

## Authors' responses (TC-2017-170)

The authors would like to thank the editor and the reviewers for their precious time and invaluable comments. The corresponding changes and refinements are highlighted in yellow in the revised paper and are also summarized in our responses below. Authors' responses are in blue. Reviewer's comments are in black. When the manuscript is cited, it is shown in italics.

### Editor's comments

I strongly agree with referee #3 comment that the discussion on the relationship between Arctic lead fraction, thermodynamics and ice dynamics is too superficial (this point was also raised by referee #1's report on the initial submission). I suggest to either remove section 5.3 (possible integrating some of the text in other parts of the manuscript) or provide a more thorough analysis. In case of the latter, I might contact an additional expert in this field to comment on your analysis

→ Thank you for your comment. As suggested, we removed section 5.3 and moved some of the text to section 4.2 (L14 P21-28).

*"The Arctic Ocean circulations have contributed to the change in the state of sea ice. The lead fraction in Northwestern Greenland in Figs. 6 and 7 is low because of the convergence of sea ice by two major circulations, which was clearly shown in Kwok (2015). Kwok et al. (2013) revealed that the currents speed of Beaufort Gyre and Transpolar Drift increased from the years of 1982 to 2009 and this makes the fraction of multi-year ice decrease. However, the increasing lead fraction from the years of 2011 to 2016 in this study was not seen due to the high inter-annual variability of a lead fraction, particularly in the spring season (Figs. 6 and 7). High sensitivities in the marginal sea ice zone might result in not catching the increasing trend of Arctic lead fraction shown in the literature. In order to properly compare the Arctic current circulations and lead fraction, long-term lead fraction data are needed."*

→ We also changed the third objective in the introduction (P2 L31-32), accordingly.

From "briefly examine the relationship between Arctic lead fraction and thermodynamics and ice dynamics."

To "3) examine the spatiotemporal distribution of lead fractions."

In addition to the suggestions of the referees, I ask you to make the following changes:

1. On p.25 remove "which implies that recent ice status has become more vulnerable to anomalous atmospheric and oceanic conditions." Given the short time span of the observations, this is speculative.

→ Thank you for your comment. We revised the sentence (P23 L10-11).

## Authors' responses (TC-2017-170)

*"In addition, this study showed the high inter-annual variability of Pan-Arctic lead fractions in recent years (i.e., 2011-2016)."*

2. In section 5.4, on novelty and limitations, justify why you present method should be used rather than the method proposed Lee et al., 2016.

➔ Thank you for your comment. We added a brief justification of the use of the waveform mixture algorithm (P23 L8-10).

*"The proposed waveform mixture algorithm would be very useful in an operational system than the threshold-based methods including Lee et al. (2016)."*

3. Include a discussion on the difficulties in detection leads in the summer season, as in your response to R3 on the P5L25 comment.

➔ We added the discussion (P14 L3-7).

*"The period from June to September is generally considered as the melting season. In this season, the presence of leads as well as melt pond in sea ice are dominant. It is difficult to accurately distinguish leads from sea ice due to the fact that waveform of the melt pond is quite similar to that of leads. Since the lead detection methods for the retrieval of sea ice thickness do not work well in the melting season, the sea ice thickness during the melting season is still unavailable (Tilling et al., 2017)."*

Technical correction:

Rephrase "CryoSat-2 observations less than 5 are in the 10 km grid around the coastline of Arctic" to, for example, "less than 5 observations are found in the 10 km grids in the marginal zones of the Arctic ocean"

➔ We revised the sentence as suggested.

## Authors' responses (TC-2017-170)

The authors would like to thank the editor and the reviewers for their precious time and invaluable comments. The corresponding changes and refinements are highlighted in yellow in the revised paper and are also summarized in our responses below. Authors' responses are in blue. Reviewer's comments are in black. When the manuscript is cited, it is shown in italics.

