Error assessment of satellite derived lead fraction in the Arctic

N. Ivanova[1], P. Rampal[1], and S. Bouillon[1]

[1]{Nansen Environmental and Remote Sensing Center, and Bjerknes Centre for Climate Research, Bergen, Norway}

Correspondence to: N. Ivanova (natalia.ivanova@nersc.no)

Abstract

Leads within consolidated sea ice control heat exchange between the ocean and the atmosphere during winter thus constituting an important climate parameter. These narrow elongated features occur when sea ice is fracturing under the action of wind and currents, reducing the local mechanical strength of the ice cover, which in turn impact the sea ice drift pattern. This makes a high quality lead fraction (LF) dataset to be in demand for sea ice model evaluation, initialization and for assimilation of such data in regional models. In this context, available LF dataset retrieved from satellite passive microwave observations (Advanced Microwave Scanning Radiometer – Earth Observing System, AMSR-E) is of great value, providing pan-Arctic light- and cloud-independent daily coverage since 2002. In this study errors in this dataset are quantified using accurate LF estimates retrieved from Synthetic Aperture Radar (SAR) images employing a threshold technique. A consistent overestimation of LF by a factor of 2–4 is found in the AMSR-E LF product. It is shown that a simple adjustment of the upper tie point used in the method to estimate the LF can reduce the pixel-wise error by a factor of 2 on average. Applying such adjustment to the full dataset may thus significantly increase the quality and value of the original dataset.

1 Introduction

In winter leads control heat transfer between the ocean and the atmosphere despite their relatively small areal coverage. For instance, sensible heat flux through leads can be of the

order of 600 W m^{-2}, compared to an annual average of about 3 W m^{-2} over ice (Maykut, 1978). This applies to leads represented by both open water and thin ice, but in winter the refreezing happens very quickly and open water leads exist only for a very short time (Weeks, 2010). Open-water leads alone, even though covering only 1–2% of the central Arctic, contribute more than 70% to the upward heat fluxes (Marcq and Weiss, 2012). Model simulations showed that even 1% change in sea ice concentration due to the increase in areal lead fraction could lead to a 3.5 K difference in the near-surface atmospheric temperature (Lüpkes et al., 2008). Studying signatures of leads and surrounding ice in the images from Moderate Resolution Imaging Spectroradiometer (MODIS) Beitsch et al. (2014) showed that difference in ice surface temperature between thicker ice and a lead covered by thin ice could be as large as 15–20 K, while open water and thin ice in leads differed in temperature by up to 10 K (Fig. 2 in Beitsch et al., 2014). This makes the surface energy budget very sensitive to the fraction of the surface covered by leads in the Arctic, where in recent years sea ice cover has become younger (Maslanik et al., 2007) and mechanically weaker (Rampal et al., 2009).

Areal fraction of leads in the Arctic sea ice can be viewed as a parameter reflecting loss in mechanical strength of the ice pack and indicating the degree of surrounding sea ice mobility. Rampal et al. (2009) reported steady increase in sea ice deformation rate and drift during 1979–2007 and argued for possible causal relation between the two. These trends still remain a challenge to capture for the current sea ice models, especially because they fail at simulating sea ice fracturing and lead opening with the correct properties. Accurate observations of lead fraction are thus of high importance for model evaluation and for being assimilated into models as initial conditions, or during a simulation. For example, Bouillon and Rampal (2015) and Rampal et al. (2015) presented recently a new sea ice model, which is able to use information on lead fraction to constrain the local mechanical response of sea ice to winds and currents, with a significant impact on performance with respect to e.g. simulated sea ice drift and deformation. In this context, using accurate estimates of lead fraction with their associated uncertainties is therefore crucial.

A method for areal lead fraction (LF) retrieval from Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E) was developed by Röhrs and Kaleschke (2012) (see also Röhrs et al., 2012) and allows to detect leads wider than 3 km. The method was able to detect 50% of leads when compared to a MODIS image and localize the leads correctly when qualitatively compared to Synthetic Aperture Radar (SAR) images and

CryoSat-2 tracks (Röhrs et al., 2012). A daily light- and cloud-independent pan-Arctic LF dataset (AMSR-E LF) for winter months November–April from 2002 to 2011 was obtained using this method and published at Integrated Climate Date Center – ICDC, University of Hamburg (http://icdc.zmaw.de/), and represents a unique and valuable dataset. It was then used to automatically obtain lead location and orientation with a success rate of 57% (Bröhan and Kaleschke, 2014). Preferred lead orientations were found typical for different regions of the Arctic.

The AMSR-E LF method is essentially a thin ice concentration retrieval method, which was adapted to identify leads by using median filtering. This filtering enhances the leads' features due to their narrow and elongated shape. Therefore, other thin ice retrieval methods based on passive microwave observations (e.g., Mäkynen and Similä, 2015, Naoki et al., 2008, Cavalieri, 1994) cannot be used directly for LF retrieval. Sea ice concentration algorithm ASI (Svendsen et al., 1987, Kaleschke et al., 2001, Spreen et al., 2008) was able to identify leads (Beitsch et al. 2014) when implemented at 89 GHz frequency of AMSR2 on-board the Global Change Observation Mission–Water satellite with resolution of 3.125 km. However, this approach is limited in time coverage because AMSR2 started to deliver the data only in 2012 (http://suzaku.eorc.jaxa.jp). Also quantitative validation work may be still needed because only qualitative assessment using MODIS images was presented in Beitsch et al. (2014).

A lead detection method based on MODIS ice surface temperature was developed by Willmes and Heinemann (2015). The method classifies a scene into leads and artefacts, where for the first class (leads) the success rate is as large as 95%. However, in the class of artefacts, which are mostly caused by ambiguity in cloud identification, there is a 50% chance of it being either a lead or an artefact. Combined retrieval error from the two classes for a daily map, obtained by averaging, is estimated to be 28%. The method gives daily lead occurrence maps at 1 km^2 resolution.

A number of classifiers applied to CryoSat-2 were tested for lead detection potential, and the most promising one identified and used to derive LF and lead width distribution (Wernecke and Kaleschke, 2015). The selected classifier was able to detect ~68% of leads correctly, and only ~3% of ice measurements were falsely identified as leads. Despite such good capability and fine resolution of 250 m, LF retrievals from CryoSat-2 are limited spatially, because the measurements are conducted by tracks making daily pan-Arctic coverage impossible; and temporally, the satellite being launched in 2010. Suggested approaches using laser altimeter

for lead detection (e.g., Farrell et al., 2009 with the Ice, Cloud and land Elevation Satellite, ICESat) have similar limitation.

Lindsay and Rothrock (1995) suggested a method for retrieval of lead widths and LF from thermal and reflected solar channels on the advanced very high resolution radiometer (AVHRR). The nominal resolution of the instrument is 1.1 km, and it is also able to resolve subpixel-sized leads due to strong contrast caused by leads and their network-like pattern. However, an AVHRR-retrieved LF dataset would be limited to cloud-free areas, and its quality would depend on the quality of cloud masking defining these areas.

Automatic classification of leads from SAR is difficult, because radar backscatter signature of leads in SAR images can be ambiguous. This is due to wind roughening of the open water in the leads and occasional presence of frost flowers when new ice has just formed in a lead (Röhrs et al., 2012). To the authors' knowledge, no method has so far been presented in literature addressing automatic LF retrievals from SAR. Existing sea ice classification methods (Berg and Eriksson, 2012; Karvonen, 2012; Karvonen, 2014; Leigh et al., 2014; Liu et al., 2015) could potentially be adapted and tested for this purpose. However, the task of identifying such narrow elongated features as leads is different from sea ice classification. For example, Korosov et al. (2015) demonstrated that these features could not be distinguished using a support vector machine (SVM) approach applied to SAR subsets (windows), while such a technique was good enough for ice/water separation in general. Also this study showed that even applying this method to segments, which significantly improved its feature-resolving capacity, was not satisfactory, and that the SVM would need to be trained specifically targeting leads.

As it is outlined above, there are a variety of available promising methods to detect leads and retrieve LF from satellites. They all have their advantages and disadvantages and, depending on these, can be used for achieving different purposes. The topic of this study is a dataset meeting the following criteria: retrieving LF (note the difference with lead occurrence), daily coverage, pan-Arctic, cloud- and light-independent, covering longest possible time period. The AMSR-E LF appears to be the only suitable dataset in this context, and therefore we find it necessary to provide quantitative error estimations of this dataset, which has not been done before. Based on analysis of the errors, we suggest a possible improvement of the AMSR-E based method. In order to achieve the goal of this study, a simple method for LF retrieval from SAR is suggested. Currently the method is specifically adapted for the purposes of this

study, but further development can give a universal approach for areal LF retrieval from SAR, which would be highly valuable.

Following the Introduction, Sect. 2 of the paper describes the data used for the study, and Sect. 3 explains the SAR-based method. The results are presented in Sect. 4 and 5 followed by Discussion and Conclusions.

