Satellite Passive Microwave Sea-Ice Concentration Data Set Inter-comparison for Arctic Summer Conditions

Stefan Kern[1], Thomas Lavergne[2], Dirk Notz[3], Leif Toudal Pedersen[4], and Rasmus Tonboe[5]

[1]Integrated Climate Data Center (ICDC), Center for Earth System Research and Sustainability (CEN), University of Hamburg, Hamburg, Germany
[2]Research and Development Department, Norwegian Meteorological Institute, Oslo, Norway
[3]Institute for Marine Research, University of Hamburg and Max-Planck Institute for Meteorology, Hamburg, Germany
[4]Danish Technical University, Lyngby, Denmark
[5]Danish Meteorological Institute, Copenhagen, Denmark

Correspondence to: Stefan Kern (stefan.kern@uni-hamburg.de)

Abstract. We report on results of a systematic inter-comparison of 10 global sea-ice concentration (SIC) data products at 12.5 to 50.0 km grid resolution from satellite passive microwave (PMW) observations for the Arctic during summer. The products are compared against SIC and net ice-surface fraction (ISF) – SIC minus the per-grid cell melt-pond fraction (MPF) on sea ice – as derived from MODerate resolution Imaging Spectroradiometer (MODIS) satellite observations and observed from ice-going vessels. Like in Kern et al. (2019), we group the 10 products based on the concept of the SIC retrieval used. Group I consists of products of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI SAF) and European Space Agency (ESA) Climate Change Initiative (CCI) algorithms. Group II consists of products derived with the Comiso bootstrap algorithm and the National Oceanographic and Atmospheric Administration (NOAA) National Snow and Ice Data Center (NSIDC) SIC climate data record (CDR). Group III consists of Arctic Radiation and Turbulence Interaction Study (ARTIST) Sea Ice (ASI) and National Aeronautics and Space Administration (NASA) Team (NT) algorithm products and group IV consists of products of the enhanced NASA Team algorithm (NT2). We find wide-spread positive and negative differences between PMW and MODIS SIC with magnitudes frequently reaching up to 20-25 % for groups I and III and up to 30-35 % for groups II and IV. On a pan-Arctic scale these differences may cancel out: Arctic average SIC from group I products agrees with MODIS within 2-5 % accuracy during the entire melt period from May through September. Group II and IV products *over-estimate* MODIS Arctic average SIC by 5-10 %. Out of group III, ASI is similar to group I products while NT SIC *under-estimates* MODIS Arctic average SIC by 5-10 %. These differences, when translated into the impact computing Arctic sea-ice area (SIA), match well with the differences in SIA between the four groups reported for the summer months by Kern et al. (2019). MODIS ISF is systematically over-estimated by all products; NT provides the smallest over-estimations (up to 25 %), group II and IV products the largest over-estimations (up to 45 %). The spatial distribution of the observed over-estimation of MODIS ISF agrees reasonably well with the spatial distribution of the MODIS MPF and we find a robust linear relationship between PMW SIC and MODIS ISF for group I and III products during peak melt, i.e. July and August. We discuss different cases taking into account the expected influence of ice-surface properties other than melt ponds, i.e. wet snow and coarse grained snow / refrozen surface, on brightness temperatures and their ratios used as input to the SIC retrieval algorithms. Based on this discussion we identify the mismatch between the actually observed surface properties and those represented by the ice tie points as the most likely reason for i) the observed differences between PMW SIC and MODIS ISF and for ii) the often surprisingly small difference between PMW and MODIS SIC in areas of high melt-pond fraction. We conclude that all 10 SIC products are highly inaccurate during summer melt. We hypothesize that the unknown amount of melt-pond signatures likely included in the ice tie points plays an important role – particularly for groups I and II – and recommend to conduct further research in this field.

1 Introduction

A considerable number of different algorithms to compute the sea-ice concentration from satellite passive microwave (PMW) brightness temperature (TB) measurements has been developed during the past decades. All exploit the fact that under typical viewing angles (50-55 degrees) the difference in microwave TB between open water and sea ice is sufficiently large to estimate sea-ice concentration.

In the Polar Regions, freezing conditions prevail during winter. During summer, melting conditions prevail or at least coexist with freezing conditions. The changes in snow and sea-ice properties in response to the melting conditions complicate the retrieval of the sea-ice concentration from microwave TB measurements. This applies in particular to the Arctic. First signs of melt are an increase in snow wetness and melt-refreeze cycles, triggered by diurnal warming and nocturnal cooling, leading to an increase in snow grain size and snow density. Wet snow is a good absorber of microwave radiation and has an emissivity close to 1. Therefore, microwave TBs measured over wet snow are often very close to the physical temperature of the melting snow, i.e. 0 °C. As a consequence, wet snow masks the radiometric difference between different ice types, e.g. first-year and multiyear ice. Above a certain wetness and snow thickness (a few centimetres), the influence of wet snow on microwave TBs in the frequency range used here (see Table 1) can be regarded as being independent of frequency and polarization. The influence of coarse-grained snow is more complex. During diurnal melting, it behaves like wet snow. During nocturnal cooling, the liquid water refreezes and absorbs considerably less microwave radiation. This allows for volume scattering from within the snow, which – in contrast to the absorption of microwave radiation by wet (coarse-grained) snow is both frequency and

polarization dependent. More details about the influence of these parameters on microwave TBs relevant for retrieval of sea-ice concentration during summer are given, e.g., in Kern et al. (2016).

60   Continued melting results in increasing snow wetness until it becomes saturated with melt water – at which stage melt ponds start to form. The fraction of the ice surface covered by melt ponds formed from melting snow and sea ice typically varies between 10 and 40 % but it can exceed 50 %, e.g., early in the melt season and on particularly level sea ice such as land-fast sea ice (e.g. Webster et al., 2015; Divine et al., 2015; Landy et al., 2014). The fraction of liquid water due to melt ponds on the sea ice poses a particular challenge for the sea-ice concentration retrieval using microwave TB measurements because

65 the penetration depth of microwaves in water at the frequencies listed in Table 1 is of the order of 1 mm (Ulaby et al., 1986). Thus, a water layer with a depth of only a few millimetres is sufficiently opaque to block the thermal microwave emission of the sea ice underneath completely. In addition, the emissivity of fresh water in the ponds and the emissivity of saline water in the leads are the same at most of the microwave frequencies above 10 GHz that we use here. Therefore, during summer liquid water in form of melt ponds on the sea ice is indistinguishable from liquid water in the cracks and leads between the ice floes

70 in the microwave frequency range used here (e.g. Gogineni et al., 1992; Grenfell and Lohanick, 1985). This has direct consequences for the sea-ice concentration retrieval using satellite TB measurements.

       Several studies have revealed various degrees of under-estimation of the sea-ice concentration during summer conditions in the Arctic (e.g. Ivanova et al., 2015; Rösel et al., 2012b; Markus and Dokken, 2002; Comiso and Kwok, 1996; Steffen and Schweiger, 1991; Cavalieri et al., 1990). A natural explanation of this observed under-estimation would therefore

75 be that those satellite products are rather a good measure of the "net ice surface fraction", that is 1 minus the area fraction of all surface water in the satellite field-of-view. We illustrate the typical summer sea-ice concentration retrieval by a simple example. Consider two grid cells A and B observed during the summer melt season. Grid cell A has 100 % sea-ice cover with 40 % melt-pond fraction. Grid cell B has 75 % sea-ice cover with 15 % melt ponds. Based on physical principles a sea-ice concentration retrieval algorithm should provide a value of 60 % in both cases, i.e. the so-called "net ice surface fraction".

80 There is evidence from literature (e.g., Comiso and Kwok, 1996; Kern et al., 2016) that this is however not the case. It is rather very likely that an algorithm would provide a value of, for instance, 85% for both grid cells A and B, because melt-induced changes in the surface emissivity of the visible part of the sea ice are often insufficiently taken into account, yielding to an over-estimation of the actual net ice surface fraction. Providing a value of 85 % this algorithm would under-estimate the actual sea-ice concentration in grid cell A by 15 % while it would over-estimate it by 10 % in grid cell B. If we interpret the provided

85 value as a net ice surface fraction, it is an over-estimation by 25 % in both cases. In other words, for this quite typical example the retrieved value is highly inaccurate and biased compared to either the actual sea-ice concentration or the net ice surface fraction. The magnitude of this bias is largely unknown and it appears not to be reflected by an appropriate increase in retrieval uncertainty estimates which – when at all provided – are a measure of the precision, i.e. the interval within which the reported sea-ice concentration estimate typically varies, and not of the bias. In Fig. 1 we show the seasonal cycle of the sea-ice

90 concentration algorithm standard error – aka the precision – of the OSI-450, SICCI-25km and SICCI-50km products (Lavergne et al., 2019) for illustration; OSI-450 and SICCI are product names for SIC CDRs derived from a collaboration of European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI SAF) and European Space Agency (ESA) Climate Change Initiative (CCI) programs (see Lavergne et al., 2019). To summarize: we do not know what the sea-ice concentration algorithms actually measure during summer (actual sea-ice

95 concentration or net sea-ice surface fraction), and whichever they measure the accuracy is poor compared to the winter conditions.

       The unknown accuracy makes it difficult if not impossible to use summer satellite PMW sea-ice concentration maps and sea-ice area (SIA) for the evaluation of numerical models (e.g., Notz, 2014; Burgard et al., 2019), or to assimilate such data into numerical models for a quantitative improvement of, e.g., sea-ice forecast for shipping (e.g., Melia et al., 2017). As

100 a consequence, studies about the long-term development of the Arctic sea-ice cover prefer to use sea-ice extent (SIE) over SIA. The SIE is computed as the sum of all grid cells with more than 15% sea-ice concentration: $SIE = \sum A_{SIC>15\%}$, with grid-cell area A. The SIA includes the actual SIC as weight $SIA = \sum(A \times \frac{SIC}{100\%})$; a SIC threshold of 15% is not applied regularly. Consequently, biases in the sea-ice concentration certainly have a small influence on the summer SIE while the impact on SIA can be quite large (see for example Kern et al. 2019 Fig. 6 and Figs. G2 and G3). However, it has been found that the SIE and

105 its trend provide a limited metric for the performance of numerical models (e.g., Notz, 2014), and that prediction of the minima in September Arctic SIA and SIE would benefit from giving more weight to SIA (e.g. Petty et al., 2018). One way to overcome the SIA biases in summer could be to focus on its trends (as opposed to its absolute value) (e.g. Comiso et al., 2017; Ivanova et al., 2014) but this cannot be the solution since there is no guarantee that these biases are stable along the whole time series.

       With this study, we aim to give more information about the accuracy of current satellite PMW sea-ice concentration

110 products during summer. We present a systematic inter-comparison of 10 such products (see Sect. 2, and Kern et al., 2019) with independent estimates of the summertime Arctic sea-ice concentration, net ice surface fraction, and melt-pond fraction derived from observations of the MODerate resolution Imaging Spectroradiometer (MODIS) aboard the Earth Observation Satellite (EOS) TERRA (Rösel et al., 2011, 2012a). We show the pan-Arctic sea-ice concentration biases with respect to MODIS sea-ice concentration and ice surface fraction for the 10 products for the period 2003 through 2011, illustrate the

115 spatiotemporal variability of these biases, and quantify the biases as a function of melt season progress and melt-pond fraction. We describe the data and inter-comparison methods used in Sect. 2. In Sect. 3 we give an overview about the pan-Arctic results of our inter-comparison. Section 4 focuses on more detailed comparisons to MODIS sea-ice concentration and ice surface fraction, and illustrates the potential of a bias correction as well as the impact on the computation of the sea-ice area. Our paper closes with a discussion and concluding remarks in Sect. 5.

2 Data & Methodologies

2.1 Sea-ice concentration data sets

Like in Kern et al. (2019), we consider 10 different sea-ice concentration products which we very briefly summarize in Table 1. More information about these products is given in Kern et al. (2019, Appendix 7.1 -7.6). There are many more algorithms and products available than we are using here, see e.g. Ivanova et al. (2015). The main criteria for our choice of algorithms and products are 1) length of the product time series, 2) grid resolution, 3) accessibility and sustained extension, and 4) overlap with the melt-pond fraction evaluation data set. Due to these criteria we have not selected products with < 10 years coverage or with a grid resolution < 12.5 km.

The algorithms used to generate the 10 products can be distinguished by their general approach to derive the SIC. We took advantage of this fact and assigned products to four different groups: I to IV (see Table 1). Algorithms of group I: OSI-450 and the three SICCI algorithm versions, employ a self-optimizing hybrid approach combining an algorithm with superior accuracy over open water with an algorithm with superior accuracy over consolidated sea ice (see e.g. Ivanova et al., 2015, Lavergne et al., 2019). In addition, all products of group I utilize an optimized tie-point retrieval scheme mitigating inter-sensor inconsistencies. Algorithms of group II employ an advanced bootstrap technique to compute the SIC in a 2-dimensional space of either dual-polarization 37 GHz TBs or vertically polarized TBs at 19 and 37 GHz (see Comiso and Nishio, 2008); the group hence includes data derived with the Comiso Bootstrap algorithm from two different PMW TB data sets. Group II includes the National Oceanographic and Atmospheric Administration (NOAA) National Snow and Ice Data Center (NSIDC) SIC CDR because it is largely determined by the input Comiso Bootstrap algorithm. Group III includes those algorithms of our set of 10 that derive the SIC mainly from the TB polarization difference at 19 GHz (NT) or at near-90 GHz (ASI). Also, these two products use constant tie points. Even though the results of this paper suggest, that ASI appears to fit better into group I, as will be briefly discussed later, we keep that product in group III to avoid confusion with Kern et al. (2019). Finally, group IV includes one product only, the enhanced NASA-Team (NT2) algorithm. While the NT2 advances the NT algorithm, e.g. by adding near-90 GHz TBs, the SIC retrieval itself differs fundamentally from that of all other nine algorithms. We refer to Kern et al. (2019) and the references therein for further reading.

In the following few paragraphs we provide some more general remarks to the satellite PMW data products used. We refer to Lavergne et al. (2019) and Kern et al. (2019) for further information.

The difference in microwave TBs observed over open water (low) and land (high) combined with the diameter of the field-of-view of several kilometres to a few tens of kilometres can cause spurious sea-ice concentrations to appear along coasts (e.g. Lavergne et al., 2019). In this paper, we do neither further correct potential differences between the 10 products caused by this effect nor do we pay particular attention to this effect.

Atmospheric moisture and wind-induced roughening of the ocean surface can cause spurious sea-ice concentrations in areas that are actually ice free. To mitigate this noise, different kind of weather filters are applied in the 10 products used. These and their effects on SIA and SIE estimated from the sea-ice concentration data are discussed in Kern et al. (2019). The focus of this paper is on the performance of the 10 products during summer conditions over consolidated ice, where the weather filters have no effects. Therefore, we do not further discuss weather filters in this paper.

In near-100 % and near-0 % sea-ice concentration conditions, most retrieval algorithms will naturally retrieve a bell-shaped distribution of sea-ice concentration values, returning values both below and above 100 % or 0 % sea-ice concentration (e.g. Ivanova et al., 2015). While the EUMETSAT-OSISAF – ESA-CCI products (group I, see Table 1) allow using the naturally retrieved sea-ice concentration on either side of 100 % the others do not. In those other products any sea-ice concentration values retrieved as being larger than 100 % are set to 100 % and lost to the user. The availability of these "off-range" estimates in the four group I products was used in Kern et al. (2019) to demonstrate how the "off-range" distribution can effectively be reconstructed a-posteriori for most of the other products from their truncated sea-ice concentration distributions. Kern et al. (2019) illustrated that products with over-estimated sea-ice concentration (modal value of the non-truncated distribution larger than 100 %) would obtain better validation statistics (smaller bias and RMSE) than products with no over-estimation (modal value of the non-truncated distribution exactly at 100 %). The larger the over-estimation, the better the statistics would be. We briefly discuss this issue and its relevance for our comparison with the MODIS data set in Sect. S3.1 in the supplementary material.

We will mainly focus our discussion of the results obtained (Sect. 3 and 4) on products that we selected to be representative of the four groups in Kern et al. (2019). These products are: OSI-450 for group I, CBT-SSMI for group II, NT1-SSMI for group III and NT2-AMSRE for group IV (see Table 1 for the product acronyms used). We refer to the supplementary material where, starting with Figure S03, we show some of the results for all 10 products.

2.2 The MODIS data set

We use the MODerate resolution Imaging Spectroradiometer (MODIS) Arctic melt-pond fraction data set developed by Rösel et al. (2011; 2012a): Rösel et al. (2015), https://doi.org/10.1594/WDCC/MODIS__Arctic__MPF_V02, last accessed May 19, 2020. This data set is provided for the Arctic Ocean north of 60°N with 8-daily temporal resolution on the NSIDC polar-stereographic grid with 12.5 km x 12.5 km grid resolution at 70 degrees north. It extends from day-of-year (DOY) 129, i.e. May 9 (May 8 for leap years), to DOY 256, i.e. September 13 (September 12 for leap years) and hence covers pre-melt, melt advance, peak melt and end-of-melt conditions.