### Reviewer #2:

The authors have worked hard to address my original comments. The paper is much improved as:

- The method is expanded and better structured
- The authors no longer oversell the evaluation results from comparing different lead classification methods with MODIS imagery. For example, I now agree with the abstract statement that waveform mixture analysis “comparable and promising performance: compared with other lead detection methods, rather than the “better performance” stated before.
- Display items are clear

However, I have a few remaining comments that should be addressed. Please see below.

P6 L12-14: The highlighted addition still doesn't explain which particular orbit files are selected. First of the month? A consistent date each time?

➔ The CryoSat-2 orbit files were selected on the 15<sup>th</sup> of each month 2011-2016 (P6 L14).

*“The selection of endmembers is essential in the framework of waveform mixture analysis. Among CryoSat-2 orbit files between 2011 and 2016, a total of 48 orbit files were selected to extract endmember samples by month (15<sup>th</sup> from Jan. to May and from Oct. to Dec.), which fully transverse the broad Arctic Ocean (Fig. 2).”*

P7 L6: “...months and years should be compared \*\*using visual analysis\*\* with the waveforms... i.e. introduce the method of comparison here rather than L9.

➔ Thank you for your comment. We added “through visual analysis” (P7 L7) and deleted the phrase from the original location as suggested.

P7 L9: The authors should comment on the limitations of such a visual analysis, rather than a statistical one, to compare waveforms.

➔ Thank you for your comment. We commented on the limitations of such visual analysis in P7 L12-13.

*“However, such visual analysis cannot guarantee how the waveforms are quantitatively different by month and year.”*

P10 L10-13: Please explain in the manuscript that as the percentage of permuted observations increases, the grid sensitivity also increases but the difference is not significant, hence 30% was chosen.

## Authors' responses (TC-2017-170)

➔ Thank you for your comment. We added the explanation in P10 L10-12.

*"We tested various percentage values to identify which percentage is appropriate to represent grid sensitivity. As the percentage increased, the grid sensitivity (i.e., standard deviation) also increased but the spatial difference was not significant, hence a 30 % was chosen."*

## Authors' responses (TC-2017-170)

The authors would like to thank the editor and the reviewers for their precious time and invaluable comments. The corresponding changes and refinements are highlighted in yellow in the revised paper and are also summarized in our responses below. Authors' responses are in blue. Reviewer's comments are in black. When the manuscript is cited, it is shown in italics.

### Reviewer #3:

#### General Comments:

I acknowledge the efforts of the authors to revise the manuscript according to my comments. Some of my major concerns have been addressed sufficiently from my point of view, especially the lack of detailed information in some of the sections and certainly regarding the WMA algorithm.

However, two major concerns remain:

- 1) To me, it seems that the proposed waveform mixture algorithm does not represent a significant improvement compared to former lead detection algorithms methods, also given the limitations of the evaluation due to the resolution of the MODIS images. This is shown by the evaluation shown in Figure 5. The overall accuracy seems to be even slightly lower than in the method proposed by the author in a previous paper (Lee et al., 2016). Except the potential independence of CryoSat baseline changes, the authors cannot convincingly show why one should use the WMA algorithm instead of former used algorithms. However, since the method itself used here is novel, this might merit publication.
- The differences of the lead user's accuracy and overall accuracy between Lee et al., (2016) and the waveform mixture algorithm are 2.61% and 0.19 %. We agree with your point but we believe that the waveform mixture algorithm will be worthwhile in an operational system because the proposed approach does not require to change any parameters when CryoSat-2 baseline is updated.
- 2) Objective 3) is still treated superficial. Although in the revision you now write "briefly examine", it is still remains an objective. But what you show is rather a discussion of the state of the art. I am missing a detailed analysis. It is also now shown how the forming of leads is related to atmospheric forcing. On the other hand, this objective is not mentioned neither in the title nor in the abstract. Therefore, I would suggest to either expand this analysis in this paper, e.g. considering atmospheric circulation patterns, or to do it in a different study.
- Thank you for your comment. According to your comment and Editor's, we removed section 5.3 and moved some relevant text to section 4.2 (L14 P21-28). We also revised the third objective, accordingly.

