Response to the comments of reviewer #1 to

Assessment of error in satellite derived lead fraction in Arctic
by N. Ivanova, P. Rampal, and S. Bouillon

In the following we provide detailed answers (marked by "**A**") to your comments ("**R**") and describe what changes are done in the revised manuscript ("**Changes**").

R: The manuscript presents a valuable contribution to sea ice remote sensing techniques, the chosen analysis methods are sound, and the authors discuss the most important limitations and open questions of their study. I can therefore recommend it's publication in TC after some minor issues are resolved.

A: Thank you for such a positive response about our results, we value it greatly.

Minor Issues:

R: Both AMSR-E and SAR LF algorithms heavily rely on empirical tie-point or threshold values. How do you justify that one can be used as ground truth for validating the other?

A: This is a good point, and it should indeed be brought up more clearly in the revised version of the manuscript. We justify the use of the SAR-based method, firstly, by the combination of the fine resolution of this sensor (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. This allowed us to perform visual quality control of every pair SAR image – SAR lead fraction retrieval. An example is demonstrated in Fig. A1.1. Ambiguous images or images where many of leads were missed by the method were discarded.

However, even in the selected images there would still be some leads missed by the method presumably because the ice in them is thicker than the threshold allows. An example of lead fraction comparison is shown in the Fig. A1.1. Panel a) shows the original SAR image, panel b) shows the lead identification by the SAR-based method, where the red color corresponds to the identified leads, and the panel c) shows the SAR image overlaid by AMSR-E lead fraction original product (in %) with the color scale to the right. The features that are missed by SAR are relatively small, and are usually not captured by the coarse-resolution AMSR-E either (Fig A1.1-c). Please also note 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. In addition, we only performed the validation of AMSR-E lead fraction in locations where both datasets (SAR and AMSR-E) gave lead fraction > 0%, which excludes the cases of missed features in the reference SAR dataset.

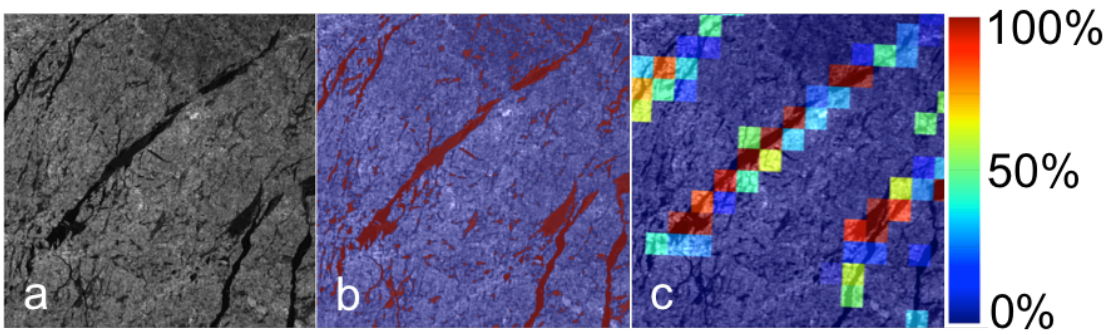

Figure A1.1. a) A subset of the original SAR image (backscatter) on the 8 March 2009, b) respective classified SAR-image (water/ice) with a) as background, c) AMSR-E lead fraction in % with a) as background.

Changes: The message provided above will be expressed more clearly in the revised manuscript.

R: Also, the data presented in Figure 4a raises the question if the SAR and AMSR-E LF algorithms measure the same quantity at all, or if both algorithms retrieve different sea ice properties.

A: This is a very relevant comment and a point that deserves to be discussed. First we would like to show one example that shows that SAR and AMSR-E LF algorithms measure similar quantity to some extent. And then we will discuss the possible differences in measured quantities.
To start, it should be mentioned that what we observed from comparison of SAR LF retrievals overlaid by AMSR-E LF product was that both identify the leads in the same locations. However, the amount of the lead area (lead fraction in percent) is a different story. Fig. A1.2 (left) shows similar scatterplot to Fig. 4 of the manuscript, except that it is only one-day example (8 March 2009 shown in the Fig. 8 of the manuscript).

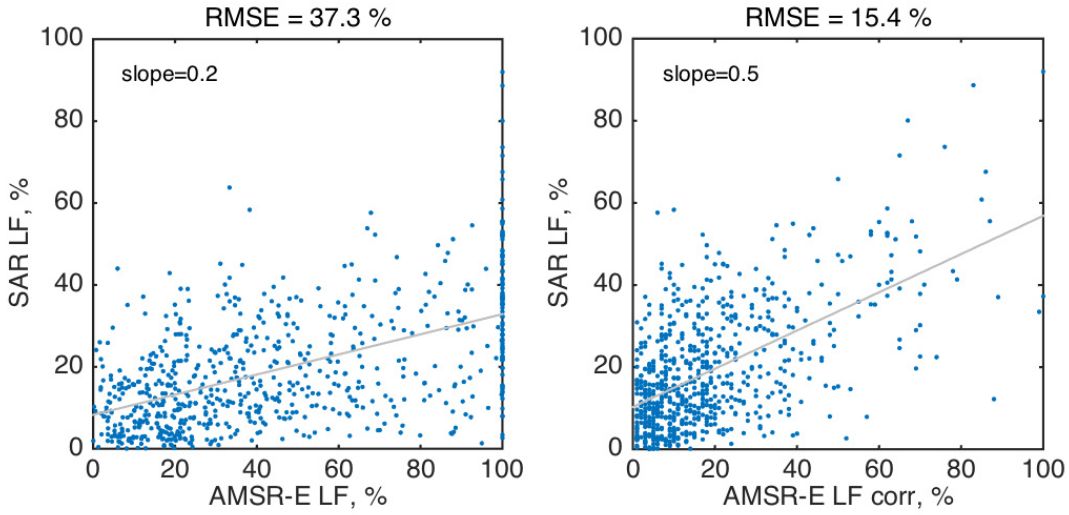

Figure A1.2. Left: scatterplot of SAR LF and AMSR-E LF using the original method of Rörs et al. (2012). Right: same but using the corrected AMSR-E method (new tie point).

The left panel of Fig. A1.2 compares the original AMSR-E lead fraction to that of the reference SAR dataset, while the right panel compares the AMSR-E lead fraction obtained with the new tie-point suggested in the study. The tie-point adjustment made the AMSR-E and SAR datasets agree significantly better: the RMSE improved from 37.3% to 15.4% (please note the numbers in the original manuscript are somewhat larger, this is a result of a mistake which is corrected now) and the slope of the regression line became closer to 1 (increased from 0.2 to 0.5). We believe this is an indirect indication of that the two methods do observe similar quantity.

However, as we mentioned in the beginning, this is the case only to some extent. The mechanisms that form the signal from an area with leads, represented by either open water or thin ice, are substantially different for SAR and AMSR-E. SAR is sensitive to roughness, while AMSR-E is sensitive to emitted brightness – these are two different physical parameters and they transfer into sea ice related quantities (thickness, type or concentration) in different ways. 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).

What we can see from visual inspection of SAR retrievals (including comparison to MODIS) is that if SAR identifies a feature, it retrieves its limits correctly. This is natural because ice in leads stays relatively smooth for a while and therefore SAR is still capable of interpreting it as a lead (except some cases discussed in the manuscript, which were excluded from the comparison).