The melt-pond fraction retrieval is based on the calibrated and atmospherically corrected reflectance values measured by MODIS channels 1, 3 and 4 available in the MOD9A1 8-day product. For this product, reflectance values measured during

eight consecutive days were re-projected from the original MODIS tiles into the NSIDC polar stereographic grid with 500 m x 500 m grid resolution, composited over the 8-day period, and combined with the cloud- and land masks provided with the MOD09 product. Composited means that a cloud-free section in a MODIS tile of a more recent satellite overpass is preferred over an older satellite overpass within the 8-day period. For the retrieval, it is assumed that each 500 m grid cell is solely covered by fractions of three surface types: open water in leads and openings between ice floes, melt ponds, and sea ice and snow. The sum of these fractions is assumed to equal 1. Via a spectral un-mixing approach and an artificial neural network the measured reflectance values are converted into the fractions of these three surface types per grid cell, followed by the interpolation onto the 12.5 km grid used for the final product.

The product contains the melt-pond fraction (MPF), the open water fraction (OWF), the standard deviation of the MPF values at 500 m grid resolution, and the number of valid 500 m MPF estimates. This latter number is a measure of the number of clear-sky 500 m grid cells. In addition the product contains so-called "clear-sky" versions of the 12.5 km gridded MPF and OWF data computed only for those 12.5 km grid cells where more than 90 % of the input 500 m grid cells are denoted clear-sky. We note that the MPF is a measure of the melt-pond fraction per grid cell. No MPF values are provided for 12.5 km grid cells with an OWF larger than 85 %.

Rösel et al. (2012a) reported root-mean-squared errors (RMSE) between ~4 % and 11 % compared to air-borne data. Kern et al. (2016) compared daily estimates of the MPF for June to August 2009 with ship-based observations of the MPF and reported RMSE values between ~6 % and ~15 %. Istomina et al. (2015) and Marks (2015) confirmed the validity of the MODIS MPF data set with different independent observations of melt ponds. Zhang et al. (2018) evaluated the MODIS MDF data set with independent high-resolution satellite observations. Experience working with this data led us to conclude that the MPF estimates are accurate to within a few percent. MODIS SIC and MODIS ISF share the same accuracy as MPF and are taken to be as accurate as 5 % in our study (see also Kern et al., 2016).

Besides unaccounted cloud influence there is another limitation that needs to be kept in mind when using this data set. The used approach is based on three channels, which limits the maximum number of surface types to be discriminated to three. Ponds on first-year ice, however, have different spectral characteristics than ponds on multiyear ice: while the latter appear and remain bluish and relatively bright, the former become darker with advancing melt season until they eventually melt through the ice. As a consequence, towards the end of the melt season melt ponds on first-year ice might be assigned to the class open water. During the same time of the melting season, melt ponds might be covered by a slush or thin ice layer. Depending on the properties of this layer, the melt pond is either still assigned to the class melt pond, or it is assigned to the class ice. In addition, new ice forming between the ice floes in the high Arctic towards the end of the melt season could be classified as melt ponds (see also Rösel et al., 2012a). Because of these ambiguities in the retrieval of the melt-pond fraction it is likely that the accuracy of the parameters derived is poorer towards the end of the melt season, i.e. September.

We note that – due to the 8-day compositing (see above) – this product is less well suited to define melt-onset or the length of the melt period with daily temporal resolution. Later-on we will work with four distinct phases of the summer melt period. Our definition of theses phases is not driven by exact dates but by changes in the overall pan-Arctic melt-pond fraction evolution (see below). Therefore we are confident that eventual biases that might occur due to the 8-day compositing, e.g. a melt-pond fraction map of an 8-day period is not representative of the entire 8-day period but of the last 1-2 days of it, does not influence our results – except potentially increasing the noise. The same applies to issues such as melt-refreeze cycles.

For this paper, we used the clear-sky versions of MPF and OWF. In addition, we exclude all those MODIS data set grid cells where the ratio between the 12.5 km gridded MPF value and the standard deviation of the 500 m MPF values is smaller than 1. While this step filters out grid cells with an actual true large MPF variability, at the same time it reduces the influence of cloud cover artefacts. A similar filtering effect could have been achieved by increasing the percentage of 500 m grid cells required to consider a 12.5 km grid cell value clear-sky from 90 % to, for instance, 95 %. However, in that case the number of valid MODIS product data would have reduced drastically.

We are interested in the fraction of ice detectable with PMW sensors. We call this the net ice surface fraction (ISF). ISF is related to OWF and MPF as follows: ISF + MPF + OWF = 1. We thus derive two parameters from the MODIS data set: the MODIS sea-ice concentration, MODIS SIC, which is 1 – OWF, and the MODIS ice surface fraction, MODIS ISF, which is 1 – OWF – MPF.

Sea-ice concentrations of the 10 products (Sect. 2.1) are co-located with the MODIS parameters via finding the grid cell pairs with the minimum difference (in kilometres) between the grid-cell centres. For this step, we converted the latitude and longitude coordinates of both data sets, i.e. the PMW products and the MODIS products, into metric coordinates using the WGS84 ellipsoid allowing to compute the minimum distance via simple geometry. We do not interpolate any of the data sets. We do not perform any averaging in case that multiple (small) grid cells of one product fall into one (large) grid cell of the other product. All comparisons are carried out at the native grid resolution. When compared to the 25 km products, this results in a lower number of co-located grid cells for SICCI-50km and a higher number for SICCI-12km and ASI-SSMI. Finally, the collocated PMW SIC data are averaged in time over the same eight days used in the respective 8-daily MODIS product, i.e. for the MODIS product of DOY=129 we average over data from DOY 129 through 136. If valid sea-ice concentrations of fewer than three days within this 8-day period are available, this grid cell is discarded from further analysis.

In Fig. 2 we illustrate the melt-pond development in the Arctic Ocean as relevant for this paper. The maps show melt-pond distributions of a DOY representative of the four periods considered in this paper: pre-melt, melt advance, peak melt and end-of-melt in the maps of Fig. 2a-d, respectively. For these maps, we selected years where the data coverage is particularly good, i.e. with only few grid cells discarded as potentially cloud contaminated. Below each map we show histograms of the melt-pond fraction of the respective DOY of years 2003 to 2011, to illustrate the inter-annual variability.

3   Pan-Arctic summer-time sea-ice conditions

We begin our inter-comparison with an illustration of the sea-ice conditions in the Arctic during summer as seen by the satellite products. For this step, we compute an ensemble multi-annual (2002-2011) median of the monthly mean sea-ice concentration from the 10 sea-ice concentration products and subtract it from the respective sea-ice concentration of the individual product. This computation is carried out at 50 km grid resolution using a common land mask (see Kern et al., 2019). In Fig. 3 we show the differences between the products representative of groups I to IV (see end of Sect. 2.1) and the ensemble median as an example for the month of July (Fig. 3a-d) along with a map of the ensemble median sea-ice concentration (Fig. 4e). Figure 3 illustrates considerable differences between the four groups. Sea-ice concentration differences for July are particularly negative (mean sea-ice concentration smaller than ensemble median) for group III and particularly positive (mean sea-ice concentration larger than ensemble median) for group II. We refer to Fig. S04 in the supplementary material for difference maps of all 10 products.

Differences between the individual products' multi-annual pan-Arctic monthly mean sea-ice concentration and the ensemble median increase from winter (Table 2, top row, Jan/Feb) to summer (Table 2, bottom row, July/Aug). Group I products show less sea ice than the ensemble median; group II and group IV products show more sea ice than the ensemble median; this applies to winter and summer. The absolute sea-ice concentration differences between the individual products and the ensemble median mostly increase from winter to summer.

These findings document, together with the results presented and discussed in Kern et al. (2019, Fig. 11 and Appendix G), that the 10 PMW SIC distributions differ considerably in summer. These findings agree with results from previous inter-comparisons of PMW SIC products (e.g., Comiso et al., 2017; Ivanova et al., 2014, 2015; Spreen et al., 2008; Meier, 2005).

4   Results

4.1   Inter-comparison against MODIS sea-ice concentration (SIC)

Figure 4 exemplifies how the difference: PMW SIC minus MODIS SIC changes as a function of the stage of melt for the four groups. For this illustration, we select the same 8-day periods as used in the maps shown in Fig. 2. The four rows represent stages of melt: pre-melt (DOY 129, May 9-16), melt-advance (DOY 169, June 18-25), peak melt (DOY 201, July 20-27) and end of melt (DOY 241, Aug. 29-Sep. 5). The four columns represent groups I to IV by showing results of OSI-450 (group I), CBT-SSMI (group II), NT1-SSMI (group III) and NT2-AMSR-E (group IV). These examples are taken from different years, chosen because of a relatively small number of invalid or missing data. Figure 5 shows two-dimensional (2-D) histograms of PMW SIC (y-axis) versus MODIS SIC (x-axis) corresponding to the SIC maps used for the differences shown in Fig. 4. Note the logarithmic scale of the count. Similar histograms but based on SIC data of years 2003 to 2011 are shown in Fig. S05 in the supplementary material. We omit the pre-melt examples in Fig. 5 (and Fig. S05) because they exhibit limited additional information but show them for completeness in Fig. S06 in the supplementary material.

4.1.1   Pre-melt

For the pre-melt example (Fig. 4a-d), only groups I and III exhibit notable areas of larger over- and under-estimation of MODIS SIC (Fig. 4a, c), e.g. north of the Laptev Sea and the Fram Strait where group III exhibits negative differences above 15 % in magnitude. Apart from these patches of larger differences we can state that for pre-melt PMW SIC and MODIS SIC mostly agree within their uncertainties.

4.1.2   Melt advance

For the melt advance example (Fig. 4e-h), all four groups over-estimate MODIS SIC by 5-10 % south of the pole facing Greenland and the Greenland, Barents and Kara Sea. In most other regions, MODIS SIC is either under-estimated, e.g. in much of the central Arctic Ocean (Fig. 4e, g), or over-estimated, e.g. in the Beaufort and Chukchi Sea (Fig. 4f, h). Absolute differences remain mostly below 15 %. At this stage, melt has commenced everywhere and we observe a range of MPF values (see Fig. 2 b). However, we find no unique correspondence between SIC differences and the MPF. On the one hand, MPF values below 15 % (Fig. 2b) correspond well to areas with only small absolute differences for groups II and IV (Fig. 4f, h). On the other hand, near-0 % differences between PMW SIC and MODIS SIC co-exist with MPF values ranging from below 10 % to above 30 % for the same groups. Likewise, for groups I and III, the spatial variability of SIC differences in the central Arctic Ocean (Fig. 4e, g) is not reflected by the spatial variability in the MPF (Fig. 2b). The above-mentioned differences between groups I and III on the one hand and groups II and IV on the other hand, are also evident in the 2-D histograms matching the maps of Fig. 4e-h (Fig. 5, left column) . Over the period 2003-2011 (Fig. S05, left column, supplementary material), all four groups have the majority of SIC value pairs concentrated at values above 90 %. The differences in the distribution of the SIC value pairs between the groups are well reflected in the slight differences in linear regression line slope and intercept.

4.1.3   Peak melt

For the peak melt example, groups II and IV over-estimate MODIS SIC almost everywhere by up to 20% (Fig. 4j, l). This over-estimation is confirmed well by the respective 2-D histograms (Fig. 5e, k). Regions with only 5 to 10 % over-estimation of MODIS SIC by these groups correspond to a range of different MPF value: up to 15 % in the central Arctic Ocean, around 30 % in the southern Beaufort Sea and 40 % in the Canadian Arctic Archipelago (see Fig. 2c). Spatial patterns of SIC differences of groups I and III (Fig. 4i, k) are relatively similar to each other but differ considerably from those of the other two groups. We find that of all groups, group I has the highest linear correlation (0.84) and the smallest root-mean-squared difference (RMSD) of 7.8 % between PMW SIC and MODIS SIC (see Fig. 5, middle column). Over the period 2003-2011 (Fig. S05, middle column, supplementary material), groups I and III provide a quite symmetric distribution with linear correlations of 0.86 and 0.87, respectively. For the other two groups, over-estimation of MODIS SIC dominates – in agreement with Fig. 4j and 4l. Even though the linear correlation of 0.85 of group II is as high as those of groups I and III the distribution of values suggests two separate linear regressions.

4.1.4   End-of-melt

For the end-of-melt example, spatial patterns of MODIS SIC over-estimation by groups II and IV are very similar (Fig. 4n, p) as are the respective 2-D histograms (Fig. 5f, l and Fig. S05f, l, supplementary material). MODIS SIC over-estimation is largest where the melt-pond fraction is largest and vice versa (compare with Fig. 2d) – except in the southern Beaufort Sea.Overall, group I has the smallest differences to MODIS SIC (Fig. 4m), the most symmetric SIC distribution around the identity line (Fig. 5c) and the smallest RMSD of 7.8 % of all groups. Over the period 2003-2011 (Fig. S05, right column, supplementary material), group I and also group III have a quite symmetric SIC distribution, similar to peak melt.

4.1.5   Summary of the comparison to MODIS SIC

We summarize the average values of the statistical parameters: regression line slope and intercept (or offset), linear correlation and RMSD in Table 3. The averages are computed separately for the four stages of melt as the arithmetic mean over all parameter values of the respective group's products and 8-day periods within years 2003 to 2011. For example, for the average RMSD of group I for pre-melt, we average over four (products in group I) times three (three 8-day periods within pre-melt: DOY 129, 137 and 145) times nine (years) values. We do not further interpret the values given for pre-melt and refer the reader to Figs. S06 and S09 in the supplementary material. Table 3 shows an increase in correlation, RMSD and slope from melt advance to peak melt for all four groups. Overall, highest correlations between PMW SIC and MODIS SIC are obtained for groups I and III: 0.75 as a mean over melt advance to end-of-melt. If we take the RMSD as a measure of how accurate PMW SIC match MODIS SIC, group I products are the most accurate ones with a mean RMSD from melt-advance to end-of-melt of 8.3 %.

Figure 6 summarizes our results about the pan-Arctic (Arctic Ocean and Canadian Arctic Archipelago) multi-annual mean melt-season development of PMW SIC of the four products representative of the four groups in comparison to MODIS SIC and MPF. We refer to Fig. S07 in the supplementary material for results of the all 10 products. Temporal sampling is eight days. The mean MODIS MPF is smaller than 5 % until the end of May (pre-melt), gradually increasing (melt advance) to a mean MPF of 20-25 % between DOY 180 and DOY 235, i.e. between end of June and the 3rd/4th week of August (peak melt). We find near-0 % differences between PMW SIC and MODIS SIC for group I (Fig. 6a), orange symbols). These result from positive and negative biases cancelling out (see Fig. 4, Fig. 5b and Fig. S05, supplementary material). Group II (Fig. 6b) exhibits near-0 % differences until the end of June but shows up to 10 % *more* sea ice than MODIS afterwards. Group III (Fig. 6c) first exhibits differences close to zero but shows *less* sea ice than MODIS during peak melt.

Of the four groups of products investigated, we get three different kinds of agreement between PMW SIC and MODIS SIC. Most importantly, instead of the under-estimation commonly reported in the literature (e.g. Rösel et al, 2012b: Comiso and Kwok, 1996; Steffen and Schweiger, 1991; Cavalieri et al., 1990) our results suggest that an over-estimation of the actual sea-ice concentration is common for several PMW SIC products. With that this study agrees with the findings in Kern et al. (2016) but we note that the latter study is based i) on data of one summer season only; ii) on data of a sub-region of the Arctic Ocean only; iii) on PMW SIC values computed from re-implementations of PMW SIC algorithms using fixed winter ice-tie points and allowing SIC values larger than 100 %.

4.2   Inter-comparison against MODIS net ice-surface fraction (ISF)

Melt ponds have the largest impact on PMW SIC because of the inability – at the microwave frequencies used – to discriminate between open water in the form of melt ponds on the ice floes and open water in the form of leads and openings between the ice floes. Like detailed in the introduction, based on physical principles a sea-ice concentration retrieval algorithm should provide a value of 60 % for a case A: 100 % sea-ice concentration with 40% melt ponds, and a case B: 75% sea-ice concentration with 15 % melt ponds; in other words the algorithm should provide the net ice-surface fraction. Therefore, a logical next step to better understand the causes of the SIC differences reported in Sect. 4.1 is to investigate how PMW SIC compares to ISF, as derived, e.g., from MODIS (Sect. 2.2). The inter-comparison between PMW SIC and MODIS ISF is carried out similarly to the inter-comparison to MODIS SIC (Sect. 4.1). We organize the results exactly in the same structure as in Sect. 4.1. We present in Fig. 7 a set of maps of the differences PMW SIC minus MODIS ISF for the previously selected

8-day periods (compare Fig. 4), complemented by the respective 2-D histograms shown in Fig. 8 (compare Fig. 5) and extended to the entire period 2003-2011 in Fig. S08 in the supplementary material (compare also Fig. S05).

4.2.1 Pre-melt

For the pre-melt example (Fig. 7a-d), almost no melt ponds are observed in the Arctic Ocean (see Fig. 2a); MODIS ISF equals MODIS SIC. The maps of the differences PMW SIC minus MODIS ISF shown in Fig. 7a-d are almost identical to the maps shown in Fig. 4a-d. The differences between the 2-D histograms of this example are small, but when considering the entire period 2003-2011 we observe a tail of near-100 % PMW SIC values which spreads over a range of MODIS ISF between 60-70 % and 100 % (compare Figs. S06 and S09, left with right column, supplementary material). The values in this tail are from locations where MODIS ISF is smaller than MODIS SIC and PMW SIC over-estimates MODIS ISF, e.g. in the Laptev Sea and the East Siberian Sea (Fig. 7a-d). At these locations MPF is ~ 10 % (Fig. 2a).