From "briefly examine the relationship between Arctic lead fraction and thermodynamics and ice dynamics."

## Authors' responses (TC-2017-170)

To “3) examine the spatiotemporal distribution of lead fractions.”

(L14 P21-28). “*The Arctic Ocean circulations have contributed to the change in the state of sea ice. The lead fraction in Northwestern Greenland in Figs. 6 and 7 is low because of the convergence of sea ice by two major circulations, which was clearly shown in Kwok (2015). Kwok et al. (2013) revealed that the currents speed of Beaufort Gyre and Transpolar Drift increased from the years of 1982 to 2009 and this makes the fraction of multi-year ice decrease. However, the increasing lead fraction from the years of 2011 to 2016 in this study was not seen due to the high inter-annual variability of a lead fraction, particularly in the spring season (Figs. 6 and 7). High sensitivities in the marginal sea ice zone might result in not catching the increasing trend of Arctic lead fraction shown in the literature. In order to properly compare the Arctic current circulations and lead fraction, long-term lead fraction data are needed.*”

In general, I suggest to streamline the manuscript to improve readability, see therefore some comments below.

Specific Comments:

P1 L10-12: I suggest to delete these first two sentences. They rather belong to the Introduction.

➔ As suggested, we removed the two sentences in the abstract.

P1 L15: delete “based on the waveform mixture analysis” – repetition.

➔ We removed the phrase.

P1 L20-21: “which show a strong inter-annual variability of recent sea ice cover during 2011-2016, excluding the summer season (i.e., June to September)” – Can you quantify this? What means “strong” in this context? Otherwise this sentence is not really valuable.

➔ The inter-annual variability of averaged lead fraction is identified in the revised manuscript (Fig. 8). We removed the word, “strong” from the abstract and conclusion.

P3 L1: “is not easily affected” – What does it mean? How is it affected? Be more precise here. You might consider to shift that part to the discussion.

➔ Thank you for your comment. We added more explanation (P2 L 28-30).

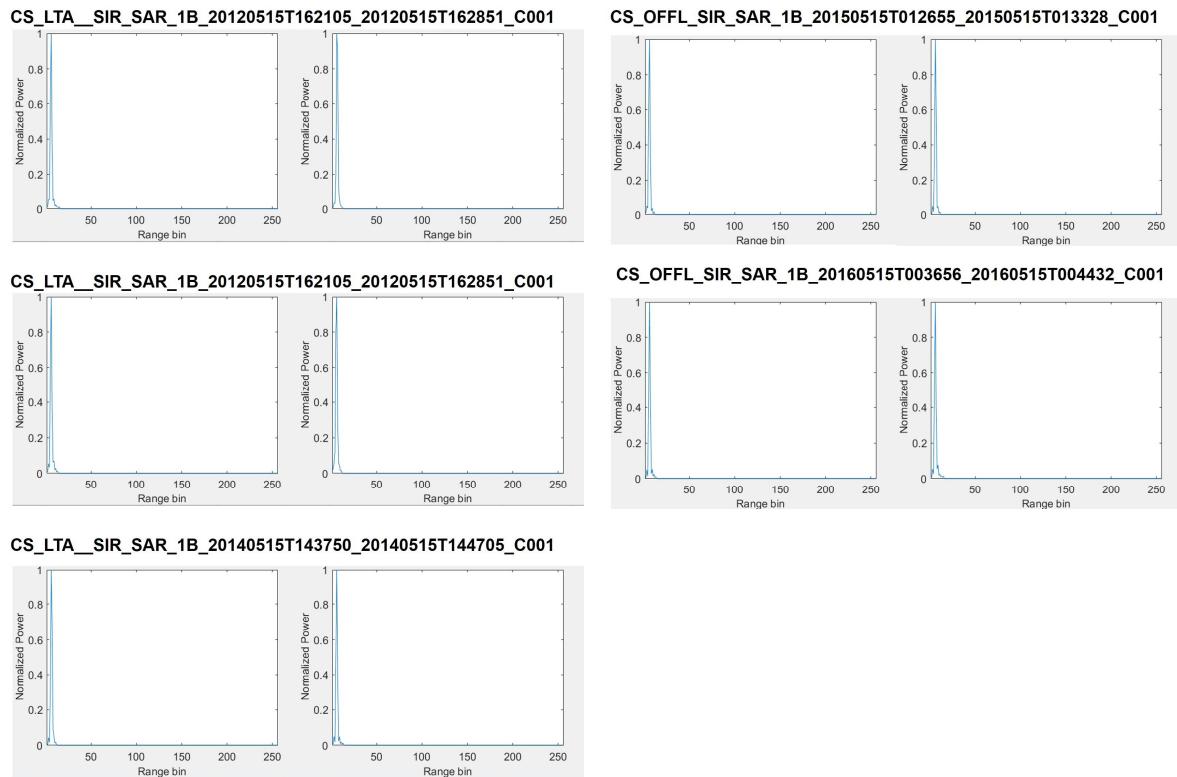
*“The lead detection using the proposed waveform mixture algorithm does not need to change any parameters to detect leads when the CryoSat-2 baseline is updated, which is a significant advantage compared to the existing threshold-based lead detection methods.”*