For AMSR-E the situation is different: a relatively large area of ~40 km^2 (the grid spacing of the final product is 6.25 km × 6.25 km) is included in each grid cell and within it neither the borders of the leads nor different surface types (nilas, new ice, open water, etc.) are resolved, but a total amount of open water + thin ice in percent found in the grid cell is retrieved. The signal is a sum of all the surfaces present in the grid cell. The signature of thin ice (new ice, nilas, and pancake ice), expressed by the ratio of brightness temperatures used in the method (and based on Eppler et al., 1992), is quite distinct from thicker ice, and therefore the method is capturing the leads in correct locations, but calculation of the areal fraction of lead in the grid cell is based purely on an empirical constant obtained using MODIS images. The upper tie point, corresponding to LF = 100%, is found by selecting the ASMR-E grid cells that contain 100% thin ice + open water according to MODIS image and similarly the lower tie point is found for the cases with no leads present in the footprint. But there will be certain variance around the tie points due to variance in the brightness temperature ratio caused by presence of different surface types within the category of "lead" defined by both several types of thin ice and open water (Fig. 1 in Rörs et al., 2012), and sometimes it will be a mix of them. For example, the anomaly of this ratio (relative to the surrounding thicker ice) will be quite different between nilas and new ice, and the value of the anomaly for open water is even further away. They are all indeed above the background value, but the amount of this difference would vary depending on what type of thin ice is there (or open water).

Changes: We will address the issue in the discussion.

R: The second half of the manuscript lacks good structure. For example, the adjustment of the ASMR-E LF retrieval is somehow hidden in the discussion part. I would recommend to dedicate an own sub-section to the determination of the new tie-points, and give recommendations on how the algorithm should be used in the future. This is an important result that must be easy to find for someone that skims through

the paper.

A: The reason for presenting the AMSR-E-based method adjustment in the discussion section was that this adjustment was only a one-day case study, a test to demonstrate the possibility. We did not consider it as one of the main results of the paper because it was not completed on (and thus confirmed by) larger amount of data. It was rather a suggestion for future developments. However, we agree with the reviewer that it would be useful to make it more visible and easy to find.

Changes: The method adjustment will be moved to a new subsection with a title that will tell the reader that this is not final result. For example "future possibilities…", "suggestion…", "case study", etc.

R: page 6319 line 21: what do you mean by the phrase in parenthesis? Please reformulate this sentence and be clear what criteria the desired data set should meet, or which criteria is doesn't need to meet if that is important here.

A: We meant that our target is retrieving lead area (lead fraction) per grid cell as opposed to other studies that have already performed the validation of the lead locations and orientations in the AMSR-E dataset. They have shown that the AMSR-E product has leads in the correct places, but they have not looked at the amount of lead water/thin ice per grid cell as compared to other data sources quantitatively.

Changes: Will be clarified in the revised manuscript.

R: Equation 1: What is the scale of the median filter?
A: It is 5-by-5 pixels (p. 6322, l. 11)
Changes: "size of five pixels" will be changed to "5-by-5 pixels"

R: Is there a possibility that the SAR LF algorithm could oversee any leads?
A: please see our reply to the first item in the "Minor issues".

R: page 6330 line 8: Can you give a reference or other evidence for the power law behaviour?
Changes: The references that were meant here originally were Wernecke and Kaleschke (2015), and Marcq and Weiss (2012). However, this sentence (and other referring to the power law) will be removed from the revised manuscript because we do not find it justified connecting our reasoning to the distribution of lead width while addressing the lead fraction.

Technical corrections:

R: page 6317 line 14: instead of correlations, I think you mean a causal relation, or dependence.
A: Yes, this is correct. Thanks for notifying this. We will change the text to "causal relation".

R: A couple of articles ("The") are missing in the text.
Changes: The manuscript will be given to a native English-speaker for proofreading.

R: Change the title to “Error assessment of satellite derived lead fraction in the Arctic”
Changes: This is a good suggestion, we will do that.

R: Equation 1: If there is no vector cross product involved, use a simple dot instead of the cross.
Changes: Will be adjusted accordingly.

R: page6324, line 17: Remove “While”. It is not possible to form a sentence from to dependent clauses with no main sentence.
Changes: Will be adjusted accordingly.

R: page 6330 line 17: “define” is not suitable here. I would argue that you have determined the value for r100.
Changes: Will be adjusted accordingly.

References:
(References listed in the manuscript are not duplicated here)

Eppler, D., Farmer, L., Lohanick, A., Anderson, M., Cavalieri, D., Comiso, J., Gloersen, P., Garrity, C., Grenfell, T., Hallikainen, M., Maslanik, J., M¨atzler, C., Melloh, R., Rubinstein, I., and Swift, C.: Passive microwave signatures of sea ice, in: Microwave Remote Sensing of Sea Ice, edited by: Carsey, F. D., no. 68 in Geophysical Monograph, American Geophysical Union, 1992.

Response to the comments of reviewer #2 to

Assessment of error in satellite derived lead fraction in Arctic
by N. Ivanova, P. Rampal, and S. Bouillon

In the following we provide detailed answers (marked by "**A**") to your comments ("**R**") and describe what changes are done in the revised manuscript ("**Changes**").

R: General comments
In general, the paper is well written and structured, and easy to read and understand. The data processing and analysis methods are scientifically sound and discussed in needed detail.

A: Thank you for this positive feedback on our work, we appreciate it very much.

R: My main critic and concern in the paper is the method used for SAR based estimation of LF:
1) It is very simple, not up to current state-of-art in SAR based sea ice classification. The algorithm is basically a backscattering coefficient threshold method to separate leads with calm open ocean or thin ice (low sigma0) from sea ice (high sigma0). Leads with rough ocean or deformed thin ice are not identified, as the authors point out. The threshold method results a binary mask of leads vs. sea ice at 100 m pixel size which can be aggregated over a larger area to LF estimates. The threshold algorithm uses as input median filtered (5x5 pixel window) sigma0 values, which reduces the effect of radar speckle, but I think that still the radar speckle has some effect on the results if a sigma0 value is close to threshold causing randomness to which class, lead or sea ice, a pixel is assigned.
I suggest that the authors estimate the equivalent number of looks in their SAR imagery rectified to 100 m pixel size, yielding an estimate for radiometric resolution of the sigma0 (~10*log10(1+1/sqrt(ENL))), and study possible effects in the lead-sea ice classification.
The authors claim that no method has so far have been presented in literature addressing automatic LF retrievals from SAR. I would say that this may not be true as recent years papers have published on SAR based sea ice concentration (SIC) retrieval which may also be applied here, see e.g.:
Anders Berg and Leif E. B. Eriksson, "SAR Algorithm for Sea Ice Concentration—Evaluation for the Baltic Sea", IEEE Geoscience and Remote Sensing Letters, Vol. 9 (2012), 5, p. 938 - 942.
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. 2
Steven Leigh, Zhijie Wang, David Clausi. "Automated Ice-Water Classification Using Dual Polarization SAR Satellite Imagery," IEEE Transactions on Geoscience and Remote Sensing, vol 52, no 9, 2014.
Liu, Huiying; Guo, Huadong; Zhang, Lu, "SVM-Based Sea Ice Classification Using Textural Features and Concentration From RADARSAT-2 Dual-Pol ScanSAR Data", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,

8(4), pp 1601-1613, 2015.
I suggest that the authors conduct a thorough review of SAR based sea ice classification methods relevant to their study, and see if they can develop or apply a better method for the LF estimation. If you decide to continue use your current method then a firm detailed justification is needed.

