land (distland), spatial variability (Sobel operator), and temporal variability for: SICCI (left column) and OSISAF (right column) products.

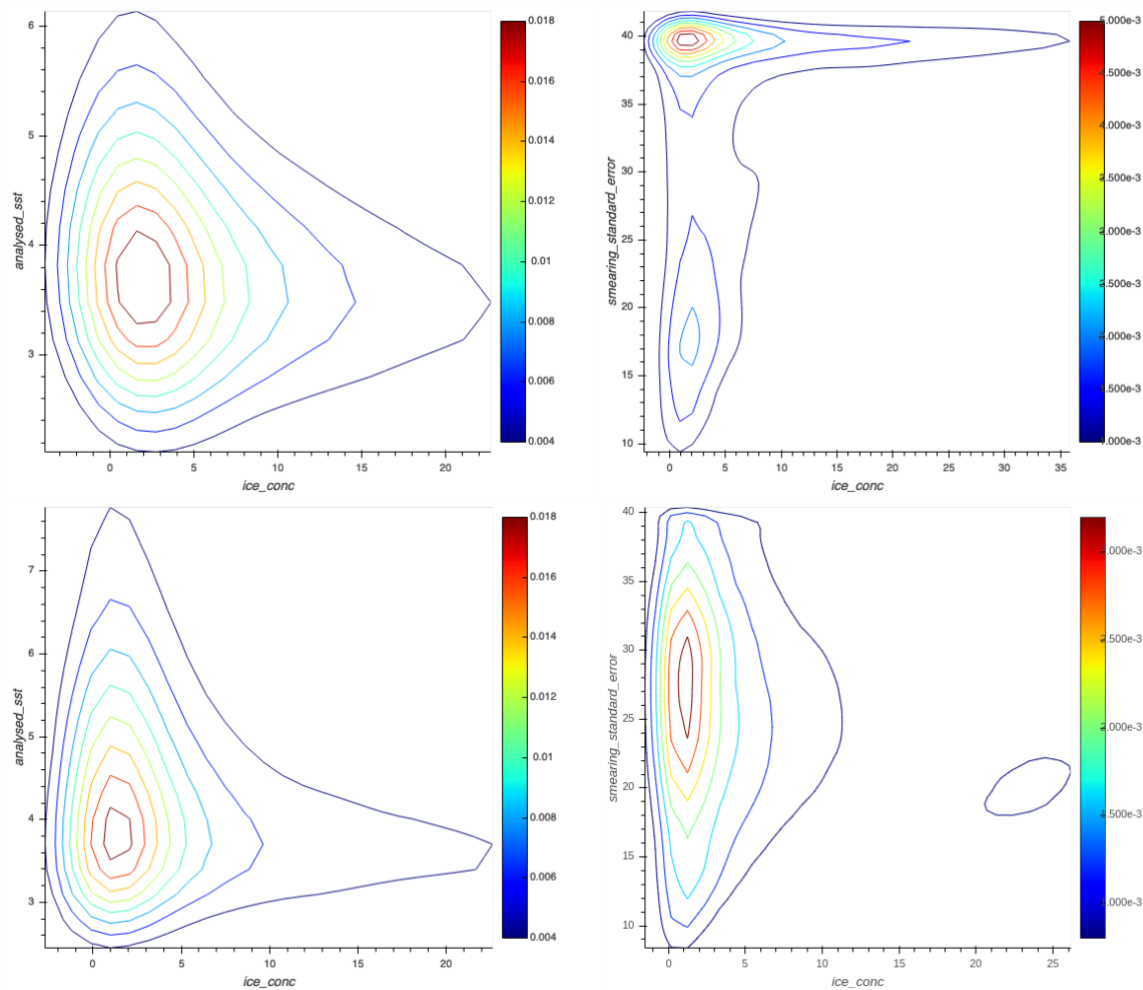


Figure 14. Joint pdfs of SIC inconsistencies and associated OSTIA SST (left column), and SIC inconsistencies and associated smearing (SSE) uncertainties (right column) for SICCI (top row) and OSISAF (bottom row) products.

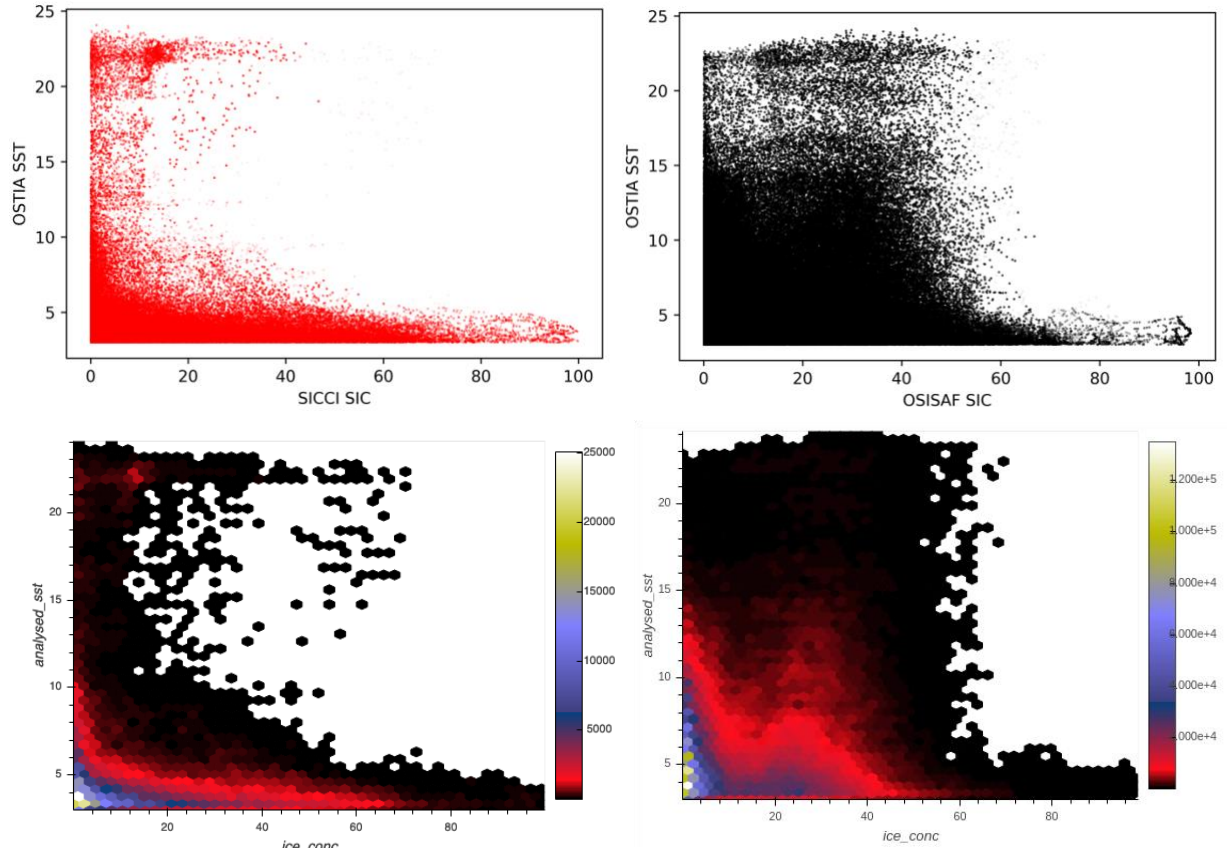


Figure 15. SIC vs. SST dependence for SICCI (left) and OSISAF (right). Top: scatter plots; bottom: hexbin plots with colors indicating density of the bins for the SIC and SST dependence displayed in the scatter plots above.

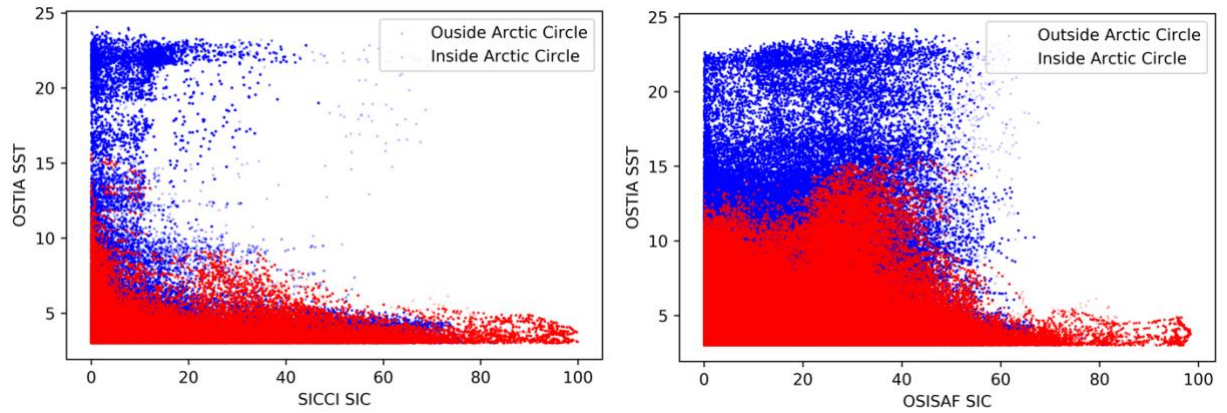


Figure 16. Scatter plot of the 2014 inconsistencies vs. the OSTIA SST, color-coded by location relative to the Arctic Circle (Latitude = 66.5 °N). The SICCI SIC-SST dependence is shown on the left and the OSISAF on the right. Inconsistencies inside the Arctic circle are colored in red and inconsistencies outside the Arctic circle are colored in blue. Inconsistencies from outside the Arctic circle have more scatter, which is not surprising since this is where the marginal seas of the Arctic are located (Figs. 5 and 6) and where the SIC retrieval algorithms appear to have more issues as suggested by the abundance of inconsistencies in these areas (Fig. 2).

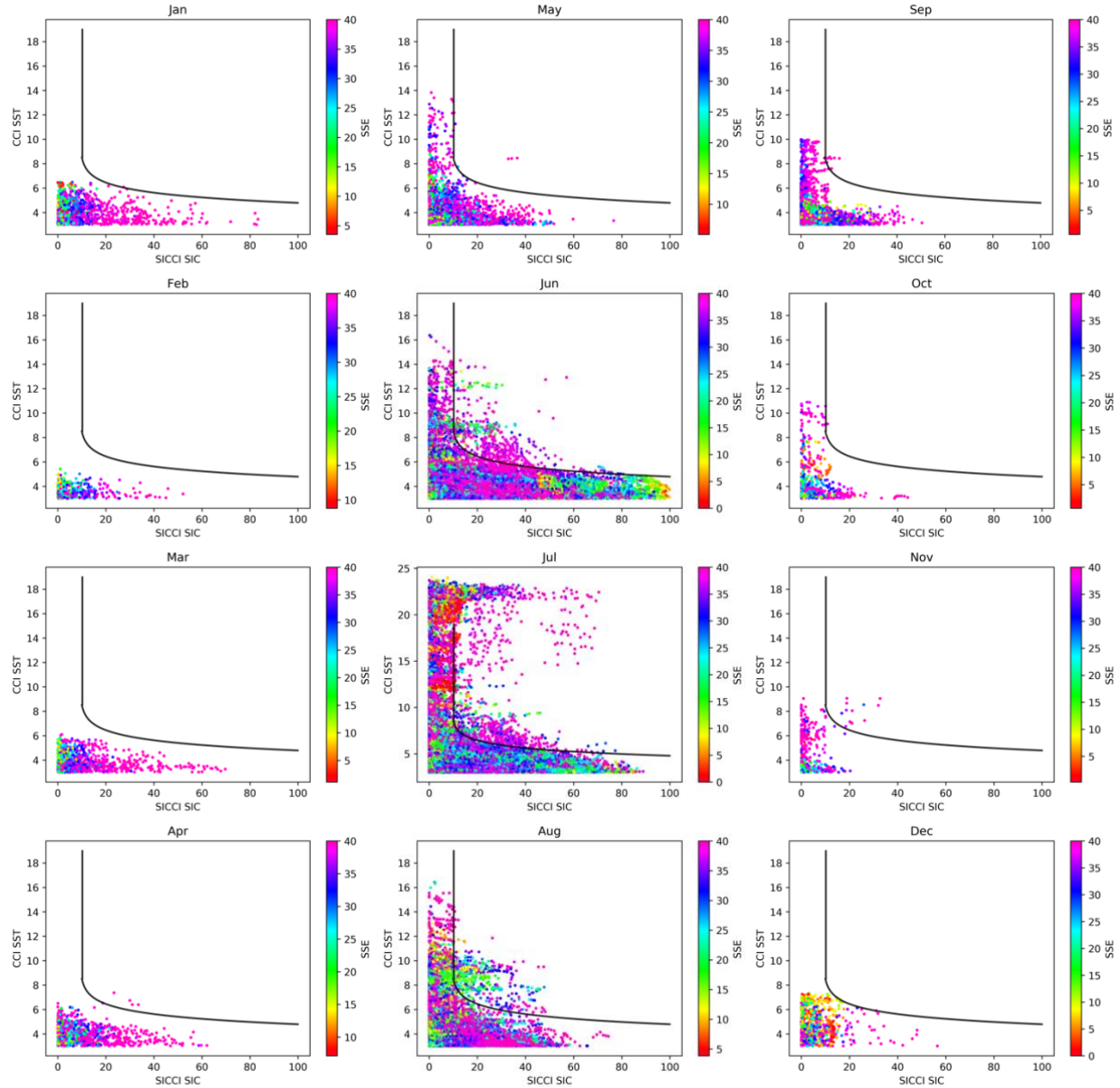


Figure 17. Monthly scatter plots of SIC vs. OSTIA SST for SICCI 2014, color-coded by the SSE. The curve in black is a combination of $SIC = 10\%$ (vertical line) and the SST_{lim} function (curve) for $SIC > 10\%$. Together, they illustrate filter function F5. Only data to the right of the function is filtered out by F5. With few exceptions, the filter eliminates outliers mostly from June through August. The July plot illustrates why the F5 filter needs to be used in combination with a requirement for SST_{max} . Similarly, the plot for June shows the need for a filter that bounds the function at a SIC_{max} .

5. Filter Evaluation

We need to choose the values of the thresholds for DIS_{min} , SIC_{max} , SST_{max} , SSE_{min} , and SSE_{max} for the proposed filters. As an initial approach to narrowing the range of the threshold values identified from the marginal pdfs in Section 4, we used the standard deviation (SD) of the SIC values remaining after elimination of suspected retrievals. The assumption was that the SD of the remaining set should be smaller in magnitude due to reduced noise following removal of the inconsistencies that exceed the thresholds. To evaluate the impact of the different filters using the finalized thresholds, a more detailed metric employing a combination of SIC and SST values was used and is described in Section 5.2. Ultimately, the final judgement on the success of the different filters was also subjectively based on the appearance of the pdfs for the filtered SIC retrievals and corresponding joint pdfs for SIC vs. SST and SIC vs. SSE (Section 5.3), as density distributions for the denoised sets should not exhibit the discontinuities observed in the original inconsistency set.