2 Data

2.1 The AMSR-E LF dataset

The daily gridded AMSR-E LF dataset for the time period of November 2003–April 2011 was used (downloaded in February 2015, http://icdc.zmaw.de/1/daten/cryosphere/lead-area-fraction-amsre.html). It covers winter months of November through April and is provided on a polar-stereographic grid with 6.25 km resolution distributed by National Snow and Ice Data Center (NSIDC). LF is expressed as the percentage of a grid cell covered by leads, which are represented by either open water or thin ice. Since openings refreeze very quickly in winter, the majority of the data entries are thin ice concentrations. Following the original paper of Röhrs et al. (2012), thin ice is defined as new ice, nilas, and pancake ice, according to the classification of the World Meteorological Organization (WMO, 1989). The dataset is limited to areas where sea ice concentration is above 90%, as retrieved by the ASI algorithm.

The method used to retrieve LF from AMSR-E (Röhrs et al., 2012) relies on the unique signature of thin ice and open water defined by brightness temperature ratio in the 89 GHz and 19 GHz vertically polarized channels of the radiometer. Further, median filtering is applied to exclude the part of the signal coming form the atmosphere and enhance the features of leads due to their narrow and elongated shape so different from the more homogeneous background.

The AMSR-E LF dataset is shown in Fig. 1 by the number of measurements in each bin expressed in % of the total number of measurements (relative frequency), where each bin has a width of 5% except the first one, which excludes LF < 1%. These very small values of LF in the dataset appeared rather random on the daily maps and therefore were excluded assuming the method's precision would not have allowed resolving them anyway. All the grid cells close to land were also removed (2 grid cells away from land) because these areas contained

large amount of near 100% LF values, which may be caused by either real presence of the coastal polynyas/leads or an artefact due to the vicinity of land. Fig. 1 shows the full dataset covering all the winters from November 2003 through April 2011 (~26 millions measurements) by blue bars, and each month from November 2008 to April 2009 (varying from ~430 to ~600 thousands measurements) by different colours. The histograms for these months reflect the tendency observed in the full dataset, thus allowing us to limit the analyses presented in this paper to only this one winter. The last bin (LF 95–100%), characterised by significant amount of measurements in comparison to the other bins with high LF values, will be addressed in later sections.

For the validation by SAR images the AMSR-E LF dataset was re-projected on the domain defined in Sect. 2.2 using Nansat – an open source Python toolbox for processing 2D satellite earth observation data (Korosov et al., 2015, Korosov et al., 2016).

2.2 The SAR images

ENVISAT ASAR WSM (advanced SAR wide swath mode) images at HH-polarisation acquired during the winter of November 2008–April 2009 were used in this study. The area of interest is defined by the geographical coordinates (83N, 20W), (87N, 36W), (87N, 34E), (83N, 15E) and is shown in Fig. 2 by the red rectangle. This area located north of Fram Strait was chosen due to relatively large amount of leads occurring in this particular region (see e.g. Bröhan and Kaleschke, 2014) so that sufficient amount of AMSR-E LF retrievals would be available for validation, and because this region is well covered by SAR data. The SAR images originally provided at spatial resolution of 150 m × 150 m (pixel spacing: 75 m × 75 m), were re-projected using the Nansat toolbox onto a polar stereographic projection with nominal resolution of 100 m × 100 m with latitude of origin and central meridian defined by the central coordinates of the selected area. Calibrated surface backscattering coefficient (ASAR Product Handbook, 2007) normalized over ice was used for this study (we will refer to this value as backscatter). The procedure of normalization represents compensation for incidence angle variation, established empirically, and is described in more detail in Zakhvatkina et al. (2013).

3 SAR-based threshold technique

A threshold technique similar to the one developed for lead detection from MODIS-derived ice surface temperature (Willmes and Heinemann, 2015) is suggested for automatic lead identification in SAR scenes. Visual inspection of SAR images shows that leads, in most cases, have lower backscatter than surrounding thicker ice. The transition is defined by a threshold, which is not constant from one image to another, as we find from automatic lead detection tests conducted on a number of SAR images. Therefore, we use characteristics of backscatter distributions for each SAR scene instead. Before the threshold can be applied to a SAR scene (a subset is shown in Fig. 3a and respective distribution in Fig. 3d, beige bars) the image is undergone median filtering with a window size of 5×5 pixels (found experimentally), corresponding to spatial scale of 500 m × 500 m, which reduces the noise while preserving the edges of the features. One such filtered subset of a SAR image is shown in Fig. 3b (distribution in Fig. 3d, blue bars), where dark blue areas correspond to leads. Comparison of distributions before filtering (wider) and after shows the noise-reducing effect of the median filtering. After applying the threshold, so that all the backscatter values below its value are classified as leads and the rest – as ice, a binary map (Fig. 3c) is retrieved. The threshold (σ_0^t) is defined as

$$\sigma_0^t = \sigma_0^P - n_\delta \cdot \delta, \quad (1)$$

where σ_0^P is the backscatter value at the peak of the distribution (blue line in Fig. 3d), δ is the standard deviation of the distribution, and n_δ is a number of standard deviations to move away from the peak, that enables automatic identification of leads. The threshold was first tried with $n_\delta = 1$ and $n_\delta = 2$ (dashed red lines), but it was found that an intermediate value $n_\delta = 1.5$ (solid red line) worked better and therefore was chosen. This was established by visual comparison of the lead fraction retrievals with different threshold values. The mean of the distribution is shown by dashed grey line for reference.

Next, SAR-based LF is calculated for each AMSR-E grid cell where LF value is above 1%. All the pixels classified as lead by SAR within such grid cell are added together and divided by the total number of SAR-pixels in it, which gives a percentage after multiplying it by 100.

Suggested approach is rather simplistic, but it is sufficient for our purpose (more details in the Sect. 4.1.2 and 4.3), while for a wider application one must consider the limitations addressed in the Discussion section.

4 Results

4.1 Reference Lead Fraction Datasets Retrieved from SAR

Using the approach described in Sect. 3, we produced two SAR-based reference datasets: one with manual quality control of each SAR subset of 1000×1000 pixels (MQC SAR LF) and one based on automatic threshold where quality control is done by discarding images with obviously unsuccessful LF retrievals (SAR LF).

4.1.1 MQC SAR LF

This high-quality dataset was produced in order to verify the larger SAR LF dataset (Sect. 4.1.2). Significantly larger amount of measurements in the SAR LF allows robust statistical analysis, but visual quality control of each image, given that leads are numerous small features, is hardly achievable. For the MQC SAR LF two criteria need to be verified: 1) whether the classification is successful and 2) whether leads are identified in exactly the same locations in the SAR- and AMSR-E-derived datasets. The latter was mostly the case, however sometimes a lead in AMSR-E LF was misplaced by a distance large enough so that the two datasets mismatch. We believe this misplacement is caused by cases of relatively fast sea ice drift in the area. If we consider an AMSR-E grid cell of 6.25 km × 6.25 km size, a SAR image is taken at a certain time of the day in this grid cell, while ASMR-E LF is a gridded daily product and thus provides an average over all the swaths covering this grid cell collected during 24 hours. During a few hours the lead could have moved fast enough to disappear from the given grid cell. From visual analysis of the images we could say that this situation did not happen very often, however a quantitative estimate of how much it affects the validation was needed. Thus, we make an assumption that if the distribution of SAR LF is similar to that of MQC SAR LF, where we made sure every lead was located correctly, the misplacements were indeed seldom the case also in the SAR LF dataset.

To produce the MQC SAR LF, 5 SAR scenes acquired in March 2009 with sufficient amount of easily distinguishable leads were selected. It was found that the quality of LF retrieval

increases when dividing SAR scenes into subsets, and the subset size of 1000×1000 pixels showed to be sufficient. Using such small subsets rather than a full SAR image provides more accurate thresholds because it limits possible variability in conditions within the subset. Such conditions can be wind speed or ice surface properties (wet or dry ice, for example). Defining a threshold locally not only eliminates significance of these effects, but it takes advantage also of smaller variety of surfaces in general. For example, presence of open water, land, consolidated ice, wet ice, dry ice, and marinal ice zone in one image will make it difficult to find a threshold that will only identify leads. Using a smaller subset, on the other hand, where only consolidated ice with leads is present, will give clearer threshold.

The threshold was thus calculated individually for each 1000×1000 pixels subset using Eq. (1) with n_δ selected manually, and used to calculate LF in corresponding AMSR-E grid cells. The classification in each subset was then inspected visually, comparing the two collocated maps: backscatter and MQC SAR LF, in order to make sure it was successful. This procedure gave 1645 high-quality MQC SAR LF retrievals, which were then used to verify the findings based on a larger SAR LF dataset.