4.2.2 Melt advance

For the melt advance example (Fig. 7e-h), we find widespread over-estimation of MODIS ISF by all groups. Highest over-estimations occur in the Chukchi and Beaufort Sea: up to 25 % for groups I and III (Fig. 7e, g), up to 35 % for groups II and IV (Fig. 7f, h). North of this region an area extending across the Arctic Ocean towards Fram Strait has the smallest over-estimation of MODIS ISF. Overall, the spatial pattern of the differences PMW SIC minus MODIS ISF matches reasonably well with the respective MPF map (Fig. 2b). For example does the area with differences ≤ 10 % in the central Arctic Ocean coincide with a MPF of ~10 %. However, areas with highest over-estimation of MODIS ISF (e.g. by groups II and IV) do not necessarily coincide with the highest MPF. We get back to this issue in Sect. 5.1. The respective 2-D histograms reveal a bi-modal distribution of the value pairs and magnitude of counts, which is similar for groups I and III (Fig. 8a, g) on the one hand, and groups II and IV (Fig. 8d, j) on the other hand – like we found in Sect. 4.1 for MODIS SIC. The locations of the modes agree well with the differences shown in Fig. 7e-h. The findings from Fig. 8, left column, appear to be typical for the entire period 2003-2011 as illustrated by the 2-D histograms shown in Fig. S08, left column, in the supplementary material.

4.2.3 Peak melt

For the peak melt example (Fig. 7i-l), PMW SIC over-estimates MODIS ISF everywhere. The spatial distribution of the differences PMW SIC minus MODIS ISF matches relatively well with the observed MPF (Fig. 2c). The over-estimation is particularly high for group II (Fig. 7j): 20-25 % in the central Arctic Ocean and up to 45 % in the Chukchi and Beaufort Sea. The over-estimation is lowest for group III (Fig. 7k): 20-25 % in most areas and is relatively homogeneous with respect to the SIC range as is evident in the respective 2-D histogram (Fig. 8h). This applies also for group I (Fig. 8 b). Accordingly, the highest linear correlation coefficient and lowest RMSD values are obtained for groups I and III. The findings from Fig. 8, middle column, appear to be typical for the entire period 2003-2011 as illustrated by the 2-D histograms shown in Fig. S08, middle column, in the supplementary material.

4.2.4 End-of-melt

For the end-of-melt example (Fig. 7m-p), we find reasonable agreement between the distribution of the MPF (Fig. 2d) and the difference PMW SIC minus MODIS ISF for all groups. Areas of MPF less than 5 % coincide with differences between -10 % and 10 %. Areas with high MPF, however, do not necessarily coincide with areas of a large difference PMW SIC minus MODIS ISF across the groups, as for instance the region with MPF ~35 % in the Beaufort Sea (Fig. 2d) for which we find differences between ~15 % (Fig. 7o) and ~40 % (Fig. 7n, p). The respective 2-D histograms (Fig. 8, right column) reveal very different relationships between PMW SIC and MODIS ISF for groups I and III on the one hand and groups II and IV on the other hand. When considering the entire period 2003-2011 (Fig. S08, right column, in the supplementary material), values scatter much more and areas with high counts are much less confined than for the example shown in Fig. 8. This could be the result of i) substantial inter-annual variation in the late summer co-existence and extent of melting and freezing conditions and of ii) a larger uncertainty in the surface type classification in the MODIS product due to an unknown fraction of already re-frozen melt ponds (see Sect. 2.2).

4.2.5 Summary of the comparison to MODIS ISF

Values entering Table 4 are computed similarly as those in Table 3 (see section 4.1.5). They show an increase in mean values of correlation, RMSD and slope from melt advance to peak melt for all four groups. Overall, we find highest correlations between PMW SIC and MODIS ISF for groups I and III: 0.76 as the mean from melt advance to end-of-melt. For these groups, we also find the smallest mean RMSD values: 19.0 % and 19.5 % as the mean from melt advance to end-of-melt, and 23.2 % and 24.5 % during peak melt. These values can be taken as a measure of MODIS ISF over-estimation by PMW SIC. Slopes of the linear regressions get closest to 1 for groups I and III during peak melt, suggesting a solid linear relationship between PMW SIC and MODIS ISF – also in view of the distributions of values and counts in the 2-D histograms.

Agreement between the MPF and the magnitude of the difference PMW SIC minus MODIS ISF differs among the four
410   groups (compare Figs. 2 and 7). It appears that MODIS ISF is over-estimated by group III by an amount smaller than the MPF
while for groups II and IV the MODIS ISF over-estimation is often larger than the MPF. This observation is confirmed by Fig.
9 (compare with Fig. 6). For group I (Fig. 9a), MPF values (in cyan) agree with the difference PMW SIC minus MODIS ISF
(in orange) within 2% for the entire melt season. Hence, on a pan-Arctic scale, averaged over years 2003 to 2011, group I
products' over-estimation of MODIS ISF *equals* the MPF. The over-estimation of MODIS ISF by group II (Fig. 9b) and group
415   IV (Fig. 9d) is *smaller* than the MPF during peak melt and end-of-melt by up to 10 %, while for group III (Fig. 9c) this over-
estimation is *larger* than the MPF by up to 10 %. We refer to Fig. S10 in the supplementary material for results obtained for
all 10 products.

We note in this context, that we carried out an inter-comparison between ship-based visual sea-ice observations, providing
independent estimates of SIC, MPF and ISF, and all 10 products' SIC data. The results of this inter-comparison support our
420   findings from Sect. 4.1 and this Section (see Sect. S1.1, S2.1 and Fig. S01 and S02 in the supplementary material).

4.3   Bias correction as a potential way forward

The main motivations for this paper are to evaluate the performance of PMW SIC products during summer conditions
and to better understand why PMW SIC products usually do not provide the net ice surface fraction – which they should
425   following physical principles. The results presented so far document that none of the 10 PMW SIC products provide a faithful
picture of the ISF, nor are accurate measures of the summer SIC. The discussion given further below in Sect. 5 does not only
reveal possible explanations of the diversity of evaluation results but demonstrate the complexity involved in a potentially
planned improvement of the used algorithms – be it by further development of the algorithm itself, be it via application of
more advanced ice tie-point retrieval approaches. Here we discuss potential ways forward in the short to medium term (using
430   existing PMW products) and in the longer term (preparing and using improved PMW SIC products).

For groups I and III, our comparison between PMW SIC and MODIS SIC (Sect. 4.1) and between PMW SIC and
MODIS ISF (Sect. 4.2) suggests linear functional relationships. In the short term, these offer the prospect for users of the
existing PMW SIC datasets from these two groups to perform bias corrections of the PMW SIC towards either true SIC
(representative of the sea ice area fraction of the geophysical model at hand), or net ISF (representative of the ice surface clear
435   of melt ponds). The mean slope and intercept values prepared in Table 3 and Table 4 but also slope and intercept of the
individual linear regression lines (e.g. Figs. 5 and 8) could allow such a bias correction, noting all the limitations of these
parameter values that are derived at a pan-Arctic and, as presented in Tables 3 and 4, multi-year scale. For example, we find
values of the linear correlation larger than 0.85 and slope close to 1 (see Table 4) with respect to MODIS ISF. With such a
bias correction one might be able to get closer to the physically more meaningful result of a PMW SIC which equals the net
440   ISF.

We do not explore or comment at length a bias correction of PMW SIC towards true SIC with values in Fig. 5 (or
Table 3). As predicted by physics, the bias correction towards true SIC is less skilled than towards net ISF, as can be assessed
by the lower correlation values R in Fig. 5 (SIC) compared to Fig. 8 (ISF). A bias correction towards SIC would attempt to
force the PMW SIC product to represent open water in two different ways: as sea ice when it is a melt pond and as true open
445   water when it is a lead / opening between the ice floes, despite the fact that the surface emissivity and hence the observed TB
is determined by the overall total amount of liquid water at the surface.

We hence only use linear regression equations obtained from the comparison between PMW SIC and MODIS ISF
for a bias correction of PMW SIC towards MODIS ISF. We first test how well the bias correction works in comparison to
MODIS ISF, i.e. investigate whether the difference between PMW SIC and MODIS ISF is reduced to zero, and subsequently
450   compare the bias-corrected PMW SIC to MODIS SIC. This bias correction is exemplarily carried out for OSI-450 (group I),
CBT-SSMI (group II) and NT1-SSMI (group III) for peak melt (DOY 201, year 2009) in Fig. 10 and for melt advance (DOY
169, year 2010) in Fig. S11 in the supplementary material. Note that we use slope and intercept values obtained exactly for
these examples, i.e. from Fig. 8, and not from Table 4.

The bias correction works well with respect to MODIS ISF for peak melt. The majority of the differences bias-
455   corrected PMW SIC minus MODIS ISF has a magnitude less than 5 % (Fig. 10d-f). The linear correlations are as high as for
the un-corrected case and the RMSD reduces to around 6-7 % (compare Fig. 10a-c with Fig. 8b, e, h); the slope is almost
identical with the identity line for OSI-450 and NT1-SSMI. We note that if the results of this bias correction prove to be of
equal quality for other parts of the peak-melt period and other years, one could use the respective equations to obtain an
independent estimate of the ISF from the entire PMW SIC data record, i.e. from 1979 to today. This could serve as an important
460   boundary condition for the estimation of the surface albedo independent of daylight and cloud cover, complementing existing
data sets and aiding in their evaluation (e.g. Riihela et al., 2010, 2017). For melt advance (Fig. S11a-f, supplementary material),
differences between bias-corrected PMW SIC and MODIS ISF are considerably larger than for peak melt, especially for CBT-
SSMI and NT1-SSMI. While the slopes all agree quite well with the identity line, RMSD values are much larger than for peak
melt. This is also well evident from the larger scatter of value pairs in the respective 2D-histograms. During melt advance it
465   appears advisable to use non-truncated SIC values if available, because the fraction of SIC larger than 100 % is the highest
during the summer melt cycle (see Sect. S3.1 in the supplementary material); at this stage we did, however, not further quantify
the effect this may have on the results of the bias correction performed.

As expected, the difference bias-corrected PMW SIC minus MODIS SIC is negative all over and has a magnitude of
about 25 %. We find a relatively homogeneous distribution of differences (Fig. 10j-l). We find value pairs in the respective
470   2D-histograms to be confined below the identity line around a linear regression line with a slope slightly larger than 1 and

linear correlations comparable to the un-corrected PMW SIC (Fig. 5b, e, h). Most striking is the similarity of the distributions in the maps and 2D-histograms across the three products and the fact that the RMSD between bias-corrected PMW SIC and MODIS SIC not only agrees within 1 % among the three products but also agrees with the modal MPF of 25 % for DOY 201 of the year 2009 (see Fig. 2c). Thus, the bias correction towards ISF reconciles the various PMW SIC products towards a consistent difference PMW SIC minus MODIS SIC being of the same order of magnitude as the average MPF. In contrast, during melt advance with melting and frozen, wet and dry surfaces co-existing the results of the bias correction of PMW SIC appear less convincing (Fig. S11g-l, supplementary material). Here CBT-SSMI provides a difference bias-corrected PMW SIC minus MODIS SIC which in the central Arctic Ocean is uniform at about 10 %. While this value matches well with the MPF map and the first mode (9 %) of the bi-modal MPF distribution for DOY 169 of the year 2010 (Fig. 2b), the other differences range between -40 % and 10 % demonstrating that during melt advance a bias correction as proposed is potentially of limited value.

A logical next step would be to use the mean values of slope and intercept (Table 4) instead of the individual values (see above). We note that the mean values (Table 4) differ considerably from the individual values. For example, for OSI-450 during the 8-day period beginning at DOY 201 in year 2009 we have individual values of intercept and slope of 27.3 % and 1.01, respectively, while the respective mean values for group I during peak melt are 30.3 % and 0.86 (Table 4). We did not yet apply these, however, because we regard these first attempts as a feasibility study only. To carry out a comprehensive study about the potential of a bias correction of PMW SIC towards MODIS ISF would require a well-thought concept about how to adequately evaluate the bias-corrected SIC; this is beyond the scope of this paper.

After such a study, in the short term, values given in Table 4 could allow users of the existing PMW SIC datasets to bias correct these products towards net ISF. Such a bias correction is however not necessarily useful in practice. Indeed, users must now rely on additional source of information to link their SIC (e.g. from a geo-physical model) to a measure of the ISF. This for example requires a trustworthy representation of the evolution of melt ponds on sea ice in their model. Several such melt pond schemes are being developed (e.g. Pedersen et al., 2009; Flocco et al., 2010; Scott and Feltham, 2010; Holland et al., 2012; Skyllingstad et al., 2015; Popović and Abbot, 2017) but their application and evaluation reveal some challenges remain (e.g. Light et al., 2015; Tsamados et al., 2015; Zhang et al., 2018; Burgard et al., 2019; Dorn et al., 2019). Still, in the long run, using PMW SIC as an observation of net ISF should be favoured, as it is more meaningful and will be more accurate. This will especially be the case when producers of PMW SIC datasets put additional effort in improving their algorithms and/or ice tie-point selection schemes to actually retrieve un-biased observations of the net ISF. There is furthermore no doubt that both improving melt pond schemes in models and designing better PMW-based SIC algorithms in summer will benefit from better accuracy and availability of EO-based melt-pond fraction CDRs from visible / infrared imager instruments such as NASA MODIS (e.g., Rösel et al., 2012), the European Space Agency's MEdium Resolution Imager Sensor (MERIS) (e.g., Istomina et al., 2015; Zege et al., 2015), or the Copernicus Ocean and Land Colour Imager (OLCI). There is a critical EO-observation gap to be filled here in order to further improve the Sea Ice ECV.

4.4 The impact on sea-ice area

Independent of the way forward and future attempts to get closer to what appears to be physically more correct when using satellite PMW data for SIC retrieval we note that there might be applications which require an accurate SIC *including* the melt ponds on top, i.e. without the need to understand why the PMW SIC does not match the actual net ISF. The classical application would be the computation of the SIA, the sum of the area of all ice-covered grid cells weighed by SIC. We demonstrated in Sect. 4.1 which groups over- and/or under-estimate MODIS SIC where and by which amount (Figs. 4, 5 and S05, supplementary material). We illustrated that on a pan-Arctic scale, averaged over years 2003-2011 group I exhibits a near-0 % bias, while group III appears to under-estimate MODIS SIC by 5-10 % during peak melt and end-of-melt, and group II appears to over-estimate MODIS SIC by around 10 % (see Fig. 6). This finding holds for melt-pond fractions up to 30 %, for NT1-SSMI and group II products even up to 40 % (Fig. 11a). In addition, Fig. 11b further illustrates how well the difference PMW SIC minus MODIS ISF can be seen as a linear function of the MPF for group I – at least up to a MPF of ~ 30 %.

Coming back to the computation of SIA and the potential influence of melt ponds: As shown in Kern et al. (2019) and Ivanova et al. (2014), the choice of the product for the computation of SIA from PMW SIC data makes a difference. For months July through September of years 2002-2011, the SIA computed from PMW SIC of group I products is ~400 000 km² larger than SIA computed from NT1-SSMI (group III) and ~600 000 km² smaller than SIA computed from group II products (Kern et al., 2019, Fig. G2g-i). The average pan-Arctic MODIS SIC for these months is between 85 % and 90 % (Fig. 6 and Fig. S07, supplementary material). Considering a value of 90 % and assuming an extent of 6 million square kilometres to be covered by some amount of sea ice on average for these months, we end up with a SIA of about 5.4 million square kilometres based on MODIS. Group I products, exhibiting zero bias to MODIS SIC (Fig. 6a, Fig. S07a-d) yield the same SIA estimate. NT1-SSMI, exhibiting a negative bias of 5-10 % (Fig. 6c, S07i), say 7 %, i.e. a pan-Arctic average SIC of 83 %, yields a SIA of 5.0 million square kilometres. Group II products, exhibiting a positive bias of ~10 % (Fig. 6b, Fig. S07e-g), i.e. a pan-Arctic average SIC of 100 %, yield a SIA of 6.0 million square kilometres. Based on these considerations we can conclude that the summer-time differences between the SIA estimates of the 10 products presented by Kern et al. (2019) can be explained well with the differences between PMW SIC and MODIS SIC presented in this paper.

5   Discussion and Conclusions

5.1   Discussion

5.1.1   *Understanding our observations*

Our results demonstrate that the different products respond quite differently to the changes in the sea-ice cover during summer melt, and that none of them is doing things quite right. This is not surprising given the variety of different sea-ice and snow physical properties relevant for satellite PMW sensing of sea ice – open and re-frozen melt ponds, slush, saturated or wet snow, new snow, coarse-grained melting or re-frozen snow, bare melting ice, bare dry ice, submerged ice and various forms of new ice – co-existing during summer at pan-Arctic scale but possibly even within one satellite footprint. These physical properties not only undergo substantial changes during the melt season, they also have a large spatiotemporal variability. The net surface energy balance driving the melting or freezing is very sensitive to variations in the cloud cover and to precipitation events, which can vary on short temporal and local spatial scales. Melting and re-freezing of coarse-grained snow or formation of a thin ice cover at the melt-pond surface can occur within a few hours.