A: The SAR-based LF retrieval method used in the paper is indeed very simplistic, and we understand the concern raised by the reviewer. We agree that we might have been not enough clear in explaining why we think the method can be used as it is for the goals set in this work. Therefore here and in the revised manuscript we will provide a more clear and detailed justification, as suggested by the reviewer. The two main points raised by the reviewer are: a) more sophisticated and up to current state-of-the-art SAR sea ice classification methods, available from literature, are not used; b) the effect of the speckle noise on the classification.
a) As the reviewer has pointed out there have been a few SAR-based sea ice classification methods published. What we wanted to say in our manuscript was that none of them has been applied to perform automatic lead fraction retrieval. We believe that every specific application of SAR-imagery processing requires individual approach. This stems from the fact that SAR provides indirect information about the surface properties we are interested in, which requires an individual approach for conversion of this indirect information (backscatter) to each physical parameter. Therefore one cannot assume that the mentioned methods will work for the case of identifying leads. For example, Korosov et al. (2015) have demonstrated that these thin features cannot be distinguished using an SVM-based method 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 the resolving capacity of the method, was not good enough, and it was found that the SVM needed to be trained specifically targeting leads. The same would apply to the method suggested by Berg and Eriksson (2012) where a neural network would need to be trained to fit the purpose of lead fraction retrieval and to be suitable for the Arctic sea ice, which is different in properties from that of the Baltic Sea. The method of Karvonen (2012) is also developed for the Baltic Sea, and its aim is to retrieve sea ice concentration, so it is not designed to separate open water from thin ice, which is important when working with leads in Arctic winter. The methods by Karvonen (2014), Leigh et al. (2014) and Liu et al. (2015) are all designed for dual-polarization SAR images, which we do not have access to. Sentinel-1 with its free access to the SAR images was not considered because of lack of overlap in time between Sentinel-1 (launched in 2014) and AMSR-E (stopped operating in 2011). We agree that it is possible to implement these methods, adapt them for our purpose and test their performance, but a task like this belongs to method development for SAR-based lead fraction retrieval, which is out of scope of this study. It would deserve a separate study and publication, and would definitely have been of great value. The focus of our paper is passive microwave retrieval of lead fraction. Such a dataset is highly valuable and therefore we set a goal to put uncertainty bars on it to the best of our knowledge and to look for possibilities to improve it. We consider our SAR lead fraction dataset to be good enough for this purpose for the reasons explained in the following answers.

b) Speckle noise is sufficiently reduced by the median filter applied in the suggested method, and the features are enhanced by it, as demonstrated in the Fig. 3a, b of the

manuscript. The panel b of the figure shows that the features are outlined quite precisely. From our visual inspection of the resulting retrievals (including comparison to MODIS) we saw that the noise did not affect the outcome largely. It is however still possible that a few pixels on the border of each feature are indeed misclassified – something we cannot capture with our eyes – and an uncertainty is introduced. However, if we take into account the fact that the AMSR-E-based method is expected to resolve relatively wide leads (> 3 km, Röhrs et al., 2012) and the retrieval has a grid size of 6.25 km × 6.25 km, while the grid size of our SAR-based retrieval is 100 m × 100 m, we assume this uncertainty is small in comparison to the difference between the AMSR-E and SAR retrievals especially for the bin of lead fractions near 100% for AMSR-E, which represent the main issue with the dataset. Another way speckle noise could affect the retrieval is providing single-pixel (or very small cluster of pixels) false leads, but such cases are excluded from the study by excluding AMSR-E LF < 1%. Such cases were not observed in the considered SAR imagery either: if SAR had such a small feature identified as lead, AMSR-E did not have a value there. And if AMSR-E had had a value, this would have indicated that it was not a result of speckle noise in the SAR image, but a true feature.

Changes: In the revised version of the manuscript we will provide the clarifications given above (shortened) and include an overview of the publications on SAR sea ice classification methods in the Introduction.

R: 2) The AMSR-E LF product accuracy is studied using an another LF product based on satellite data, this case SAR, and further, the SAR based method itself is not fully validated. The authors constructed a manually quality controlled sub-set of SAR based LF data for the AMSR-E LF accuracy studies. I suggest to study if you could add visual spectrum MODIS and MERIS data from March-April to further validate your SAR based LF estimation. At least larger leads should be visible in the optical imagery. During nighttime also 1 km MODIS ice surface temperature data could also useful (e.g. as by Willmes and Heinemann 2015). My work with these data have shown that manual quality control of cloud masking is unfortunately needed. The authors could also check availability of fine resolution Landsat etc. data over their study area.
In general, I am concerned about validating one remote sensing product with another product, which itself is fully mature and validated. Your study is more about intercomparison of two RS products and finding their differences, and you may not get solid information about the AMSR-E LF product accuracy against the true (in-situ) lead fraction. You claim that AMSR-E LF product is overestimating LF, but are you sure that your SAR based LF method is not underestimating, and thus, leading to this conclusion?

A: The reviewer questions the reliability of the reference SAR lead fraction dataset and suggests using MODIS, MERIS or Landsat products to validate the reference dataset further. As a consequence, the reviewer also points out that with our data we should not be able to quantify the errors in the AMSR-E dataset. Therefore, we would like to provide a clear step-by-step description of our approach, which will explain why we think that our reference dataset is acceptable for the purpose and the error estimations are useful. We might have been unclear in our original manuscript in this regard.

a) We conclude that the AMSR-E method overestimates lead fraction first from the visual analysis of the overlays of image pairs – SAR image and AMSR-E lead fraction retrieval – an example is shown in the Fig. 5 of the manuscript (the zoom-in) and explained in the text. We refer particularly to the last bin of the AMSR-E LF distribution, which contains large amount of observations close to 100% (Fig. 1 of the manuscript), 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.
b) In order to quantify this overestimation we retrieve LF from SAR. The SAR-based approach may indeed underestimate lead fraction, but we performed visual inspection of every SAR image, which indicated that we detected most of the leads. In addition, quality control is performed to exclude the ambiguous cases (see below). The residual underestimation seemed much smaller in comparison to the overestimation by AMSR-E clearly seen from the visual analysis of image pair overlays (AMSR-E LF and SAR images).
c) We make an assumption that the grouping of large amount of lead fraction values around 100% in the AMSR-E dataset (Fig. 1 of the manuscript) can be fixed by dividing the incorrect values by a certain factor. But some values are assigned 100% correctly, and since we do not know which ones, we approach the problem from the other side. We imitate the problem by multiplying the lead fraction retrieved from SAR (in all bins) instead, and we retrieve a distribution similar to that of the original AMSR-E dataset, which confirms our assumption. We find the best match for such factor varying from 2.5 to 3.7, but to allow for some uncertainty, we say it is between 2 and 4. So, already here we account for possible underestimation/overestimation by SAR.
d) The last clarification that needs to be provided here is the quality control of our SAR-based reference dataset.
We will first mention that we have performed a visual comparison of the SAR dataset to a MODIS image, and the SAR method captured the majority of features correctly. However the time difference between the images was 12 hours, and thus some of the features have changed their width. This will be an issue in general, if one would like to compare lead fraction from SAR and MODIS – it is extremely hard to find sufficient amount of image pairs which will be close in time, cover the same spatial area, have leads present in the footprint, and be cloud-free (MODIS) – all at the same time. In our visual comparison of this one pair of images there were indeed some smaller features that were missed by the SAR method. We chose however to restrain ourselves from performing of a validation of our SAR dataset versus MODIS (MERIS or Landsat), because quantification of lead fraction from such measurements requires another threshold technique, which will need validation too.
Therefore we did following to check the quality of the reference dataset:
- Making sure that our main dataset provides the same distribution as high quality dataset where every subset was visually analyzed and only certain cases selected (please see the description of the MQC SAR LF dataset).
- Visual evaluation of every triple of images: original SAR image, SAR lead fraction, AMSR-E lead fraction. Ambiguous cases were discarded (please find the details in the manuscript).
- We were only considering the AMSR-E grid cells (6.25 km by 6.25 km size) that had LF value > 1% and where the SAR LF for this grid cell returned a value of LF > 1% too. This means that we have excluded all the leads the SAR has eventually

missed. This is what we mean by that our SAR dataset is tailored to the purpose. It is not dependent on the AMSR-E dataset, but we only validate AMSR-E grid cells, which contain a non-zero lead fraction value.

Changes: We will make the description of our approach clearer in the revised version of the manuscript according to what is provided above.

R: 3) SAR based LF estimation is difficult, especially using single channel SAR imagery as in your case (ENVISAR C-band HH-pol WSM). I propose that you check if finer resolution ENVISAT APP images are available over your study are in Nov 2008 – Apr 2009. If that is the case then use APP based LF estimates to validate those from coarser resolution WSM imagery. In addition, are there any RADARSAT-2 dual-pol ScanSAR images available (through MyOcean or metno's Ice Service data archive)? Addition of cross-pol data could enhance lead detection.
How about ICESat data for SAR lead detection validation? ICESat data were acquired in Nov 2008 – Apr 2009?