4.1.2 SAR LF

To produce this dataset, SAR subsets of 3500×3500 pixels each (on average) were used: the full SAR images were cut to match the region of interest (Fig. 2). The quality control of this validation dataset was done by visual inspection of every classified subset together with the original SAR subset (backscatter) as demonstrated in Fig. 4 (a subset of 875×875 pixels is shown). Panel a) shows the original SAR image, panel b) shows the lead identification by the SAR-based method, where the red colour corresponds to the identified leads, and the panel c) shows the SAR image overlaid by AMSR-E lead fraction original product (in %) with the colour scale to the right. We rely on the combination of the fine resolution of SAR (pixel spacing of original product: 75 m × 75 m, geometric resolution: 150 m × 150 m) and its capability to separate smooth surfaces such as open water or thin ice in leads (appear darker than the background) from the rough surfaces (surrounding thicker ice). In addition, leads have a characteristic shape: they are narrow elongated features. These three factors put together make it possible to visually recognize the leads in SAR images. The features that were missed by SAR were relatively small, and were usually not captured by the coarse-resolution AMSR-E either (Fig. 4c). Note also that the AMSR-E based method was found to identify only leads wider than 3 km (Röhrs et al., 2012), and such features were normally

identified successfully by SAR in our study. We have also performed a visual comparison of a SAR dataset sample to a MODIS image (2500×2500 pixels at 250 m resolution), and saw that the SAR method captured the majority of features correctly.

In this process images were discarded in cases of unsuccessful lead identification, which is when features that appear like leads were missed by the method in significant amount. This was of particular importance in cases when AMSR-E LF identified a feature in the respective location, to secure proper error estimation for the AMSR-E LF product.

The majority of subsets contained leads represented by signatures darker than the surrounding background, while subsets containing large amount of leads with brighter signature were discarded. This means that the majority of the leads in the selected subsets were either composed by thin ice or calm open water. Therefore, the wind speed is not taken into account in this study, but for a more general application this would have been necessary to account for wind roughening of the open water areas in leads. As a result we obtained a dataset for the period of November 2008–April 2009, made of 21–47 subsets (3500×3500 pixels each) per month, with number of measurements varying from about 8 000 to 19 500 (Table 1) depending on the month.

4.2 Comparison of the AMSR-E LF and MQC SAR LF

Before any analysis of the AMSR-E LF and MQC SAR LF datasets could be performed, they were filtered so that only those AMSR-E grid cells (6.25 km by 6.25 km size) were used, which had LF value > 1% and where the SAR LF for this grid cell returned a value of LF > 1% too. Thus, we only analyse the non-zero values of the AMSR-E LF dataset, and exclude all the leads that the SAR method has eventually missed. The same applies to the next section, where AMSR-E LF is compared to SAR LF.

The AMSR-E LF and MQC SAR LF datasets are shown in Fig. 5 as a scatterplot (left) and histograms (right). The scatterplot shows that the majority of the points are located below the 1-to-1 line (dashed grey line), which means that in most cases AMSR-E LF overestimates the LF as compared to the SAR retrievals. The linear regression line (red) has a slope of 0.2. Note that for the value of AMSR-E LF 100% there is wide range of MQC SAR LF values covering almost the full scale from 0% to 100%. The point-wise root mean square error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(LF_{AMSRE\,i} - LF_{SAR\,i}\right)^2}, \quad (2)$$

where n is the total number of measurements (1645 in this case), is equal to 33%. The determination coefficient R^2, retrieved as squared correlation coefficient and expressed in percent, is 13% (significant). The right panel of Fig. 5 shows histograms of the two datasets representing number of measurements per each 5%-bin expressed in % of total number of measurements. The distributions of the two datasets look principally different, characterized by steep decrease in number of cases with increasing LF for SAR and wide distribution of values in the AMSR LF. Thus, for LF>20% AMSR LF seems to overestimate largely the number of cases and underestimate this number for lower LF values. Similarly to the full AMSR-E LF dataset (Fig 1.) the near 100% bin contains relatively large amount of measurements. In fact, about 94% of all the data in this bin in the full AMSR-E dataset are above 99.9%. In order to understand the origin of such large amount of LF near 100% we compare spatial maps of LF obtained from AMSR-E and SAR. As an example of such analysis, Fig. 6 shows part of a SAR image overlaid by collocated AMSR-E LF product, where one can see general overestimation of LF by AMSR-E (larger grid cells shown as percentage by different colours). But in particular it is clear for the LF 100% cases (red grid cells): these often correspond to a smaller amount of water/thin ice in the SAR image. Four neighbouring AMSR-E grid cells are shown in a close-up inset, where three of them have a LF value of 100% (the fourth one has no value), while the SAR image in the background clearly contains one lead that covers only about 25% of the right grid cell, 40% of the upper grid cell and about 60% of the left one, where also smaller cracks are present.

4.3 Error estimations of the AMSR-E LF based on SAR LF

Same procedure as in Sect. 4.2 is now applied using the large SAR LF dataset. Histograms for collocated datasets AMSR-E LF and SAR LF are produced for each month of the considered period (Fig. 7). They show the same tendency as when using the shorter high-quality dataset. The distributions here are much smoother because of the significantly larger number of measurements. The similarity of the distributions coming from high-quality MQC SAR LF and SAR LF allow us to base our conclusions on the larger dataset (SAR LF) thus providing more accurate estimates of errors.

Having this significant amount of collocated SAR and AMSR-E retrievals of LF we can confirm that the peak in AMSR-E LF dataset near 100% represents an artefact. This is also supported by the visual analysis of overlay of every image pair: AMSR-E LF and SAR LF. AMSR-E LF had relatively large amount of observations close to 100%, while in SAR images the area covered by leads in such grid cells was obviously smaller in almost all the cases. The cases where one lead width would take the full AMSR-E grid cell or even more (lead width larger than one grid cell) were extremely rare in our selection. We believe that this grouping of large amount of measurements near the value of 100% is a result of the assumption lying behind the AMSR-E method for LF retrieval. The method is based on the ratio of the brightness temperatures (r) in 89 GHz and 19 GHz channels (Röhrs et al., 2012). The assumption is that all the values of this ratio above a certain constant value (a tie point) will give LF 100%. All the other values are linearly interpolated between a tie point for LF 0% ($r0$) and a tie point for LF 100% ($r100$). If the upper tie point $r100$ is too low, a significant amount of LF values assigned to a value of 100% by this cut-off may actually correspond to a variety of LF much lower than 100%. This is reflected in Fig. 5 (left) and Fig. 6, where values of LF 100% in AMSR-E dataset correspond to variety of values from SAR dataset. Ideally, an improvement of ASMR-E LF method is needed, for example, by adjusting the upper tie point so that the full range of LF values are covered. We address this further in the Sect. 5.

Since production of a new improved AMSR-E LF dataset is out of scope of this study, we suggest imitating the same problem with the SAR LF dataset instead. Introduction of a new upper tie point $r'100$ would be equivalent to dividing of all the AMSR-E LF values by a certain factor, defined as $f = (r'100 - r0)/(r100 - r0)$, because the method is based on linear interpolation of all the values between the limits of the range. Since the LF values in the near 100% bin for AMSR LF are unknown, we suggest multiplying the SAR LF dataset by such factor instead. In order to define the value of f (also referred to as AMSR-E factor) we vary its value from 1 to 5 and calculate respective RMSE as a measure of difference between the histograms of AMSR-E LF and SAR LF datasets for each month (Fig 7.):

$$RMSE_h = \sqrt{\frac{1}{n_b}\sum_{i=1}^{n_b}\left(RF_{AMSRE\,i} - RF_{SAR\,i}\right)^2}, \quad (3)$$

where RF stands for relative frequency in each bin, and n_b is the number of bins. Obtained $RMSE_h$ is plotted as a function of f in Fig. 8 (left), where each month is assigned different

colour and March 2009 is highlighted by bold line to illustrate the principle. By minimizing $RMSE_h$ we find optimal f value for each month, which amounts to 3.3, 2.5, 2.8, 3.7, 2.8, and 2.7 for the months from November 2008 to April 2009 respectively. Multiplying the SAR LF dataset for each month by respective factor gives a histogram with similar issue at 100% as the AMSR-E LF dataset has (yellow bars in Fig. 8, right). The values in other bins also redistribute in a way that is similar to the AMSR-E LF dataset. Original histograms of AMSR-E LF and SAR LF (same as Fig. 7, but for the full winter) are also shown for reference.

The systematic overestimation of AMSR-E LF data also affects the mean value of the distribution. For winter 2009, the mean value of AMSRE LF ($\overline{LF_{AMSRE}}$) is equal to 31%, whereas it is equal to 13% for the SAR LF ($\overline{LF_{SAR}}$). The absolute relative difference $100 \times \left| \left(\overline{LF_{AMSRE}} - \overline{LF_{SAR}} \right) \Big/ \overline{LF_{SAR}} \right|$ decreases from 140% with no correction to 17% when using the correction factors found here.