Besides melt ponds wet snow and melting and re-frozen coarse-grained snow are the most relevant surface parameters. At the microwave frequencies used in this paper the emissivity of the wet snow cover is close to 1, resulting in a microwave TB close to 273.15 K – the melting temperature of snow. Typical increases in microwave TB due to an increase in snow wetness range between 10-15 K and 60 K (Kern et al., 2016, Table 1). The magnitude of this TB increase depends on the sea-ice emissivity being a function of frequency and polarization. The increase is higher for multiyear than first-year ice. It is higher at horizontal (H) than vertical (V) polarization and at higher (near-90 GHz) than lower (19 GHz) frequency (f). This is all in accordance with the lower TB of multiyear ice than first-year ice and the lower TB at horizontal than vertical polarisation of winter sea ice. Concomitant is a decrease in the normalized TB polarization difference: $PR(f) = \frac{TB_V(f) - TB_H(f)}{TB_V(f) + TB_H(f)}$, at a frequency of 19 GHz: PR19, or 89 GHz: PR89, as well as a decrease in the magnitude of normalized TB gradient ratios: $GR(f_1 f_2) = \frac{TB_V(f_1) - TB_V(f_2)}{TB_V(f_1) + TB_V(f_2)}$, between 19 GHz and 37 GHz: GR3719, or 19 GHz and 89 GHz: GR8919, quantities that are used in the NT1-SSMI and NT2-AMSR-E algorithms (see Kern et al. (2019) for a summary of relevant technical aspects of the 10 algorithms used). Typical decreases in microwave TB due to an increase in snow grain size, e.g. due to surface refreezing or surface crust formation, are around 15-35 K (Kern et al., 2016, Tables 1 and 3). The magnitude of such a decrease is larger at horizontal than vertical polarization, and larger at higher than lower frequencies. Concomitant is an increase in the magnitude of, e.g., PR19 and GR3719 by 0.02 and 0.05, respectively. Such increases correspond to about 10 % in SIC. Melting of a coarse-grained snow cover reverts the above-mentioned changes, causing diurnally changing microwave TB values should melt-refreeze cycles commence.

In summary, whenever the surface conditions become wetter, microwave TB increase, while polarization and frequency differences decrease. Whenever surface conditions become drier, microwave TB decrease, while polarization and frequency differences increase. This view is certainly a simplification of the true conditions which are more complex due to the vertical structure of the snow cover, the different near-surface properties of first-year ice compared to multiyear ice, melt-pond drainage and other processes. However, this view allows us to understand that during summer, differences between an actually observed microwave TB or TB difference and an ice tie point can be caused by the mismatch between actual and tie-point conditions with respect to the representation of: i) melt ponds; ii) snow wetness; iii) snow grain size / surface type; iv) ice type; v) a mixture of all these.

The implications for the SIC retrieval depend on the type and update interval of the tie points of pure sea ice, i.e. 100 % sea-ice concentration (see e.g. Lavergne et al., 2019). The ASI algorithm (group III) uses one global fixed sea ice tie point value (Kaleschke et al., 2001). NT1-SSMI (group III) uses one fixed set of fixed TB values for first-year ice and multiyear ice. NT2-AMSR-E (group IV) uses sets of 12 fixed TB values of all involved channels (see Table 1) of three different ice types: thin ice, ice type A (merges first-year and multiyear ice) and ice type C (sea ice with a thick snow cover); the 12 fixed TB values are based on the 12 different atmospheric states used to compute the look-up tables for the SIC retrieval (Markus and Cavalieri, 2009). All other products (groups I and II), except the contribution of NT1-SSMI to the NOAA CDR product, use an ice line which interpolates between signatures of first-year and multiyear ice and which is updated daily (Lavergne et al., 2019, Comiso and Nishio, 2008). For group I products this ice line is computed from TB measurements over closed ice within a moving 15-days interval centred at the day of the actual SIC retrieval; closed ice is defined as grid cells with more than 95% NASA-Team algorithm SIC. A post-processing step optimizes the location of the ice line with respect to the different TB values encountered as function of ice type. The Comiso bootstrap algorithm (CBT-SSMI and CBT-AMSR-E) derives the ice line via linear regression analysis of the respective TB value cluster. This is done in both TB spaces, i.e. TB37V / TB37H used for SIC larger than 90 % and TB37V / TB19V used for SIC below or equal to 90 % (Comiso et al., 1997). The offset (or intercept) of the obtained linear regression line is increased by a few Kelvin to account for the presence of some open water (2-3 %) in closed ice areas (Comiso and Nishio, 2008; Comiso, 2009). These differences in the ice tie points already suggest that the different products represent the actual sea-ice conditions with different levels of accuracy. None, to our best knowledge, of the algorithms used in the 10 products employ regionally varying ice tie points notwithstanding the large spatial variability of the relevant physical properties during the melt season.

585   For all groups we observe small areas of elevated positive differences PMW SIC minus MODIS ISF (Fig. 4a-d). These areas can be explained with the concurrent melt-pond fraction. An influence by elevated snow wetness is unlikely, because this would cause an increase in PMW SIC which in turn would result in an overestimation of both MODIS SIC *and* MODIS ISF – which is not observed. However, group I and III reveal patches of MODIS SIC and ISF under-estimation (Fig. 4b, d, see also Figs. S07 and S10 in the supplementary material) not found for the other groups. As *one possibility* these patches

590   could be explained by a re-frozen surface or coarse-grained snow not represented in the ice tie points. NT1-SSMI (group III) PMW SIC is based on PR19 and GR3719 and the above-mentioned surface conditions would cause an under-estimation of the SIC (see Kern et al., 2016, Fig. 6a: respective data pairs would move away from the red ice line towards the open water tie point). The algorithms of group I use NT1-SSMI SIC > 95 % as a-priori information for the computation of the ice tie point (Lavergne et al., 2019). Grid cells with an actual near-100 % SIC, where such an under-estimation by NT1-SSMI occurs under

595   the mentioned surface conditions, are possibly excluded from the tie-point estimation for group I algorithms. As a consequence such regions are not represented by the ice tie point, we have a mismatch between actual and tie-point conditions and the retrieved SIC is biased low.

5.1.1.2   Example 2: melt conditions

We find cases where near-100 % MODIS SIC coincides with a near-0 % difference PMW SIC minus MODIS SIC,

600   an over-estimation of MODIS ISF by 10-15 % and a MPF of 10-15 %, e.g. for CBT-SSMI (group II) in the central Arctic Ocean (Figs. 4f and 7f). One would expect that the open water associated with the melt ponds (non-zero MPF) lowers the actually observed TB and that therefore the actual PMW SIC is smaller than the MODIS SIC. This is not the case. We offer three explanations. Explanation A: the ice tie point includes some influence of melt ponds. In that case the ice tie point (see e.g. the ice line in Kern et al., 2016, Fig. 6c, d) would be located at a lower TB value slightly closer to the open water tie point.

605   The observed TB would then match with this ice tie point – provided that actual ice surface properties between the melt ponds match the conditions represented by the ice tie point – and the retrieved SIC would be close to 100 %. We hypothesize that this is one of the most likely reasons for the over-estimation of MODIS ISF by group I products. These products use NT1-SSMI SIC > 95% to define regions for ice-tie point retrieval (Lavergne et al., 2019), regions which according to the results of our paper exhibit a non-zero melt-pond fraction. Explanation B: the surface between the melt ponds is wet but this is not

610   represented by the ice tie point. In that case the observed TB is lowered by the melt ponds but at the same time increased by the wet surface. Both effects could compensate such that the observed TB is close enough to the ice tie point to yield near-100 % SIC. Evidence for an increase in TB during summer melt in June is given, e.g. in Kern et al. (2016, Fig. 8a-c); the cluster of increased TB values is located considerably above the winter-time ice line concomitant with near-100 % MODIS ISF (Kern et al., 2016, Fig. 6c). Explanation C: the ice tie point represents a refrozen surface or multiyear ice and because of this is

615   located similarly closer to the open water tie point as in explanation (A). The observed TB would match the ice tie point for the wrong reason and the algorithm would provide near-100 % SIC.

We find other cases where 100 % PMW SIC coincides with a MODIS SIC of 85% and a melt-pond fraction of 25-30%, e.g. for CBT-SSMI (group II) in the Chukchi Sea (Figs. 4f and 7f). Here, despite the large open water fraction of 40-45 %, PMW SIC is 100 %, which corresponds to an over-estimation of MODIS SIC by 15% and of MODIS ISF by 40-45 %. All

620   explanations suggested in the previous paragraph might apply here – very likely in combination with each other. Such a large over-estimation of MODIS ISF would, if we use only explanation B, require un-physical sea-ice surface emissivities larger than 1 (not shown). Using different algorithms Kern et al. (2016) computed the SIC based on elevated summer-time microwave TB values. They found that – theoretically – SIC values would need to be as high as 140 % for the fraction of the grid cell not covered by water to explain their observed differences between PMW SIC and MODIS ISF for MODIS SIC values above 90

625   %. The way ice tie points are derived in the Comiso bootstrap algorithm suggests i) inclusion of melt ponds in the tie point, ii) unaccounted wet snow / wet surface between the melt ponds and iii) ice type mismatches to be the most likely combination leading to the observed over-estimation.

5.1.1.3   Summary

The co-existence of different surface properties during summer adds complexity to the SIC retrieval using satellite

630   PMW TB observations. Our attempts to explain the observations suggest that an adequate understanding of – on the one hand – the actually encountered sea-ice and snow properties and – on the other hand – the properties represented by the ice tie points is required. The influence exerted by different surface properties on the actually measured TB or TB differences like PR19 or GR3719 can cancel out. Examples of such properties are the co-existence of melting and re-frozen coarse-grained snow or the co-existence of wet snow and melt ponds. A consequence of this is that despite the actual surface conditions do not match the

635   ice tie-point conditions, retrieved PMW SIC sometimes appear to be accurate. It needs to be better understood how ice tie points are derived during summer conditions and how their validity can be assessed as a function of location and time. Based on our findings, one of the largest issues could be the inclusion of an unknown amount of melt ponds into the ice tie point.

5.2   Conclusions

Following up on the release of three new global sea-ice concentration (SIC) climate data records (CDRs) described

640   in Lavergne et al. (2019), this paper focuses on an inter-comparison of these three CDRs and seven other SIC products (see

Kern et al., 2019) with estimates of the SIC and the net ice-surface fraction (ISF) in the Arctic during summer (May through September) for years 2003 through 2011 obtained from satellite observations by the MODIS sensor. The motivation for this dedicated paper is the fact that it is particularly challenging to derive the SIC of melting sea ice. It is impossible with the current microwave radiometer sensors to distinguish water in melt ponds on top of the sea ice from the water in the leads between the sea-ice floes. What we expect to measure with the microwave radiometer sensor is therefore the ISF even though this is not what scientists normally relate to the term "sea-ice concentration". The ISF radiometric signature and especially its variability is difficult to characterise when the sea ice is melting and this results in large random and systematic uncertainties when retrieving the SIC and ISF from PMW observations. Our study employs 10 SIC products which we assign to four groups based on their retrieval algorithm (see Table 1) – like we did in our previous study (Kern et al., 2019). Our results document that none of the ten PMW SIC products provide an accurate measure of the summer SIC, or a faithful picture of the ISF.

Overall we find group I products (SICCI and OSI-450, see Table 1) exhibit a near-0 % bias to the MODIS Arctic average SIC – independent of melt-pond fractions up to ~35 % (Fig. 11a). Group II (CBT-SSMI, CBT-AMSRE and NOAA-CDR) and IV (NT2-AMSRE) products have a positive bias of 5-10 % and NT1-SSMI (group III) has a negative bias of 5-10 %. However, these small overall biases are the result of wide-spread, spatiotemporally varying positive and negative differences of substantial magnitude which cancel out in a pan-Arctic mean for some of the products. Magnitudes of these biases frequently reach up to 20-25 % for groups I and III and up to 30-35 % for groups II and IV. Unlike what is generally accepted and reported in the literature, we observe that several PMW SIC products significantly over-estimate true SIC.

By comparing PMW SIC with MODIS ISF and the MODIS melt-pond fraction (MPF) we find that SIC products of all four groups substantially over-estimate MODIS ISF. This over-estimation is generally related to the MPF but the degree of over-estimation varies between the groups. On pan-Arctic scale, group I products over-estimate MODIS ISF by almost exactly the overall mean MPF for values below ~30 % (Fig. 11b). In contrast, groups II and IV over-estimate MODIS ISF by an amount 5-10 % higher than the mean MPF while NT1-SSMI (group III) over-estimate MODIS ISF by an amount 5-10 % smaller than the mean MPF. ASI-SSMI, albeit assigned to group III based on its SIC retrieval approach, performs considerably different to the other product assigned to group III: NT1-SSMI; results for ASI-SSMI actually resemble a combination of our findings for group I and group IV (see Fig. 11 and Figs. S07 and S10 in the supplementary material).

The observed differences between PMW SIC and MODIS SIC or ISF cannot be explained by the presence of melt ponds alone. Often regions exhibiting high MPF values do not coincide with regions of large over-estimation of MODIS ISF by PMW SIC. For example, we find regions with near-100 % PMW SIC coinciding with MODIS ISF and MODIS SIC of 55 % and 85 %, respectively. The associated MPF is only ~30 % and with that over-estimation of MODIS ISF is 15 % larger than the MPF. By taking into account the different PMW data used by the various algorithms and relevant surface properties other than melt ponds – such as wet snow, (refrozen) coarse grained snow or a frozen surface in general – we discuss potential reasons for our observations. For the near-19 GHz and near-37 GHz frequency channels often used for SIC retrieval (see Table 1), open water in the form of melt ponds reduces the observed PMW brightness temperatures. However, during summer the surface / snow between melt ponds is likely wet which results in an increase of their brightness temperatures compared to typical winter and spring conditions. The two effects counteract and might result in a PMW SIC close to 100 % despite a considerably smaller ISF of, e.g., 80 % caused by the melt ponds. This applies in particular when the ice tie point does not well represent wet snow conditions. Another scenario that would have the same effect is that the ice tie point does represent wet snow conditions between the melt ponds well but does *in addition* include an un-known amount of open water due to melt ponds, i.e. does not represent 100 % melt-pond free sea ice. Our results illustrate that of the two groups of products employing advanced ice tie point retrieval methods, i.e. groups I and II, the methods of group I are considerably more successful in mitigating the unwanted influence of surface properties other than melt ponds.

Our inter-comparison reveals similarity in the results between group I and III products on the one hand and between group II and IV products on the other hand. This similarity is particularly interesting because ice tie-point estimation differs substantially between group I and III, and between group II and IV. While groups I and II have comparably advanced schemes to derive the ice tie point at daily temporal scale, ice tie points are fixed throughout the year for algorithms of the other two groups. This appears to call for a revision of the currently used concepts to derive and use ice tie points during summer. One potential solution to reduce SIC biases reported in this paper would be to add regional variation by, e.g. incorporating the different regionally varying stages of melt into the ice tie-point estimation. This could be done, e.g., by using maps of melt onset derived from PMW observations (e.g. Stroeve et al., 2014; Markus et al., 2009) – possibly in combination with atmospheric re-analysis data or observations of the ice surface temperature. Another solution could be to simplify the entire SIC retrieval process by, e.g., assuming melt conditions globally and derive a global ice tie point for melting conditions. The smaller sensitivity of lower frequency channels, i.e. near 7 GHz or near 1 GHz, to changes in snow grain size or snow wetness make these particularly good candidates for ice tie-point optimization during summer conditions. Such channels are for example offered by AMSR-E, AMSR2, the Soil Moisture and Ocean Salinity (SMOS) mission and the planned Copernicus Imaging Microwave Radiometer (CIMR) mission.

One goal of such an optimization could be to further support the switch from sea-ice extent (SIE) to sea-ice area (SIA) as the main parameter to monitor long-term changes of the Arctic sea-ice cover. Sea-ice covers sharing the same sea-ice edge provide the same SIE regardless of how open the sea-ice cover inside the ice edge actually is. In contrast, SIA allows one to distinguish between a comparably open sea-ice cover = low SIA and a highly compact sea-ice cover = large SIA. Hence the SIA clearly outperforms SIE in terms of providing information about the status of the sea-ice cover inside the ice edge (e.g. Notz, 2014), but is also much more affected by systematic summer biases as documented here for all 10 algorithms. Another goal could be to support moving away from retrieving sea-ice concentrations during winter and some highly inaccurate, ill-defined quantity during summer, still called sea-ice concentration, and rather retrieve net ice-surface fraction year-round – the quantity which by physical means it the one accessible with these microwave radiometry frequencies year-round. A starting

point of such a switch could be based on the obtained linear relationships between PMW SIC and MODIS ISF which appeared to be robust enough for our attempt to perform a bias correction of the PMW SIC during peak-melt conditions. In any case, such a switch will require improved algorithms for PMW observations, and improved melt-pond formulations in geophysical models. In turn, this will require preparing improved, error-characterized, longer and sustained melt-pond fraction datasets derived from satellite observations in the visible/near-infrared frequency range (MODIS, MERIS, OLCI, etc.). We recommend that the sea-ice communities work towards such a switch to improve sea ice as an Essential Climate Variable (ECV).