A: For the time and area of interest we had neither APP or dual-pol images available. The ICESat data in the winter 2008–2009 were of low quality (K. Khvorostovsky, radar and laser expert at NERSC, personal communication) and were not used by other experts in the field either (e.g., Kwok et al: http://rkwok.jpl.nasa.gov/icesat/download.html; Yi and Zwally: http://nsidc.org/data/nsidc-0393).

R: In summary, I would like to see you study further accuracy of your SAR LF estimation method, as it is new one and presented for the first time in your paper.

A: Above we have provided justifications on the choices we made with regard to the accuracy of the SAR-based method. Here we would only like to add that this work was not aiming at development of a method for lead fraction retrieval from SAR, but rather focused on validating of non-zero measurements available in the AMSR-E dataset.

R: In the paper you mention/discuss about thin ice, but you don't define exact thickness range or WMO ice classes for the thin ice. 3

A: Since we were evaluating the existing lead fraction dataset from AMSR-E, we follow the definition of thin ice given in the original paper of Röhrs et al. (2012). There thin ice is defined according to World Meteorological Organization (WMO, 1989) as including new ice, nilas, and pancake ice.

Changes: the definition of thin ice will be included in the revised manuscript.

R: Specific comments
Page and line numbers refer here to the printer-friendly version of the article.

Introduction

R: Page 6317, line 1: "Model simulations showed that even 1% change in sea ice concentration due to the increase in areal lead fraction can lead to a 3.5 K difference

in the surface temperature"
Does this refer to sea ice surface temperature or surface air temperature?
A: Near-surface atmospheric temperature is meant here.
Changes: Adjusted accordingly.

R: p. 6317, l. 17: "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.
How about the exact lead locations, especially larger ones, are they important?
A: From a purely forecast perspective, we investigated that question and found that having the right location of leads at time zero is definitely important if one wants to get e.g. the right drift pattern in the following few days. However, the statistics of drift and deformation can be captured by only inserting statistical information about the leads like their amount and their sizes, almost independently of their location. This was only tested so far on simulations ran over the central Arctic, not for regions covering only the marginal seas like e.g. the Barents Sea.

R: l. 6318, l. 16: "However, this approach is limited in time coverage because AMSR2 started to deliver the data only in 2012 (http://suzaku.eorc.jaxa.jp), and quantitative validation work is still needed."
Quantitative validation work is needed - is this your opinion or found in a journal paper?

A: The paper of Beitsch et al. (2014) where this method is presented is targeting sea ice concentration and is not adapted to lead detection as for example is done (by deliberate enhancement of thin elongated features) in Röhrs et al. (2012) using similar instrument (AMSR-E). On the other hand, we understand the concern of the reviewer – the AMSR2 resolution used in the Beitsch et al. (2014) is finer than that of AMSR-E and this may mean that the enhancement is not necessary. However, our opinion is that quantitative validation is needed because Beitsch et al. (2014) only do visual assessment using three consecutive MODIS images (true color and ice surface temperature), which gives qualitative assessment, but no number reflecting the difference between the sea ice concentration at 3 km resolution and the MODIS image is obtained. A quantitative assessment is only provided for sea ice concentrations retrieved by different algorithms and at different resolutions from passive microwave data.

Changes: This point will be clarified in the Introduction.

R: p. 6318, l. 28: "The selected classifier was able to detect 68% of leads correctly, and only 3% of ice measurements were falsely identified as leads."
Why quite high fraction of leads were missed with CryoSat-2 data? This could be of interest to a reader.

A: Since our work is not addressing methods of lead fraction retrievals from CryoSat-2, we would prefer not to go into much detail on this aspect. We believe it is a combined result of several reasons, and interested readers can refer to the original paper of Wernecke and Kaleschke (2015), which we cite in the manuscript.

R: p. 6319, l. 26: "Based on analysis of the errors we introduce a correction factor for

the existing dataset and suggest an improvement of the AMSR-E based method itself."
The AMSR-E LF data cannot be corrected, but you propose a tie point correction to the LF algorithm itself? Please correct me, if I am wrong.
A: This is correct.
Changes: Adjusted accordingly.

2.1 The AMSR-E LF dataset

R: p. 6320, l. 12: "LF is expressed as the percentage of a grid cell covered by leads, which are represented by either open water or thin ice."
What is the thickness range of thin ice here?
Please explain shortly in this sub-Section the method for the AMSR-E based LF estimation.
The AMSR-E LF is a daily gridded product? This is not explicitly mentioned.
A: On the thickness please see our answer in the end of the general comment. Yes, it is a daily gridded product.
Changes: We will add 1-2 sentences describing the method (some details are already provided in the Introduction), and add that it is a daily gridded product.

2.1 The SAR images

R: p. 6321, l. 19: "The SAR images originally provided with spatial resolution of 75m.75m,"
This is pixel size of the WSM images, not the resolution. Provide resolution and noise floor, available from ESA docs. Investigate many equivalent of number looks are in your rectified SAR images, see my comments above.
A: Geometric resolution is 150 m ×150 m, pixel spacing 75 m × 75 m. On the noise, please see our reply to the general comment 1), with the noise floor the same reasoning applies.
Changes: The resolution will be specified.

R: p. 6321, l. 22: "Calibrated surface backscattering coefficient (ASAR Product Handbook, 2007) normalized over ice was used for this study (we will refer to this value as backscatter)."
Unclear what this normalization means, incidence angle variation compensation?
A: yes, it is empirical incidence angle variation compensation.
Changes: Will be specified in the manuscript.

3 SAR-based threshold technique

R: p. 6322, l. 22: "The threshold was first tried with nd=1 and nd = 2 (dashed red lines), but it was found that an intermediate value nd = 1.5 (solid red line) worked better and therefore was chosen."
How this was determined? Visually analyzing results with different nd? 4
A: Yes, visual analysis comparing the lead fraction retrievals with different threshold values.
Changes: Will be specified in the manuscript.

R: p. 6323, l. 3: "This method is developed strictly for the purpose of the AMSR-E

LF dataset validation and therefore does not represent an independent LF retrieval method from SAR."
I don't understand this, to my opinion your SAR LF should be independent of the AMSR-E LF dataset if you want to validate the AMSR-E LF.
What is the effect of SAR noise floor in your sigma0 threshold based LF method? Measured sigma0 is sum of target true sigma0 and SAR noise equivalent sigma0. Can the threshold in (1) be lower than the noise equivalent sigma0? If so then leads are not detected at all.
A: This is a problem of our formulation; we will adjust it in the revised version. What is meant here is that we only validate AMSR-E grid cells where a non-zero lead fraction is available. We also discard all the SAR images where the classification did not work because we are not developing a SAR-based method and therefore we do not need to make it work for all possible scenarios. We take each non-zero AMSR-E grid cell and "look" what is in there in a SAR image. The procedure is performed automatically on many images.

We know that the leads are detected because we can see it visually. The leads that are missed by SAR do not play any role because such areas are excluded (they do not participate in the comparison with AMSR-E). Mixed cases where only part of a lead is identified by SAR and the rest is not (which will make this AMSR-E grid cell eligible for the comparison), has almost not been encountered. Please note that the AMSR-E method is identifying only relatively wide leads (>3 km), for such size of the leads SAR was either identifying the whole lead or not identifying anything, but no partial identification has occurred.

Changes: The issue will be clarified in the manuscript.

4.1.1 MQC SAR LF

R: p. 6324, l. 13: "Defining a threshold locally not only eliminates significance of these effects, but it takes advantage also of less variety of surfaces in general."
How threshold definition was done, manually?
A: yes.
Changes: Will be clarified in the manuscript.