Finally the agreement between SAR LF and AMSR-E LF datasets can be estimated by the point-wise RMSE of LF for the whole winter 2009 as defined by the Eq. (2), the total number of measurements n being 64 063 here. Here $LF_{SAR\,i}$ are the LF values obtained when multiplying by the correction factor, so that point-wise RMSE is relatively independent of the systematic bias in AMSR-E LF. The point-wise RMSE is equal to 43% and is an estimate of the standard deviation of the difference between AMSRE-E LF and SAR LF. However, similar computation of RMSE using $LF_{SAR\,i}$ without correction gives a value of 33%, suggesting the need for a more physically justified approach, e.g. by improving the AMSR-E based method.

5 Suggested improvement of the AMSR-E-based method

In the Sect. 4.3 we made an assumption that the upper tie point in the AMSR-E-based method should be increased in order to cover the full range of LF values. To test this assumption we implement the method according to Röhrs et al. (2012) and calculate LF from the AMSR-E brightness temperatures on the 8 March 2009 with the original tie points (a subset is shown in Fig. 9, upper left), i.e. with the upper tie point $r100 = 0.05$. Such calculations give similar distribution of LF values (Fig. 9, upper right) as was found in the full AMSR-E LF dataset (Fig. 1). Using the linear relationship between $r100$ and f, and the optimal value of f for

March 2009 ($f = 2.8$), we calculate that $r100$ should be increased to 0.113 ($r'100$). This new tie point value gives a distribution closer to that of the SAR LF dataset (Fig. 9, bottom right) – the value of $RMSE_h$ (Eq. (3)) decreasing from 5.4% (corresponding to $f = 1$ in Fig. 8, left) to 0.9%. The close-up insets similar to the one in Fig. 6 show that the leads are identified in the same locations as before, but the LF values are lower (Fig. 9, bottom left). Fig. 10 (left) shows similar scatterplot to Fig. 5 (left), but for this one-day example. The right panel of Fig. 10 compares the AMSR-E lead fraction obtained with the new tie-point to that of the reference SAR LF dataset. The tie-point adjustment made the AMSR-E and SAR datasets agree significantly better: the point-wise RMSE (Eq. (2)) for this one-day dataset of 750 collocated LF measurements decreased from 37% to 15% and the slope of the regression line became closer to 1 (increased from 0.2 to 0.5). The determination coefficient R^2 showed also slight improvement increasing from 27% to 33% (both coefficients are significant).

We thus believe that implementation of such an adjustment to the full AMSR-E LF dataset will lead to a much better agreement with the SAR LF dataset. The new tie point $r'100$ retrieved for the other months amounts to 0.131, 0.103, 0.113, 0.145 for November 2008 – February 2009 respectively, and 0.110 for April 2009. The average value of the new tie point $r'100$ weighted by the number of observations for each month is 0.117 and is therefore our best estimate for winter 2008–2009.

6 Discussion

A method to retrieve LF from SAR backscattering coefficient is introduced. This simple threshold technique is only suitable for the purposes of this study, and is thus not universal. However, its potential is shown, and the limitations are identified allowing further developments of such a method, which is out of scope of this study.

One of the limitations is ambiguity of SAR signatures corresponding to leads. When a lead is represented by calm open water or thin ice, it has lower backscatter values than surrounding thicker ice and therefore can be identified by a threshold. However, in cases when wind is roughening the open water surface in the lead, its signature becomes brighter. Another case of such ambiguity is presence of frost flowers on the newly refrozen lead, which also causes brighter signatures (Röhrs et al., 2012). Such leads with brighter signature than the background are not identified by the presented SAR method, but are sometimes (but not always) identified by the AMSR-E method. These cases did not occur much in the considered

examples and were discarded from the analysis thus not affecting the conclusions. For a more universal SAR-based method such cases can be included by introducing two thresholds – one for the leads appearing darker than the background and one for the ones appearing brighter. In that case two different sides of the backscatter distribution will be used independently.

Another limitation of the approach used here is presence of areas with presumably wet snow/ice, which appear rather dark on a SAR image and therefore are classified as leads by the threshold method. These cases did not occur often in our selection, and they did not influence the comparison because AMSR-E LF usually does not identify leads in such areas, and we only included the grid cells where AMSR-E LF dataset had any value above 1%. The threshold is also sensitive to the sea ice thickness. At a given threshold only leads with ice thin enough will be identified as leads. Since we do not know how thick the ice is, it adds to the ambiguity of such method. In other words, by selecting a threshold we indirectly set the sea ice thickness limit. When the distribution is bimodal (one mode for leads and one for thicker ice), a value between the peaks can be used as threshold, as suggested by Lindsay and Rothrock (1995) for distributions of temperature or brightness. However, such cases were so rare in the selected SAR images that this approach was discarded. To achieve bimodal distribution, the LF calculation procedure can be applied to SAR scenes divided into sub-scenes (size of approximately 1000×1000 pixels), which will demand more processing time. Such definition of threshold could serve as a more robust approach when developing an independent method for automatic SAR LF retrieval. For the purposes of this study the quality of the suggested simple threshold method was considered sufficient because of the quality control steps that were undertaken. Firstly, visual inspection of every SAR LF subset was performed; it indicated that we detected most of the leads. Secondly, the ambiguous cases were excluded. And finally, the analysis was limited to only those AMSR-E LF grid cells where both datasets give a non-zero value of LF.

Analysing the results of the comparison between AMSR-E and SAR one should keep in mind that the surface parameters these two instruments are sensitive to are not exactly the same. The mechanisms that form the signal from an area with leads, represented by either open water or thin ice, are substantially different for SAR (sensitive to roughness) and AMSR-E (sensitive to emitted brightness). In addition, they have different resolution: 150 m ×150 m for SAR and 6 km × 4 km for AMSR-E (the footprint size of the 89 GHz channel). Thus, SAR is capable of identifying the leads in accurate locations and resolving their limits correctly (this

was concluded from visual inspection of SAR retrievals, including comparison to a MODIS image). For AMSR-E the signal is an aggregated effect of all the surfaces present in the grid cell: open water, thin ice and thicker ice, from which the percentage of open water/thin ice per grid cell needs to be deduced.

It should be noted that even an improved AMSR-E LF method would still have its limitations. For example, it would not be able to capture leads narrower than 3 km due to its resolution, while leads as narrow as a few meters transmit turbulent heat more than two times as efficient as the ones hundreds of meters wide (Marcq and Weiss 2012). For studies like e.g. assessing the integrated heat fluxes through leads in wintertime, the AMSR-E LF dataset alone will thus not be sufficient and other methods should be used in addition. Another limitation of such a method would be retrieval of LF in summer, when interpretation of passive microwave observations is challenging.

7 Conclusions

This work was partly motivated by the need of an accurate pan-Arctic lead fraction (LF) dataset for initialisation and evaluation of regional sea ice models. One such dataset was identified as having good potential for the purpose – daily pan-Arctic LF retrieved from Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E), a passive microwave instrument independent of cloud cover and light conditions. In this study we set a goal to evaluate the AMSR-E LF dataset and provide quantitative estimate of eventual errors. These can serve as a measure of uncertainty of the product and background for a correction.

After analysis of the AMSR-E LF dataset and comparison to LF retrievals from Synthetic Aperture Radar (SAR) we identified an issue with the near 100% LF values in this dataset. More specifically, we concluded that the tie points used in the AMSR-E method were located too closely to each other, which caused a truncation of the real LF range. This means that LF values obtained with such tie points represent a range of values erroneously stretched over larger range (e.g., 0–250%) and are cut off at 100%, where all the values above 100% are converted to 100% thus causing the loss of all the values above. A larger distance between the tie points would accommodate all the real LF values and give the correct range of 0–100% as output. Such an adjustment of tie points is equivalent to dividing of AMSR-E LF by a certain

factor. Since the information about LF>100% is lost in the AMSR-E LF dataset, we imitated the issue by multiplying SAR LF by this factor instead. In this manner we found that the current AMSR-E LF dataset overestimated LF by a factor of ~2–4 over the winter 2008-2009 depending on the month considered. The absolute relative difference between the datasets expressed by $100 \times \left| \left(\overline{LF_{AMSRE}} - \overline{LF_{SAR}} \right) \Big/ \overline{LF_{SAR}} \right|$ decreased from 140% with no correction to 17% when using this correction factor. However, this approach is not suitable for correction of local values, but rather reflects statistical characteristics of the dataset over the whole Arctic (e.g., mean), which is confirmed by increase in the point-wise root mean square error (RMSE) between the AMSR-E LF and the SAR LF dataset with correction from 33% to 43%.

We argued that an adjustment of the AMSR-E LF method needed to be done before more accurate error estimation could be retrieved. We therefore tried out such an adjustment by implementing the AMSR-E-based method using higher value of the upper tie point, and found that indeed the AMSR-E LF distribution became similar to that of SAR LF. The $RMSE_h$ used as measure of difference between the two histograms decreased from 5.4% to 0.9%, while the point-wise RMSE for this one-day test dataset of 750 collocated LF measurements decreased from 37% to 15%, or by a factor of ~2. We observed that leads were still placed in the same locations, while the LF values became lower, which corresponded to what we observed from the SAR LF dataset. We estimated the new upper tie point for each months of the winter of 2008–2009 and found the values in the range from 0.103 to 0.145, or 0.117 for the full winter as an average weighted by the number of measurements for each month. We believe that similar simple adjustment applied to the full AMSR-E LF dataset will lead to significantly lower errors when evaluated using SAR, making this dataset more valuable for e.g. assimilation into models or model evaluation.