5.1 Threshold Determination

To help us determine the optimal thresholds for the final filter combinations, we apply the individual filters using a range of possible thresholds determined from the pdfs in Section 4 and evaluate the statistics shown on Tables 3 (SICCI) and 5 (OSISAF). We then look for the threshold, from the range of realizations, that produces the subset of filtered data with the minimum standard deviation, since this subset has the least spread or variability introduced by the noise source being addressed by the specific filter. This is indicated by the columns shaded in blue in Tables 3 and 5. Once we have narrowed in on a threshold based on the outcome with the lowest SD, we evaluate the next filter over the truncated data, i.e., the data that remains after the cut-offs from the previous filter. For clarity, we carry on the statistics from the previous finalized correction (the shaded blue column) at the beginning of each table, since they become the new input data to the next filter. This facilitates comparing the statistical properties of the new corrections at each step of the chain and its effectiveness at removing additional variability (the larger the difference between the highest and lowest SD, the greater the deviation and the higher the variability). The headings of the columns in Tables 3 and 5 indicate the valid ranges of the parameters being tested, since the tabulated statistics are for the truncated sets that passed the filter; for instance, the sensitivity analyses to determine the final SSE threshold for SICCI (Table 3, F4) indicate a heading of “ $3 \leq SSE \leq 33$ ” for the column shaded in blue. This is the range of valid SSEs corresponding to the remaining set, after the filter F4 eliminated SIC inconsistencies with $SSE < 3\%$ and $SSE > 33\%$. The reduced data sets that remain at the end of the chain of sequential corrections tested in Tables 3 and 5 (i.e., what is left after F4) are used as the common input to the final filters, F5 and F6. Statistics for the two ensuing filtering schemes are shown in Tables 4 and 6 for SICCI and OSISAF, respectively.

Selected thresholds for the proposed filters tested in Tables 3 and 5 for SICCI and OSISAF are as follows:

SICCI filters:

F1: $DIS < 50$ km

F2: $SIC > 75\%$

F3: SST > 10 °C
F4: SSE < 3% and SSE > 33%.

OSISAF filters:

F1: DIS < 60 km
F3: SST > 10 °C
F4: SSE < 3% and SSE > 37%.

The final thresholds shown above are consistent with the observations made in Section 4 based on the pdfs of the parameters used in the filters (Fig. 13). For instance, since the spikes in the pdf of the SSE were significantly larger for SICCI than OSISAF, it was expected that SICCI would need a more restrictive SSE (F4) filter, especially for high smearing uncertainties, which is reflected in the findings of Tables 3 and 5 with $SSE_{max} = 33\%$ for SICCI and $SSE_{max} = 37\%$ for OSISAF. Similarly, OSISAF appears to be more affected by proximity-to-land issues needing a more aggressive filter in terms of minimum distance from land. Once again, the sensitivity analyses led to $DIS_{min} = 50$ km for SICCI and $DIS_{min} = 60$ km for OSISAF.

Note that there is no F2 filter for OSISAF since the maximum SIC observed after the OSISAF inconsistencies are filtered by distance is 73.7% (Table 5, F1, last row), which is below the final concentration threshold used with SICCI (F2: SIC max = 75%) to eliminate the high concentrations remaining in the F1-filtered data (Table 3, F1, last row shows maximum concentrations > 90%). Even if we settled on the same distance-from land threshold used with OSISAF ($DIS_{max} = 60\%$), the truncated SICCI data would still have SIC > 85% (Table 3, F1, DIS < 60, last row), requiring a high concentration filter. Maps of MVC of the remaining inconsistencies from the finalized thresholds, at each step of the filtering sequence, are shown in Figs. 19 and 20 for SICCI and OSISAF, respectively.

The two filtering chains (F5 and F6 in Tables 4 and 6) reduce the original pool of inconsistencies to a subset less than 2% the original size with what appears to be marginal improvements in terms of the mean and SD: for the SICCI inconsistencies (98.3% of original data was cut off), the mean decreased from 15% in the original set to ~13% in the final sets (Table 4) and the SD decreased from 17.0% to ~15% in both filtering schemes (F5 and F6); for the OSISAF inconsistencies (99.1% of the original data was cut off), the mean SIC decreased from 15.0% to ~12% (Table 6) and the SD decreased from 14% to ~12% in both filtering schemes. This implies a difference in SD between the original and final sets is of the order of 2% for both products. Even though the SDs of the filtered data (12 – 15%) suggest that there is a lot of variability left in the remaining inconsistency sets after all the filtering is done, the SDs of the denoised sets are representative of the variability within the ice edge. This is supported by maps of the filtered data (Figs. 19 and 20) showing that the great majority of the remaining SICs are located in areas of low SIC (around the ice edge), with the exception of the Hudson Bay and some small areas of intermediate-high concentrations around the Canadian Archipelago, and the Beaufort and Laptev Seas.

Table 3. Sensitivity analyses of the SICCI corrections to incremental changes in the filtering thresholds. Columns shaded in blue mark the thresholds that produce the most effective filters at reducing the variability of the remaining set (lowest SD). Since filters are applied sequentially, the first column of the next filter is the shaded column from the previous correction. This is done to facilitate identification of further improvement by subsequent filters

SICCI F1	No filter	DIS \geq 40 km	DIS \geq 50 km	DIS \geq 60 km	DIS \geq 70 km
count	10503580	5530229	4967974	4330457	3695539
mean	14.89	14.51	14.56	14.54	14.48
SD	16.99	14.79	14.72	14.79	14.83
min	0.01	0.01	0.01	0.01	0.01
25%	2.25	3.14	3.19	3.12	3.01
50%	8.27	9.59	9.69	9.60	9.46
75%	21.89	21.06	21.20	21.22	21.20
max	99.92	92.32	91.64	86.87	86.87

SICCI F2	DIS \geq 50km	SIC \leq 75%	SIC \leq 80%	SIC \leq 85%	SIC \leq 90%
count	4967974	4965795	4967368	4967805	4967951
mean	14.56	14.53	14.55	14.55	14.56
SD	14.72	14.67	14.71	14.72	14.72
min	0.01	0.01	0.01	0.01	0.01
25%	3.19	3.19	3.19	3.19	3.19
50%	9.69	9.68	9.68	9.69	9.69
75%	21.20	21.18	21.20	21.20	21.20
max	91.64	74.99	79.98	84.98	89.96

SICCI F3	SIC \leq 75%	SST \leq 8°C	SST \leq 10°C	SST \leq 12°C	SST \leq 14°C
count	4965795	306486	308929	309537	310453
mean	14.53	16.05	16.00	15.97	15.94
SD	14.67	15.39	15.35	15.35	15.34
min	0.01	0.01	0.01	0.01	0.01
25%	3.19	3.71	3.70	3.69	3.69
50%	9.68	11.11	11.06	11.04	10.99
75%	21.18	23.95	23.83	23.81	23.77
max	74.99	74.26	74.26	74.26	74.26

SICCI F4	SST \leq 10°C	3 \leq SSE \leq 31	3 \leq SSE \leq 33	3 \leq SSE \leq 35	3 \leq SSE \leq 37
count	308929	165671	183966	203756	225137
mean	16.00	12.58	12.92	13.33	13.75
SD	15.35	15.32	15.28	15.30	15.28
min	0.01	0.01	0.01	0.01	0.01
25%	3.70	2.0	2.19	2.42	2.65
50%	11.06	6.21	6.71	7.27	7.92
75%	23.83	16.87	17.40	17.96	18.73
max	74.26	74.26	74.26	74.26	74.26

Table 4: Statistics for the SICCI product after final filters F5 and F6 are applied to the sequential chain of corrections shown in Table 3

SICCI Final	$3 \leq \text{SSE} \leq 33$	F5	F6
count	183966	182308	180526
mean	12.92	12.89	12.97
SD	15.28	15.33	15.39
min	0.01	0.01	0.01
25%	2.19	2.16	2.15
50%	6.71	6.61	6.65
75%	17.40	17.37	17.53
max	74.26	74.26	74.26

Table 5: Sensitivity analyses of the OSISAF corrections to incremental changes in the filtering thresholds. Columns shaded in blue mark the thresholds that produce the most effective filters at reducing the variability of the remaining set (lowest SD). Since filters are applied sequentially, the first column of the next filter is the shaded column from the previous correction. This is done to facilitate identification of further improvement by subsequent filters

OSISAF F1	No filter	$\text{DIS} \geq 40 \text{ km}$	$\text{DIS} \geq 50 \text{ km}$	$\text{DIS} \geq 60 \text{ km}$	$\text{DIS} \geq 70 \text{ km}$
Count	87888090	12131750	10104590	8347062	6678692
Mean	15.03	12.99	12.58	12.09	11.86
SD	14.14	11.71	11.54	11.47	11.49
Min	0.01	0.01	0.01	0.01	0.01
25%	2.42	3.47	3.39	3.08	2.85
50%	10.74	10.32	9.68	8.93	8.61
75%	25.37	18.95	18.28	17.70	17.44
max	98.67	79.97	73.70	73.70	73.70

OSISAF F3	$\text{DIS} \geq 60 \text{ km}$	$\text{SST} \leq 8^\circ\text{C}$	$\text{SST} \leq 10^\circ\text{C}$	$\text{SST} \leq 12^\circ\text{C}$	$\text{SST} \leq 14^\circ\text{C}$
Count	8347062	477448	479125	479534	480095
Mean	12.09	13.17	13.15	13.14	13.13
SD	11.47	12.23	12.22	12.22	12.22
Min	0.01	0.01	0.01	0.01	0.01
25%	3.08	3.44	3.44	3.44	3.43
50%	8.93	9.96	9.94	9.92	9.91
75%	17.70	19.25	19.21	19.20	19.19
Max	73.70	72.10	72.10	72.10	72.10

OSISAF F4	$\text{SST} \leq 10^\circ\text{C}$	$3 \leq \text{SSE} \leq 33$	$3 \leq \text{SSE} \leq 35$	$3 \leq \text{SSE} \leq 37$	$3 \leq \text{SSE} \leq 38$
count	479125	417583	432408	445868	452039
mean	13.15	11.74	11.94	12.16	12.30
SD	12.22	11.87	11.80	11.76	11.77
min	0.01	0.01	0.01	0.01	0.01
25%	3.44	2.85	2.99	3.13	3.18
50%	9.94	8.15	8.52	8.92	9.09

75%	19.21	16.60	16.90	17.31	17.60
max	72.10	72.10	72.10	72.10	72.10

Table 6: Statistics for the OSISAF product after final filters F5 and F6 are applied to the sequential chain of corrections shown in Table 5

OSISAF	$3 \leq \text{SSE} \leq 37$	F5	F6
count	445868	441560	439761
mean	12.16	12.12	12.15
SD	11.76	11.80	11.81
min	0.01	0.01	0.01
25%	3.13	3.08	3.07
50%	8.92	8.76	8.81
75%	17.31	17.25	17.32
max	72.10	72.10	72.10

5.2 Proposed Metrics to Evaluate Filter Impact

Although the improvement of the proposed filters suggested by the statistics in Tables 4 and 6 seems marginal at first sight, the decrease in standard deviation before and after the filter is not the correct metric to assess the impact of the corrections. In order to understand why this is the case, we just need to look at the binned density plots of SIC vs. SST for the truncated sets after a new filter application (Figs. 23–24 for SICCI and OSISAF, respectively). As can be seen from the two plots on the second row of Fig. 23 for SICCI, the spread of the bins, and hence the SD, increases considerably after the F3 filter is applied. This is indeed the case as shown in Table 3 for F3, with the SD increasing from 14.7% (first column) to 15.4% (blue column). The F3 filter cuts off a significant portion of the low concentration values with high SSTs. The red hexbins at the end of this tail, visible in the left plot for the input data (F1-F2 subset) to F3, indicate that there are bins with a high number of counts in this cluster. After their elimination, the SD of the remaining set is “seeing” more spread because it is measuring the impact of removing the densely-compacted cluster at the end of the high SST-low SIC tail. In a purely statistical sense, the increase in SD might suggest that the F3 filter made things worse resulting in a truncated set with increased variability, yet we have a high degree of certainty that these values are wrong (they correspond to false SIC retrievals in the Baltic and Okhotsk Seas where open water pixels were misclassified as ice likely due to atmospheric influences and land spillover effects) and thus are justified in eliminating them. This goes to show that metrics based solely on the SD are not meaningful here when we are trying to assess the impact of the proposed filters.

In order to evaluate the impact of the suggested corrections, we tested two metrics that combine the SIC and the SST in one term, so that changes in one variable reflect the impact on the other in a more insightful way for the problem at hand. Those metrics are the sum of the SDs for both SIC and SST, $M1 = \text{SD}_{\text{sic}} + \text{SD}_{\text{sst}}$, and a new metric M2 we define as follows:

$$M1 = \sqrt{SD_{sic}^2 + SD_{sst}^2} \quad (2)$$

$$M2 = \frac{\sum (SST - 3) * SIC}{N}$$

where the standard deviations add in quadrature in M1. For M2, N is the size of the unfiltered SIC inconsistency set (i.e., the sample size without cut offs). We subtract $SST = 3 \text{ }^\circ\text{C}$, since this is the minimum temperature attained by the inconsistencies, according to our definition of what constitutes an inconsistency ($SIC > 0\%$ and $SST \geq 3 \text{ }^\circ\text{C}$). The equation for M2 is a sort of “energy conservation” metric, but instead of energy it looks at the SIC-SST conservation in a mixed pixel whereby, in a physically realistic scenario, if the SSTs are high in the open ocean portion of the pixel, the concentrations should be low in the ice-covered fraction and, vice versa, if the SICs are high, the corresponding SSTs should be low. This implies that when the SIC and SST are “consistent” in the expected fashion, M2 should be close to 0. The units of M2 are $^\circ\text{C}$. In thermodynamics, thermal energy, κT , where κ is the Boltzmann’s constant and T is absolute temperature, has units of Kelvin. Although consistent in a physical sense with units of thermal energy conservation, the units of M2 are meaningless in the present context, so we have opted to ignore them hereafter. By averaging M2 by the unfiltered population size, we are simply bringing the sum of the product of the two SDs to a manageable number that can otherwise be very large and difficult to grasp. The proposed metrics are evaluated in Tables 7 – 10 for the SICCI product and in Tables 11 – 13 for OSISAF. We also include the SD_{sic} , the standard precision metric, for completeness. As shown in Tables 7 and 11, the original (unfiltered) set of inconsistencies has M2 metrics which are far from the expected zero-target, with $M2 = 1.6$ for the SICCI inconsistencies and $M2 = 4.0$ for the OSISAF inconsistencies.

Before we proceed with the assessment of the impact of the corrections, we need to take a look at Fig. 18. The different panels in Fig. 18 show the SICCI data remove by each filter when applied in isolation. The takeaway message from this illustration is that there is an overlap in the parameter space of corrections with different filters removing portions of the same data. The impact of each individual filter, therefore, depends in the order in which they are applied. Since some filters remove significant portions of the data (i.e., F1, F3) and others remove very little (F2), it may seem that the last filters in the sequence have little or no impact at all (F4, F5, and F6). For this particular reason, we first look at the overall impact of the filters in isolation and then we explore the impact of the sequencing order, especially in connection with the order of the filtering schemes evaluated in the previous section. A note of caution with regards to filters F5 and F6 is that F5 cannot be applied individually without F3, whereas F6 and F3 are mutually exclusive. This is exemplified in the last two panels of Fig. 18. It can be seen in the left panel that if F5 is not bounded by an SST_{max} in the interval for $SIC < 10\%$, then it is not going to filter out the low concentrations with extremely high SST inconsistencies. Leaving that fraction of data would have a measurable impact in the SD of the truncated set; therefore, F5 has to be used in conjunction with F3. The right panel, on the other hand, shows that F6 does not need an SST max bound since the function for SST_{lim} (Eq. 1) crosses $SIC = 0.01\%$ (the minimum observed SIC inconsistency according to the statistics presented in Tables 3 and 5) at $SST = 12.7 \text{ }^\circ\text{C}$. Leaving the small fraction of inconsistencies with SSTs between 10 and $12.7 \text{ }^\circ\text{C}$ does not have a significant impact in the SD. Hence, we can skip using the F3 with applying the filtering sequence ending with F6.