R: p. 6324, l. 20: "The classification in each subset was then inspected visually, comparing the three collocated maps: backscatter, MQC SAR LF and AMSR-E LF, in order to make sure it was successful."
Now your MQC SAR LF dataset is not independent of AMSR-E LF? If so then how this effects your SAR LF vs. AMSR-E LF comparison? I think your SAR LF should be totally independent of AMSR-E LF.
A: The MQC SAR LF dataset is not dependent on AMSR-E LF. This comparison of the triplets of the images was done with main focus on the checking of the MQC SAR LF against the original backscatter, while we also looked at the AMSR-E LF solely in order to exclude triplets where for example large areas of AMSR-E LF subset were missing. This last part was only for purposes of the dataset cleaning.
Changes: Will be clarified in the manuscript.

4.1.2 SAR LF

R: p. 6325, l. 5: "The majority of subsets contained leads represented by signatures

darker than surrounding background, while if those with brighter signature were present in large amount such images were discarded."
This is a serious drawback of your SAR based LF method, leads with high sigma0 are not detected. You have circumvented this by discarding SAR images with leads having high sigma0. Another, and better, way would be to further develop your algorithm.
A: Yes, if our purpose were to develop a SAR-based lead fraction retrieval method, this would have to be done in a better way.

5 Discussion

R: p. 6238, l. 20: "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, which allows further developments of the method."
It is good that you admit many limitations of your SAR LF method, but I am not very happy about circumventing this by stating that it is still suitable for your study. You could target more universal SAR LF algorithm to increase value of your paper. See my General comments.
A: Please see our answers above.

R: p. 6329, l .22: "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."
Yes, also to my experience cases when a sigma0 has a clear bimodal distribution are rare, as there are leads with low or high sigma0, smooth FYI with low sigma0 etc., i.e. the overall sea ice sigma0 distribution can have many local peaks.

R: p. 6331, l. 2: "For example, it would not be able to capture leads narrower than 3 km due to its resolution,...
This limitation should also mentioned in Introductory and 2.1 The AMSR-E LF dataset Sections.
A: It is mentioned in the Introduction (p. 6317 l. 28).

References:
(References listed in the manuscript are not duplicated here)

Korosov, A., Zakhvatkina, N., and Muckenhuber, S. Ice/water classification of Sentinel-1 images. EGU 2015, 17 April 2015.

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

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

Natalia 10/2/2016 16:11
Deleted: A

Natalia 10/2/2016 16:11
Deleted: of error

Natalia 10/2/2016 16:11
Deleted: in

Natalia 11/2/2016 14:45
Deleted: ,[2]

Natalia 11/2/2016 14:45
Deleted: ,[2]

Natalia 11/2/2016 14:46
Deleted: ,[2]

Natalia 11/2/2016 14:45
Deleted: [2]{Bjerknes Centre for Climate Research, Bergen, Norway}

Natalia 11/2/2016 14:46
Deleted: s

Natalia Ivanova 11/2/2016 18:18
Deleted: the

Natalia 11/2/2016 14:47
Deleted: Here we

Natalia 11/2/2016 14:48
Deleted: quantify

Natalia 11/2/2016 14:48
Deleted: , also introduced in this work

Natalia 11/2/2016 14:48
Deleted: We find a

Natalia 11/2/2016 14:49
Deleted: by a factor of 2–4 of the LF estimates

Natalia 11/2/2016 14:50
Deleted: We

Natalia 11/2/2016 14:50
Deleted: show for a data sample from the AMSR-E LF dataset

Natalia 11/2/2016 14:50
Deleted: s

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

Natalia 10/2/2016 12:52
Deleted: can

Natalia 10/2/2016 12:52
Deleted: surface

Natalia Ivanova 11/2/2016 18:29
Deleted: in the Arctic

Natalia Ivanova 11/2/2016 18:29
Deleted: ich

Natalia Ivanova 11/2/2016 18:30
Deleted: changed

Natalia Ivanova 11/2/2016 18:30
Deleted: in recent years with the shift towards

Natalia Ivanova 11/2/2016 18:29
Deleted: sea ice cover

Natalia 10/2/2016 16:16
Deleted: correlation

Natalia Ivanova 11/2/2016 18:33
Deleted: has been

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).

Natalia 10/2/2016 12:54
Deleted: , and

Natalia 10/2/2016 12:55
Deleted: is still

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

Natalia 11/2/2016 15:06
Deleted: (

Natalia 11/2/2016 15:06
Deleted:)

Natalia 11/2/2016 14:32
Deleted: not only

Natalia 11/2/2016 15:09
Deleted: , location or orientation

Natalia 10/2/2016 12:59
Deleted: introduce a correction factor for the existing dataset and

Natalia 10/2/2016 13:00
Deleted: n

Natalia 10/2/2016 12:59
Deleted: itsel

Natalia 10/2/2016 12:59
Deleted: f.

Natalia Ivanova 11/2/2016 18:49
Deleted: SAR-based

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

Natalia Ivanova 11/2/2016 18:51
Deleted: the

Natalia Ivanova 11/2/2016 18:51
Deleted: characteristic for

Natalia Ivanova 11/2/2016 18:51
Deleted: thin ice and open water

Natalia 11/2/2016 15:34
Deleted: as a

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).

Natalia 26/1/2016 14:42
Deleted: https://github.com/nansencenter/nansat

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).

Natalia 10/2/2016 13:14
Deleted: with

Natalia 10/2/2016 13:13
Deleted: 75

Natalia 10/2/2016 13:13
Deleted: 75

Natalia 10/2/2016 13:13
Deleted: ,

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.

Natalia 11/2/2016 14:34
Deleted: five

Natalia 11/2/2016 12:29
Comment [1]: Please note change in the equation

Natalia 11/2/2016 15:37
Deleted: For reference

Natalia 11/2/2016 15:37
Deleted: t

Natalia 11/2/2016 15:37
Deleted: 0.

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.

Natalia 10/2/2016 14:38
Deleted: This method is developed strictly for the purpose of the AMSR-E LF dataset validation and therefore does not represent an independent LF retrieval method from SAR. Its limitations and potential of further development for wider applications are 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).

Natalia Ivanova 11/2/2016 19:01
Deleted: validation

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.

Natalia Ivanova 11/2/2016 19:03
Deleted: located

Natalia Ivanova 11/2/2016 19:03
Deleted: places

Natalia Ivanova 11/2/2016 19:03
Deleted: both

Natalia Ivanova 11/2/2016 19:03
Deleted: retrieved

Natalia Ivanova 11/2/2016 19:04
Deleted: LF

Natalia Ivanova 11/2/2016 19:06
Deleted: Therefore

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

Natalia 11/2/2016 15:47
Deleted: five

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). 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. 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

Natalia Ivanova 11/2/2016 19:07
Deleted: less

Natalia 10/2/2016 16:18
Deleted: While

Natalia 10/2/2016 14:45
Deleted: three

Natalia 10/2/2016 14:45
Deleted: ,

Natalia 10/2/2016 14:45
Deleted: and AMSR-E LF

Natalia 10/2/2016 14:45
Deleted:

Natalia 10/2/2016 11:24
Deleted: and collocated AMSR-E LF product

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.

Natalia Ivanova 11/2/2016 19:12
Deleted: if

Natalia Ivanova 11/2/2016 19:13
Deleted: those

Natalia Ivanova 11/2/2016 19:13
Deleted: were present in large amount such images

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.

Natalia Ivanova 11/2/2016 19:14
Deleted: two

The AMSR-E LF and MQC SAR LF datasets are shown in Fig. 4 as a scatterplot (left) and histograms (right). The scatterplot shows that the majority of the points are located below the 1-to-1 line, which means that in most cases AMSR-E LF overestimates the LF as compared to the SAR retrievals. 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 right panel of Fig. 4 shows histograms of the two datasets representing number of measurements per each 5%-bin expressed in % of total number of measurements (1645 in this case). 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. 5 shows part of a SAR image overlaid by collocated AMSR-E LF product, where one can see general overestimation of LF by AMSR-

Natalia 11/2/2016 15:52
Deleted: largely

Natalia 10/2/2016 11:54
Deleted: From Fig. 1 one could assume that the amount of measurements in this bin should be smaller than in the previous bin following the gradual decline of the distribution (accordingly to the power-law distributions suggested by Wernecke and Kaleschke, 2015 and Marcq and Weiss, 2012), so that there is a much larger amount of smaller leads as compared to large ones.