Acknowledgements

The research was mainly supported by the Centre for Climate Dynamics at the Bjerknes Centre. Additional support was provided by Nansen Environmental and Remote Sensing Centre, Bergen, Norway (NERSC). Sylvain Bouillon was supported by the Research Council of Norway through the post-doc project SIMech, Sea Ice Mechanics: from satellites to numerical models (No. 231179/F20, 2014-2016). The authors would like to thank Lars

Kaleschke (Institut für Meereskunde, KlimaCampus, University of Hamburg, Germany) for useful discussions and Anton Korosov (NERSC) for thorough technical support.

References

ASAR Product Handbook, online. Issue 2.2, European Space Agency, 27 February 2007. Available: https://earth.esa.int/handbooks/asar/CNTR.html, last access: January 2014, 2007.

Beitsch, A., Kaleschke, L., and Kern, S.: Investigating High-Resolution AMSR2 Sea Ice Concentrations during the February 2013 Fracture Event in the Beaufort Sea, Remote Sensing, 6(5), 3841-3856, doi:10.3390/rs6053841, 2014.

Berg, A., and Eriksson, L. E. B.: SAR Algorithm for Sea Ice Concentration—Evaluation for the Baltic Sea, IEEE T. Geosci. Remote Lett., 9(5), 938–942, 2012.

Bouillon, S., and Rampal, P.: Presentation of the dynamical core of neXtSIM, a new sea ice model, Ocean Mod., 91(C), 23–37, doi:10.1016/j.ocemod.2015.04.005, 2015.

Bröhan, D. and Kaleschke, L.: A Nine-Year Climatology of Arctic Sea Ice Lead Orientation and Frequency from AMSR-E, Remote Sensing, 6(2), 1451-1475, doi:10.3390/rs6021451, 2014.

Cavalieri, D. J.: A microwave technique for mapping thin sea ice, J. Geophys. Res., 99(C6), 12561–12572, doi:10.1029/94JC00707, 1994.

Farrell, S. L., Laxon, S. W., McAdoo, D. C., Yi, D., and Zwally, H. J.: Five years of Arctic sea ice freeboard measurements from the Ice, Cloud and land Elevation Satellite, J. Geophys. Res., 114, C04008, doi:10.1029/2008JC005074, 2009.

Kaleschke, L., Lupkes, 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. Remote Sens. 27, 526–537, 2001.

Karvonen, J.: Baltic Sea ice concentration estimation based on C-band HH-Polarized SAR data, IEEE J. Sel. Top. Appl., 5, 1874-1884, doi:10.1109/JSTARS.2012.2209199, 2012.

Karvonen, J.: Baltic Sea ice concentration estimation based on C-band Dual-Polarized SAR data, IEEE T. Geosci. Remote, 52, 5558-5566, doi:10.1109/TGRS.2013.2290331, 2014.

Korosov, A., Hansen, M. W., and Yamakava, A.: Nansat – scientist friendly toolbox for processing satellite data, World Ocean Scientific Congress, Cochin, India, 2–8 February, TS-13/130, 2015.

Korosov, A., Zakhvatkina, N., and Muckenhuber, S.: Ice/water classification of Sentinel-1 images, Geophysical Research Abstracts, 17, EGU2015-12487-1, EGU General Assembly 2015, Vienna, Austria, 17 April, 2015.

Korosov, A., Hansen, M. W., Yamakawa, A., Dagestad, K., Vines, A., Riechert, M., Myasoedov, A., Morozov, E. A., and Zakhvatkina, N.: Nansat v0.6.7 stable. URL: http://dx.doi.org/10.5281/zenodo.45188, doi: 10.5281/zenodo.45188, 2016.

Leigh, S., Wang, Z., and Clausi, D.: Automated Ice-Water Classification Using Dual Polarization SAR Satellite Imagery, IEEE T. Geosci. Remote, 52(9), 2014.

Lindsay, R. W. and Rothrock, D. A.: Arctic sea ice leads from advanced very high resolution radiometer images, J. Geophys. Res., 100(C3), 4533–4544, doi:10.1029/94JC02393, 1995.

Liu, H., Guo, H., and Zhang, L.: SVM-Based Sea Ice Classification Using Textural Features and Concentration From RADARSAT-2 Dual-Pol ScanSAR Data, IEEE J. Sel. Top. Appl, 8(4), 1601–1613, 2015.

Lüpkes, C., Vihma, T., Birnbaum, G., and Wacker, U.: Influence of leads in sea ice on the temperature of the atmospheric boundary layer during polar night, Geophys. Res. Lett., 35, L03805, doi:10.1029/2007GL032461, 2008.

Mäkynen, M. and Similä, M.: Thin Ice Detection in the Barents and Kara Seas With AMSR-E and SSMIS Radiometer Data, IEEE T. Geosci. Remote, IEEE Early Access Articles, doi: 10.1109/TGRS.2015.2416393, 2015.

Marcq, S. and Weiss, J.: Influence of sea ice lead-width distribution on turbulent heat transfer between the ocean and the atmosphere, The Cryosphere, 6, 143-156, doi:10.5194/tc-6-143-2012, 2012.

Maslanik, J. A., Fowler, C., Stroeve, J., Drobot, S., Zwally, J., Yi, D., and Emery, W.: A younger, thinner Arctic ice cover: Increased potential for rapid, extensive sea-ice loss, Geophys. Res. Lett., 34, L24501, doi:10.1029/2007GL032043, 2007.

Maykut, G.A.: Energy exchange over young sea ice in the central Arctic, J. Geophys. Res., 83(C7), 3646-3658, 1978.

Naoki, K., Ukita, J., Nishio, F., Nakayama, M., Comiso, J. C., and Gasiewski, A: Thin sea ice thickness as inferred from passive microwave and in situ observations, J. Geophys. Res., 113, 2156-2202, 2008.

Rampal, P., Weiss, J., and Marsan, D.: Positive trend in the mean speed and deformation rate of Arctic sea ice, 1979–2007, J. Geophys. Res., 114(C5), C05013, doi:10.1029/2008JC005066, 2009.

Rampal, P., Bouillon, S., Ólason, E., and Morlighem, M.: neXtSIM: a new Lagrangian sea ice model, The Cryosphere Discuss., 9, 5885-5941, doi:10.5194/tcd-9-5885-2015, 2015.

Röhrs, J. and Kaleschke, L.: An algorithm to detect sea ice leads by using AMSR-E passive microwave imagery, The Cryosphere, 6, 343-352, doi:10.5194/tc-6-343-2012, 2012.

Röhrs, J., Kaleschke, L., Bröhan, D., and Siligam, P. K.: Corrigendum to "An algorithm to detect sea ice leads by using AMSR-E passive microwave imagery" published in The Cryosphere, 6, 343–352, 2012, The Cryosphere, 6, 365-365, doi:10.5194/tc-6-365-2012, 2012.

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

Svendsen, E., Matzler, C., and Grenfell, T. C.: A model for retrieving total sea ice concentration from a spaceborne dual-polarized passive microwave instrument operating near 90 GHz. Int. J. Remote Sens. 8, 1479–1487, 1987.

Weeks, W. F.: On Sea Ice.University of Alaska Press, Fairbanks, Alaska, 2010.

Wernecke, A. and Kaleschke, L.: Lead detection in Arctic sea ice from CryoSat-2: quality assessment, lead area fraction and width distribution, The Cryosphere, 9, 1955-1968, doi:10.5194/tc-9-1955-2015, 2015.

Willmes, S., and Heinemann, G.: Pan-Arctic lead detection from MODIS thermal infrared imagery, Annals of Glaciology 56, 29–37, doi: 10.3189/2015AoG69A615, 2015.

WMO: The World Meteorological Organization Sea Ice Nomenclature (WMO No. 259, TP-145, Supplement No. 5), 1989.

Zakhvatkina, N. Yu., Alexandrov, V. Yu., Johannessen, O. M., Sandven, S., and Frolov I. Ye.: Classification of Sea Ice Types in ENVISAT Synthetic Aperture Radar Images, IEEE T. Geosci. Remote, 51, 2587-2600, doi:10.1109/TGRS.2012.2212445, 2013.