Data availability. All sea-ice concentration products except SICCI-12km are publicly available from the sources provided in the reference list or in Kern et al. (2019). The SICCI-12km product is available upon request from T. Lavergne. The MODIS data set of sea-ice concentration, melt-pond fractions and net ice surface fraction is available from: https://doi.org/10.1594/WDCC/MODIS__Arctic__MPF_V02 , the standardized ship-based observations are available from: https://doi.org/10.26050/WDCC/ESACCIPSMVSBSIO .

Author contributions. SK wrote the manuscript. All co-authors contributed to the concept and work presented in the paper and also assisted in the writing. SK performed the data analysis and inter-comparison with contributions in the interpretation of the results from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The work presented here was funded by EUMETSAT (through the 3rd Continuous Developments and Operation Phase of OSI SAF) and ESA (through the Climate Change Initiative Sea_Ice_cci project), and the German Research Foundation (DFG) Excellence Initiative CLISAP under Grant EXC 177/2. The publication contributes to the Cluster of Excellence 'CLICCS – Climate, Climatic Change, and Society' and to the Center for Earth System Research and Sustainability (CEN) of the University of Hamburg.

6   References

Alekseeva, T. Tikhonov, V., Frolov, S., Repina, I., Raev, M., Sokolova, J., Sharkov, E., Afanasieva, E., and Serovetnikov, S.: Comparison of Arctic sea ice concentration from the NASA Team, ASI, and VASIA2 algorithms with summer and winter ship data, Rem. Sens., 11(21), 2481, http://doi.org/10.3390/rs11212481, 2019.

Beitsch, A., Kern, S., and Kaleschke, L.: Comparison of SSM/I and AMSR-E sea ice concentrations with ASPeCt ship observations around Antarctica, IEEE Trans. Geosci. Rem. Sens., 53(4), 1985-1996, http://doi.org/10.1109/TGRS.2014.2351497, 2015.

Burgard, C., Notz, D., Pedersen, L. T., and Tonboe, R. T.: The Arctic Ocean observation operator for 6.9 GHz (ARC3O) – Part 2: Development and evaluation. The Cryosphere Disc., http://doi.org/10.5194/tc-2019-318, 2019.

Cavalieri D. J., Gloersen, P., and Campbell, W. J.: Determination of Sea Ice Parameters with the NIMBUS 7 SMMR. J. Geophys. Res., 89(D4), 5355-5369, 1984.

Cavalieri, D. J., J. Crawford, M. Drinkwater, W. J. Emery, D. T. Eppler, L. D. Farmer, M. Goodberlet, R. Jentz, A. Milman, C. Morris, R. Onstott, A. Schweiger, R. Shuchman, K. Steffen, C. T. Swift, C. Wackerman, and R. L. Weaver. 1992. NASA sea ice validation program for the DMSP SSM/I: final report. NASA Technical Memorandum 104559. National Aeronautics and Space Administration, Washington, D.C. 126 pages.

Cavalieri, D. J., Burns, B. A., and Onstott, R. G.: Investigation of the effects of summer let on the calculation of sea ice concentration using active and passive microwave data, J. Geophys. Res., 95(C4), 5359-5369, 1990.

Cavalieri, D. J., Parkinson, C. L., Gloersen, P., Comiso, J. C., and Zwally, H. J.: Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets, J. Geophys. Res., 104(C7), 15803–15814, http://doi.org/10.1029/1999JC900081, 1999.

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

Comiso, J. C.: Enhanced sea ice concentrations and ice extents from AMSR-E data, J. Rem. Sens. Soc. Japan, 29(1), 199-215, 2009.

Comiso, J. C. and Kwok, R.: Surface and radiative characteristics of the summer Arctic sea ice cover from multisensory satellite observation, J. Geophys. Res., 101(C12), 28,397-28,416, 1996.

Comiso, J. C., and Nishio, F.: Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data, J. Geophys. Res., 113, C02S07, http://doi.org/10.1029/2007JC004257, 2008.

Comiso, J. C., Cavalieri, D. J., Parkinson, C. L., and Gloersen, P.: Passive microwave algorithms for sea ice concentration: A comparison of two techniques, Rem. Sens. Environ., 60(3), 357-384, 1997.

Comiso, J. C., Cavalieri, D. J., and Markus, T.: Sea ice concentration, ice temperature, and snow depth, using AMSR-E data, IEEE Trans. Geosci. Rem. Sens., 41(2), 243-252, https://doi.org/10.1109/TGRS.2002.808317, 2003.

Comiso, J. C., Meier, W. N., and Gersten, R.: Variability and trends in the Arctic Sea ice cover: Results from different techniques, J. Geophys. Res.-Oceans, 122, 6883–6900, http://doi.org/10.1002/2017JC012768, 2017. Divine, D. V., Granskog, M. A., Hudson, S. R., Pedersen, C. A., Karlsen, T. I., Divina, S. A., Renner, A. H. H., Gerland, S.: Regional melt-pond fraction and albedo of thin Arctic first-year drift ice in late summer. The Cryosphere, 9(1), 255-268, http://doi.org/10.5194/tc-9-255-2015, 2015.

Dorn, W., Rinke, A., Köberle, C., Dethloff, K., and Gerdes, R.: Evaluation of the sea-ice simulation in the upgraded version of the coupled regional atmosphere-ocean-sea ice model HIRHAM-NAOSIM 2.0. Atmosphere, 10(8), 431, http://doi.org/10.3390/atmos10080431, 2019.

Ezraty, R., Girard-Arduin, F., Piollé, J.-F., Kaleschke, L., and Heygster, G.: Arctic and Antarctic sea ice concentration and Arctic sea ice drift estimated from special sensor microwave data – Users's Manual, Version 2.1, IFREMER, Brest, France, February 2007.

Flocco, D., Feltham, D. L., and Turner, A. K.: Incorporation of a physically based melt pond scheme into the sea ice component of a climate model. J. Geophys. Res., 115, C08012, http://doi.org/10.1029/2009JC005568, 2010.

Gogineni, S. P., Moore, R. K., Grenfell, T. C., Barber, D. G., Digby, S., and Drinkwater, M.: The effects of freeze-up and melt processes on microwave signatures. In: Microwave Remote Sensing of Sea Ice, ed.: F. Carsey, American Geophys. Union, pp. 329-341, 1992.

Grenfell, T. C., and Lohanick, A. W., Temporal variations of the microwave signatures of sea ice during the late spring and early summer near Mould Bay, Northwest Territories. J. Geophys., Res., 90(C3), 5063-5074, 1985.

Holland, M. M., Bailey, D. A., Briegleb, B. P., Light, B., and Hunke, E.: Improved sea ice shortwave radiation physics in CCSM4: The impact of melt ponds and aerosols on Arctic sea ice. J. Climate, 25(5), 1413-1430, http://doi.org/10.1175/JCLI-D-11-00078.s1, 2012.

Hutchings, J. K., and Orlich, A. R.: A standardized protocol and database for discrete visual observations of Arctic sea ice. Earth System Sci. Data, submitted, 2018.

Istomina, L., Heygster, G., Huntemann, M., Marks, H., Melsheimer, C., Zege, E., Malinka, A., Prikhach, A., and Katsev, I.: Melt pond fraction and spectral sea ice albedo retrieval from MERIS data – Part 2: Case studies and trends of sea ice albedo and melt ponds in the Arctic for years 2002–2011, The Cryosphere, 9, 1567-1578, http://doi.org/10.5194/tc-9-1567-2015 , 2015.

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, http://doi.org/10.5194/tc-9-1797-2015, 2015.

Ivanova, N., Johannessen, O. M., Pedersen, R. T., and Tonboe, R. T.: Retrieval of Arctic sea ice parameters by satellite passive microwave sensors: A comparison of eleven sea ice concentration algorithms, IEEE Trans. Geosci. Rem. Sens., 52(11), 7233-7246, http://doi.org/10.1109/TGRS.2014.2310136, 2014.

Kaleschke, L., Lüpkes, C., Vihma, T., Haarpaintner, J., Bochert, A., Hartmann, J., and Heygster, G.: SSM/I sea ice remote sensing for mesoscale ocean-atmosphere interaction analysis, Can. J. Rem. Sens., 27(5), 526-537, 2001.

Kern, S., Lavergne, T., Notz, D., Pedersen, L., Tonboe, R., Saldo, R., and Sørensen, A. M.: Satellite Passive Microwave Sea-Ice Concentration Data Set Intercomparison: Closed Ice and Ship-Based Observations, The Cryosphere, 13(12), 3261-3307, http://doi.org/10.5194/tc-13-3261-2019, 2019.

Kern, S., Rösel, A., Pedersen, L. T., Ivanova, N., Saldo, R., and Tonboe, R. T.: The impact of melt ponds on summertime microwave brightness temperatures and sea-ice concentrations, The Cryosphere, 10, 2217-2239, http://doi.org/10.5194/tc-10-2217-2016, 2016.

Landy, J., Ehn, J., Shields, M., and Barber, D.: Surface and melt pond evolution on landfast first-year sea ice in the Canadian Arctic Archipelago, J. Geophys. Res. - Oceans, 119, 3054-3075, http://doi.org/10.1002/2013JC009617, 2014.

Lavergne, T., Macdonald Sørensen, A., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjær, Eastwood, S., Gabarro, C., Heygster, G., Killie, M., Brandt Kreiner, M., Lavelle, J., Saldo, R., Sandven, S., and Pedersen, L.: Version 2 of the EUMETSAT OSI SAF and ESA-CCI sea-ice concentration climate data records, The Cryosphere, 13(1), 49–78, http://doi.org/10.5194/tc-13-49-2019, 2019.

Light, B., Dickinson, S., Perovich, D. K., and Holland, M. M.: Evolution of Arctic sea ice albedo in CCSM4 simulations: Episodic summer snowfall and frozen summers. J. Geophys. Res.-Oceans, 120(1), 284-303, http://doi.org/10.1002/2014JC010149, 2015.

Marks, H.: Investigation of algorithms to retrieve melt pond fraction on Arctic sea ice from optical satellite observations, Master Thesis, Institute of Astronomy & Astrophysics Computational Physics, Eberhard Karls Universität Tübingen and Institute of Environmental Physics, Universität Bremen, 82 pp., June 2015.

Markus, T., and Cavalieri, D. J.: An enhancement of the NASA Team sea ice algorithm, IEEE Trans. Geosci. Rem. Sens., 38(3), 1387-1398, 2000.

Markus, T., and Cavalieri, D. J.: The AMSR-E NT2 sea ice concentration algorithm: its basis and implementation. J. Rem. Sens. Soc. Japan, 29(1), 216-225, 2009.

Markus, T., Stroeve, J. C., and Miller, J.: Recent changes in Arctic sea ice melt onset, freezeup, and melt season length. J. Geophys. Res., 114, C12024, http://doi.org/10.1029/2009JC005436, 2009.

Meier, W. N.: Comparison of passive microwave ice concentration algorithm retrievals with AVHRR imagery in Arctic peripheral seas, IEEE Trans. Geosci. Rem. Sens., 43(6), 1324-1337, http://doi.org/10.1109/TGRS.2005.846151, 2005.

Meier, W. N., and Windnagel, A.: Sea ice concentration – climate algorithm theoretical basis document, NOAA Climate Data Record Program CDRP-ATBD-0107 Rev. 7 (03/06/2018), available at https://www.ncdc.noaa.gov/cdr/oceanic/sea-ice-concentration, 2018.

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

Melia, N., Haines, K., Hawkins, E., and Day, J. J.: Towards seasonal Arctic shipping route predictions, Environ. Res. Lett., 12(8), 084005, http://doi.org/10.1088/1748-9326/aa7a60, 2017.

Notz, D.: Sea-ice extent and its trend provide limited metrics of model performance, The Cryosphere, 8(1), 229-243, http://doi.org/10.5194/tc-8-229-2014, 2014.

830 Pedersen, C. A., Roeckner, E., Lüthje, M., and Winther, J.-G.: A new sea ice albedo scheme including melt ponds for ECHAM5 general circulation model. J. Geophys. Res., 114, D08101, http://doi.org/10.1029/2008JD010440, 2009.

Peng, G., and Meier, W. N.: Temporal and regional variability of Arctic sea-ice coverage from satellite data, Ann. Glaciol., 59(76, part 2), 191-200, http://doi.org/10.1017/aog.2017.32, 2018.

Peng, G., Meier, W. N., Scott, D., and Savoie, M.: A long-term and reproducible passive microwave sea ice concentration data
835 record for climate studies and monitoring, *Earth Syst. Sci. Data*. 5. 311-318. http://doi.org/10.5194/essd-5-311-2013, 2013.

Petty, A. A., Stroeve, J. C., Holland, P. R., Boisvert, L. N., Bliss, A. C., Kimura, N., and Meier, W. N.: the Arctic sea ice cover 2016: a year of record-low highs and higher-than-expected lows, The Cryosphere, 12(2), 433-453, http://doi.org/10.5194/tc-12-433-2018, 2018.

Popović, P., and Abbot, D.: A simple model for the evolution of melt pond coverage on permeable Arctic sea ice. The
840 Cryosphere, 11(3), 1149-1172, http://doi.org/10.5194/tc-11-1149-2017, 2017.

Riihela, A., Key, J. R., Meirink, J. F., Kuipers-Munneke, P., Palo, T., and Karlsson, K.-G.: An intercomparison and validation of satellite-based surface radiative energy flux estimates over the Arctic. J. Geophys. Res.-Atmos., 122(9), 4829-4848, http://doi.org/10.1002/2016JD026443, 2017.

Riihela, A., Laine, V., Manninen, T., Palo, T., and Vihma, T.: Validation of the Climate-SAF surface broadband albedo
845 product: Comparisons with in situ observations over Greenland and the ice-covered Arctic Ocean. Rem. Sens. Environ. 114, 2779-2790, 2010.

Rösel, A., and Kaleschke, L.: Comparison of different retrieval techniques for melt ponds on Arctic sea ice from Landsat and MODIS satellite data. Ann. Glaciol., 52, 185–191, 2011.

Rösel, A., Kaleschke, L., and Birnbaum, G.: Melt ponds on Arctic sea ice determined from MODIS satellite data using an
850 artificial neural network. The Cryosphere, 6, 431-446, http://doi.org/10.5194/tc-6-431-2012, 2012a.

Rösel, A., Kaleschke, L., and Kern, S.: Influence of melt ponds on microwave sensor's sea ice concentration retrieval algorithms. IGARSS 2012, Munich, Germany, July 23-27, 2012b.

Scott, F., and Feltham, D. L.: A model of the three-dimensional evolution of Arctic melt ponds on first-year and multiyear sea ice. J. Geophys. Res., 115, C12064, http://doi.org/10.1029/2010JC006156, 2010.

855 Skyllingstad, E. D., Shell, K. M., Collins, L., and Polashenski, C.: Simulation of the melt season using a resolved sea ice model with snow cover and melt ponds. J. Geophys. Res.-Oceans, 120(7), 5194-5215, http://doi.org/10.1002/2014JC010569, 2015.

Spreen, G., Kaleschke, L., and G. Heygster, G.: Sea ice remote sensing using AMSR-E 89-GHz channels, J. Geophys. Res., 113, C02S03, http://doi.org/10.1029/2005JC003384, 2008.

Steffen, K. and Schweiger, A.: NASA team algorithm for sea ice concentration retrieval from Defense Meteorological Satellite
860 Program special sensor microwave imager: comparison with Landsat satellite data, J. Geophys. Res., 96(C12), 21,971-21,987, 1991.

Stroeve, J. C., Markus, T., Boisvert, L., Miller, J., and Barrett, A.: Changes in Arctic melt season and implications for sea ice loss. Geophys. Res. Lett., 41(4), 1216-1225, http://doi.org/10.1002/2013GL058951, 2014.

Tilling, R. L., Ridout, A., and Shepherd, A.: Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter
865 data. Adv. Space Res., 62(6), 1203-1225, http://doi.org/10.1016/j.asr.2017.10.051, 2018.

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

Tsamados, M., Feltham, D., Petty, A., Schroeder, D., and Flocco, D.: Processes controlling surface, bottom and lateral melt of
870 Arctic sea ice in a state of the art sea ice model. Phil. Trans. R. Soc. A, 373, 20140167, http://doi.org/10.1098/rsta.2014.0167, 2015.

Ulaby, F. T., Moore, R. K., and Fung, A. K.: Microwave remote sensing, active and passive. Volume III: From theory to applications, Addison Wesley Pub. London, U.K., 1986.

Wang, Q., Li, Z., Lu, P., Lei, R., and Cheng, B.: 2014 summer Arctic sea ice thickness and concentration from shipborne
875 observations. Int. J. Digi. Earth, Jan. 5 2018, http://doi.org/10.1080/17538947.2017.142.1720, 2018.

Webster, M. A., Rigor, I. G., Perovich, D. K., Richter-Menge, J. A., Polashenski, C. M., and Light, B.: Seasonal evolution of melt ponds on Arctic sea ice, J. Geophys. Res.-Oceans, 120, 5968-5980, http://doi.org//10.1002/2015JC011030, 2015.