To evaluate the overall impact of the corrections, we first applied the filters one at a time and looked at the change in the metrics for the remaining set (see Table 7 for SICCI and Table 11 for OSISAF). We then introduce filter combinations in different sequences following different scenarios (Tables 8 – 10 and Tables 12 – 13 for SICCI and OSISAF, respectively). For instance, in Tables 8 (SICCI) and 12 (OSISAF), we evaluate the impact of using filters based on SIC and SST only (no distance corrections applied). In Tables 9 – 10 (SICCI) and 13 (OSISAF), we show the additional impact of bringing a distance correction into the mix as we evaluate the impact of the sequential order in the F5- and F6-chains. Table 14 looks at the impact of the current DMI inconsistency flagging on SICCI and OSISAF for the overlapping domain with the proposed filters (i.e., $SIC < 15\%$).

5.2.1 Impact of the filters on the SICCI inconsistencies

The impact of the individual corrections for the SICCI product is reflected in the order of the columns shown in Table 7, with the order of the parameters varying from less impactful to more impactful according to the M2 metric. Some interesting aspects of the results condensed in Table 7 for the SICCI product are as follows. The correction for extreme SIC values (F2) removes a tiny fraction of data ($< 1\%$), yet it causes a measurable impact on both the M1 and M2 metrics adding confidence to the recommendation that these retrievals should be removed (or alternatively, set back to open water, i.e., $SIC = 0\%$). The smearing correction and distance to land (F4 and F1) corrections have a very similar impact for the SICCI product in terms of the M2 metric, which is surprising since the distance correction was initially introduced for the benefit of the OSI-450 product, which has a significantly greater number of inconsistencies relative to SICCI along the coasts, as revealed by the inconsistency maps in Figs. 5 and 6. These coastal inconsistencies are in all likelihood the result of land-contaminated pixels in the microwave frequencies. The SST corrections F3+F5 have the greatest impact when applied individually according to the M2 metric. These filters produce outputs with M2 of 0.38 and 0.36, which is a significant reduction from the $M2 = 1.6$ in the original set, although they eliminate 97% of the data. Note that the metric M1 is not sensitive to the additional impact of F5 over F3, giving the same $M1 = SD_{sic} + SD_{sst}$ for both F3 and F3+F5. As a matter of fact, the M1 metric does not discriminate much among the top three most influential filters identified by M2 (last 3 columns in Table 7). By being insensitive, however, this metric suggests that the SDs of the truncated sets are robust, which adds confidence to the direction of the filters. It is also noteworthy that $M1 = SD_{sic} + SD_{sst} \approx SD_{sic}$; thus, $SD_{sst} \rightarrow 0$. This suggests that the final error in the filtered inconsistencies is dominated by SIC contributions. The implication of this observation is that by filtering spurious SIC inconsistencies, we have also eliminated potential SST outliers. The fact that SD_{sst} is not contributing to the error is welcome news since a main cautionary argument against using (NWP) ancillary data in the atmospheric correction of SIC DCRs is that these fields might introduce trends of their own. A more reaching implication is that a combined microwave SST – SIC retrieval can be beneficial for reducing noise in both products, especially at high latitudes where the SSTs are also known to have large uncertainties, albeit introduced by different sources of noise (SST is retrieved at lower microwave frequencies ($\sim 6 - 11$ GHz) where cloud water (hence, LWC) has a very different spectral signal from the SST and can be easily removed. In other words, SST has the potential to enhance the RTM-based atmospheric correction of the T_b in the SIC algorithm retrieval and vice versa, higher MW frequencies can help discriminate water vapor effects dominating the atmospheric correction of the SSTs).

The statistics shown in Table 8 consider corrections based on the SIC-related parameters (F2 and F4) and an ancillary SST product (F3, F3+F5, and F6). Applying a filter combination based on SIC parameters only (SIC+SSE) eliminates half the inconsistencies with a significant reduction in the M2 metric, but the resulting $M2 = 0.55$ indicates marginal improvements compared to what can be achieved if one is willing to consider ancillary data for filtering purposes. By adding an SST-based filter, M2 can be reduced even more, to 0.32–0.36, as shown in Table 8, but the real impact is seen when the distance-from-land filter (F1) is also included (Tables 9 and 10). A distance filter in combination with SIC-only filters (SIC + SSE) eliminates $\frac{3}{4}$ of the inconsistencies and results in M2s as low as ~ 0.2 (Table 9). Combinations including corrections for SIC, SST and distance result in $M2 < 0.18$ (last 2 columns of Tables 9 and 10).

In terms of the F6 or F5 corrections, the two filtering schemes end up having the same impact with $M2 = 0.16$ for the F6-chain, just slightly better than the $M2 = 0.17$ for the F5-chain. Implementing one over the other has no significant difference except for the fact that the F6-chain has one less filter (the F3) than the F5-chain. These filter combinations bring us closer to the zero-target in terms of our M2 metric, indicating that the remaining SICCI values ($< 2\%$ of the original set) are “consistent” with their SST counterparts, at least in terms of the SIC vs SST dependence depicted in Fig. 15. The M1 metric, on the other hand, is not sensitive enough with basically all the sequencing realizations in Tables 9 and 10 achieving $M1 \approx 15.4$. Once again, for filter combinations with $M2 < 0.38$, $SD_{\text{sic}} + SD_{\text{sst}} = SD_{\text{sic}}$, which suggests very little spread in the truncated SSTs after a few SIC corrections, which reiterates the fact that by reducing the SIC inconsistencies we are also eliminating noise in the matching SST retrievals.

There is also something to be said about the impact of the filter combination F1+F2+F3 (DIS+SIC+SST) shown in Table 10, column 2. By eliminating only those SIC believed to be wrong/misclassified (DIS+SIC+SST in Table 10), the original inconsistency set is reduced by 97.1%, with the remaining set having $M2 \sim 0.4$. Additional filtering with F4 and F5 (the last 2 columns of Table 10) eliminates a tiny fraction of noisy SIC retrievals ($\sim 1\%$), but their removal makes all the difference by cutting M2 in half and bringing the metric much closer to zero ($M2 < 0.2$). These results, as well as the high SIC correction (F2) in Table 7, suggest that “the devil is in the details,” as these small packets of concentrated variability can be highly impactful and should not be ignored.

Table 7: Impact of the individual corrections to the SICCI product as measured by different metrics

Filter	Original	SIC (F2)	SST (F6)	SSE (F4)	DIS (F1)	SST (F3)	SST (F3+F5)
% Removed	0.0	0.6	93.9	50.2	52.7	97.1	97.0
SD_{SIC} Remain	17.0	16.0	18.5	17.5	14.7	15.4	15.4
$SD_{\text{sic}} + SD_{\text{sst}}$	17.4	16.5	18.6	17.8	15.0	15.4	15.4
M2	1.60	1.56	0.78	0.57	0.51	0.38	0.36

Table 8: Impact of the SICCI corrections in terms of the proposed metrics for filter combinations including SIC and SST terms only

Filter Combo	SIC+SSE	SIC+SSE+SST	SIC+SSE+SST+F5	SIC+SSE+F6
% Removed	50.7	50.7	97.0	97.0
SD_{SIC} Remain	16.0	17.7	17.7	17.8
$SD_{\text{sic}} + SD_{\text{sst}}$	16.4	17.7	17.8	17.8
M2	0.55	0.36	0.33	0.32

Table 9: Impact of adding the distance correction to the filter combination sequence used with SICCI that culminates with filter F6 as shown in Table 4

Filter Combo	DIS+SIC+SSE	DIS+SIC+SSE+SST	DIS+SIC+SSE+F6
% Removed	72.1	98.2	98.3
SD _{SIC} Remain	14.1	15.3	15.4
SD _{sic} +SD _{sst}	14.2	15.3	15.4
M2	0.21	0.18	0.16

Table 10: Impact of adding the distance correction to the filter combination sequence used with SICCI that culminates with filter F5 as shown in Table 4

Filter Combo	DIS + SIC	DIS+SIC+SST	DIS+SIC+SST+SSE	DIS+SIC+SSE+SST +F5
% Removed	52.7	97.1	98.2	98.3
SD _{SIC} Remain	14.7	15.4	15.3	15.3
SD _{sic} +SD _{sst}	14.9	15.4	15.3	15.4
M2	0.51	0.38	0.18	0.17

5.2.2 Impact of the filters on OSISAF inconsistencies

Once again, the impact of the individual filters when applied to OSISAF is reflected in the order of the columns of Table 11, shown in increasing degree of impactfulness. Very interesting differences with respect to SICCI are observed at once. The most obvious one is that the most impactful correction for OSISAF is the distance-from-land filter (F1). Not only that, but this filter alone is enough to clean the OSISAF inconsistencies to look a lot like SICCI as revealed in the figures illustrating the impact of F1 (compare the top-right panels in Figs. 20 vs. 19, and Figs. 25 vs. 24). Most surprisingly, the F1-ensuing $M2 = 0.07$ is less than half as much the $M2$ for the heavily redacted SICCI subsets that survive the F5- and F6-chains ($M2 \sim 0.16 - 0.17$ from Tables 9 and 10). Also, there is a large separation between this and the other $M2$ values for the other filters, implying that when applied in isolation, none is as effective as F1. In fact, adding the whole sequence of filters used in the F5- and F6-chains (Table 13), only brings the $M2$ metric down by 0.02, from $M2 = 0.07$ to 0.05, at the expense of eliminating most all inconsistencies for OSISAF (only 0.5% of the original set survives the F5 and F6 filtering schemes). It is important to bear in mind that a filtering scheme for OSISAF based solely on distance-from-land would not exclude bad retrievals affected by atmospheric effects reflected in very high SSTs; thus, a minimum amount of filtering would also have to include a correction for SST max such as F3 to eliminate unwanted ice in the Baltic Sea for July and August (see the maps of remaining inconsistencies after F1 and F1+F3 in Fig. 20). Table 13 (first column) shows that this minimum filtering effort results in a satisfactory $M2 = 0.06$, just 0.01 above the metrics for the full filtering schemes considered so far. Final SDs for SIC and SIC and SST combined (Table 13) are $\sim 11.8\%$, which implies that the remaining OSISAF inconsistencies are representative of regions closer to the open water (very low concentrations), whereas the SICCI inconsistencies (SD $\sim 15.3\%$) are representative of the ice edge (SIC = 15% is the traditional value used to determine the ice edge).

Also interesting is the fact that the OSISAF product is least impacted by the smearing uncertainty correction (F4). This filter removes a merely 8.6% of the inconsistencies with a minor improvement in the $M2$ metric ($M2$ decreases from 4 in the original set to 3.68 in the F4-

filtered data). This should not be surprising since the smearing uncertainty correction was introduced to filter out noise driven by interpolation artifacts at the boundary between open water and ice (i.e., SIC = 0%), that were most likely introduced when the level 2 products were transformed to level 3 (i.e., going from satellite swath geometry to a regular grid) as part of the processing chain of the SIC products. Since SICCI is a higher resolution product than OSISAF, it has higher smearing uncertainties at the ice edge compared to the smooth transitions in OSISAF. This is apparent from visual inspection of the SSE MVC maps of the SICCI and OSISAF inconsistencies shown for June and July 2014 (bottom rows) in Figs. 5 and 6, respectively. The smoothness of OSISAF explains why the SSE filter has less impact for OSISAF than SICCI, even when combined with the SST filters (Table 12).

Table 11: Impact of the individual corrections to the OSISAF product as measured by different metrics

Filter	Original	SSE (F4)	SST (F3)	SSE revised	SST (F3+F5)	SST (F6)	DIS (F1)
% Removed	0.0	8.6	94.1	61.1	95.0	95.4	90.5
SD _{SIC} Remain	14.1	14.1	14.4	13.8	13.9	14.3	11.5
SD _{sic} +SD _{sst}	14.6	14.5	14.5	14.2	14.0	14.4	11.6
M2	4.0	3.68	1.92	1.66	0.82	0.73	0.07

Table 12: Impact of the OSISAF corrections in terms of the proposed metrics for filter combinations including SIC and SST terms only

Filter Combo	SSE+F3	SSE+F3+F5	SSE+F6
% Removed	94.5	95.4	95.7
SD _{SIC} Remain	14.3	13.9	14.3
SD _{sic} +SD _{sst}	14.5	14.0	14.3
M2	1.80	0.77	0.68

Table 13: Impact of adding the distance correction to the filter combination sequence used with OSISAF and assessment of the two filtering schemes introduced in Table 6

	F5			F6		F6 revised
Filter Combo	DIS+F3	DIS+SSE+F3	DIS+SSE+F3+F5	DIS +SSE	DIS+SSE+F6	DIS+SSE _{rev} +F6
% Removed	99.5	99.5	99.5	91.3	99.5	99.8
SD _{SIC} Remain	12.2	11.8	11.8	10.8	11.8	11.8
SD _{sic} +SD _{sst}	12.3	11.8	11.8	10.9	11.8	11.1
M2	0.06	0.05	0.05	0.06	0.05	0.02

5.2.3 Impact of the DMI Inconsistency Definition Index

Here, we evaluate what happens if we apply a filter that eliminates all the SIC retrievals flagged by the DMI inconsistency index (i.e., SIC > 15% with SSTs > 3 °C). The purpose of this comparison is to evaluate the impact of extending the definition of the DMI inconsistency set to the range $0 < \text{SIC} < 15\%$ adopted in this study. This is equivalent to considering a new filter $\text{SIC}_{\text{max}} = 15\%$. M1 and M2 metrics for the resulting filtered set are shown in Table 14. Such a filter truncates 34.5% of the original SICCI inconsistency set considered here. Although the

remaining set has a significantly lower M2 (0.56) than the original set, the corresponding M1 metric is not consistent with the M1 values reported in Sections 5.2.1 and 5.2.2 ($M1 = 5.9$ (Table 14) vs. $M1 \sim 12 - 15$ (Tables 7 – 13)). Similar low M1 values were observed in the course of this investigation (not shown in the metrics evaluation tables) when the F5 correction was applied in isolation without the F3 filter (i.e., without an SST_{\max} requirement). We, therefore, tested the impact of adding a temperature bound to the 15% SIC_{\max} filter. We used the same $SST_{\max} = 10$ °C threshold from before. Table 14 indicates that additional filtering lowers the M2 even further to 0.24 with minimal additional truncation (the new threshold eliminates an additional 3.9%). The sum of the SDs ($M1 = 4.5$), however, is still significantly lower than the M1 results from before (Tables 7 – 13 indicate that the denoised inconsistencies have spatial variabilities representative of the sea ice edge region). This discrepancy is perhaps due to the fact that the extended (non-overlapping) inconsistency set is literally contained in a box constrained by ($SIC = 0\%$, $SIC = 15\%$, $SST = 3$ °C, and $SST = 10$ °C), whereas in the proposed filtering scheme the denoised inconsistencies are bound by SIC-SST function given in Eq. 2. In other words, the 15% SIC_{\max} filter cuts the portion of the variability outside the ice edge (and in the process, lowers the SD of the truncated set), whereas the proposed filtering scheme allows for full variability in the MIZ (the SICs are truncated at 75% and $SST_{lim} = f(SIC)$). The impact of allowing for full variability outside the ice edge can be appreciated in Figs. 23 and 24, showing the changes in distribution after each filter application. Although the filtered sets shown in these figures have the highest density of inconsistencies (yellow-blue hexbins) within the bounds of the non-overlapping SIC domain (0%, 15%), there is a significant number of potentially valid SIC retrievals (red bins) with intermediate and high concentrations. By removing all $SIC > 15\%$, this filter not only eliminates potentially valid observations in the MIZ, but also removes an important contribution to the SD from the truncated concentrations. A successful denoising algorithm should achieve both noise reduction and feature preservation.