Table 1. Number of measurements in the SAR LF dataset

Month	Subsets	Measurements
Nov 2008	27	8 097
Dec 2008	34	9 392
Jan 2009	47	10 672
Feb 2009	29	7 528
Mar 2009	47	19 460
Apr 2009	21	8 914
Total	**205**	**64 063**

1

2

3

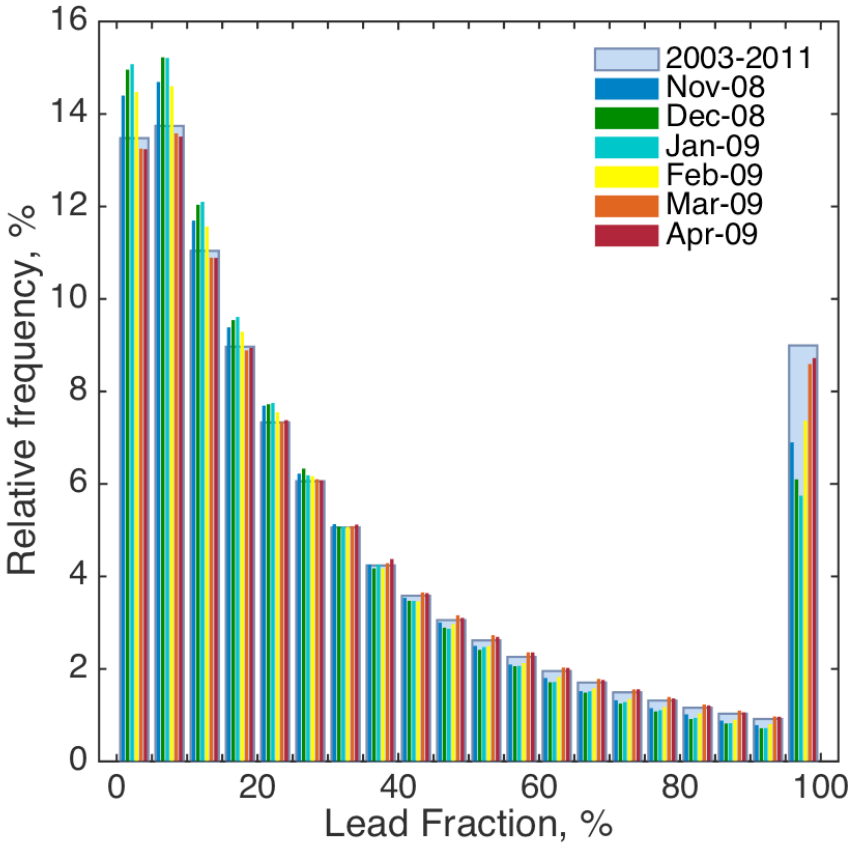


Figure 1. Histograms for AMSR-E lead fraction (LF) dataset shown as the number of measurements per each LF bin of 5% width expressed in % of the total amount of measurements (relative frequency). The blue bars show the full dataset, while each month of the winter 2008–2009 is shown by other colours (see the legend).

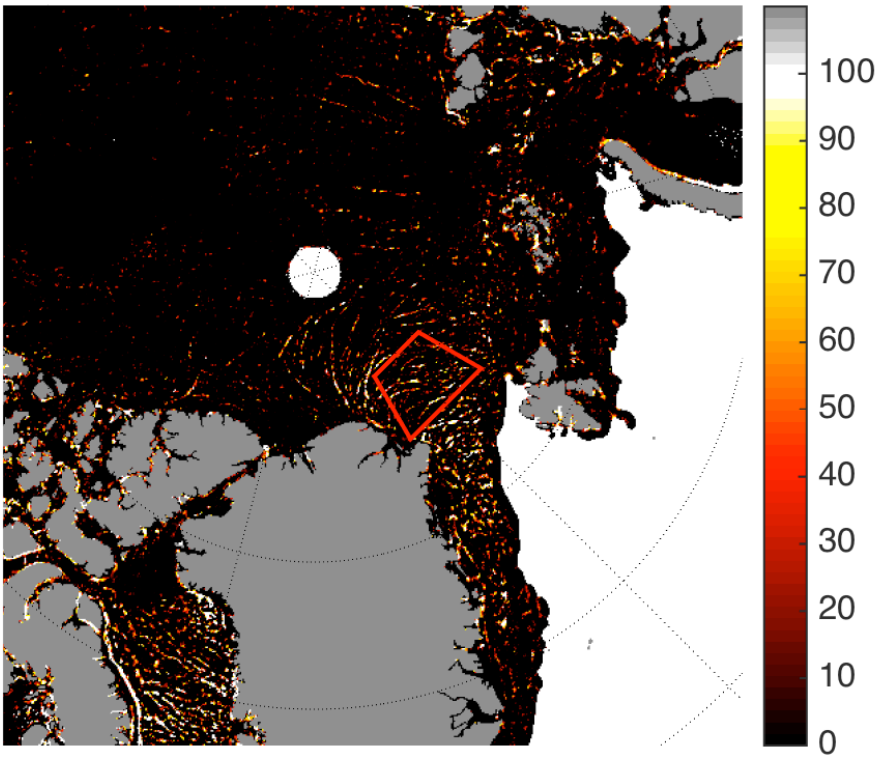


Figure 2. Area of interest is included within the red rectangle. The background map shows AMSR-E lead fraction in % (the numbers on the colour scale to the right), obtained on the 8 March 2009, and is used here only to demonstrate a sample from the product.

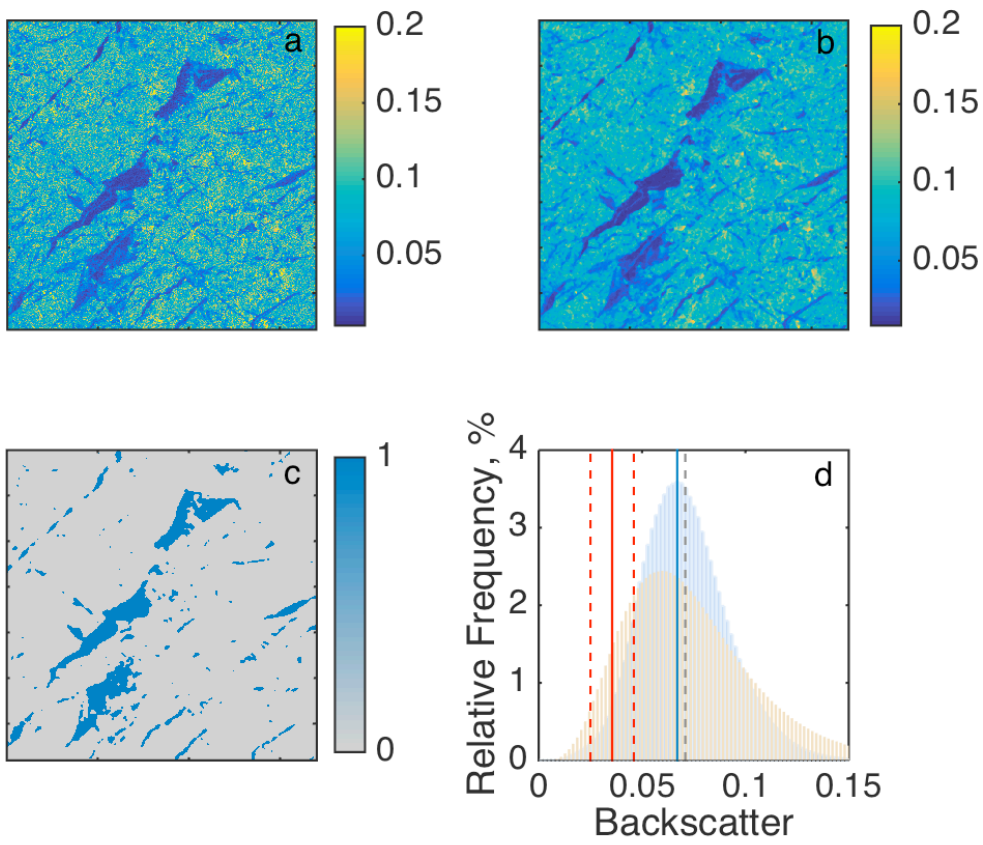


Figure 3. Threshold technique used to calculate lead fraction from SAR images: a) a subset of 680×680 pixels showing backscatter values; b) same as a) but after median filter has been applied; c) the resulting lead detection (1 – lead, 0 – ice); d) histogram of an example SAR scene taken on the 1 March 2009 (blue) with lines showing the peak (blue), threshold defined as peak minus 1.5 standard deviation (red), other thresholds (when 1 standard deviation and 2 standard deviations are used, dashed red), and mean is shown in grey dashed line. The beige histogram is for the unfiltered signal.

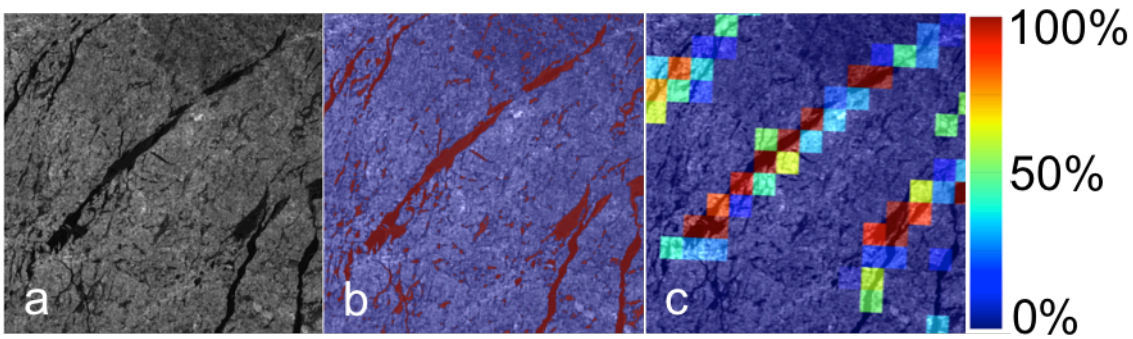


Figure 4. a) A subset of the original SAR image (backscatter) on the 8 March 2009, b) respective classified SAR-image (red – lead, blue – ice) with a) as background, c) AMSR-E lead fraction in % with a) as background.