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. 6). 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. 4 (left) and Fig. 5, where values of LF 100% in AMSR-E dataset correspond to variety of values from SAR dataset. Ideally, an

Natalia 10/2/2016 11:41
Deleted: .

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 root mean square error (RMSE) as a measure of difference between the histograms of AMSR-E LF and SAR LF datasets for each month (Fig 6.):

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

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. 7 (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. 7, 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. 6, 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 $100 \times \left|\left(\overline{LF_{AMSRE}} - \overline{LF_{SAR}}\right)/\overline{LF_{SAR}}\right|$ decreases from 140% with no correction to 17% when using the correction factors found here.

Natalia 10/2/2016 12:42
Deleted: Discussion s

Natalia 10/2/2016 12:42
Deleted: ion

Natalia 11/2/2016 15:59
Deleted: improvement of the method and

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:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(LF_{AMSRE\,i} - LF_{SAR\,i}\right)^2}, \quad (3)$$

where n is the total number of measurements (64 063). 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. 8, upper left), i.e. with the upper tie point $r100 = 0.05$. Such calculations give similar distribution of LF values (Fig. 8, 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. 8, bottom right) – the value of $RMSE_h$ (Eq. (2)) decreasing from 5.4% (corresponding to $f = 1$ in Fig. 7, left) to 0.9%, while point-wise RMSE (Eq. (3)) for this one-day dataset of 750 collocated LF measurements decreases from 37% to 15%. The close-up insets similar to the one in Fig. 5 show that the leads are identified in the same locations as before, but the LF values are lower (Fig. 8, bottom left). 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

Natalia 11/2/2016 15:56
Deleted: (see the Discussion section)

Natalia Ivanova 11/2/2016 19:25
Deleted: In order t

Natalia Ivanova 11/2/2016 19:26
Deleted: threshold

Natalia Ivanova 11/2/2016 19:29
Deleted: value of $r100$

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-

Natalia Ivanova 11/2/2016 19:29
Deleted: $r100$

Natalia 10/2/2016 12:44
Deleted: ,

Natalia 10/2/2016 12:44
Deleted: which a

Natalia 10/2/2016 12:44
Deleted: s

Natalia 10/2/2016 12:45
Deleted: the

Natalia 10/2/2016 12:45
Deleted:

Natalia Ivanova 11/2/2016 19:30
Deleted: While

Natalia 15/2/2016 08:50
Deleted: used

Natalia Ivanova 11/2/2016 19:32
Deleted: n't

Natalia Ivanova 11/2/2016 19:32
Deleted: 0.

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.

Natalia Ivanova 11/2/2016 19:34
Deleted: is

Natalia 11/2/2016 14:30
Deleted: A validation dataset is produced using this method in order to quantify errors in AMSR-E LF estimates. However, these error estimates should be considered as rather preliminary, because the AMSR-E LF product in its current form cannot be fairly compared to a validation dataset. We identify an issue related to near 100% LF values in the AMSR-E LF dataset: they occur very often, which is neither observed in the SAR datasets or conforms to the power law model usually assumed as decribing lead width distribution well. Based on these findings and the basics of the ASMR-E method, we make an assumption that the upper tie point in the method should be increased in order to cover the full range of LF values. In order 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. 8, upper left), i.e. with the upper tie point $r100 = 0.05$. Such calculations give similar distribution of LF values (Fig. 8, upper right) as was found in the full AMSR-E dataset (Fig. 1). Using the linear relationship between $r100$ and f, and the optimal value of f for March 2009 ($f = 2.8$), we define that $r100$ should be increased to 0.113. This new threshold value gives a distribution closer to SAR LF dataset (Fig. 8, bottom right) – the value of $RMSE_h$ (Eq. (2)) decreasing from 5.4% (corresponding to $f = 1$ in Fig. 7, left) to 0.9%, while point-wise RMSE (Eq. (3)) for this one-day dataset of 750 collocated LF measurements decreases from 45% to 23%. The close-up insets similar to the one in Fig. 5 show that the leads are identified in the same locations as before, but the LF values are lower (Fig. 8, bottom left). We thus believe that implementation of such an adjustment to the full AMSR-E LF dataset will lead t ... [1]

Natalia Ivanova 11/2/2016 19:35
Deleted: sum

Natalia 11/2/2016 14:08
Deleted: .

Natalia Ivanova 11/2/2016 19:38
Deleted: are

Natalia Ivanova 11/2/2016 19:39
Deleted: Additional benefit of the improved and validated AMSR-E LF dataset would be a possibility to refine sea ice concentration datasets. An improved sea ice concentration dataset for Arctic winter can be produced by implementing ASI algorithm for AMSR2 brightness temperatures, which in itself is more sensitive to the leads than other sea ice concentration algorithms (Beitsch et al. ... [2]

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.

Natalia Ivanova 11/2/2016 19:39
Deleted: n

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 $100 \times \left| \left(\overline{LF_{AMSRE}} - \overline{LF_{SAR}} \right) / \overline{LF_{SAR}} \right|$ 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.

Natalia 11/2/2016 12:28
Deleted: 45

Natalia 11/2/2016 12:28
Deleted: 23

Natalia Ivanova 11/2/2016 19:42
Deleted: were

Natalia Ivanova 11/2/2016 19:43
Deleted: –

Acknowledgements

The research was supported by the Centre for Climate Dynamics at the Bjerknes Centre. The authors would like to thank Lars Kaleschke (Institut für Meereskunde, KlimaCampus, University of Hamburg, Germany) for useful discussions and Anton Korosov (Nansen Environmental and Remote Sensing Centre, Bergen, Norway) 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.

Natalia 10/2/2016 09:54
Deleted: A.

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.

Natalia 10/2/2016 09:57
Formatted: Space Before: 6 pt, Line spacing: 1.5 lines, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

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

Natalia 11/2/2016 15:03
Deleted: Rampal, P., Bouillon, S., Olason, E., and Morlinghem, M.: neXtSIM: a new Lagrangian sea-ice model, submitted to The Cryosphere, 2015.

Natalia 8/1/2016 12:31
Deleted: Discuss.

Natalia 8/1/2016 12:31
Deleted: 2167

Natalia 8/1/2016 12:31
Deleted: 2200

Natalia 8/1/2016 12:31
Deleted: d

Natalia 8/1/2016 12:31
Deleted: 2167

Natalia 10/2/2016 12:50
Formatted: No widow/orphan control, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Natalia 10/2/2016 12:50
Formatted: Font color: Black, German

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

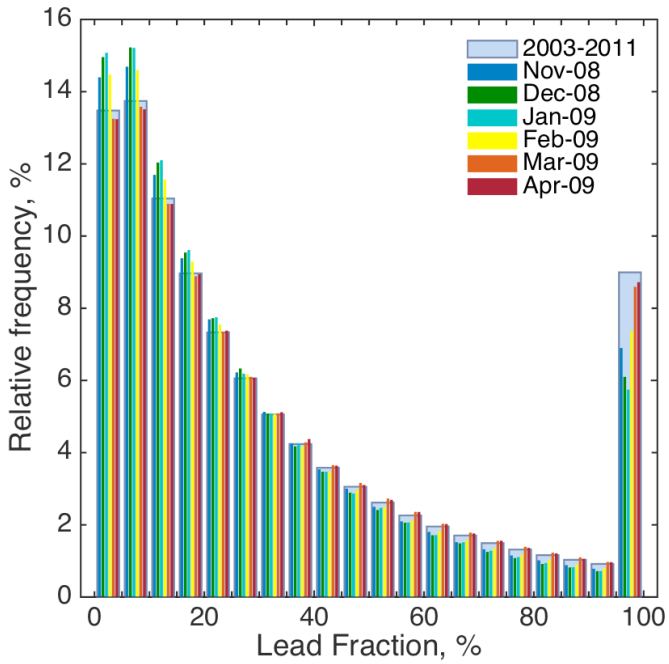

Figure 1. Histograms for AMSR-E lead fraction (LF) dataset shown as the number of measurements per each LF bin of 5% width expressed in % of the total amount of measurements (relative frequency). The blue bars show the full dataset, while each month of the winter 2008–2009 is shown by other colours (see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