Table 14: Impact of the DMI inconsistency filter

SIC Product	SICCI		OSISAF	
Filter Combo	DMI remain	DMI+SST	DMI remain	DMI+SST
% Removed	34.5	96.1	42.9	96.5
SD_{SIC} Remain	4.3	4.3	4.2	4.3
$SD_{sic} + SD_{sst}$	5.9	4.5	5.2	4.6
M2	0.56	0.24	0.57	0.36

5.3. Impact of the filters on the joint pdf of SIC vs. SST and SIC vs. SSE for the de-noised data

A final assessment of the filtering schemes just proposed is done in terms of the joint probability density functions for the fraction of data that remain unaffected after all the filters are applied. It is not the purpose of this investigation to characterize these distributions. We will dwell no further than a visual inspection to check if the de-noised data sets have distributions with similar shapes for both SIC products, as noise reduction algorithms always tend to alter the signal being de-noised to some degree. The assumption is that the remaining sets are, in fact, valid SICs, but there is always the possibility that undetected false or noisy retrievals still persist because they

result from sources other than the ones being addressed by filter design or because the selected thresholds may not match the distribution of the noise.

The joint densities of SIC and SST, $f(\text{SIC}, \text{SST})$, and SIC and SSE, $f(\text{SIC}, \text{SSE})$, for the untouched data by the F5- and F6-filter chains, are shown in Figs. 21 and 22 for SICCI and OSISAF, respectively. The SIC vs. SST dependence is characterized by an elliptical joint distribution where the isocurves (curves of constant density) adopt ellipsoid shapes. This type of distribution allows for the presence of heavy tails, as in the scatter plots of SIC vs. SST (Fig. 15) or in the marginal distributions for SIC and SST (Fig. 13). Despite the presence of tails, the distribution remains a simple linear dependence structure (i.e., the linear correlation remains the canonical measure of dependence between the two variables). If we compare the $f(\text{SIC}, \text{SST})$ for the de-noised sets with the joint distributions of the original inconsistency sets shown in Fig. 14, it can be seen that the shape of the functional dependence between SIC and SST remains invariant under the increasing transformation of the inconsistencies resulting from the sequential filters. The transformation after each filter is broken down in Figs. 23 and 24 for $f(\text{SIC}, \text{SST})$ for SICCI and OSISAF, respectively. The isocurves in the final distributions are somewhat distorted compared to the ellipsoids in Fig. 14, but the underlying population sizes have shrunk considerably as a result of the filtering.

The shape of $f(\text{SIC}, \text{SSE})$, before (Fig. 14) and after the filters (Figs. 21 and 22), is something else entirely. The $f(\text{SIC}, \text{SSE})$ transformation after each filter is also broken down in Figs. 25 and 29 for SICCI and OSISAF, respectively. The joint densities of $f(\text{SIC}, \text{SSE})$ for the de-noised (F5 and F6) SICCI sets (Fig. 21) show similar increasing linear trends that are in agreement with the smearing standard uncertainty parameterization used with all the EUMETSAT SIC products, as depicted in Fig. 3 of Tonboe et al. (2016). In that model, $\text{SSE} = 0\%$ for $\text{SIC} = 0\%$ and 100% . Inside the ice edge, $\text{SSE} = f(\text{SIC})$ with SSE increasing linearly with SIC and reaching a maximum for $\text{SIC} \sim 10\%$. Inside the MIZ, SSE remains constant with increasing SIC ($\text{SSE}_{\text{max}} \sim 12\%$) until about $\text{SIC} \sim 90\%$, when it decreases back to 0% for $\text{SIC} = 100\%$. The shape of $f(\text{SIC}, \text{SSE})$ for SICCI shows very good agreement with the theoretical model, except for a much greater SSE constant inside the MIZ ($\sim 29 - 33\%$, as by construction of the SSE_{max} threshold used in F4). The OSISAF distributions (Fig. 22), on the other hand, appear to increase smoothly over a larger range ($\text{SIC} \sim 15\%$), but the rate of increase is no longer linear. A similar curvature was also observed in the OSISAF marginal distribution for SSE (Fig. 13). Thus, despite the differences in $f(\text{SIC}, \text{SSE})$ between the original inconsistency set and the de-noised F5 and F6 OSISAF sets, the latter, at least, bear some degree of similarity with the marginal density for SSE, following the suppression of the spikes shown in Fig. 13 after application of F4. This is relevant because all marginals (univariate pdfs) of an elliptical joint distribution are also elliptical with the same characteristic generator, i.e., with the same tail function (Bradley and Taqqu, 2003). In other words, distributions with the same tail function belong to the same elliptical family. This can be corroborated in Fig. 27, showing the overlap of the marginal distribution for SICCI (pink pdf) and OSISAF (blue)'s F6-untreated SIC retrievals, where the tails of both products agree rather well for $\text{SIC} > \sim 25\%$. The equivalent plot for F5 is visually undistinguishable from the F6 chain, and hence not shown here. The fact that the tails are converging is encouraging, but the curvature in $f(\text{SIC}, \text{SSE})$ revealed in Fig. 22, suggests that there is residual noise in OSISAF, not present in SICCI, that is not being detected by our filters. We speculate that the departure from linearity in the OSISAF SIC – SSE

dependence in open water is somehow related to the smoothing impact of the 3x3 SIC operator, used in estimating the local smearing variability at the ice edge, in an already smooth product with over-smear edges (see the discussion in Section 3.2).

A closer look at the differences between the OSISAF and SICCI pdfs shown in Fig. 27 helps clarify this issue. Note that for $SIC < 10\%$, the SICCI pdf exhibits a sharper density peak than OSISAF's. However, between 10% and ~25%, OSISAF displays higher SIC densities than SICCI, showing a secondary peak centered at ~15% SIC. The exactness of these thresholds is indicative that the differences in the pdfs at low concentrations are the result of artifacts in the open water filter (OWF) used in UMETSAT's version 2 SIC products. The OWF is designed to remove spurious SIC retrievals in open water and low concentration areas affected by atmospheric influences. Lavergne et al. (2019) explain that the value of 10% is chosen well below the ice edge threshold ($SIC = 15\%$), so the filter does not interfere with sea ice extent evaluations (15% – 30% SIC). Remaining noise from deficiencies in the OWF (e.g., from atmospheric influences not considered such as LWC), however, transfer into uncertainty in SIC (Lavergne et al., 2019).

As we well know, the atmospheric correction, hence the OWF, has more impact over open-water and low-concentration values than over closed-ice conditions (Andersen et al., 2006; Lavergne et al., 2019). From our analyses, filters based on SSE have the greatest impact on ice edge artifact detection, hence our filter F4 overlaps with the OWF spatial domain. Since these two filters are correlated, we look at the SSE density distributions to infer something else about the OWF. The joint pdfs for $f(SIC, SSE)$ suggest that the 10% threshold of the OWF works well for SICCI (Fig. 21) but not so much for OSISAF (Fig. 22). Note from Fig. 21 that SICCI reaches SSE_{max} within $SIC < 10\%$, whereas in OSISAF the increasing trend between SIC and SSE expands beyond the 10%-OWF threshold, peaking around $SIC = 15\%$ (Fig. 22). This implies that the OWF leaves a significant number of noisy OSISAF retrievals untouched inside the ice edge (those with SICs between 10 and 15%). We hypothesize that these unscathed noisy SIC retrievals remaining in OSISAF are responsible for the secondary peak, centered at the ice edge threshold ($SIC = 15\%$), evident in the SIC pdf (Fig. 27) for the OSISAF de-noised data. These inconsistencies should have been masked out by our SSE filter, but stayed untouched by F4. The question is why F4 failed to detect and eliminate OWF-residual noise. The most obvious answer is that the selected thresholds for SSE_{max} and SSE_{min} do not match the distribution of the undetected noise.

As we discussed in Section 3.2, we believe that the SSE is severely underestimated for the OSI-450 CDR. We know that the averaging incurred in resampling the satellite swath (level 2) SIC retrievals onto a grid (level 3) of finer resolution than the footprint of the channels used in the retrieval, smooths and smears the gridded OSISAF product by default. This smearing is independent of the OWF applied at level 2, but has a significant impact in the computation of the OSISAF smearing uncertainties at the pixel level. The smearing uncertainty algorithm is a convolution of the level 3 gridded SICs with a 3x3 high pass filter (edge-enhancement operator). This convolution brings the value of each pixel into closer harmony with the values of its neighbors, causing additional blurring/smearing of the gradients at the sea-ice edge region in the OSISAF product, which in turn results in underestimation of the SSEs distributed with the level 3 OSISAF product. The SICCI product, on the other hand, being based on higher-frequency channels, achieves higher spatial resolution, thus the ice edge is more sharply defined in the SIC

grid cells, resulting in higher SSE pixel estimates as the uncertainty algorithm return larger uncertainty estimates where the gradients are stronger.

The fact that the SSE might be seriously underestimated for OSISAF suggests that the threshold choices for SSE_{min} and SSE_{max} used in F4 (i.e., $SSE_{min} = 3\%$ and $SSE_{max} = 37\%$) should have been more aggressive than even the ones used with SICCI (i.e., $SSE_{min} = 3\%$ and $SSE_{max} = 33\%$). As we discussed above, the joint pdfs for OSISAF's $f(SIC, SSE)$, shown in Fig. 22 for F5 and F6, indicate a departure from linearity (curvature) that extends to $\sim 14 - 16\%$ SICs. We also commented on the fact that the tails of the SIC pdfs for both CDRs converge for $SIC > 25\%$ (Fig. 27). Using these observations as a guideline to choose the more aggressive SSE thresholds, we reevaluated the OSISAF filtering using the F6-chain but with a new F4-filter that had revised thresholds: $SSE_{min} = 14\%$ and $SSE_{max} = 25\%$. A comparison of the new OSISAF SIC pdf and the previous one for SICCI is shown in Fig. 28. As can be seen from this inter-comparison, the secondary peak in the new OSISAF pdf has been eliminated as a result of the revised thresholds in the F4 filter. The smoothness of the new OSISAF pdf also reflects the smearing inherent in the OSIAF level 3 product. As for $f(SIC, SSE)$, Fig. 29 reveals that the joint density for the truncation resulting from the revised filter does, indeed, show linearity between SSE and SIC inside the 10% SIC-range prescribed in the smearing uncertainty parameterization of Tonboe et al. (2016) that matches the $f(SIC, SSE)$ for SICCI given in Fig. 21.

Despite small differences in the smoothness of the isocurves, the joint distributions for the filtered (F6)-transformed data from OSISAF (Fig. 29) not only display the same SIC-SST and SIC-SSE features of the SICCI dependences (Fig. 21), but the shape of those features matches the expected behavior from the theoretical model exceptionally well. Note, however, that to guarantee linearity between SIC and SSE for $SIC < 10\%$ in OSISAF, the SSE thresholds have to be changed from 3 to 14% for SSE_{min} and from 37 to 25% for SSE_{max} . The individual impact of the revised F4 filter is shown in Table 11. The tabulated results confirm that the new SSE thresholds have a significantly larger impact at reducing noise; thus, they are a better match for the distribution of the underlying noise. Metrics for the F6 filtering scheme using a F4 with the stricter selection criteria (Table 13, F6 revised) are: $M1 = 11.1$ and $M2 = 0.02$ at a cost of removing 99.8% of the original inconsistencies. The remaining 0.2% of the original set of inconsistencies is still a good-sized population with 439666 valid SIC retrievals salvaged (the original set is close to 88 million-large for OSISAF). Comparing these values with the metrics for the F6 filtering scheme using F4 with underestimated SSE thresholds (Table 13, F6), it is evident that the removal of the additional noise from the OWF brings the $M2$ metric ever closer to zero, which implies that, from a statistical point of view, the new F6-chain outperforms all the others. Results for the F5-chain are almost identical, so we can finally settle in one or the other, and we have chosen F6 just because it has one less filter in the chain. In summary, the proposed filters, while successful at eliminated noise, have also returned denoised signals that belong to the same elliptical family, i.e., are identically distributed, regardless of the SIC product characteristics.