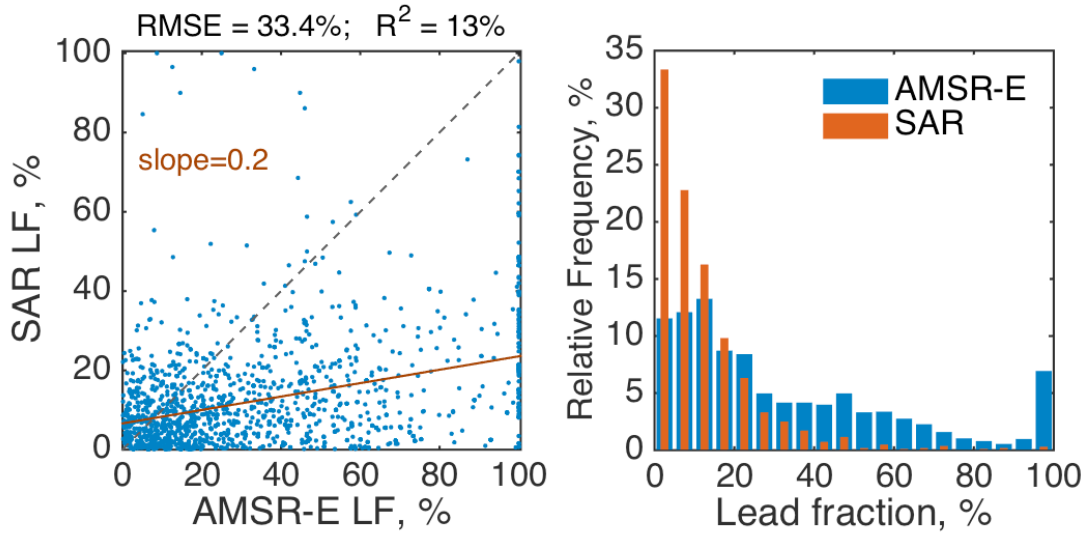


Figure 5. A comparison of AMSR-E lead fraction (LF) and SAR LF with manual quality control (MQC), the total amount of measurements is 1645. Left: scatterplot of MQC SAR LF versus AMSR-E LF (%). The 1-to-1 line is the dashed grey line, and the linear regression is shown by the red line and the slope value. Root mean square error (RMSE) and the coefficient of determination (R^2) are shown at the top of the plot. Right: histograms for the two datasets shown as percentage of measurements per each bin of 5% width (relative frequency).

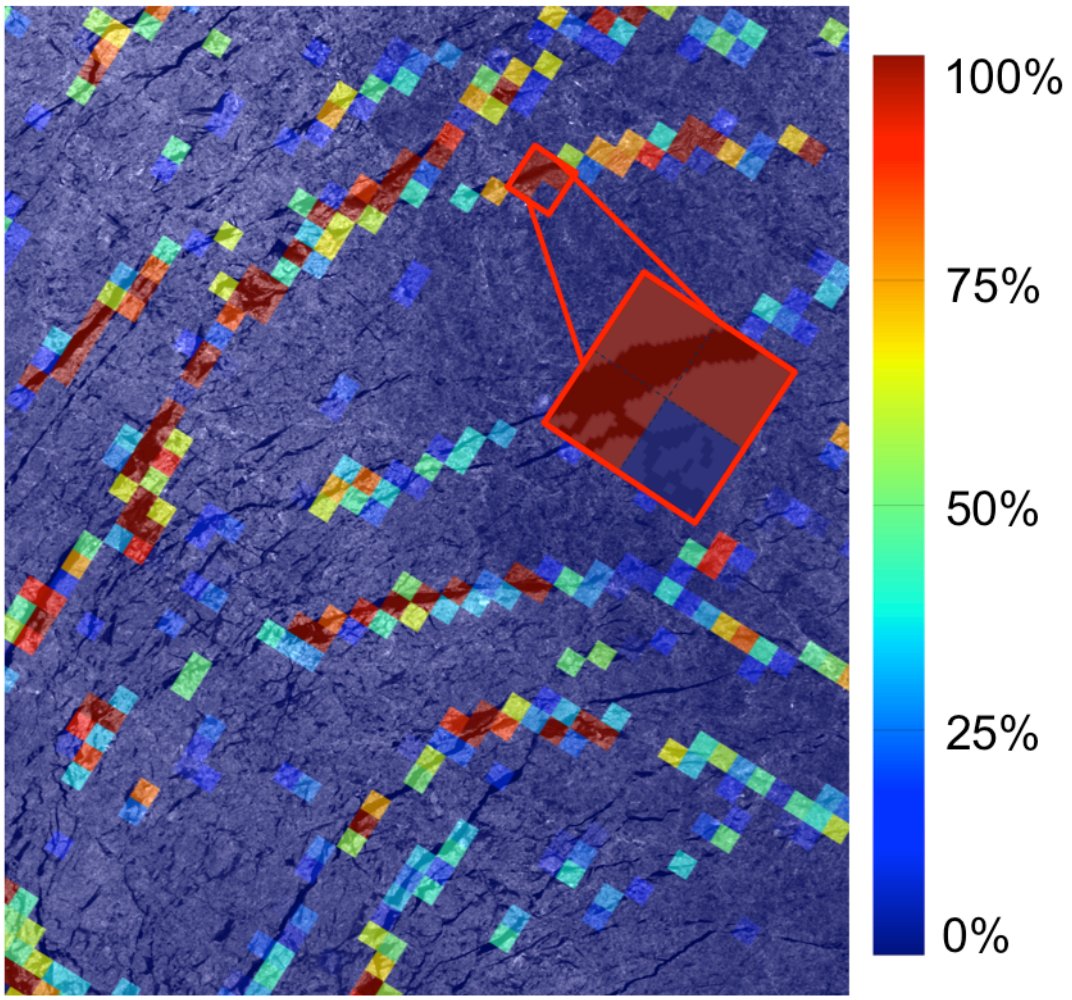


Figure 6. Subset of a SAR image taken on the 8 March 2009 overlaid by collocated AMSR-E lead fraction (LF) product, where red grid cells correspond to LF 100% (for the other values see the colour scale on the right). The zoom-in inset shows four grid cells where three of them have AMSR-E LF 100% and one has LF 0%.