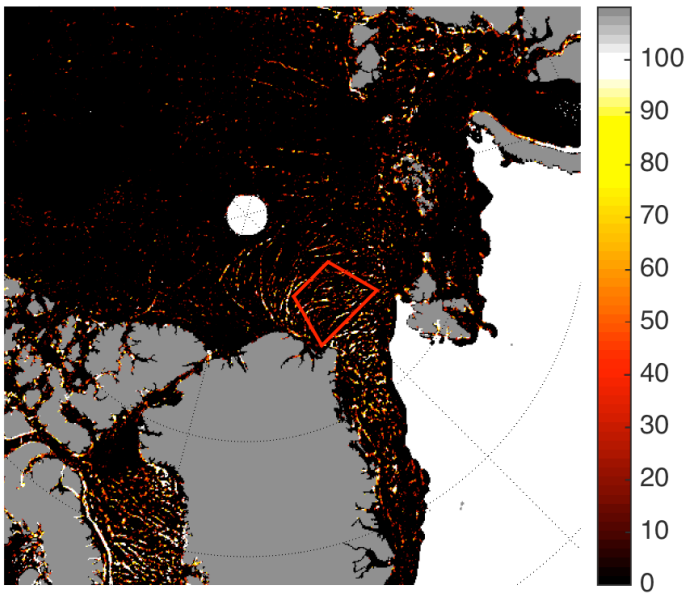

Figure 2. Area of interest is included within the red rectangle. The background map shows AMSR-E lead fraction in % (the numbers on the colour scale to the right), obtained on the 8 March 2009, and is used here only to demonstrate a sample from the product.

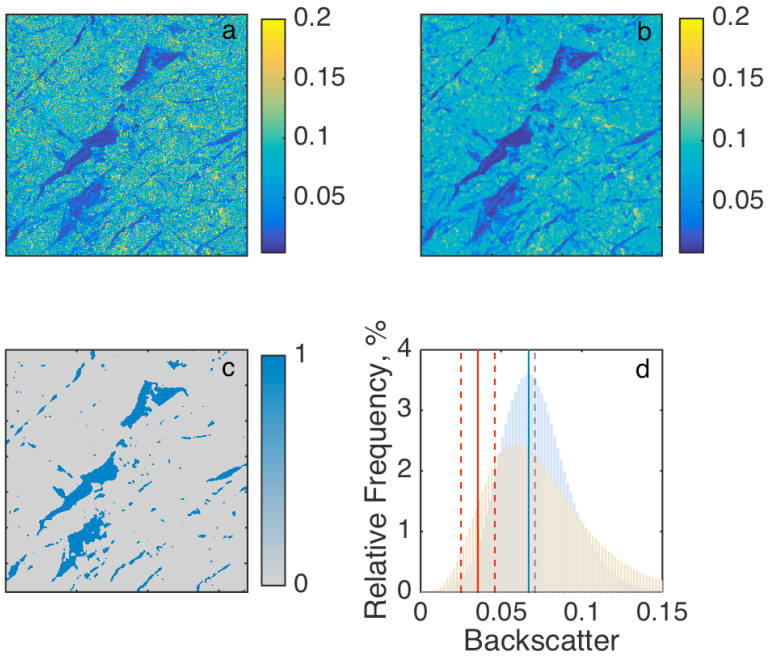

Figure 3. Threshold technique used to calculate lead fraction from SAR images: a) a subset of 680×680 pixels showing backscatter values; b) same as a) but after median filter has been applied; c) the resulting lead detection (1 – lead, 0 – ice); d) histogram of an example SAR scene taken on the 1 March 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.

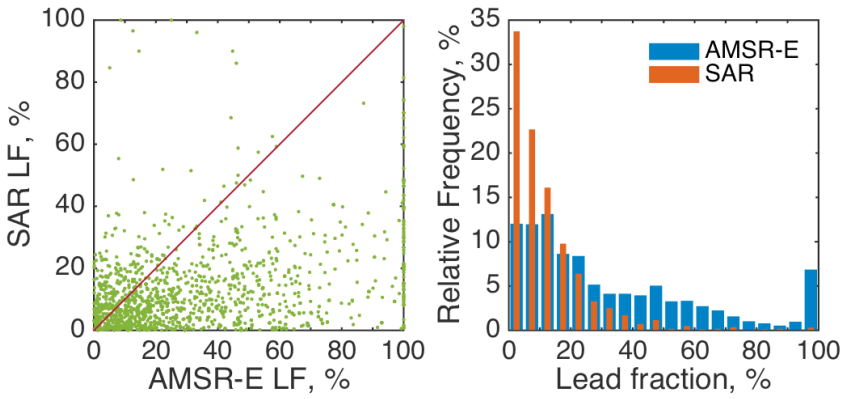

Figure 4. A comparison of AMSR-E lead fraction (LF) and SAR LF with manual quality control (MQC), the total amount of measurements is 1645. Left: scatterplot of MQC SAR LF versus AMSR-E LF (%). The 1-to-1 line is shown in dark red. Right: histograms for the two datasets shown as percentage of measurements per each bin of 5% width (relative frequency).

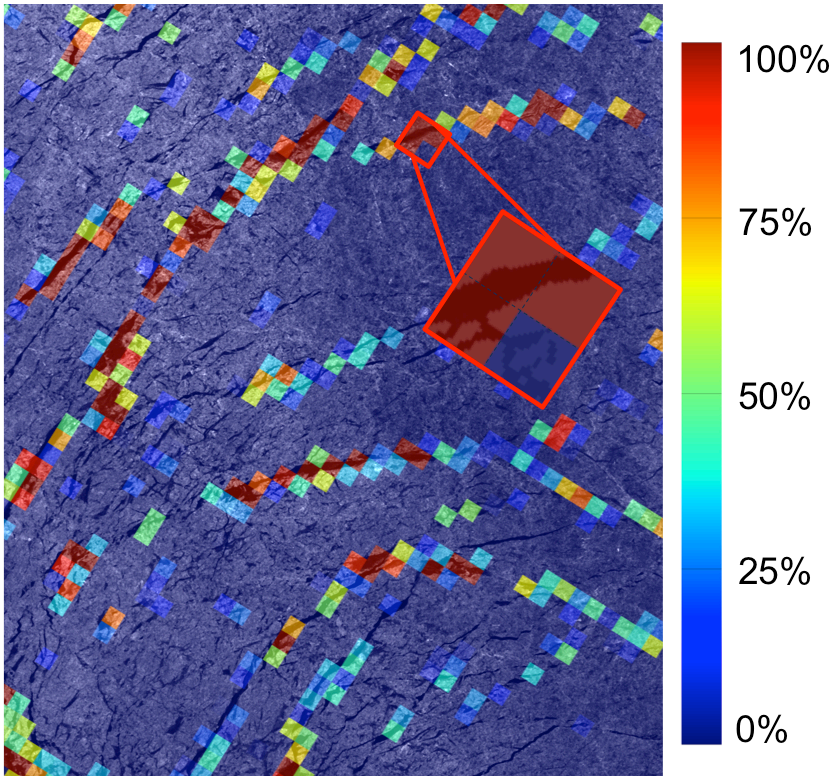

Figure 5. Subset of a SAR image taken on the 8 March 2009 overlaid by collocated AMSR-E lead fraction (LF) product, where red grid cells correspond to LF 100% (for the other values see the colour scale on the right). The zoom-in inset shows four grid cells where three of them have AMSR-E LF 100% and one has LF 0%.