Worby, A. P., Geiger, C. A., Paget, M. J., Van Woert, M. L., Ackley, S. F., and DeLiberty, T. L.: The thickness distribution of Antarctic sea ice. J. Geophys. Res., 113, C05S92, http://doi.org/10.1029/2007JC004254, 2008.

880 Xie, H., Lei, R., Ke, C., Wang, H., Li, Z., Zhao, J., and Ackley, S. F.: Summer sea ice characteristics and morphology in the Pacific Arctic sector as observed during the CHINARE 2010 cruise. The Cryosphere, 7(4), 1057-1072, http://doi.org/10.5194/tc-7-1057-2013, 2013.

Zege, E., Malinka, A., Katsev, I., Prikhach, A., Heygster, G., Istomina, L., Birnbaum, G., and Schwarz, P.: Algorithm to retrieve the melt pond fraction and the spectral albedo of Arctic summer ice from satellite optical data. Rem. Sens. Environ.,
885 163, 153-164, http://doi.org/10.1016/j.rse.2015.03.012, 2015.

Zhang, J., Schweiger, A., Webster, M., Light, B., Steele, M., Ashjian, C., Campbell, R., and Spitz, Y.: Melt pond conditions on declining Arctic sea ice over 1979-2016: Model development, validation, and results. J. Geophys. Res.-Oceans, 123, 7983-8003, http://doi.org/10.1029/2018JC014298, 2018.

7   Tables

Table 1. Overview of the investigated sea-ice concentration products. Column "ID (Algorithm)" holds to the identifier we use henceforth to refer to the data product, and which algorithm it uses. Group is an identifier for the algorithm concept used. Column "Input data" refers to the input satellite data for the data set. Columns "Tie points" and "Tie-point Update" refers to the type of tie points used and their update interval (see text for further details).

ID (algorithm)	Group	Input data & frequencies	Grid resolution & type	Tie points	Tie-point update	Reference
OSI-450 (SICCI2)	I	SMMR, SSM/I, SSMIS 19.35 & 37.0 GHz	25 km x 25 km EASE2.0	Open water, ice line	Daily	Tonboe et al., 2016; Lavergne et al., 2019
SICCI-12km (SICCI2)	I	AMSR-E, AMSR2 18.7 & 89.0 GHz	12.5 km x 12.5 km EASE2.0	Open water, ice line	Daily	Lavergne et al., 2019
SICCI-25km (SICCI2)	I	AMSR-E, AMSR2 18.7 & 36.5 GHz	25 km x 25 km EASE2.0	Open water, ice line	Daily	Lavergne et al., 2019
SICCI-50km (SICCI2)	I	AMSR-E, AMSR2 6.9 & 36.5 GHz	50 km x 50 km EASE2.0	Open water, ice line	Daily	Lavergne et al., 2019
CBT-SSMI (Comiso bootstrap)	II	SMMR, SSM/I, SSMIS 19.35 & 37.0 GHz	25 km x 25 km PolarStereo	Open water, ice line	Daily	Comiso, 1986; Comiso et al., 1997; Comiso and Nishio, 2008
NOAA-CDR (NASA Team & Comiso bootstrap)	II	SSM/I, SSMIS 19.35 & 37.0 GHz	25 km x 25 km PolarStereo	Open water, ice line & Open water, first-year ice, multiyear ice	Daily & fixed	Peng et al., 2013; Meier and Windnagel, 2018
CBT-AMSR-E (Comiso bootstrap)	II	AMSR-E 18.7 & 36.5 GHz	25 km x 25 km PolarStereo	Open water, ice line	Daily	Comiso et al., 2003; Comiso and Nishio, 2008; Comiso, 2009
ASI-SSMI (ASI)	III	SSM/I, SSMIS 85.5 GHz	12.5 km x 12.5 km PolarStereo	Open water, sea ice	Fixed	Kaleschke et al., 2001; Ezraty et al., 2007
NT1-SSMI (NASA-Team)	III	SMMR, SSM/I, SSMIS 19.35 & 37.0 GHz	25 km x 25 km PolarStereo	Open water, first-year ice, multiyear ice	Fixed	Cavalieri et al, 1984; 1992; 1999
NT2-AMSR-E (NASA-Team-2)	IV	AMSR-E 18.7, 36.5 & 89.0 GHz	25 km x 25 km PolarStereo	Open water, thin ice, ice type A, ice type C	Daily*	Markus and Cavalieri, 2000; 2009

Table 2. Overall mean difference: individual algorithm SIC minus ensemble mean SIC in percent ice concentration for the Arctic for winter (months January and February) and summer (months July (see Figure 3) and August). N denotes the total number of valid data pairs with SIC > 15.0% (see also Kern et al., 2019).

| | N | Group I | | | | Group II | | | Group III | | IV |
		SICCI12	SICCI25	SICCI50	OSI450	CBT-SSMI	NOAA-CDR	CBT-AMSRE	ASI-SSMI	NT1-SSMI	NT2-AMSRE
Jan/Feb	9821	-1.3	-1.0	-1.2	-0.6	+2.2	+2.2	+2.5	-2.8	-2.7	+2.7
July/Aug	5698	-2.7	-3.6	-4.0	-3.0	+5.6	+5.9	+5.8	+1.2	-8.1	+3.0

Table 3. Average values of linear correlation, root-mean-squared difference (RMSD), and slope as well as intercept of the linear regression between passive microwave and MODIS sea-ice concentration for product groups I to IV (see text for further information). The averages are derived as the arithmetic mean from all 8-day period values of products within one group falling into "pre-melt": DOY 129, 137, and 145, "melt advance": DOY 153 to 185, "peak melt": DOY 193 to 233, and "end-of-melt": DOY 241 and 249.

| parameter | correlation | | | | RMSD [%] | | | | slope | | | | intercept [%] | | | |
group	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
pre-melt	0.39	0.54	0.53	0.47	5.8	4.5	5.6	4.6	0.77	0.74	1.05	0.71	21.8	27.4	-6.5	29.9
melt advance	0.65	0.61	0.62	0.61	7.4	7.3	8.2	6.5	1.03	0.75	0.99	0.78	-5.4	26.0	-2.8	22.1
peak melt	0.82	0.80	0.83	0.74	8.1	11.4	9.2	10.1	1.27	1.03	1.26	1.09	-23.5	1.2	-23.8	-2.3
end of melt	0.78	0.75	0.81	0.61	9.5	11.7	12.0	11.7	1.16	1.05	1.31	0.84	-16.3	2.8	-30.7	18.1

Table 4. Mean values of linear correlation, root-mean-squared difference (RMSD), and slope as well as intercept of the linear regression between passive microwave sea-ice concentration and MODIS ice-surface fraction for product groups I to IV (see text and caption of Table 3 for further information).

parameter	correlation				RMSD [%]				slope				intercept [%]			
group	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
pre-melt	0.38	0.52	0.51	0.48	7.9	7.9	7.1	7.9	0.38	0.37	0.54	0.39	59.9	64.8	44.0	62.0
melt advance	0.72	0.62	0.70	0.62	15.8	21.1	15.9	20.0	0.59	0.40	0.58	0.43	47.0	65.1	44.7	61.2
peak melt	0.80	0.75	0.81	0.73	24.5	33.0	23.2	30.5	0.86	0.71	0.85	0.74	30.3	48.6	29.4	43.6
end of melt	0.75	0.69	0.78	0.59	18.3	26.5	17.8	25.4	0.69	0.59	0.77	0.49	34.4	50.5	26.1	55.4

910

8   Figures

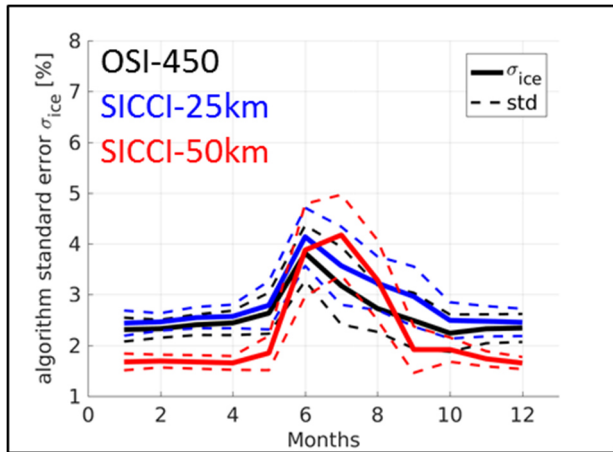
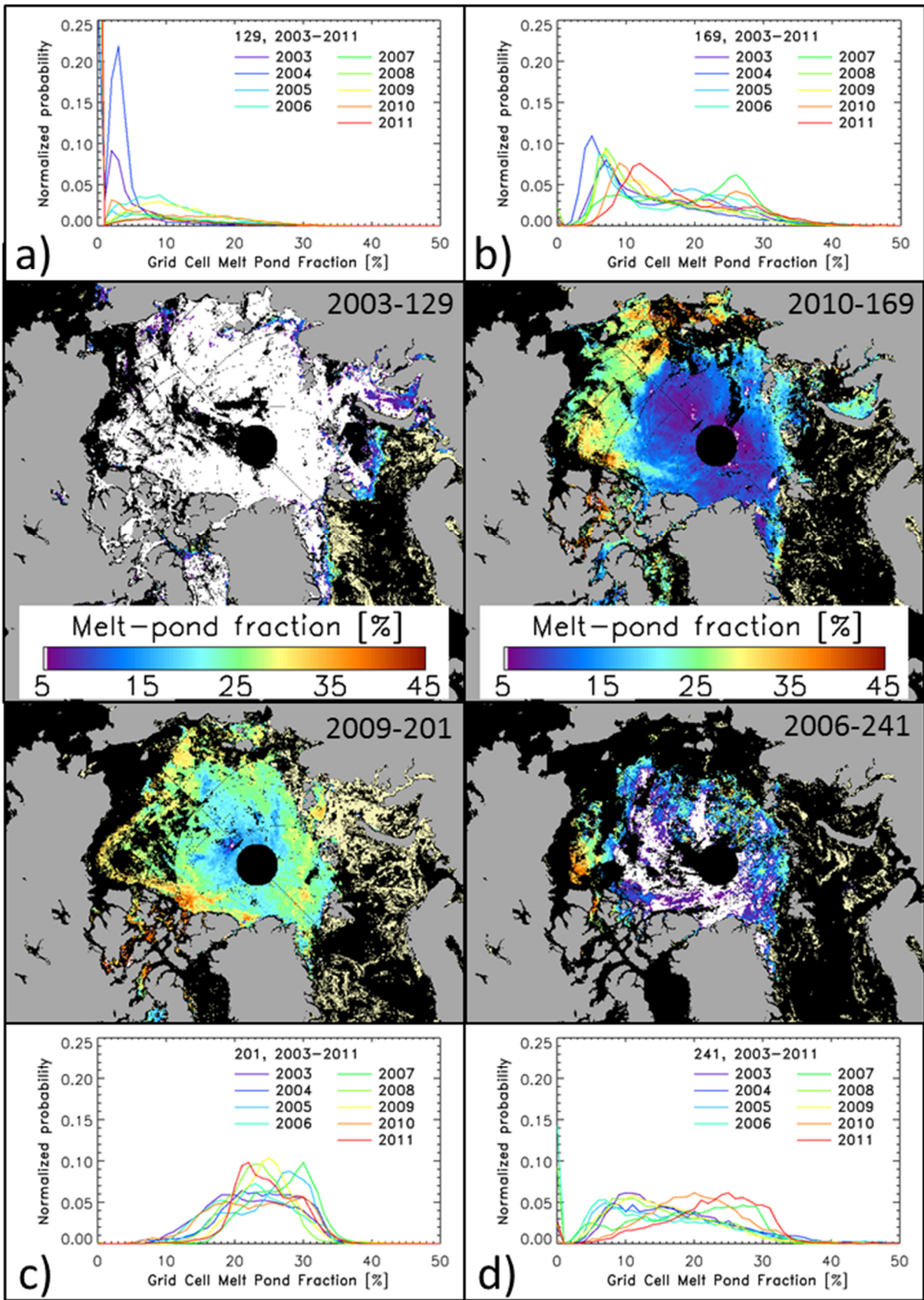



Figure 1. Seasonal cycle of the multi-annual (2002-2011) average sea-ice concentration algorithm standard error for the Arctic for grid cells
915 with more than 90 % sea-ice concentration for OSI-450, SICCI-25km and SICCI-50km products. Shown are the mean (solid line) and its standard deviation (dashed line denoted "std").

920 **Figure 2.** Sample maps of the Arctic melt-pond fraction from the MODIS data set for a) day-of-year (DOY) 129 (May 9-16) 2003, b) DOY 169 (June 18-25) 2010, c) DOY 201 (July 20-27) 2009, and d) DOY 241 (Aug. 29-Sep. 5) 2006, illustrating the conditions during pre-melt, melt advance, peak melt, and end of melt, respectively. Black color denotes open water, missing (note in this context the curvilinear one-grid-cell wide features with missing data which originate from the gridding process) or invalid data and clouds. Melt-pond fractions smaller than 5 % are displayed white. The histograms show the distribution of the melt-pond fraction for above-quoted DOY for every year of the 925 period 2003-2011.

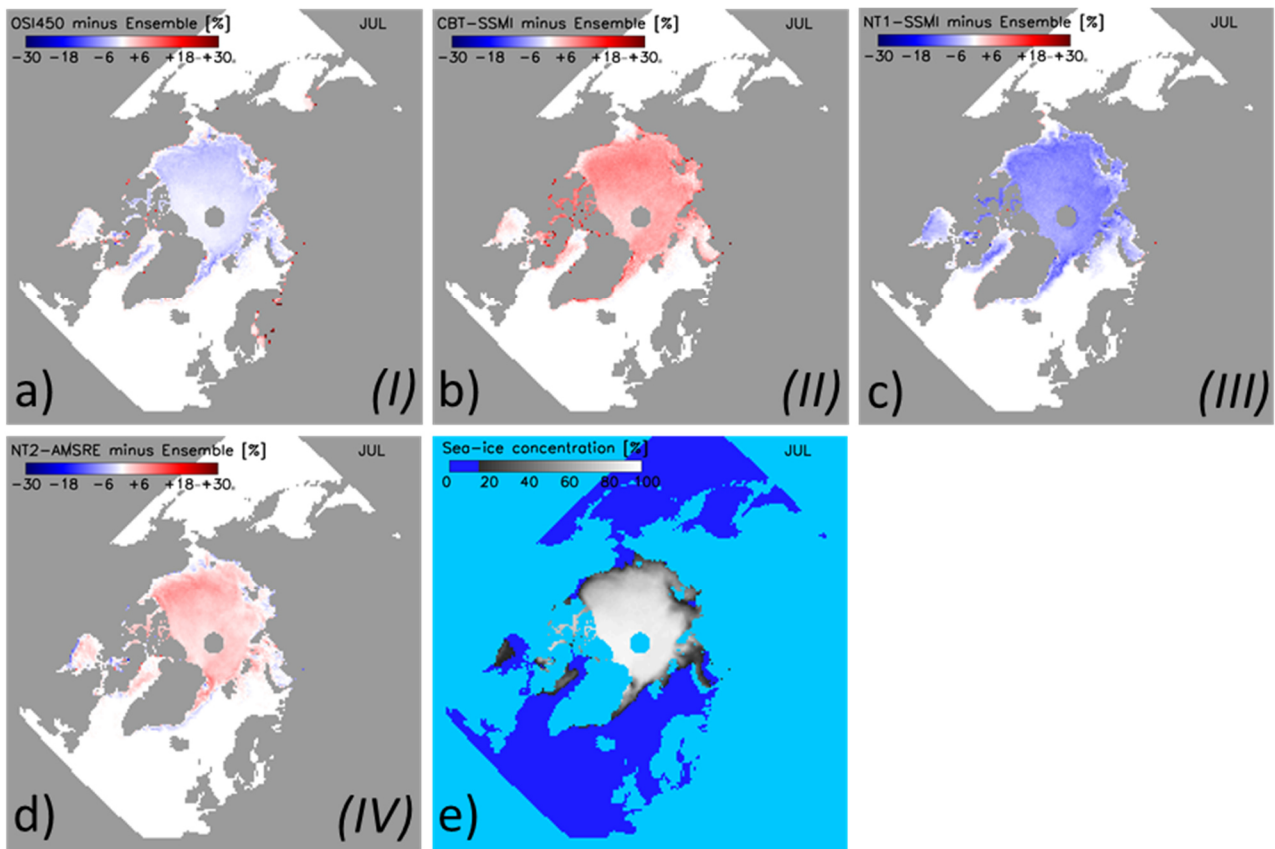



Figure 3. (a) to **(d)** Maps of the difference between the multi-annual average monthly SIC of the individual algorithms representing groups
I to IV and the 10-algorithm ensemble median multi-annual average monthly SIC **(e)** for the Arctic for July 2003-2011 (see also Fig. S04,
supplementary material). Differences are only computed for sea-ice concentration of both data sets larger than 15%. Roman numbers in bold
font denote the group (see Table 1) to which the algorithm is assigned.