P3 L15: “The period of CryoSat-2 level 1b baseline C data in this study is in Jan. – May, Oct. – Dec. 2011-2016” – Why are you using data in May also? Substantial surface melt can occur already in some regions. This will surely affect the dielectric properties of the (snow) surface

## Authors' responses (TC-2017-170)

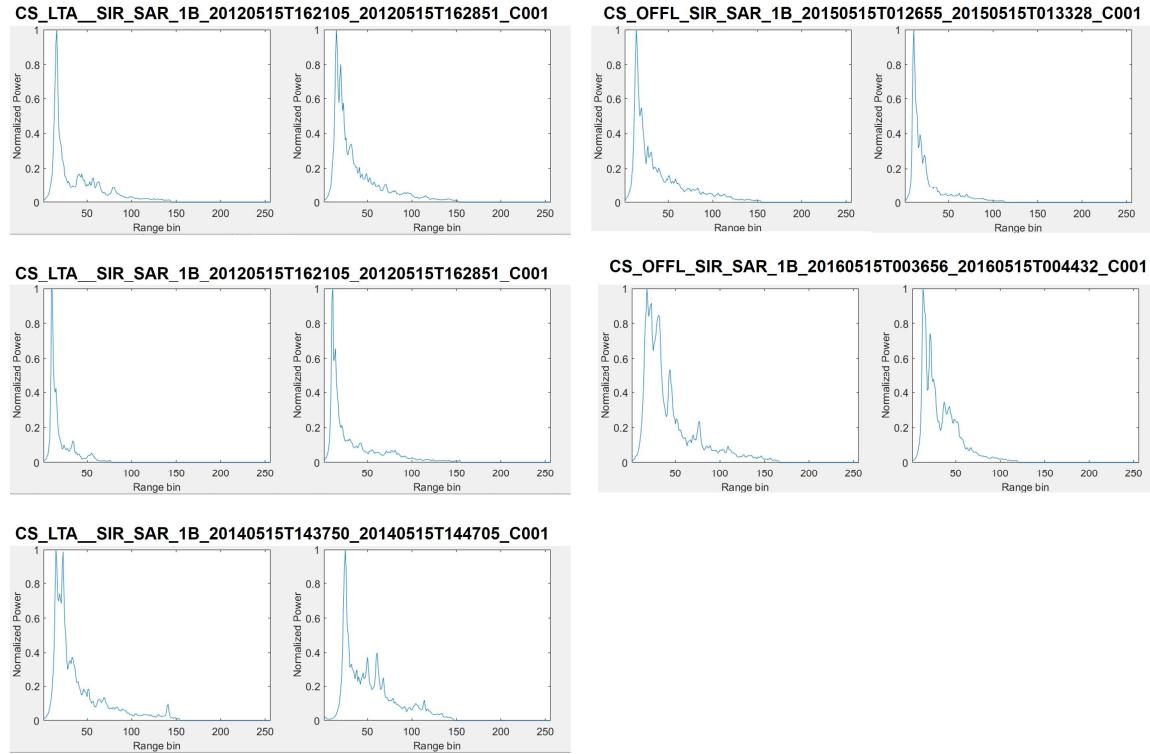
and therefore will affect the sea ice waveforms. Wouldn't that bias also the end-member selection?