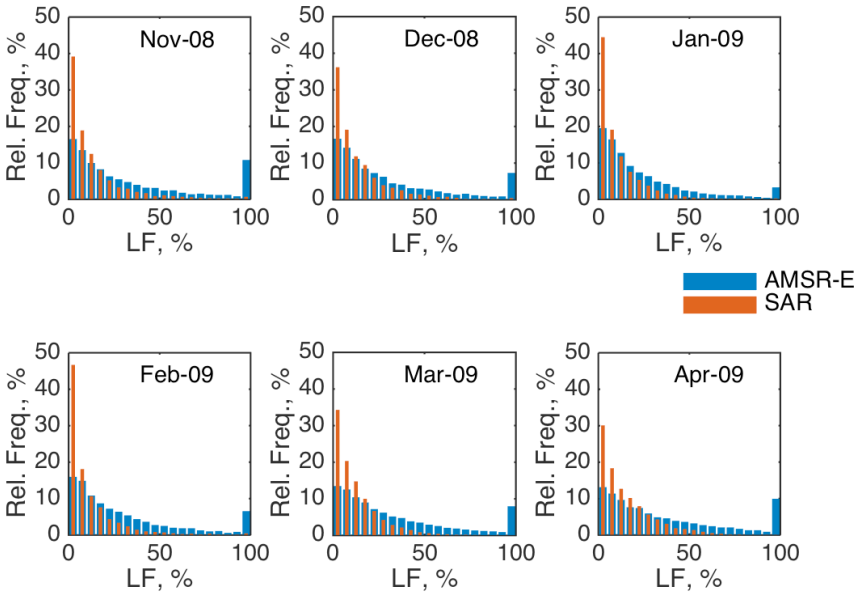

Figure 6. Histograms of the AMSR-E lead fraction (LF) and SAR LF datasets for every month from November 2008 to April 2009 shown as percentage of measurements per each bin of 5% width (relative frequency). Total amount of measurements amounts to 64 063 in 205 subsets size of 3500×3500 pixels each, and is given for each month in Table 1.

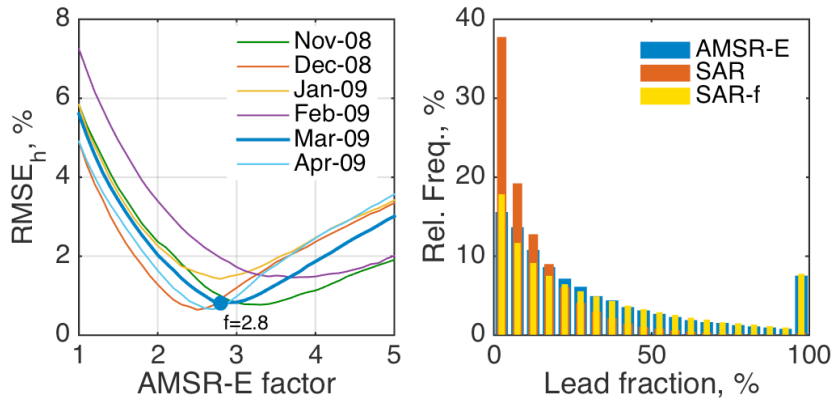

Figure 7. Left: Root Mean Square Error ($RMSE_h$, %, Eq. (2)) as a measure of difference between the histograms of AMSR-E lead fraction (LF) and SAR LF multiplied by different values of f (AMSR-E factor). To demonstrate the principle March 2009 is highlighted by bold blue line with minimum factor of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

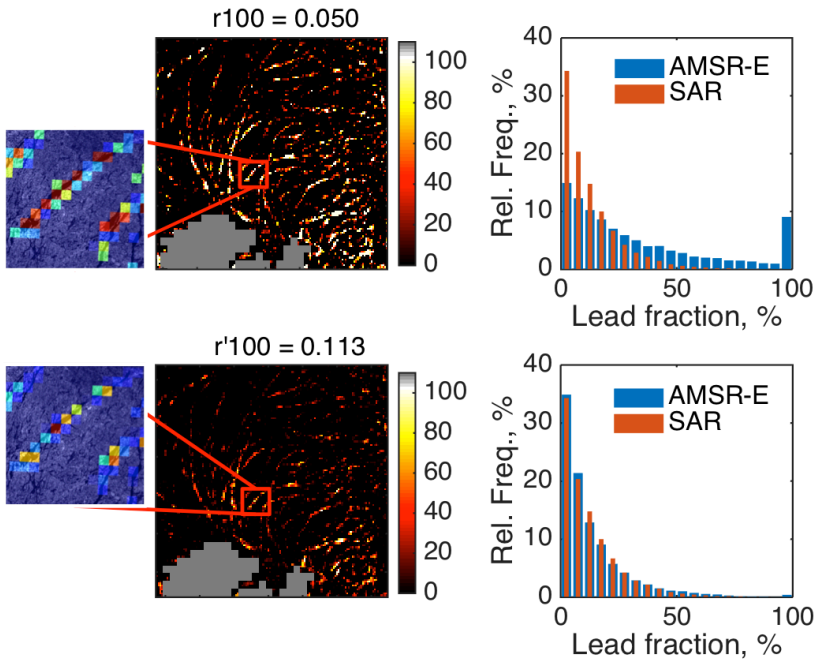

Figure 8. Adjustment of upper tie point ($r100$) of the AMSR-E-based method. Upper panels: a subset of lead fraction (LF) values located in the area of interest (Fig. 2) (left) and distribution calculated from the full LF map (entire Arctic) on the 8 March 2009 (right, blue bars). The original $r100$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

Natalia 11/2/2016 16:19

r10

r'1(

Deleted:

Page 15: [1] Deleted Natalia 11/02/16 14:30

A validation dataset is produced using this method in order to quantify errors in AMSR-E LF estimates. However, these error estimates should be considered as rather preliminary, because the AMSR-E LF product in its current form cannot be fairly compared to a validation dataset. We identify an issue related to near 100% LF values in the AMSR-E LF dataset: they occur very often, which is neither observed in the SAR datasets or conforms to the power law model usually assumed as decribing lead width distribution well. Based on these findings and the basics of the ASMR-E method, we make an assumption that the upper tie point in the method should be increased in order to cover the full range of LF values. In order 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. 8, upper left), i.e. with the upper tie point $r100 = 0.05$. Such calculations give similar distribution of LF values (Fig. 8, upper right) as was found in the full AMSR-E dataset (Fig. 1). Using the linear relationship between $r100$ and f, and the optimal value of f for March 2009 ($f = 2.8$), we define that $r100$ should be increased to 0.113. This new threshold value gives a distribution closer to SAR LF dataset (Fig. 8, bottom right) – the value of $RMSE_h$ (Eq. (2)) decreasing from 5.4% (corresponding to $f = 1$ in Fig. 7, left) to 0.9%, while point-wise RMSE (Eq. (3)) for this one-day dataset of 750 collocated LF measurements decreases from 45% to 23%. The close-up insets similar to the one in Fig. 5 show that the leads are identified in the same locations as before, but the LF values are lower (Fig. 8, bottom left). 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 value of $r100$ 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 $r100$ weighted by the number of observations for each month is 0.117 and is therefore our best estimate for winter 2008–2009.

Page 15: [2] Deleted Natalia Ivanova 11/02/16 19:39

Additional benefit of the improved and validated AMSR-E LF dataset would be a possibility to refine sea ice concentration datasets. An improved sea ice concentration dataset for Arctic winter can be produced by implementing ASI algorithm for AMSR2

brightness temperatures, which in itself is more sensitive to the leads than other sea ice concentration algorithms (Beitsch et al. 2014) and then refining it by accommodating the LF dataset. To achieve even better accuracy, the LF dataset in this case should also be implemented for the 3.125 km resolution AMSR2 brightness temperatures.