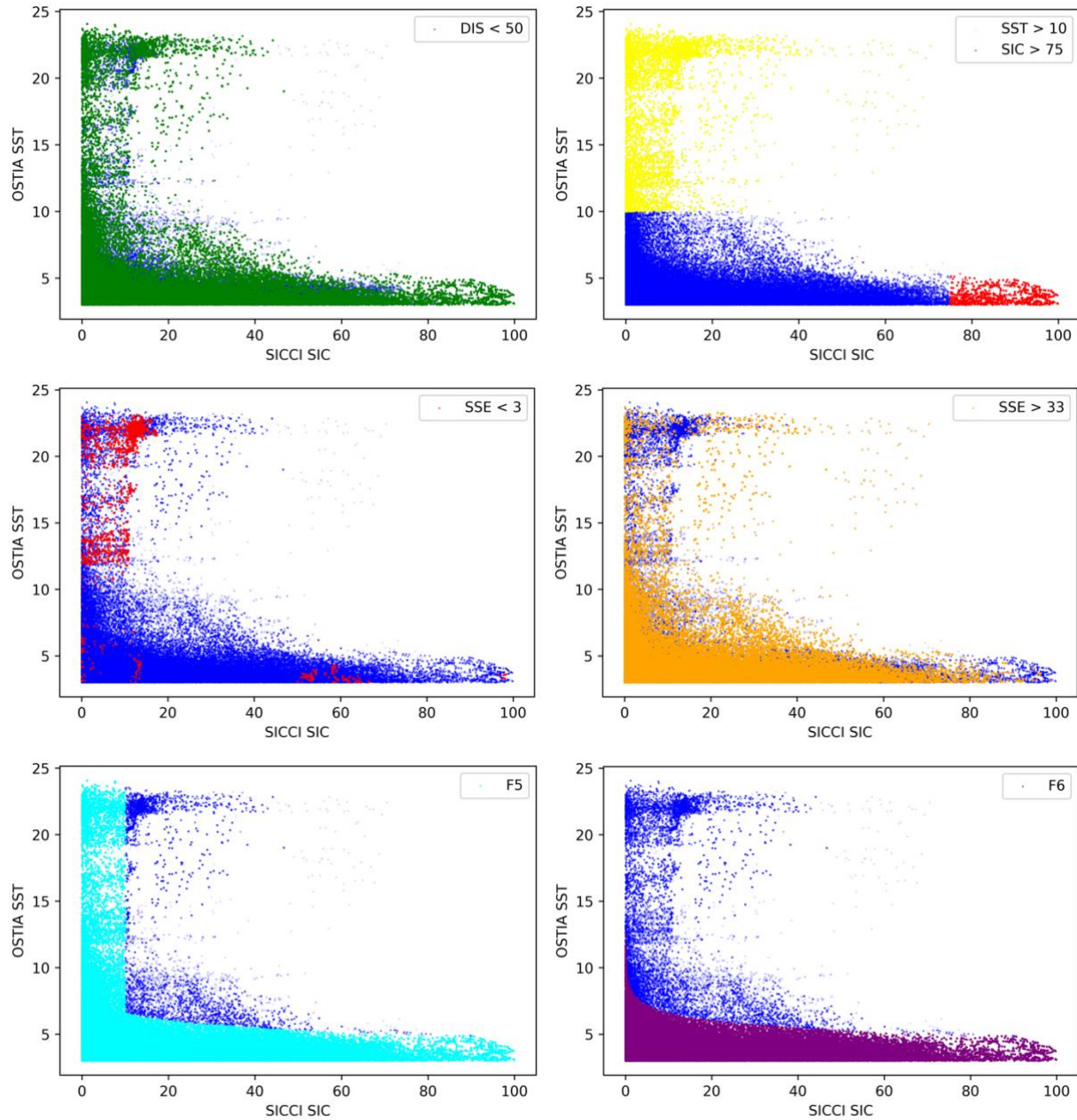
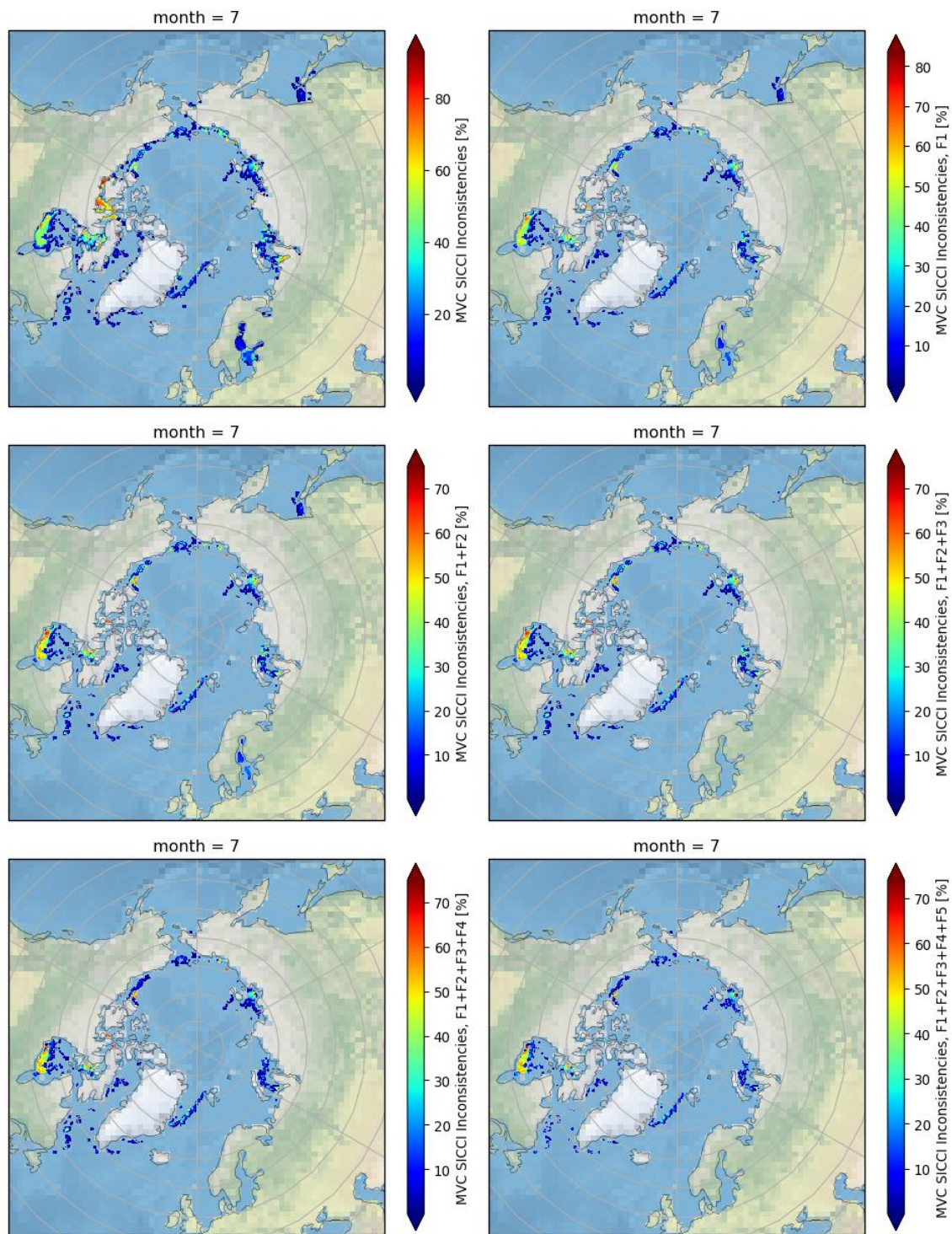


Figure 18. Scatter plot of SIC vs. SST for SICCI showing the parameter space of the SIC inconsistencies cut off by different filters. Top-left: green dots are SIC with DIS < 50 km (F1); top-right: red dots are SIC > 75% (F2) and yellow dots are SIC with SST > 10 °C; middle-left: red dots are SIC with SSE < 3%; middle-right: orange dots are SIC with SSE > 33%; bottom-left: cyan dots are SICs preserved by the F5 filter; bottom-right: purple dots are SICs preserved by the F6 filter.



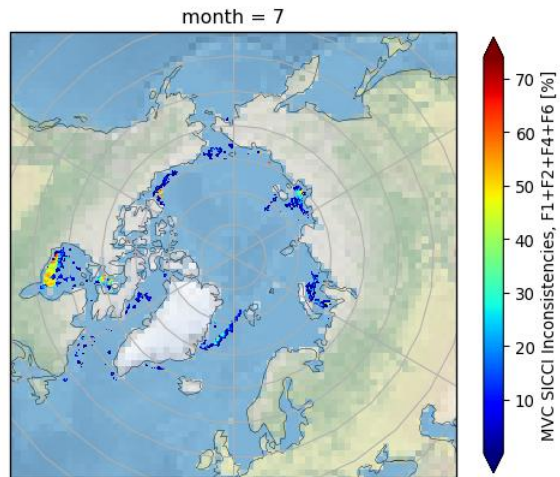


Figure 19. July 2014 – MVC for SICCI SIC inconsistencies after sequential application of different filter combinations

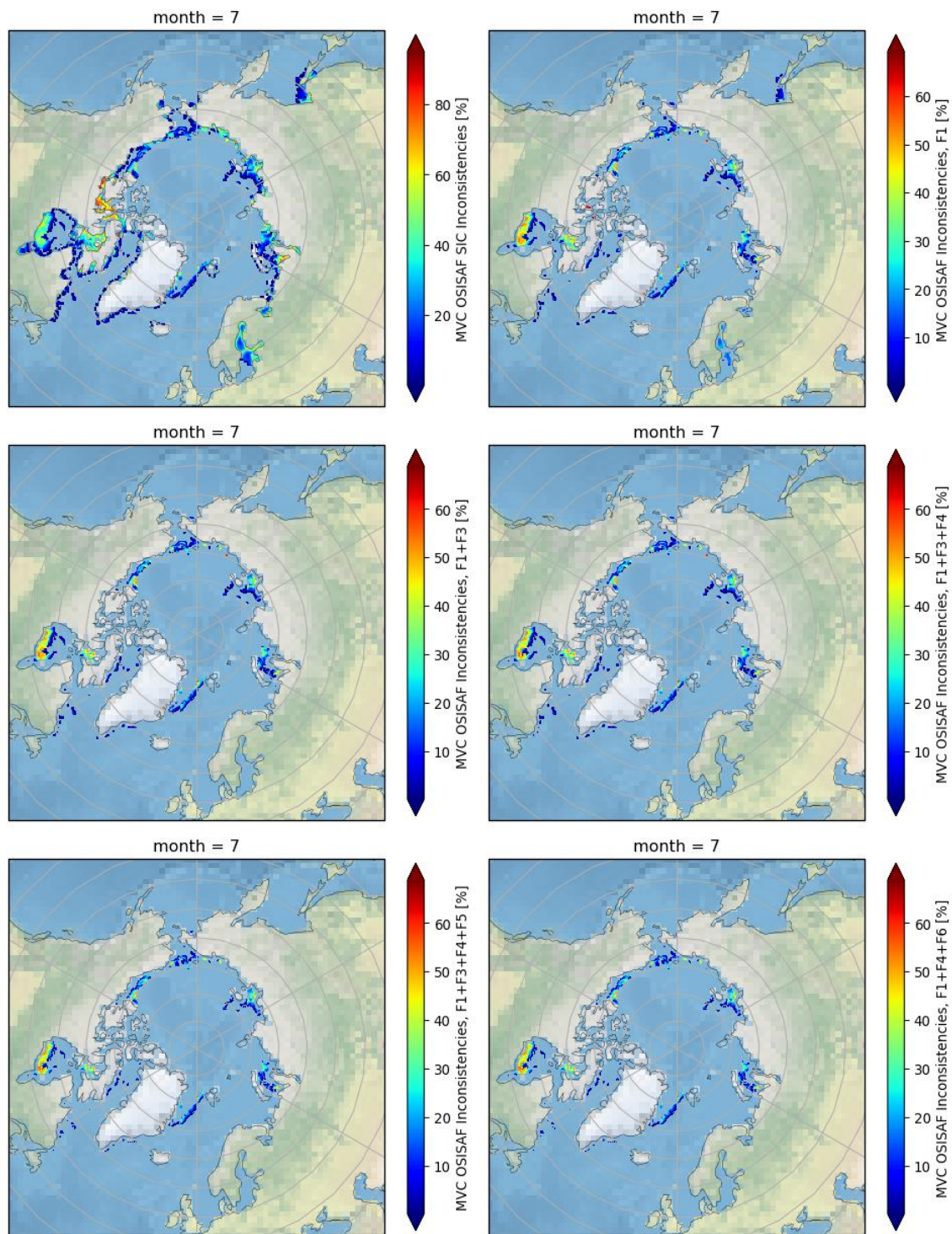


Figure 20. July 2014 – MVC for OSISAF SIC inconsistencies after sequential application of different filter combinations.

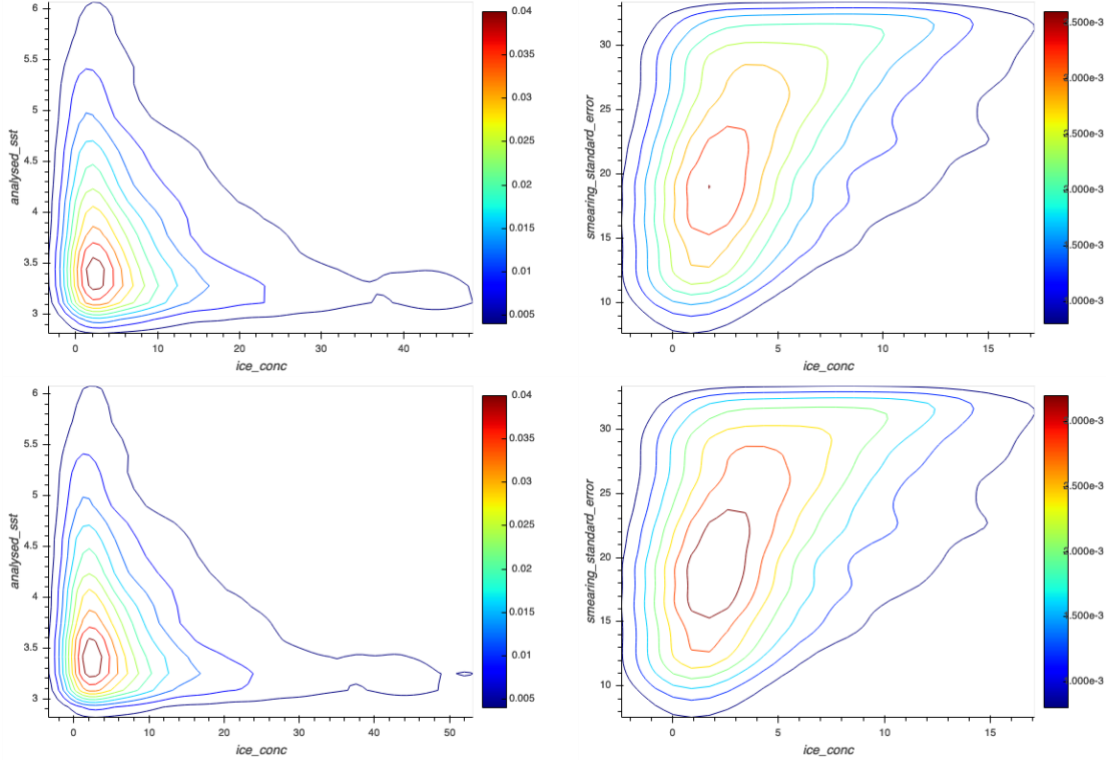


Figure 21. SICCI joint pdfs for filtered SIC vs. SST (left) and SIC vs. SSE (right) for F5 (top) and F6 (bottom) final filtering schemes.

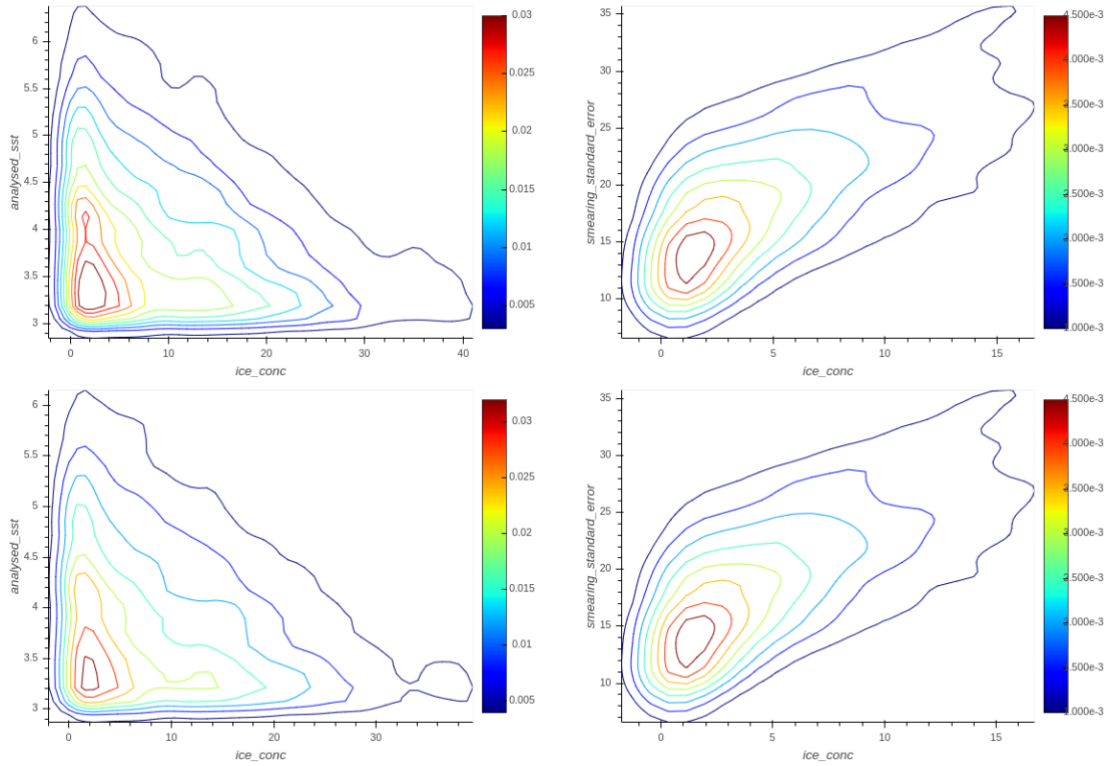


Figure 22. OSISAF joint pdfs for filtered SIC vs. SST (left) and SIC vs. SSE (right) for F5 (top) and F6 (bottom) final filtering schemes.