→ Thank you for your comment. We have checked the waveforms in May. Although surface melting occurs on the sea ice surface, the DT can select appropriate candidates for lead and ice waveform endmembers. Response Figs. 1 and 2 show that there is no significant difference in the waveforms extracted between in May and other months (Response Fig. 3). Therefore, we believe that the waveforms in May would not make a bias for endmember selection. Although we could not check all of the candidates of lead and ice waveform endmembers in May 2012-2016, a number of the waveforms randomly selected from the candidates in May suggest that the endmember selection seems appropriate based on lead and ice endmembers shown in Fig. 1 in the revised manuscript and response Fig. 3.



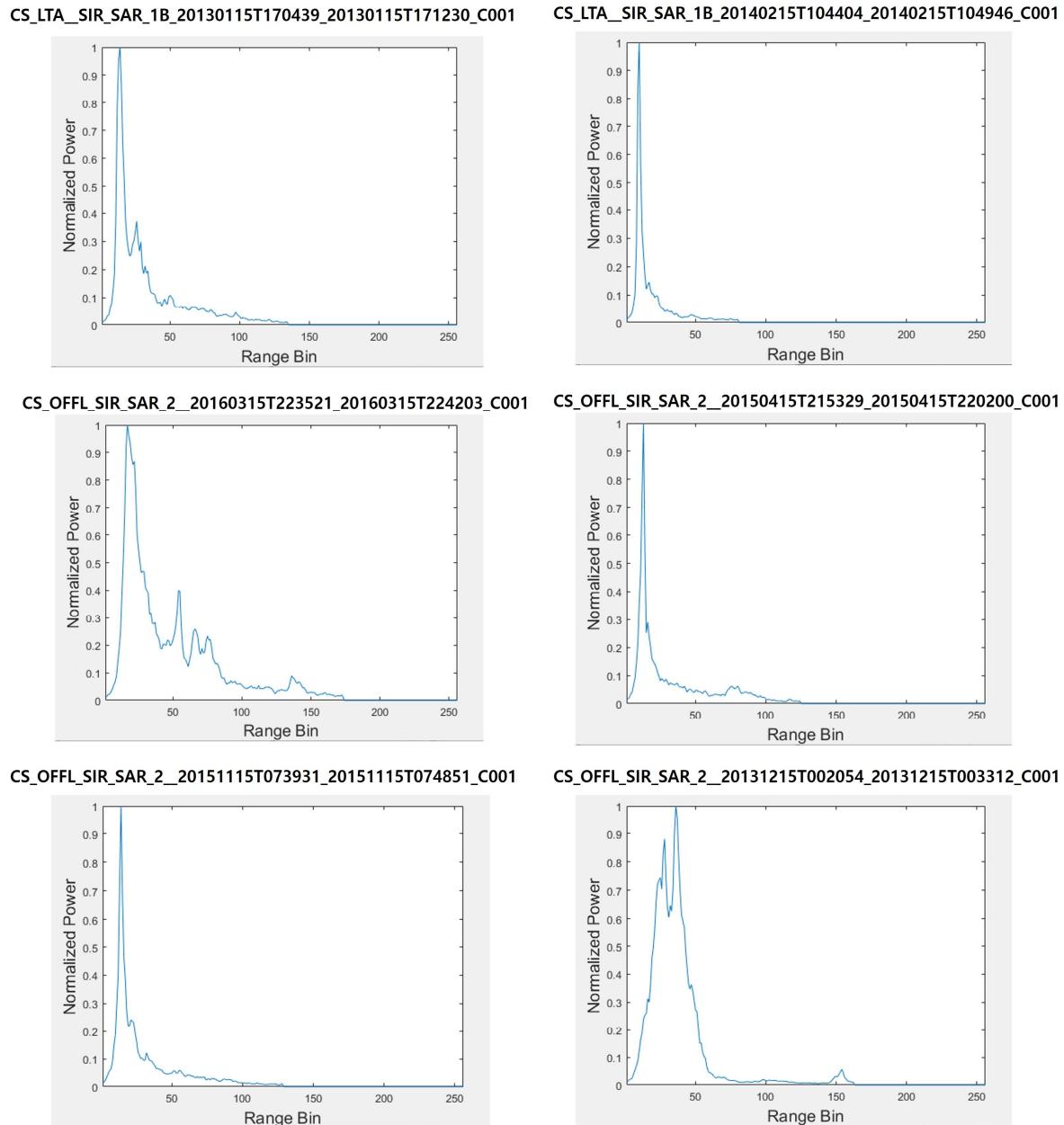
**Response Fig. 1** Examples of the lead endmember candidates selected by the DT in May 2012-2016. The text above each figure is the CryoSat-2 file name.