935

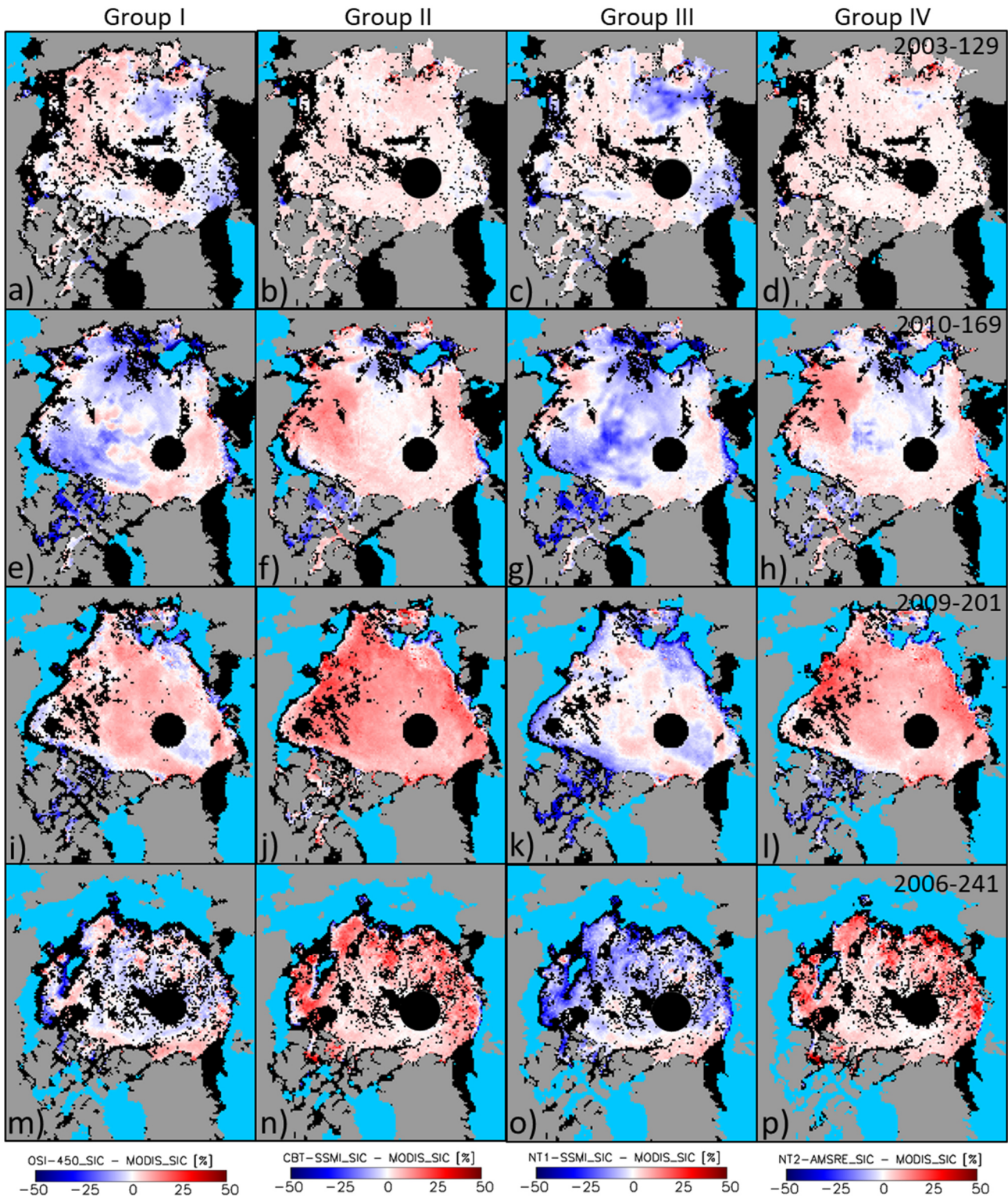



Figure 4. Maps of the difference PMW minus MODIS sea-ice concentration for a) to d): day-of-year (DOY) 129 (May 9-16) 2003, e) to h): DOY 169 (June 18-25) 2010, i) to l): DOY 201 (July 20-27) 2009, and m) to p): DOY 241 (Aug. 29-Sep. 5) 2006; these are the same periods as used in Fig. 2. The leftmost column shows OSI-450, representing group I, the second column CBT-SSMI, representing group II, the third column NT1-SSMI, representing group III, and the rightmost column NT2-AMSRE, representing group IV. Black areas denote invalid or missing data, clouds, or grid cells being ice-covered but not considered further in the analysis, e.g. in the Greenland Sea or Hudson Bay. Row starting with (a) are representative of pre-melt, row starting with (e) is melt advance, row starting with (i) is at the peak of melt, and row starting with (m) is at the end-of-melt.

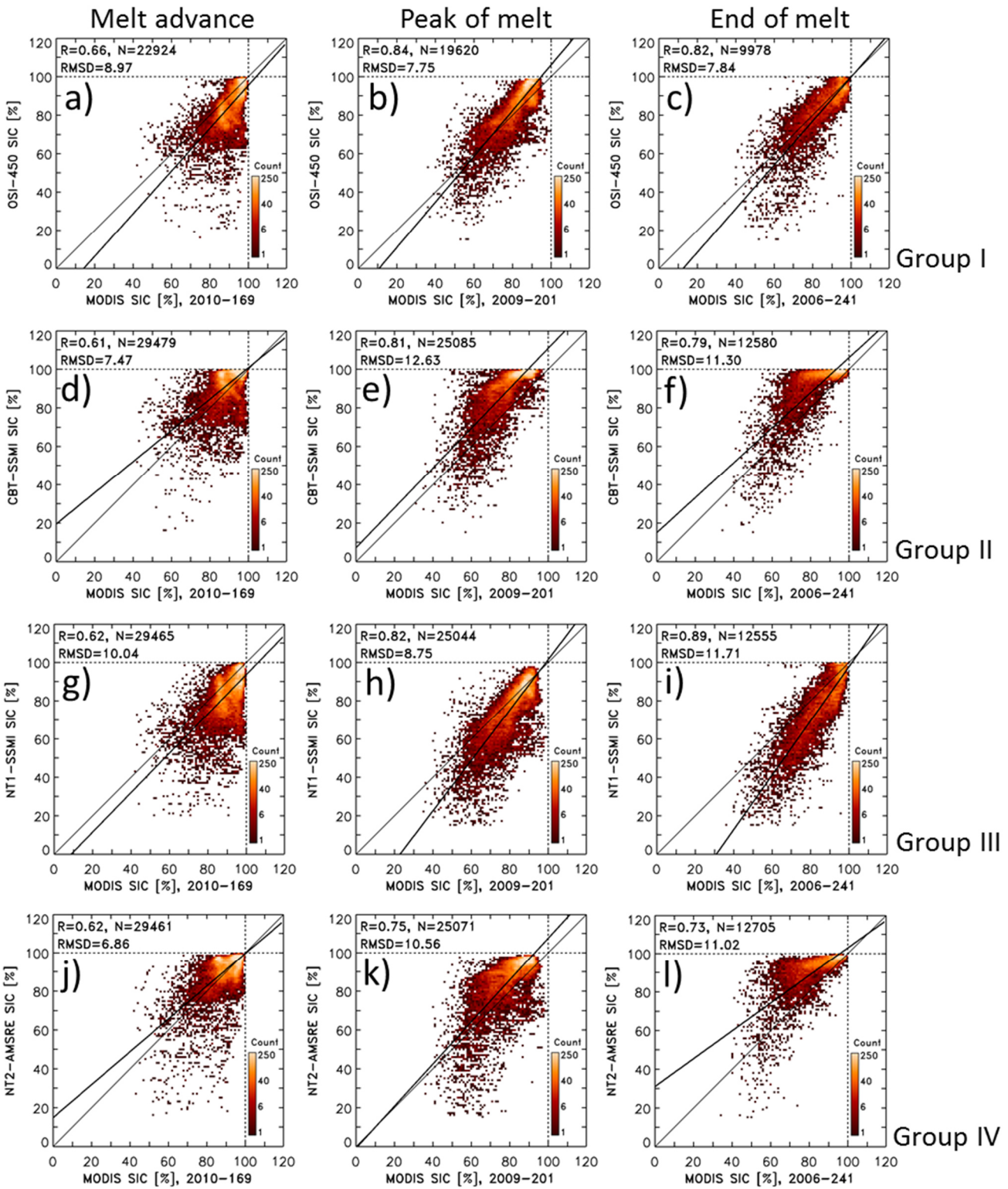



Figure 5. Two-dimensional histograms of the distribution of PMW (y-axis) versus MODIS (x-axis) SIC data pairs using a bin size of 1 % for the same 8-day periods as shown in Fig. 4e-p, i.e. melt advance, peak melt, and end of melt. The topmost row shows OSI-450 (for group I), the second row CBT-SSMI (for group II), the third row NT1-SSMI (for group III), and the bottommost row NT2-AMSRE (group IV). The thin black line is the identity line. The thick black line denotes the linear regression through the data pairs. At the top left of every image we display the linear correlation coefficient R, the number of data pairs N and the root mean squared difference RMSD; the latter is given in percent. The leftmost, middle and rightmost columns are representing melt advance, peak of melt, and end of the melt, respectively. Respective scatterplots for pre-melt are shown in Figure S06 in the supplementary material.

950

955

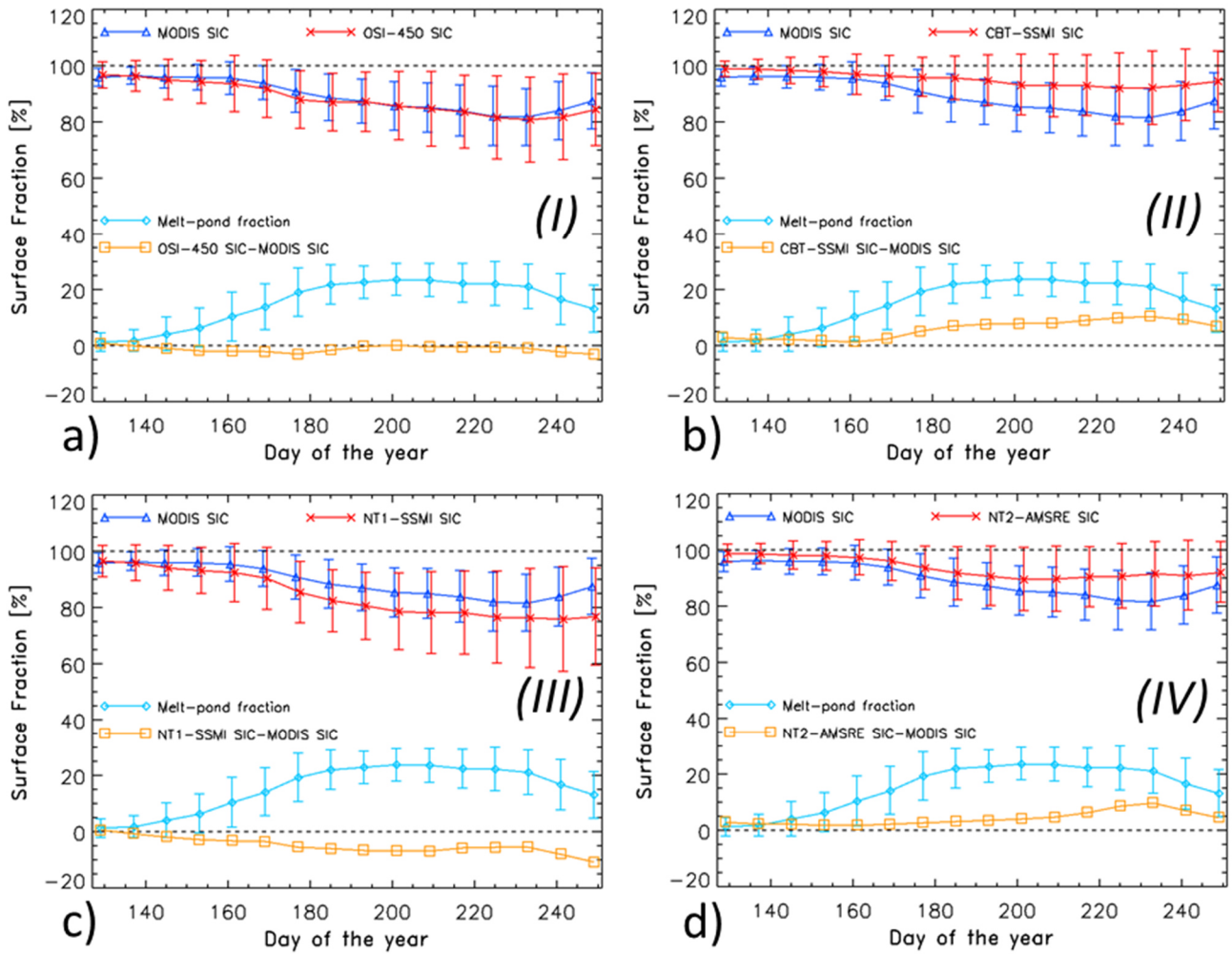



Figure 6. Average seasonal cycle of the mean (limited to Arctic Ocean and Canadian Arctic Archipelago) MODIS SIC (in blue), PMW SIC (in red), their difference PMW minus MODIS SIC (in orange), and the MODIS melt-pond fraction (in cyan), averaged for each 8-day period for the years 2003-2011 of the PMW products representing groups I to IV (see also Fig. S07, supplementary material). Error bars denote one standard deviation of the mean. Roman numbers in bold font denote the group (see Table 1) to which the algorithm is assigned.

960

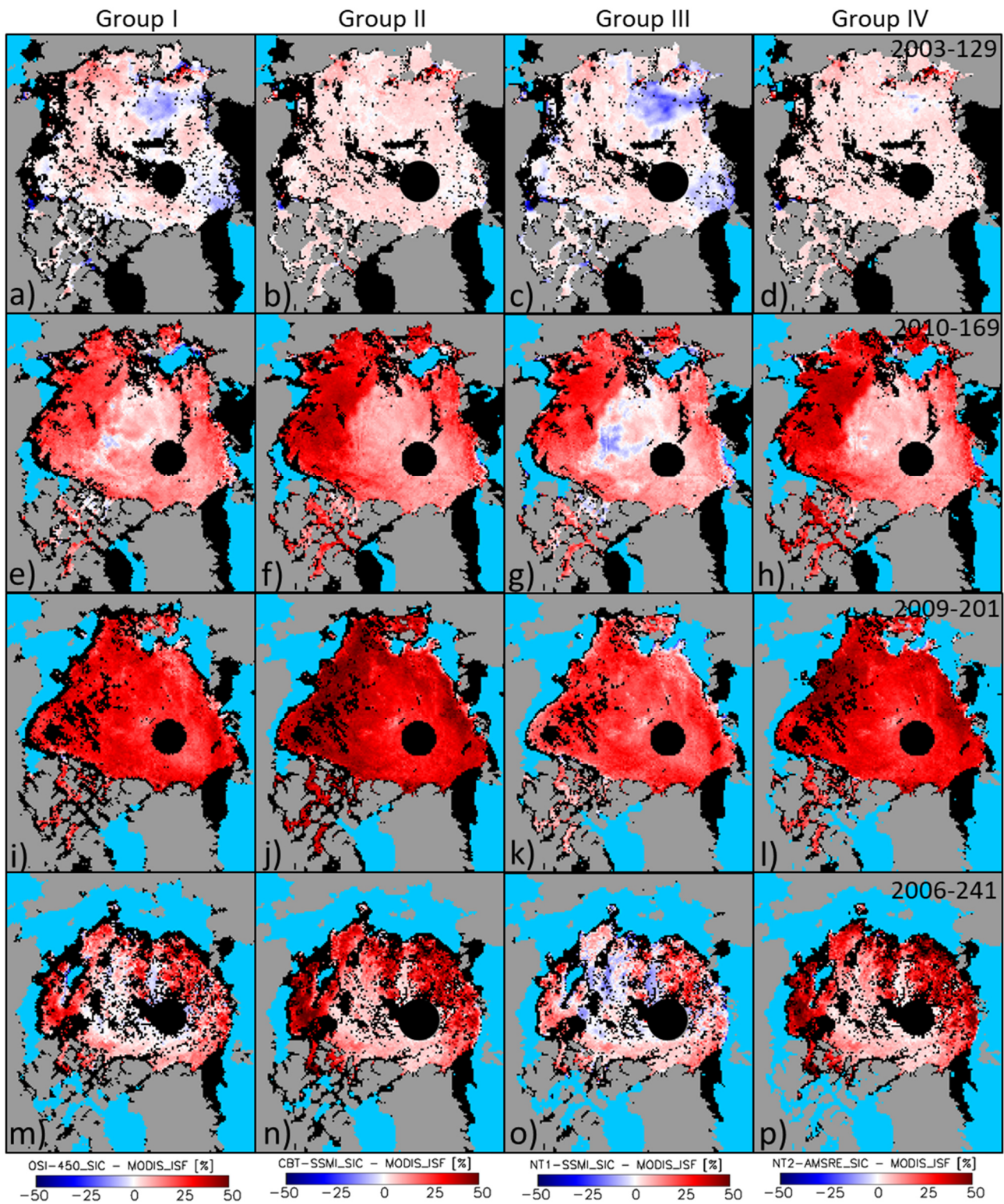



Figure 7. Maps of the difference PMW sea-ice concentration minus MODIS ice surface fraction (ISF) for the same 8-day periods as shown in Fig. 4. The leftmost column shows OSI-450, representing group I, the second column CBT-SSMI, representing group II, the third column NT1-SSMI, representing group III, and the rightmost column NT2-AMSRE, representing group IV. Black areas denote invalid or missing data, clouds, or grid cells being ice-covered but not considered further in the analysis, e.g. in the Greenland Sea or Hudson Bay. Row starting with (a) are representative of pre-melt, row starting with (e) is melt advance, row starting with (i) is at the peak of melt, and row starting with (m) is at the end of the melt.