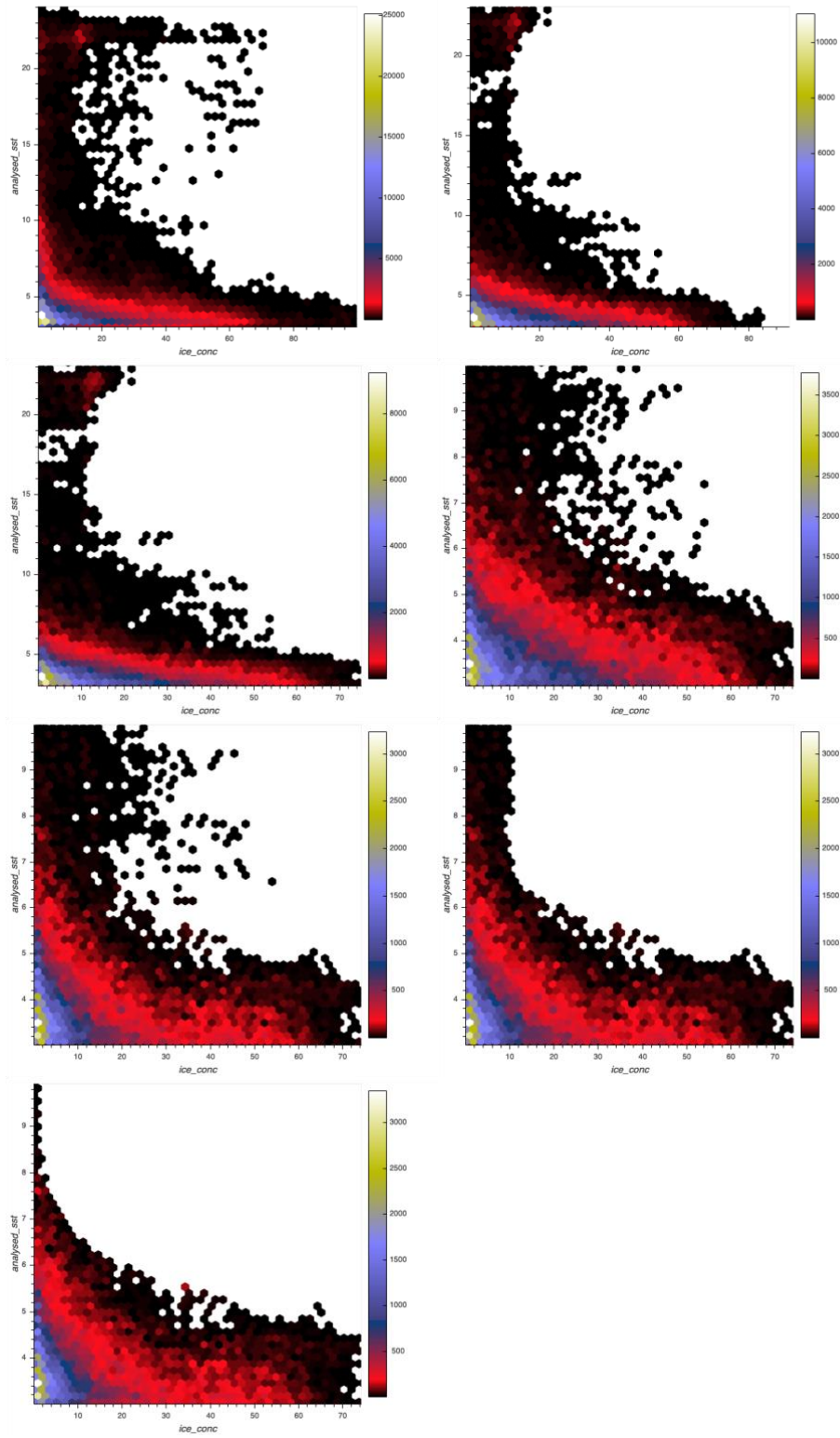


Figure 23 Hexbin density plots (counts per hexbin) of the SIC-SST dependence on truncated SICCI inconsistency sets after a new filter is applied. Top-left plot corresponds to the original inconsistency set. Next-right: F1-filtered set; row 2-left: F1-F2-filtered set; row 2-right: F1-F2-F3 filtered set; row 3-left: F1-F2-F3-F4 filtered set; row 3-right: F1-F2-F3-F4-F5 final set. Last row is for the final sets that remain after F6 filtering sequence, F1-F2-F4-F6.

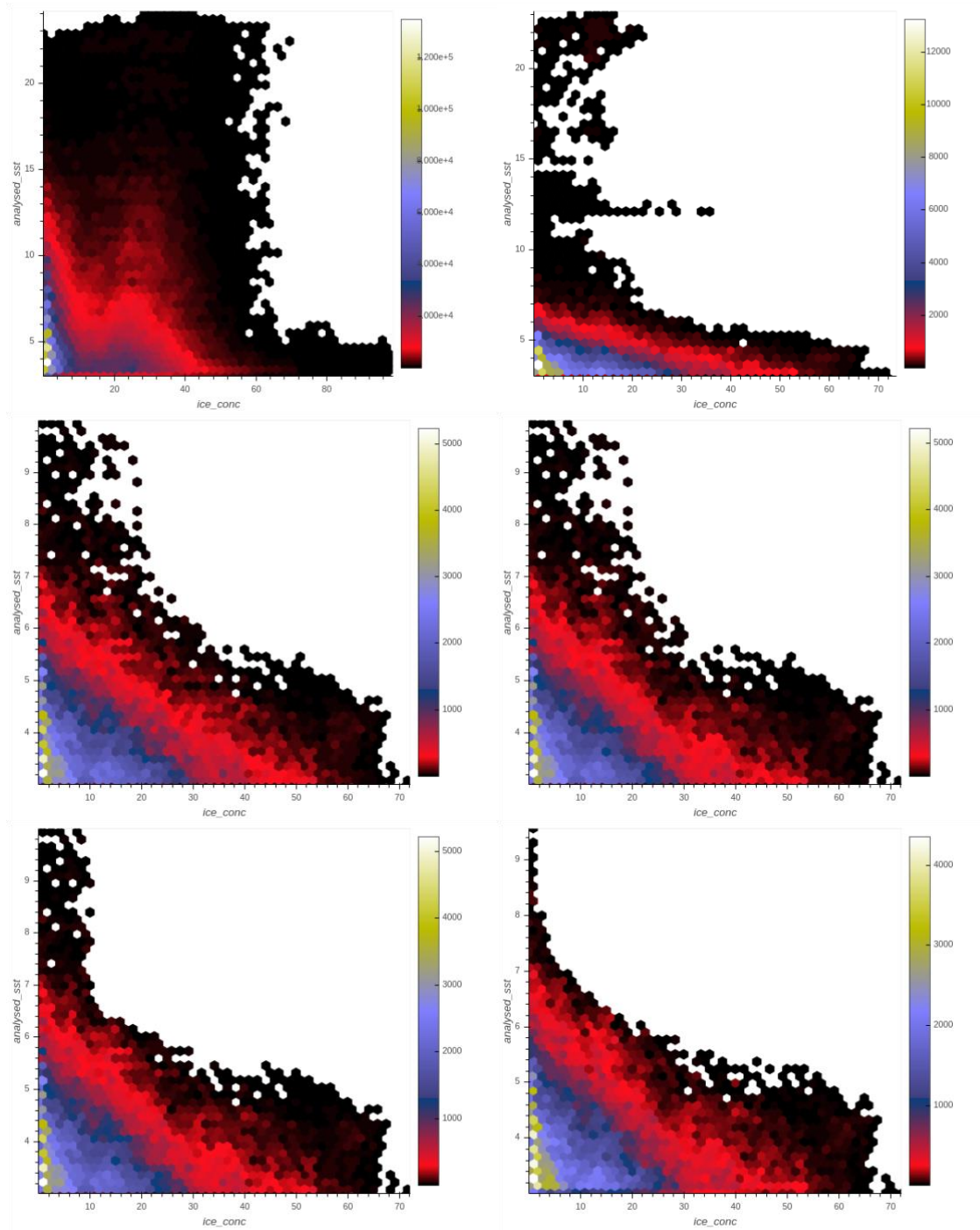


Figure 24. Hexbin density plots (counts per hexbin) of the SIC-SST dependence on truncated OSISAF inconsistency sets after a new filter is applied. Top-left plot corresponds to the OSISAF unfiltered data; top-right plot is for the F1-filtered set; middle-left is for the F1-F3 truncated set; middle-right is for the F1-F3-F4 truncated set; bottom-left plot is the final set from F1-F3-F4-F5, and bottom-right plot is the final set from F1-F4-F6 chain.

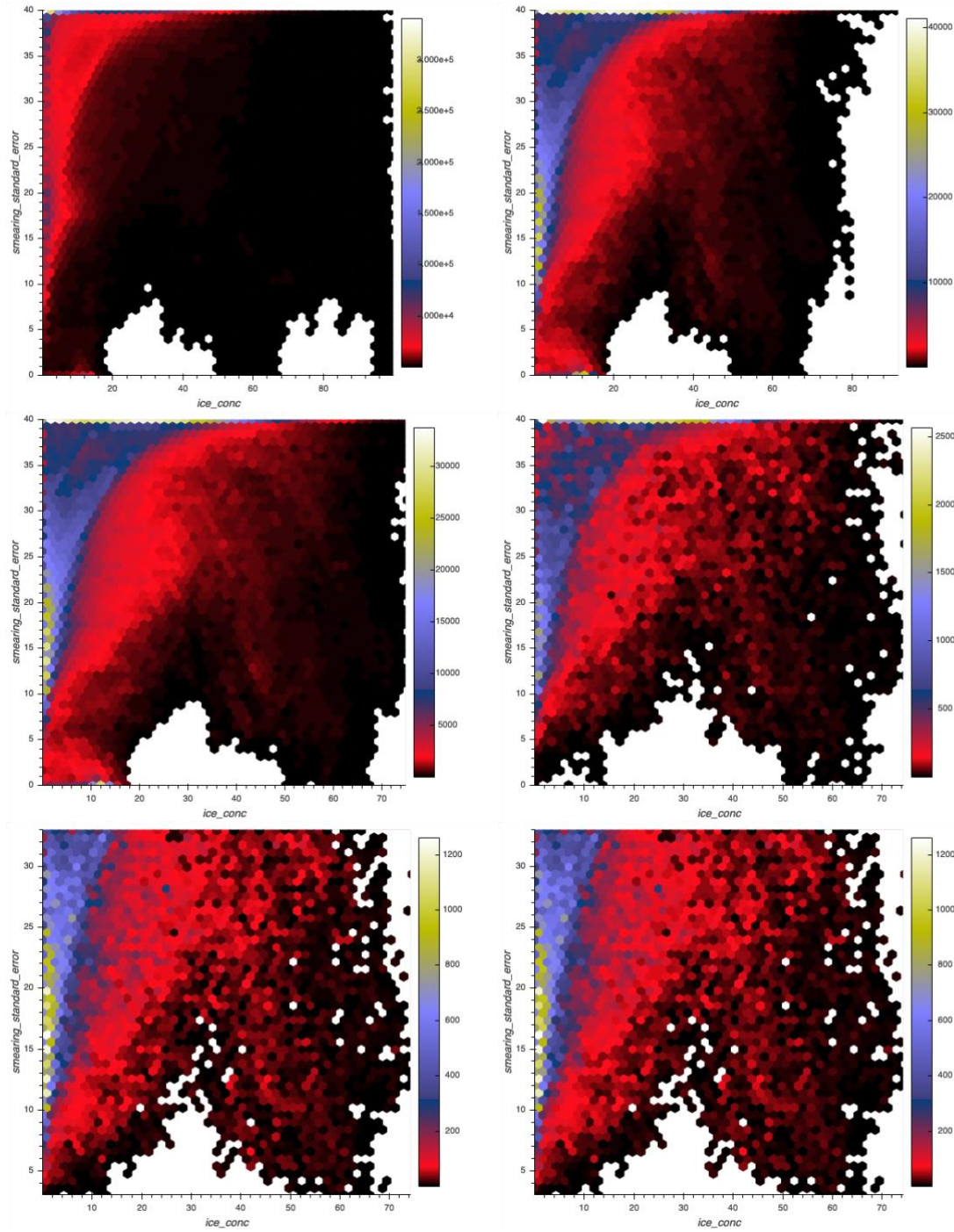


Figure 25. Bin density plots for SICCI SIC vs. SSE dependence after each filter is applied sequentially, starting with a.) no filter, b.) F1, c.) F1+F2, d.) F1+F2+F3, e.) F1+F2+F3+F4, d.) F1+F2+F3+F4+ F5