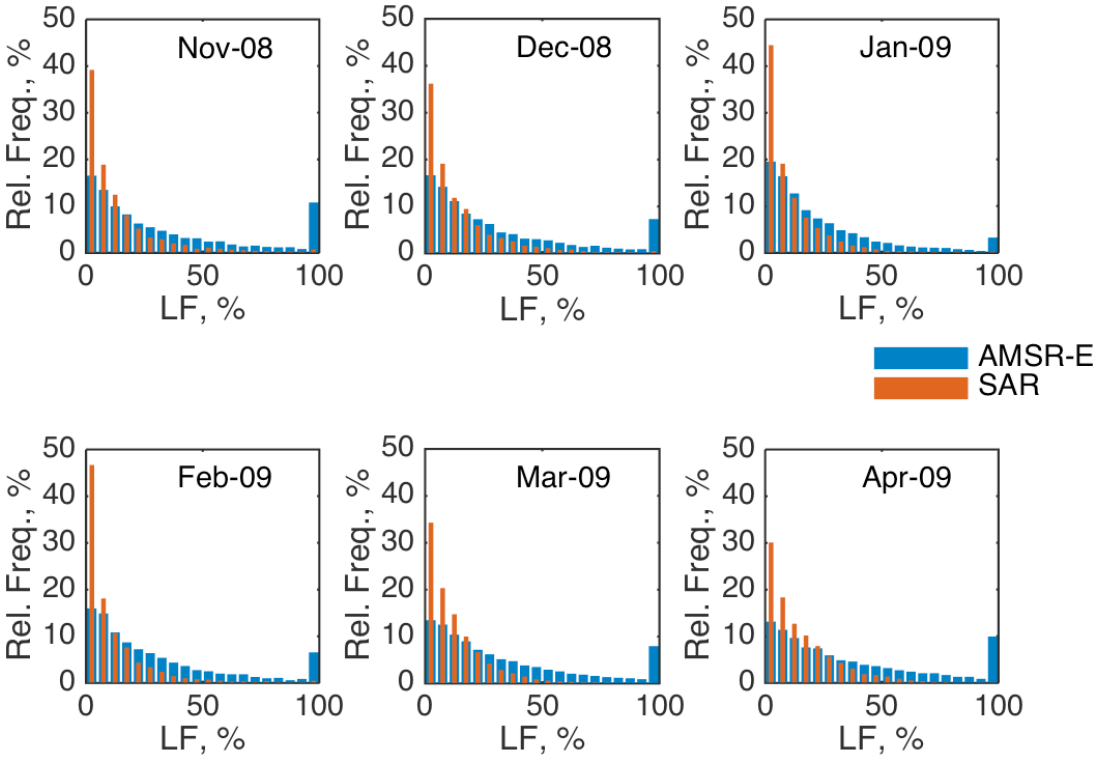
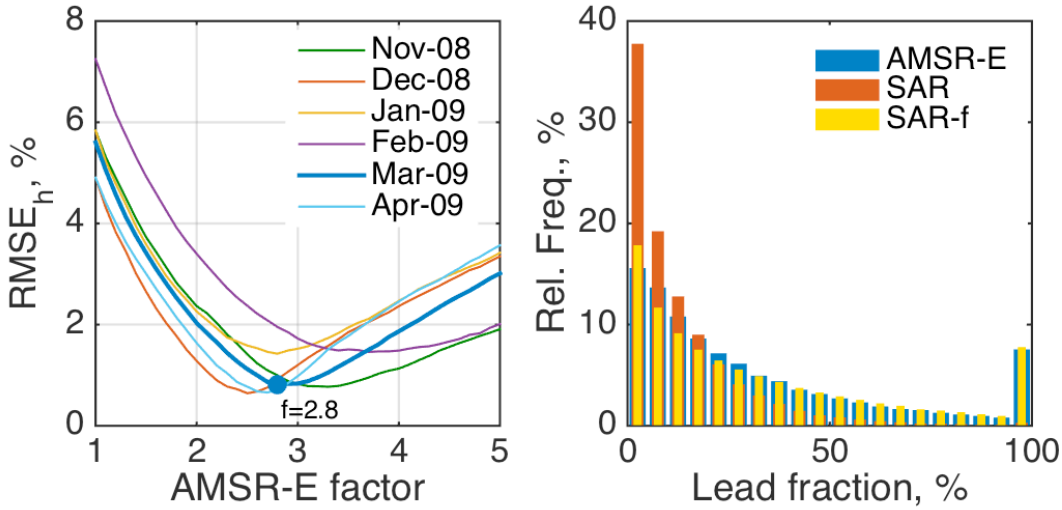


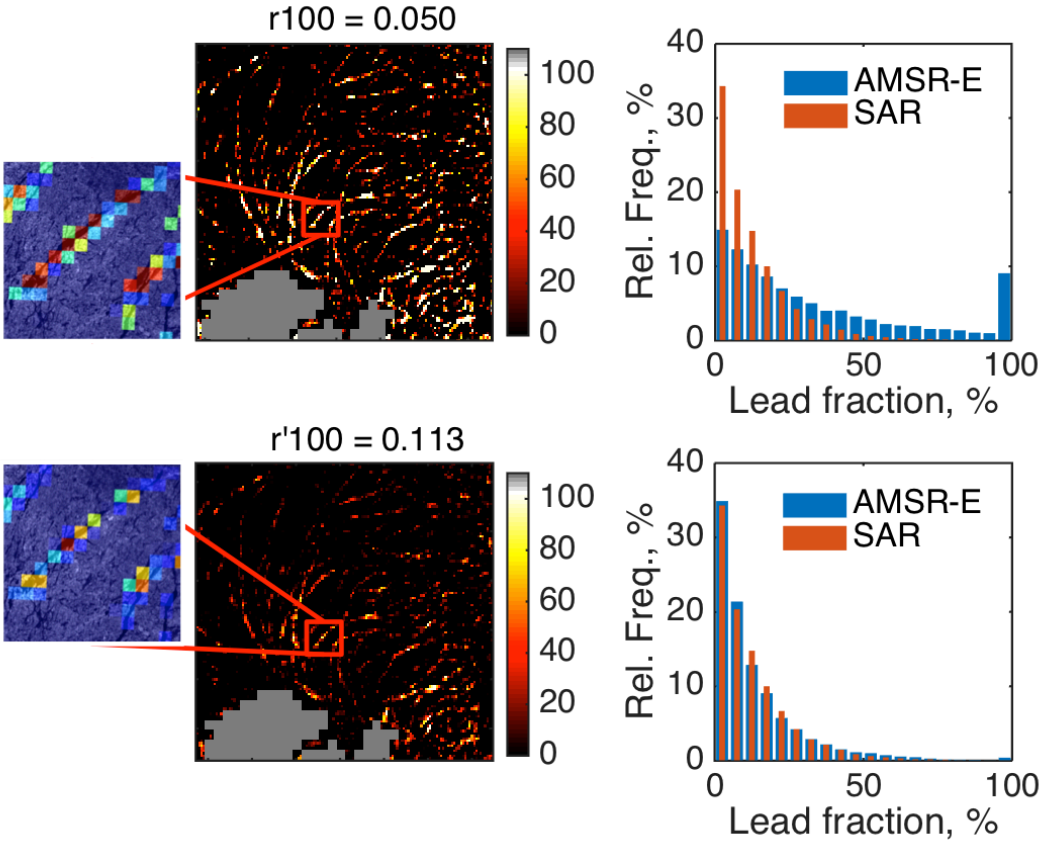
Figure 7. Histograms of the AMSR-E lead fraction (LF) and SAR LF datasets for every month from November 2008 to April 2009 shown as percentage of measurements per each bin of 5% width (relative frequency). Total amount of measurements amounts to 64 063 in 205 subsets size of 3500×3500 pixels each, and is given for each month in Table 1.



1

2 Figure 8. Left: Root mean square error ($RMSE_h$, %, Eq. (3)) as a measure of difference
3 between the histograms of AMSR-E lead fraction (LF) and SAR LF multiplied by different
4 values of f (AMSR-E factor). To demonstrate the principle March 2009 is highlighted by
5 bold blue line with minimum factor of 2.8. Right: Original histograms of AMSR-E LF and
6 SAR LF for the full winter November 2008 – April 2009, and SAR LF multiplied by
7 respective factor for each month (yellow bars).

8



1

2 Figure 9. Adjustment of upper tie point ($r100$) of the AMSR-E-based method. Upper panels:
3 a subset of lead fraction (LF) values located in the area of interest (Fig. 2) (left) and
4 distribution calculated from the full LF map (entire Arctic) on the 8 March 2009 (right, blue
5 bars). The original $r100$ value is used. The orange bars show SAR LF distribution for the
6 whole month of March 2009 for reference. Bottom panels: same, but for the adjusted $r'100$.

7

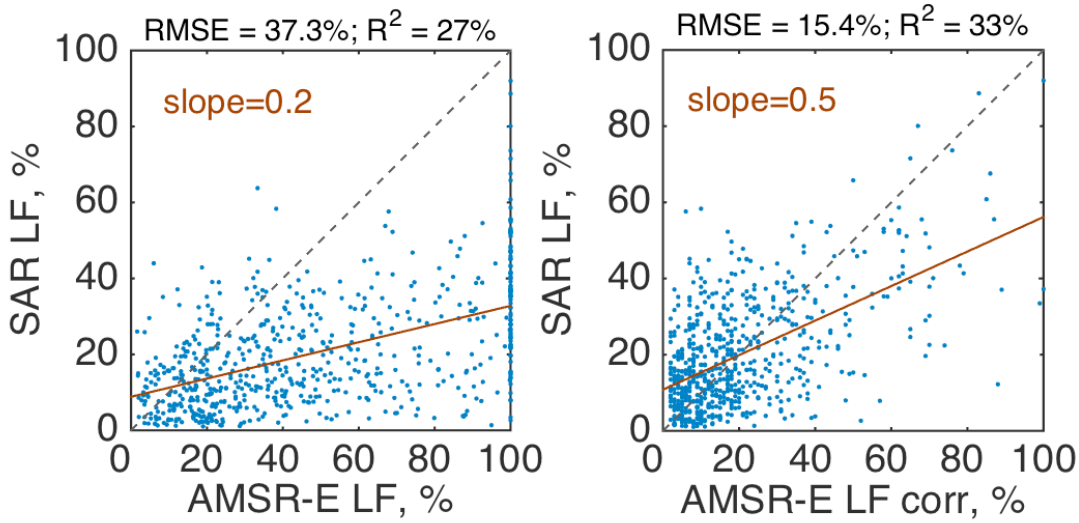


Figure 10. Left: scatterplot of SAR LF and AMSR-E LF using the original method of Röhrs et al. (2012). Right: same but using the adjusted AMSR-E method (new tie point). The 1-to-1 lines are the dashed grey lines, and the linear regressions are shown by the red lines and the values of slope. Root mean square error (RMSE) and the coefficient of determination (R^2) are shown at the top of the plots.