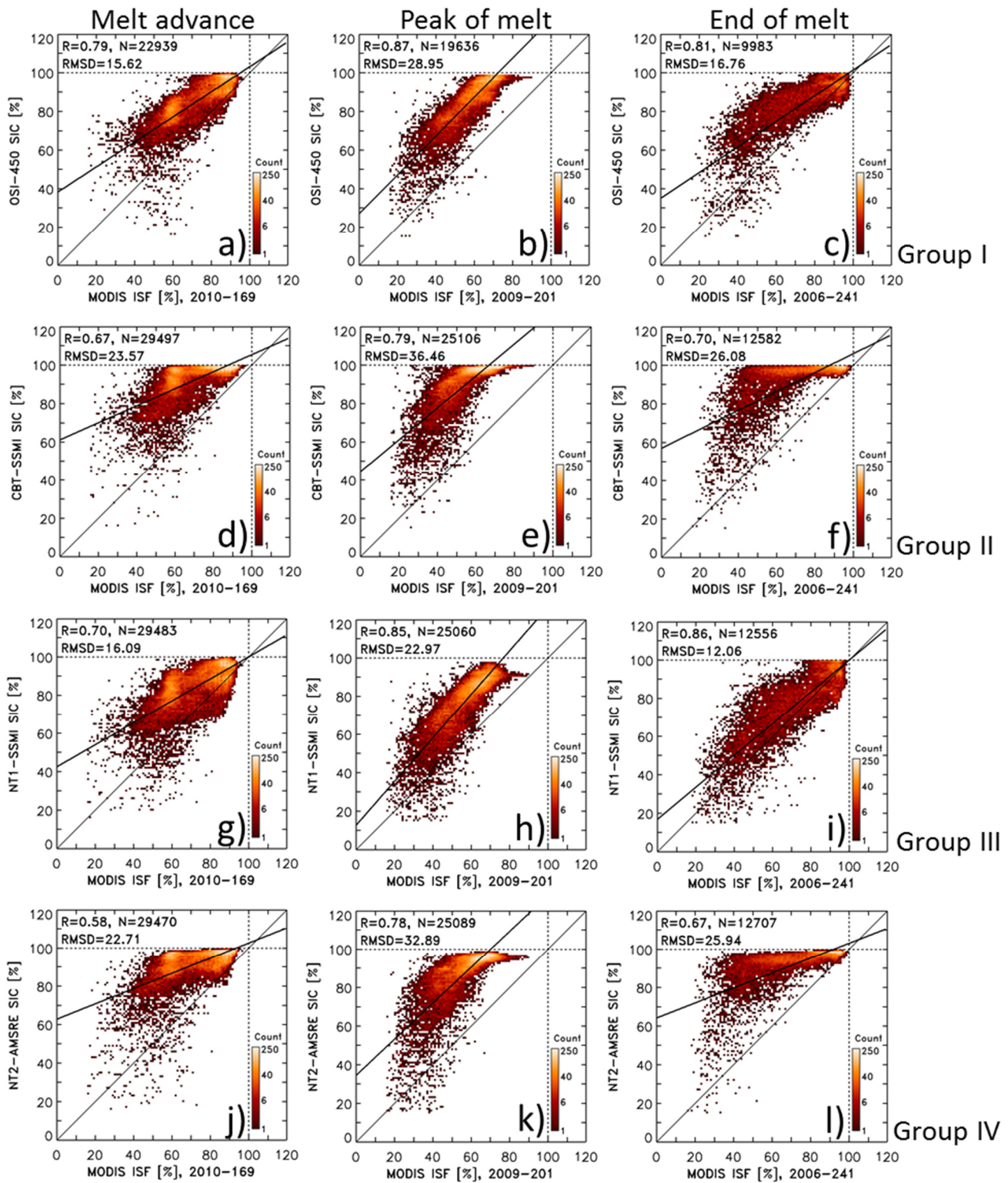
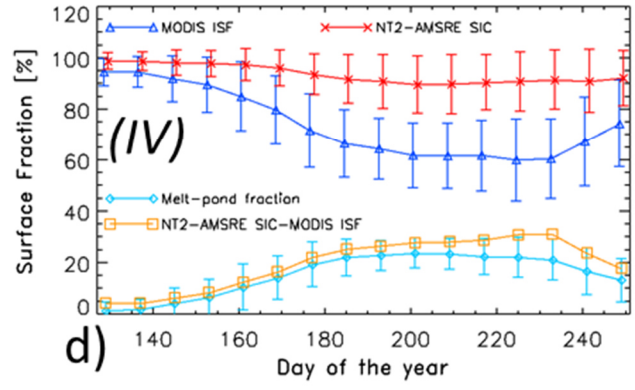
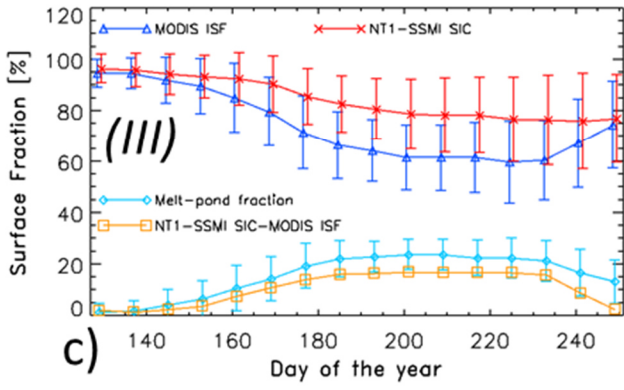
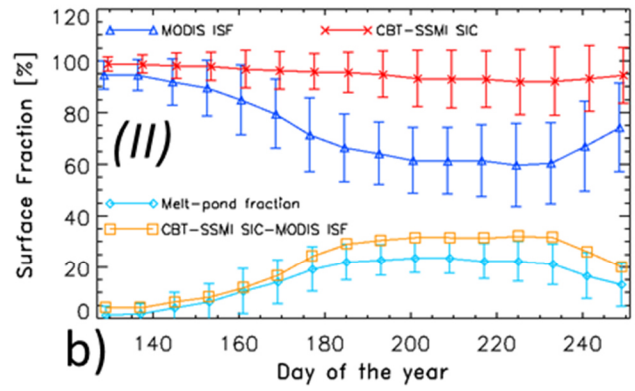
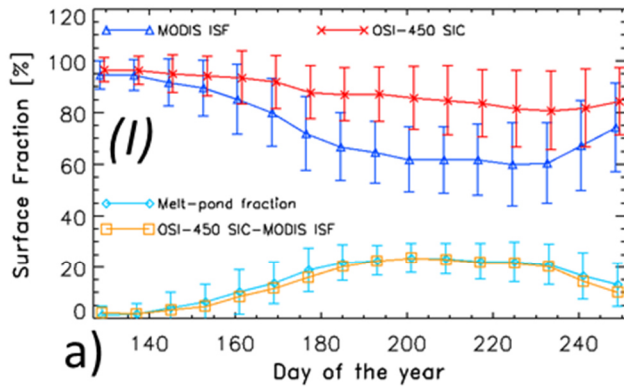



Figure 8. Two-dimensional histograms of the distribution of PMW SIC (y-axis) versus MODIS ISF (x-axis) data pairs using a bin size of 1 % for the same 8-day periods as shown in Fig. 7e-p, i.e. melt advance, peak melt, and end of melt. The topmost row shows OSI-450 (for group I), the second row CBT-SSMI (for group II), the third row NT1-SSMI (for group III), and the bottommost row NT2-AMSRE (group IV). The thin black line is the identity line. The thick black line denotes the linear regression through the data pairs. At the top left of every image we display the linear correlation coefficient R, the number of data pairs N and the root mean squared difference RMSD; the latter is given in percent. The leftmost, middle and rightmost columns are representing melt advance, peak of melt, and end of the melt, respectively. Respective scatterplots for pre-melt are shown in Figure S09 in the supplementary material.

985 **Figure 9.** Average seasonal cycle of the mean (limited to Arctic Ocean and Canadian Arctic Archipelago) MODIS ISF (in blue), PMW SIC (in red), their difference PMW SIC minus MODIS ISF (in orange), and the MODIS melt-pond fraction (in cyan), averaged for each 8-day period over the years 2003-2011 of the PMW products representing groups I to IV (see also Fig. S10 in the supplementary material). Error bars denote one standard deviation of the mean. Roman numbers in bold font denote the group (see Table 1) to which the algorithm is assigned.

990

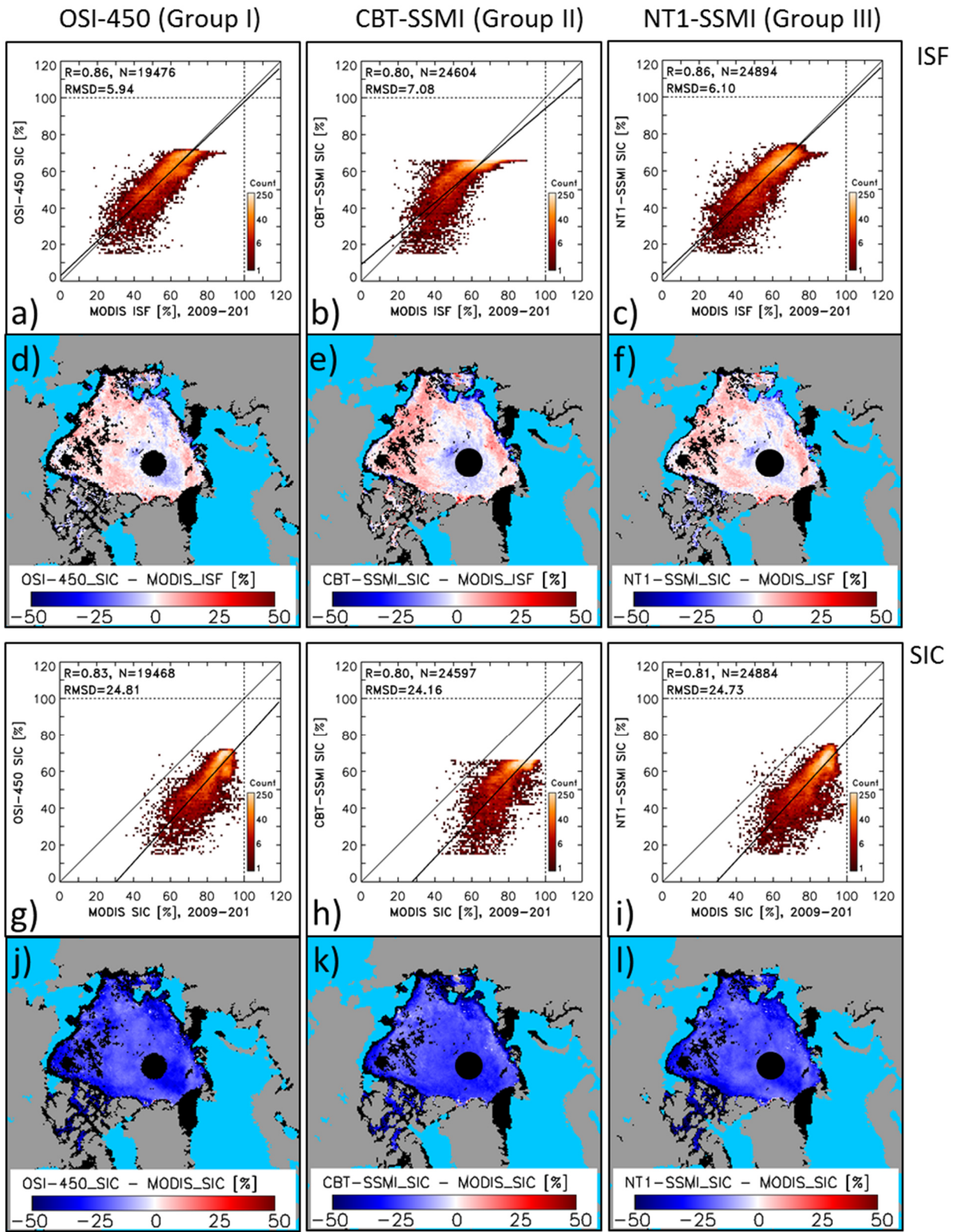



Figure 10. Illustration of the effect of a simple linear bias correction of PMW SIC towards MODIS ISF for an 8-day period during peak-melt (DOY 201, July 20-27, 2009). Topmost row, panels a) to c): Two-dimensional histograms of the distribution of bias-corrected PMW SIC (y-axis) versus MODIS ISF (x-axis) data pairs. Second row, panels d) to f): Respective maps of the difference of bias-corrected PMW SIC minus MODIS ISF. Third row, panels g) to i) Two-dimensional histograms of the distribution of bias-corrected PMW SIC (y-axis) versus MODIS SIC (x-axis) data pairs. Bottommost row, panels j) to l): Respective maps of the difference bias-corrected PMW SIC minus MODIS SIC. Leftmost, middle and rightmost columns show OSI-450 (for group I), CBT-SSMI (for group II), and NT1-SSMI (for group III). Bin size in the histograms is 1 %. The quantities given in the top left corner are R: linear correlation coefficient, N: number of valid data pairs, and RMSD: root mean squared difference. The thin black line is the identity line; the thick black line denotes the linear regression through the data pairs.

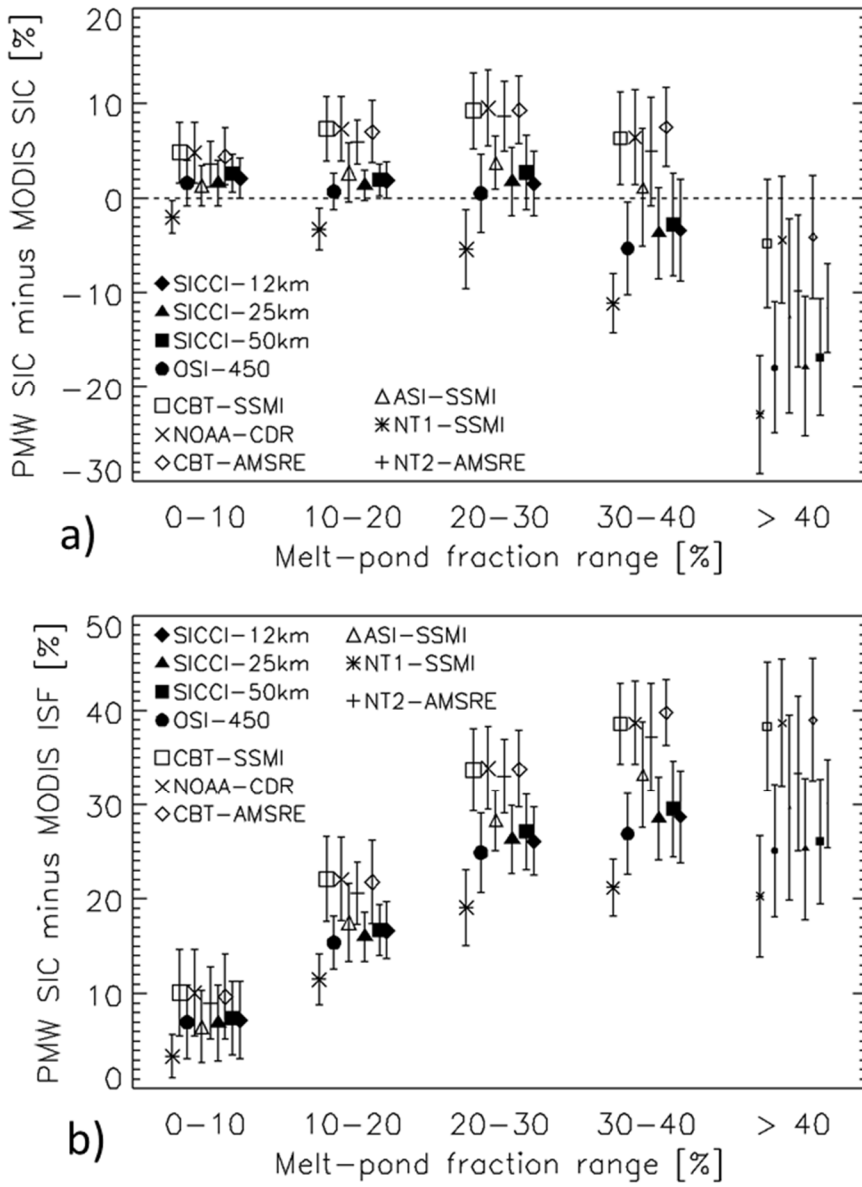



Figure 11. Mean difference PMW SIC minus MODIS SIC (image a)) and PMW SIC minus MODIS ISF (image b)) derived for all 8-day periods of the years 2003-2011 for all ten products separately for melt-pond fraction ranges 0-10 % to > 40 %. Error bars denote one standard deviation of the mean. Symbol size scales with the number of valid data pairs. The topmost four and bottommost three entries in left column of annotations denote group I (filled symbols) and group II, respectively. The topmost two entries and the last entry in right column of annotations denote group III and group IV, respectively.