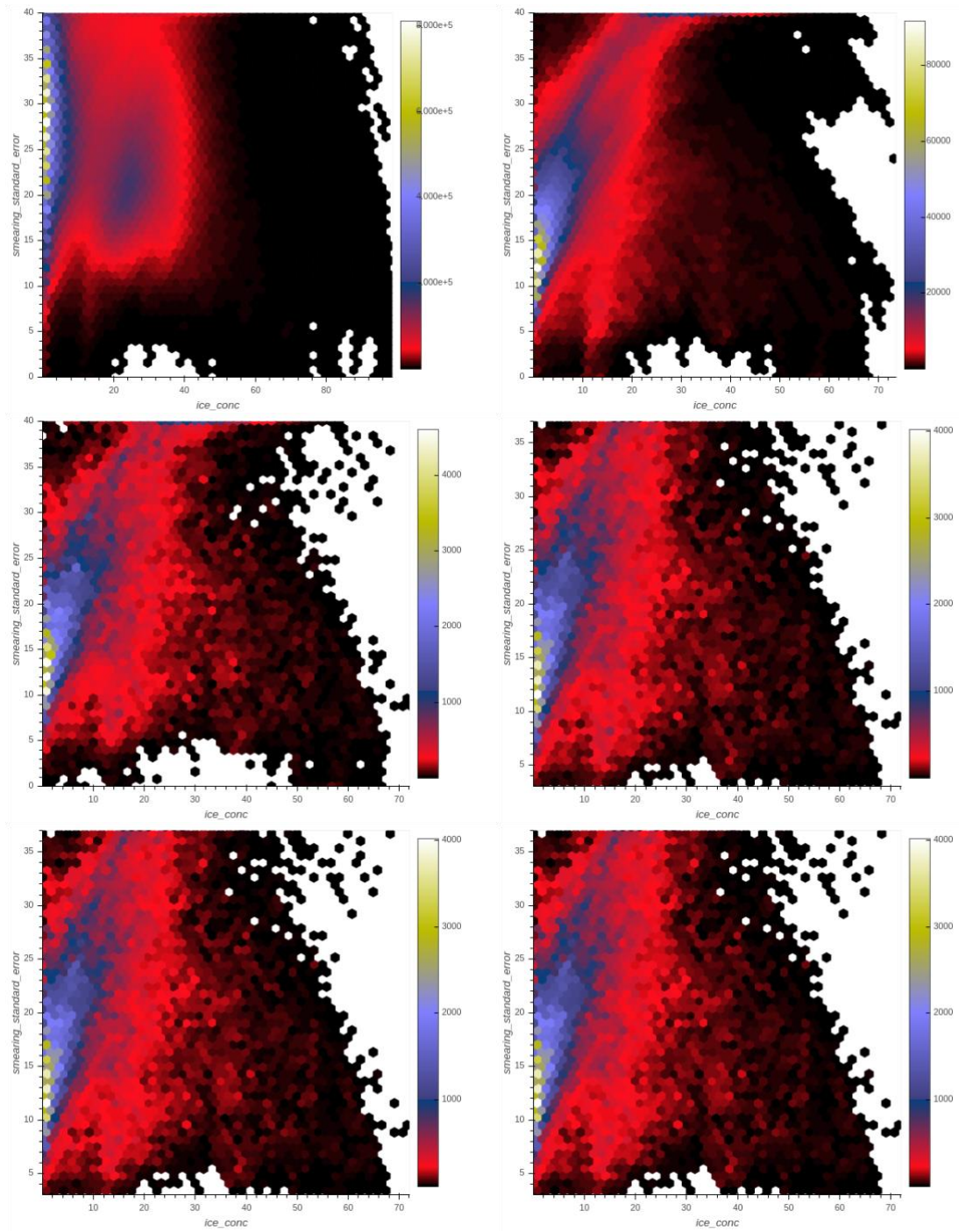


Figure 26. Bin density plots for OSISAF SIC vs. SSE after each filter is applied sequentially, starting with a.) no filter, b.) F1, c.) F1+F3, d.) F1+F3+F6

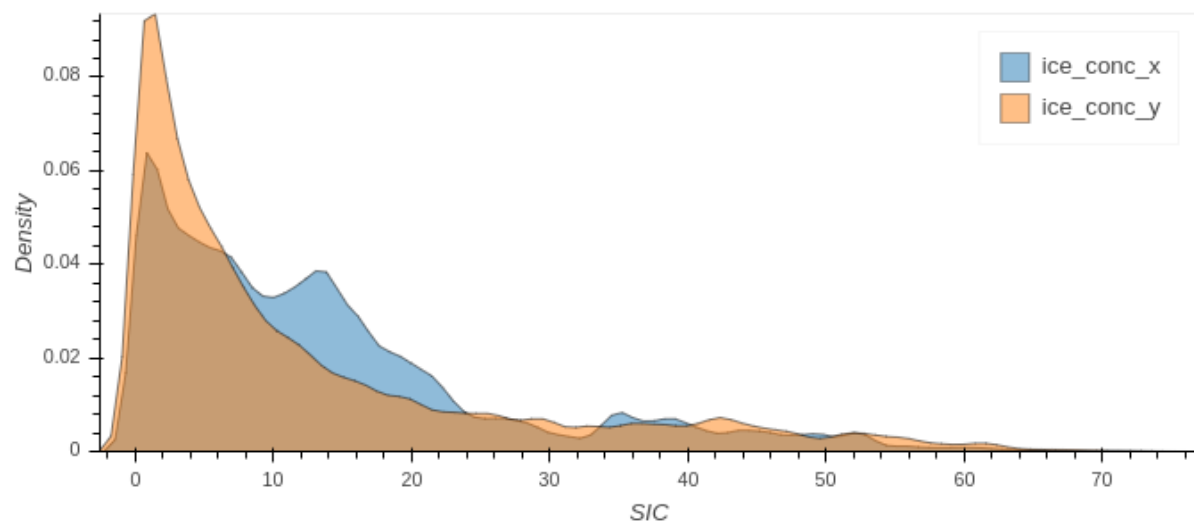


Figure 27. Overlap of the SICCI and OSISAF pdfs for the de-noised set that remains after the F6-chain filter application on the inconsistencies. Pink pdf is for SICCI and blue pdf is for OSISAF.

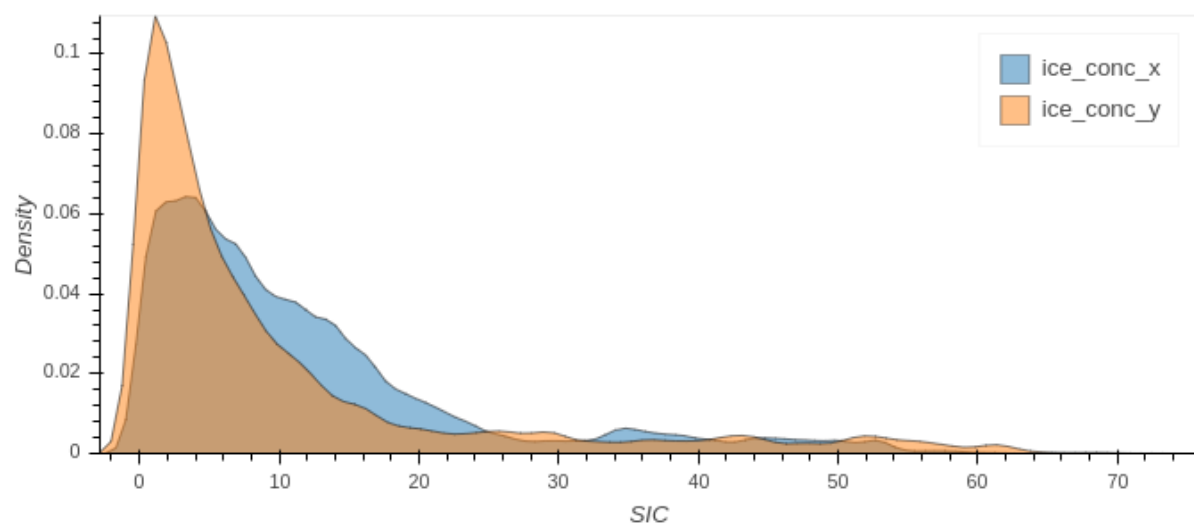


Figure 28. Overlap of the SICCI and OSISAF pdfs for the de-noised set that remains after a new F4 filter with more aggressive SSE thresholds is considered in the F6 chain. Pink SIC pdf is for SICCI and blue SIC pdf is for OSISAF.

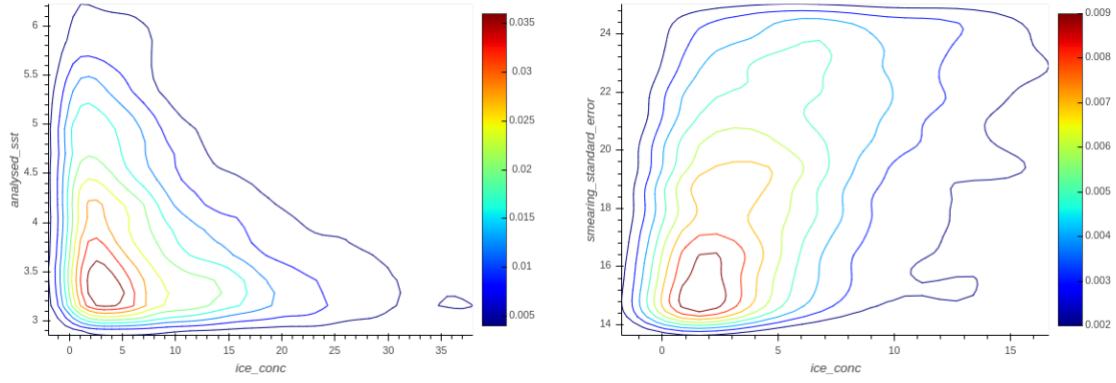


Figure 29. OSISAF joint pdfs for filtered SIC vs. SST (left) and SIC vs. SSE (right) for the F6 final filtering scheme after the intermediate F4 filter is updated for more aggressive SSE thresholds (i.e., $SSE_{\min} = 14\%$ and $SSE_{\max} = 25\%$).

6. Conclusions

In this investigation, we have explored the potential benefits of using an independent SST product to perform post-processing quality control on the EUMETSAT OSI-450 and SICCI 25-km SIC CDRs. As CDRs, the SIC products have already undergone a battery of corrections as part of their processing chain with ensuing estimates having low sensitivity to atmospheric noise including water vapor and to surface noise including wind roughening of the ocean surface, and variability of sea-ice emissivity and temperature. Estimates are never perfect, however, as there is always some unresolved uncertainty. The additional quality control is then to identify and correct for residual noise and spurious values in the retrieved SIC. The concept behind this idea of using an independent SST product to diagnose potentially wrong or incorrect SIC pixels, is that SIC and SST are predominantly dependent on different microwave frequencies and thus, have uncorrelated error sources. Hence, they can be used to mutually identify noisy or erroneous retrievals in the other, as long as the errors are introduced by one of the uncorrelated sources. The reference SST product of choice is the level 4 OSTIA 5 km-SST, which is retrieved independently from the EUMETSAT SIC CDRs. A common database of SIC and SST products have been created to facilitate these comparisons, in which the SIC products have been resampled to the same resolution of the SSTs.

As a first step, we extracted a subset of Arctic SIC data suspected of having residual noise or spurious values based solely on the fact that they corresponded to elevated SSTs. This set, termed the inconsistency set, was defined as: $\text{SIC} > 0\%$ with $\text{SST} \geq 3^\circ\text{C}$. Implicit in this selection criterion is the fact that the pixel should be partially covered with ice, so that the matching SST pixel has a surface temperature value associated with it. Hence, the inconsistency set is made entirely of mixed ice-water pixels, but in this application the two surfaces in the mixed pixel are treated independently. The size of the resulting inconsistency set was ~88 million and ~10.5 million for OSISAF and SICCI, respectively, for an entire year of matchups (in this case, 2014). Standard deviations for these two sets were 14% and 17%, respectively. These statistics are rather high, so this is supporting evidence that the inconsistency set has a large degree of spread and variability, and hence, it is dominated by unresolved noise. The overarching question of this investigation is, which retrievals from the inconsistency set are valid and which are noise or spurious values.

Maps of the geographical location of the inconsistencies reveal that SIC retrievals meeting the inconsistency criterion happen mostly at the beginning of the seasonal loss-of-ice period (June, July, and August) and are constrained to the ice edge, the coastal zones and the subpolar marginal seas, especially the Baltic Sea, the Canadian Arctic Archipelago, and the Sea of Okhotsk. In other words, the SIC inconsistencies are a phenomenon of the sea-ice boundary region, whether at the sea-ice edge in open waters or in coastal areas with retreating ice (first-year ice) during the melting season. This is a challenging set to work with because estimated uncertainties are the largest (as high as 40%) precisely at the ice edge in the summer time. The location and timing of the inconsistencies immediately points to atmospheric influences at open water and low ice concentrations and land-ocean spillovers as the most likely culprits of many of the spurious SIC inconsistencies. SIC retrievals are most sensitive to atmospheric noise at low ice concentrations and open water conditions and a significant portion of the SIC uncertainty is due to shortfalls in the atmospheric correction.

The method used for diagnosing valid and noisy SIC retrievals is simple. We use a split sample approach based on selected thresholds for variables identified as being related to the sources of the uncertainty. Below those thresholds, SIC inconsistencies are determined to be acceptable and are left untouched. Above the threshold, the SICs are determined to be spurious or noisy and are removed from the inconsistency set. Thresholds are selected based on scatter plots of SIC vs. SST, and from localized discontinuities in the marginal pdfs of the diagnostics variables showing correlation with the sources of the uncertainties. The impact of splitting the set into acceptable and noisy retrievals is evaluated using metrics based on the joint dependence of SIC and the SST as well as the overall geographical distribution of the retained values. The metrics include the combined standard uncertainty of SIC and SST, and an “energy conservation” type metric for mixed SIC-SST pixels within the MIZ in which an increase in pixel ice concentration should be balanced out by a decrease in SST and, vice versa, an increase in SST should have a corresponding pixel with a decreased ice concentration. Since the inconsistency sets are not normally distributed, the standard deviation is just a measure of the dispersion (spread and variability) of the data itself. As the acceptable set gets smaller and smaller with sequential thresholds moving more and more of the rejected pixels to the noisy set, the dispersion of the filtered data set should be reduced as part of the noise elimination process. As for the conservation metric, if the thresholds succeed at eliminating mixed pixels for which the SIC and SST retrievals are out of balance, the metric should approach zero. A successful metric should be sensitive to small changes that result from the thresholds used in the filtering process.

Our results indicate that false ice retrievals in the inconsistency set can be traced back directly to issues related to the atmospheric and land spillover corrections and to scaling effects on the retrieved smearing uncertainties at the pixel level resulting from resolution issues and interpolation artifacts when resampling over multiple scales. That the uncertainty of the inconsistency set can be explained in part by unresolved atmospheric and land contamination influences is not surprising since these two sources of uncertainty have the greatest impact over open waters and at low sea ice concentrations during the summer, which is precisely the temporal and spatial domain of the SIC inconsistencies. Additionally, the one aspect where the synergy between SIC and SST has the greatest impact in identifying disguised residual noise is by revealing flaws/shortcomings in the other product’s atmospheric correction algorithm. SSTs are retrieved at lower frequencies than SIC where the atmosphere is more transparent and sources of error such as LWC have less impact for the SSTs. A complicating factor is that unresolved errors in the inconsistency set may result from combinations of these three uncertainty sources as they share the same spatial and temporal domain. In other words, these errors may be correlated.

Resampling and interpolation artifacts should not be unexpected either as the products undergo multiple scaling transformations during the processing chain, from about 70 km and 30 km for level 2 OSISAF and SICCI to 25 km and 5 km gridded products. These artifacts affect the SIC CDRs through their uncertainty estimates, as these respond to the strength of SIC gradients near boundaries, especially at the sea ice-water edge where the interpolation artifacts also occur. Maps of gridded smearing uncertainties display strong gradients as they sharply decrease from their maximum value of ~40% inside the ice edge to 0% at the adjacent open waters. Interpolation effects interact with the smearing algorithm in two ways. When there are blocking artifacts, which result from resampling mixed pixels to resolutions finer than that of the actual observations, many of the subpixels are assigned replicated values corresponding to the coarse grid cell, smearing the zero-boundary into the open water regions. The misclassified pixels

outside the 0%-SIC boundary will then have minimum detected SICs ($SIC > 0\%$) with $SSE = 0\%$ (the open water value).