## Authors' responses (TC-2017-170)



**Response Fig. 2** Examples of the ice endmember candidates selected by the DT in May 2012-2016. The text above each figure is the CryoSat-2 file name.

## Authors' responses (TC-2017-170)



**Response Fig. 3** Examples of the ice endmember candidates selected by the DT in January, February, March, April, November, and December 2012-2016. The text above each figure is the CryoSat-2 file name.

## Authors' responses (TC-2017-170)

P15 L15/16: “the lead fraction starts to increase from April. This indicates an increasing lead fraction...” – Please rephrase.

➔ We rephrased the sentence (P14 L19-20).

*“While the lead fraction decreases from October to March (i.e., freezing season) with the minimum in March, the lead fraction starts to increase from April.”*

P26 L5: “The spatiotemporal distribution of monthly lead fraction maps were documented.” – That’s not really a conclusion.

➔ We removed the sentence in the conclusion.

P26 L6/7: “Unlike thresholds based lead detection methods, the waveform mixture analysis is less influenced on the update of baseline version of CryoSat-2 data” – Again, please, be a bit more specific here. What does “less” mean here. Indeed, this would be a unique feature probably.

➔ Thank you for your comment. We rephrased it with specific explanation (P23 L5-7).

*“Unlike thresholds based lead detection methods, since the waveform mixture algorithm solely uses waveforms, not beam behaviour parameters, it does not need to change any parameters when the CryoSat-2 baseline version is updated, which will be useful for future altimeter missions.”*

# Arctic lead detection using a waveform mixture algorithm from CryoSat-2 data

Sanggyun Lee<sup>1</sup>, Hyun-Cheol Kim<sup>2</sup>, and Jungho Im<sup>1,\*</sup>

<sup>1</sup>School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, 5 44919, South Korea

<sup>2</sup>Unit of Arctic Sea-Ice prediction, Korea Polar Research Institute (KOPRI), Incheon, 21990, South Korea

\*Correspondence to: Jungho Im ([ersgis@unist.ac.kr](mailto:ersgis@unist.ac.kr))