The smoothing nature of the 3x3 operator of the smearing uncertainty algorithm results in the underestimation of the uncertainties, especially for OSISAF. Since the OSISAF product at 25 km resolution has smeared edges resulting from gridding the level 2 SIC retrievals, it has less dynamic range at the edges compared to SICCI; thus, the low pass filter in the smearing uncertainty model returns smaller smearing uncertainties for the level 3 OSI-450 product than for the 25 km SICCI, which has sharp ice edge gradients. Therefore, the SSE estimates distributed with the level 3 OSISAF product are most likely underestimated in gradient regions. When the products are resampled from 25 to 5 km, the interpolation has a different impact on the provided SSE estimates, with the product having the strongest SSE intensity gradients being smeared across larger neighboring areas away from the ice edge than the product whose pixels have less SIC contrast (i.e., the spatial distribution of SSE_{max} is greater in the 5-km SICCI than 5-km OSISAF). These two interpolation effects manifest themselves in the SSE marginal pdf with both products showing spikes at both ends of the distribution. We reduce noise introduced by interpolation effects by using thresholds for SSE_{min} and SSE_{max} .

The smearing of the uncertainty estimates with spatial scale also appears to have repercussions when trying to remove noise left behind by the open water filter used in the CDR's atmospheric correction. As Lavergne et al. (2019) explain, "the role of the OWF is to detect and remove weather-induced false sea ice over open water while ideally preserving the true low-concentration values (typically at the ice edge)." The OWF is based on a threshold that is dynamically tuned to preserve true SIC values down to $SIC = 10\%$. The water-ice separation limit at 10% SIC is chosen to ensure that the sea ice extent, defined at $SIC = 15\%$, is not influenced by the OWF and only by the evolution of true SIC. Monthly mean time series of minimum detected SICs that are preserved by the OWF (Fig. 9 in Lavergne et al. (2019)) indicate that both products peak to 10% SIC in summer, well below the 15% threshold for the sea ice extent. The pdfs for the denoised inconsistency sets after all the thresholds under consideration were applied, however, revealed an additional peak between 10% and 20% SIC in the OSISAF pdf with a maximum at 15% SIC. This secondary peak in the OSISAF SIC distribution indicated the presence of undetected false ice concentrations beyond the 10% water-ice limit of the OWF implemented with current CDRs. It is safe to assume that the undetected false SICs in OSISAF are most likely the result of residual atmospheric contamination.

Simulations of the change in smearing uncertainty with grid resolution (Tonboe et al., 2016) indicated that the OSISAF smearing uncertainty changes from 6% at 50 km to 9% at 25 km to 15% at 5 km resolution grids. This is consistent with the SIC vs. SSE dependences observed with the 5 km-inconsistency set, with OSISAF showing increasing $SSE = f(SIC)$ up to 15%, whereas with SICCI, the SSEs were uniform after just 10% SIC. The OSISAF SSE vs. SIC increasing trend at low concentrations not only displayed a wider concentration range than SICCI, but also a departure from the (expected) linearity in the SIC vs SSE dependence observed with SICCI. This nonlinearity is another indicator of remaining dependences in the OSISAF denoised set. Since 1) we use SIC smearing uncertainties as thresholds to reduce noise along sea ice boundaries introduced by the downscaling of mixed pixels, and 2) our initial smearing thresholds failed to detect the residual atmospheric dependencies in OSISAF evidenced by the SIC pdf and joint pdf of SIC and SSE for the denoised set, then the implication is that the SSE_{min}

and SSE_{max} thresholds used in the preliminary noise elimination filters did not match the distribution of the signal and noise components for OSISAF at the 5-km resolution grid. This misrepresentation of the SSE thresholds most likely resulted from neglecting to take into account the underestimation of the OSISAF uncertainties at finer spatial scales. Similarly, the underestimation of the OSISAF uncertainties at fine spatial scales results in nonlinear structures in the SIC vs. SSE dependence at low concentrations that expand the extent of the ice edge (i.e., beyond 10% SIC ice-water threshold). The OSISAF product, therefore, requires larger uncertainty bounds to eliminate downscaling edge effects at 5 km grids. Since these effects were only observed in OSISAF, thresholds based on SIC uncertainty (e.g., SSE_{min} and SSE_{max}) should be more aggressive in OSISAF resulting in a much narrower uncertainty range for the true smearing uncertainties than in SICCI. After implementing new SSE thresholds, the non-linear structures in the SIC-SSE dependence were no longer discernible in the joint pdf, giving credence to our suspicions that the OSISAF uncertainties are indeed underestimated.

The proximity to land influenced the inconsistency set in different ways: as false SIC retrievals resulting from land-to-ocean spillovers; i.e., open water pixels mistakenly classified as ice, and as overestimated SIC retrievals as when the ice started to retreat from the coasts; i.e., a pixel with intermediate ice concentrations and mild SSTs appears to be compacted ice. The former effect was dominant in OSISAF while the latter effect only impacted SICCI. The SIC CDRs undergo extensive corrections for land contamination. Since spurious SICs due to land spillovers are more numerous in OSISAF, perhaps they are the result of geolocation errors or larger sidelobes in SSM/I. Although geolocation errors can be significant along the ice edge and coastlines, their contribution is not included in the sea ice concentration uncertainty estimate (Tonboe et al., 2016). Thresholds for a minimum distance-from-land requirement and a maximum SIC within the MIZ were implemented to eliminate SIC retrievals affected by land contamination.

False SIC retrievals influenced by the atmosphere manifested themselves in two different ways in the inconsistency set. In the first case, we found significant numbers of summer, open-water pixels misclassified as ice in subpolar arctic seas (e.g., The Baltic Sea and Sea of Okhotsk) showing corresponding SSTs in excess of 20 °C. We speculate that these spurious retrievals are directly linked to the presence of dense clouds with high LWC responsible for thunderstorms and heavy rains to these basins during summer months. Even though the microwave atmospheric correction is not representative of atmospheric conditions corresponding to cloudy atmospheres, a proxy correction can be obtained from the synergism between the SIC and SST by simply imposing a maximum SST criterion for SIC retrievals in the MIZ. This type of error impacted both products in equal measure. Since this type of error affected pixels in open waters far from the coast where there were no strong gradients/fronts, there was nothing in the associated uncertainties to indicate that these pixels were affected by uncorrected atmospheric effects, unlike the pixels near the ice edge and along coastal margins described above in connection with the OWF.

The final threshold that tied all these multiple criteria that had to be met was based on the SIC vs. SST dependence. Scatter plots of SIC inconsistencies vs. matching SSTs reveal a well-defined pattern of behavior in which increases in SIC are accompanied by decreases in SST and, vice versa, large temperature increases are only possible for areas with low ice concentration. A mixed coarse resolution pixel corresponding to the microwave BTs can have only a finite range

of allowable SIC and SST values. By devising a parametric function that describes the joint variations of SIC and SST within the MIZ, we seek to preserve the features of the SIC-SST dependence, while eliminating false SIC retrievals; i.e., we avoid allowing densely sampled regions within the ice edge to dominate the results, while simultaneously allowing for the full range of variability in the SDs. This aspect is important as the inconsistency set is not sampled randomly, but is more densely populated around the confines of the ice edge. Note that the precise functional dependence will also be a function of the product resolution. The “energy conservation” metric that looked at the continuity of the joint variations of SIC and SST proved to be remarkably sensitive to changes introduced by the thresholds. Not surprisingly, the impact of the noise reducing filters in descending order of “impactfulness” according to the conservation metric was for thresholds based on $SST = f(SIC)$, distance-from-land and SSE for SICCI, while the order for OSISAF was distance-from-land, $SST = f(SIC)$, and SSE. While most of the inconsistency set ended up being rejected by the filters, the data that remained in the valid set exhibited patterns in the distributions consistent with the theoretical expectations. The alternative metric considering the combined SD for SIC and SST was not so successful, but it gives confirmation that using the synergy of the retrievals, can improve the accuracy of both products. It also confirmed that the variability in the ice edge is dominated by SIC and not the SST. The SDs for the valid, denoised sets were 12% and 15% for OSISAF and SICCI, respectively.

In summary, exploiting the synergy between SIC and SST proved to be a good diagnostic tool to identify and reduce uncertainty due to unaccounted atmospheric influences, land contamination and interpolation artifacts. Not only does SST provide a simple solution to the elusive LWC correction, but it does so without introducing additional trends of its own. Quite the opposite, the SSTs also benefit from this synergy as they also get denoised in the split processes with the SST set corresponding to the valid SIC retrievals having $SD \approx 0^\circ\text{C}$. It is by happenstance that the evaluation of the SIC inconsistencies devolved into a validation of the SIC uncertainties. Our studies show that the estimated uncertainties in the SIC CDR products are designed with the ice edge in mind, but they underestimate changes in the uncertainties themselves as a result of scaling and spatial resolution issues associated with delivering a high-resolution product from a coarse resolution sensor. Resolution issues appear to have a significant impact in the OSI-450 uncertainty estimates delivered with the gridded products. Additionally, because the smearing uncertainty algorithm is designed to respond to the presence of strong gradients, if the source of error introducing uncertainty happens in open waters (e.g., errors associated with retrievals under wet, dense clouds, precipitation or strong winds, which can happen anywhere in the ocean), the smearing uncertainty algorithm will not be able to cover these atmospheric contributions. This is the case with false SIC retrievals in the Baltic Sea during the summer, detected via SST synergy, when LWC can have a significant impact on the 37 GHz. There was nothing in the associated uncertainties to indicate that these were questionable SIC retrievals. Consequently, the uncertainty model has functional form misspecification, as it considers all the explanatory sources of error but fails to account for all the relationships between them and potential outcomes.

The noise reduction filters tested here can be easily implemented in a binary quality control mask where different bits are set based on the thresholds, such that educated users can decide whether to use flagged inconsistencies. It would be best, however, to try to use the SIC-SST synergy in

the atmospheric correction directly at the brightness temperature level. The SICCI-50 km already uses the 6 GHz channel, and there is an RTM in place as part of the atmospheric correction, so incorporation of the SST as part of the RTM atmospheric correction should be straightforward. Besides, an improved atmospheric correction that takes into account regional atmospheric effects for selected geographical regions would be highly beneficial for improving the overall accuracy of the SIC retrievals. A post-processing pan arctic filter based on a threshold for SST_{max} , as the one proposed here, eliminates enormous amounts of data just to guarantee that there are no anomalous SIC in the Baltic Sea and Sea of Okhotsk during the summer. It is plausible that the SST_{max} threshold is masking out valid subpolar SIC retrievals with slightly higher SSTs that are not impacted by the atmosphere in the same way it affects these two specific regions. The pdf for OSISAF's associated SSTs (Fig. 13), for instance, suggests valid SSTs up to $\sim 12 - 15$ °C. Truncating valid SSTs above $SST_{max} = 10$ °C outside the Baltic Sea and Sea of Okhotsk might negatively impact estimates of ice retreat based on the maximum SST for the summer (e.g., Steele and Dickinson, 2016) and regional weather forecasts from NWP models that assimilate both SIC and SST.

A final recommendation is that future matchup data sets, created with the purpose of noise identification, avoid converting resolutions to such fine scales. The 25-km SICCI and OSI-450 products can be directly compared to the OSTIA level 4 SST $0.25^\circ \times 0.25^\circ$ resolution product without incurring in additional regridding and thus minimizing smearing of spurious retrievals at the sea ice – water boundaries (atmospheric influences) and coastal regions (atmospheric and land contamination). If additional downscaling is required, edge detection interpolation techniques that cause little blurring of edges should be used instead.

Acknowledgments

This work (Sea surface temperature metrics for evaluating OSI SAF sea ice concentration products (OSI_AS19_04)) was completed as part of the OSI-SAF Visiting Scientist Program to the Danish Meteorological Institute in June 2019. I would like to thank Dr. Rasmus Tonboe and Dr. Jacob Hoeyer for extending me the invitation to participate in this project and for sharing their knowledge. Special thanks to Pia Nielsen-Englyst for the many discussions and suggestions about the work. To the staff at OSI-SAF, specially Cecile Hernandez, for her help with all the logistics establishing the fellowship, and to the Arctic Group at DMI for welcoming me and for their kindness during my visit. It was an amazing experience and I am forever grateful.

References

- Andersen, S., Tonboe, R., Kaleschke, L., Heygster, G., and Pedersen, L. T.: Intercomparison of passive microwave sea ice concentration retrievals over the high-concentration Arctic sea ice, *J. Geophys. Res.*, **112**, C08004, doi:10.1029/2006JC003543, 2007.
- Bradley, B. O., and Taqqu, M.S.: Risk and heavy tails, in *Handbook of Heavy Tailed Distributions in Finance*, Ed. University of California, 35–103, doi:10.1016/B978-0-444-50896-6.X5000-6, 704 pp, 2003.
- Cavalieri, D. J., Parkinson, C. L., Gloersen, P., Comiso, J. C., and Zwally, H. J.: Deriving long-term time series of sea ice cover from satellite passive-microwave multi-sensor data sets, *J. Geo-phys. Res.*, **104**, 15803–15814, 1999.
- Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, *Remote Sens. Environ.*, **116**, 140–158, doi:10.1016/j.rse.2010.10.017, 2012.
- Fernandez, P., Kelly, G., and Saunders, R.: Use of SSM/I ice concentration data in the ECMWF SST analysis, *Meteorol. Appl.*, **5**, 287–296, 1998.
- Ivanova, N., Johannessen, O.M., Pedersen, L.T., and Tonboe, R.T.: Retrieval of Arctic sea ice parameters by satellite passive microwave sensors: a comparison of eleven sea ice algorithms, *IEEE T. Geosci. Remote*, **52**, 7233–7246, 2014.
- Ivanova, N., Pedersen, L.T., Tonboe, R.T., Kern, S., Heygster, G., Lavergne, T., Sørensen, A., Saldo, R., Dybkjær, G., Brucker, L., and Shokr, M.: Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations, *The Cryosphere*, **9**, 1797–1817, doi:10.5194/tc-9-1797- 2015, 2015.
- Lavergne, T., Sørensen, A.M., Kern, S., Tonboe, R.T., Notz, D., and coauthors.: Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *The Cryosphere*, **13**, 49–78, doi: 10.5194/tc-13-49-2019, 2019.
- Lu, J., Heygster, G., and Spreen, G.: Atmospheric Correction of Sea Ice Concentration Retrieval for 89 GHz AMSR-E Observations, *IEEE J-STARS.*, **11**, 1442–1457, <https://doi.org/10.1109/JSTARS.2018.2805193>, 2018.
- Nomura, A.: Global sea ice concentration data set for use with the ECMWF re-analysis system. *ECMWF Technical Report No. 76*. March 1995 (available from the ECMWF librarian).
- Peng, G., Steele, M., Bliss, A.C., Meier, W.N., and Dickinson, S.: Temporal means and variability of Arctic sea ice melt and freeze season climate indicators using a satellite climate data record. *Remote Sens.*, **10**, 1328; doi:10.3390/rs10091328, 2018.
- Steele, M., and Dickinson, S.: The phenology of Arctic Ocean surface warming, *J. Geophys. Res. Oceans*, **121**, 684 –6861, doi:10.1002/2016JC012089, 2016.
- Tonboe, R. T., Eastwood, S., Lavergne, T., Rathmann, N., and coauthors.: The EUMETSAT sea ice concentration climate data record. *The Cryosphere*, **10**, 2275–2290, doi: 10.5194/tc-10-2275-2016, 2016.