**Abstract.** We propose a waveform mixture algorithm to detect leads from CryoSat-2 data, which is novel and different from the existing threshold-based lead detection methods. The waveform mixture algorithm adopts the concept of spectral mixture 10 analysis that is widely used in the field of hyperspectral image analysis. This lead detection method was evaluated with high resolution (250m) MODIS images and showed comparable and promising performance in detecting leads when compared to the previous methods. The robustness of the proposed approach also lies in the fact that it does not require the rescaling of 15 parameters (i.e., stack standard deviation, stack skewness, stack kurtosis, pulse peakiness, and backscatter sigma), as it directly uses L1B waveform data unlike the existing threshold-based methods. Monthly lead fraction maps were produced by the waveform mixture algorithm, which shows an inter-annual variability of recent sea ice cover during 2011-2016, 20 excluding the summer season (i.e., June to September). We also compared the lead fraction maps to other lead fraction maps generated from previously published data sets, resulting in similar spatiotemporal patterns.

## 1 Introduction

Sea ice leads (hereafter referred to as “leads”), linearly elongated cracks in sea ice, are a common feature in the Arctic 20 Ocean. Leads facilitate an amount of heat and moisture exchanges between the atmosphere and the ocean because of the temperature differences between them (Maykut. 1982; Perovich et al., 2011). Although leads occupy a small portion of the Arctic Ocean, there is much more heat transfer between the atmosphere and ocean through leads than sea ice (Maykut, 1978; Marcq and Weiss, 2012). Furthermore, Lüpkes et al. (2008) showed that a 1% change in sea ice concentration owing to an 25 increase of lead fraction could increase near surface temperature in the Arctic by 3.5 K. Thus, the detection and monitoring of leads in the Arctic Ocean are crucial because they are closely related to the Arctic heat budget and the physical interaction between the atmospheric boundary layers and sea ice in the Arctic.

Satellite sensors have been the most efficient way to monitor leads in the entire Arctic region since the 1990s (Key et al., 1993; Lindsay and Rothrock, 1995; Miles and Barry, 1998). Advanced Very High Resolution Radiometer (AVHRR) and Defense Meteorological Satellite Program (DMSP) satellite visible and thermal images were used to detect leads in the early 30 1990s. Recently, the Moderate Resolution Imaging Spectroradiometer (MODIS) Ice Surface Temperature (IST) product with

the 1km spatial resolution was used to detect leads to map pan-Arctic lead presence (Willmes and Heinemann, 2015; Willmes and Heinemann, 2016). They mitigated cloud interference using a fuzzy cloud artefact filter and investigated lead dynamics based on a comparison between pan-Arctic lead maps and the characteristics of the Arctic Ocean such as shear zones, bathymetry, and currents. While optical sensors have a finer spatial resolution, they are not pragmatic in the dark 5 regions during polar nights (from December to February). In addition, leads are easily contaminated by clouds. Microwave instruments such as passive microwave sensors and altimeters have been used to detect leads and to produce lead fractions. Röhrs and Kaleschke (2012) utilized the polarization ratio of the Advanced Microwave Scanning Radiometer for EOS (AMSR-E) channels and retrieved daily thin ice concentration. With the help of the thin ice concentration, lead orientations and frequencies were derived using an image analysis technique (i.e., Hough transform) (Bröhan and Kaleschke, 2014). 10 Airborne and spaceborne radar altimeters can detect leads as well. Zygumtowska et al. (2013) used Airborne Synthetic Aperture and Interferometric Radar Altimeter System (ASIRAS), similar to CryoSat-2, to identify leads based on waveform characteristics and a Bayesian classifier. Zakharova et al. (2015) and Wernecke and Kaleschke (2015) used the spaceborne altimeters Satellite with Argos and Altika (SARAL) and CryoSat-2 to identify leads, respectively. While Zakharova et al. 15 (2015) applied simple thresholds to identify leads along with Satellite with Argos and Altika (SARAL/Altika) tracks and estimated regional lead fractions, Wernecke and Kaleschke (2015) optimized thresholds to detect leads and produced pan-Arctic lead fraction maps using CryoSat-2 with an analysis of lead width, and sea surface height.

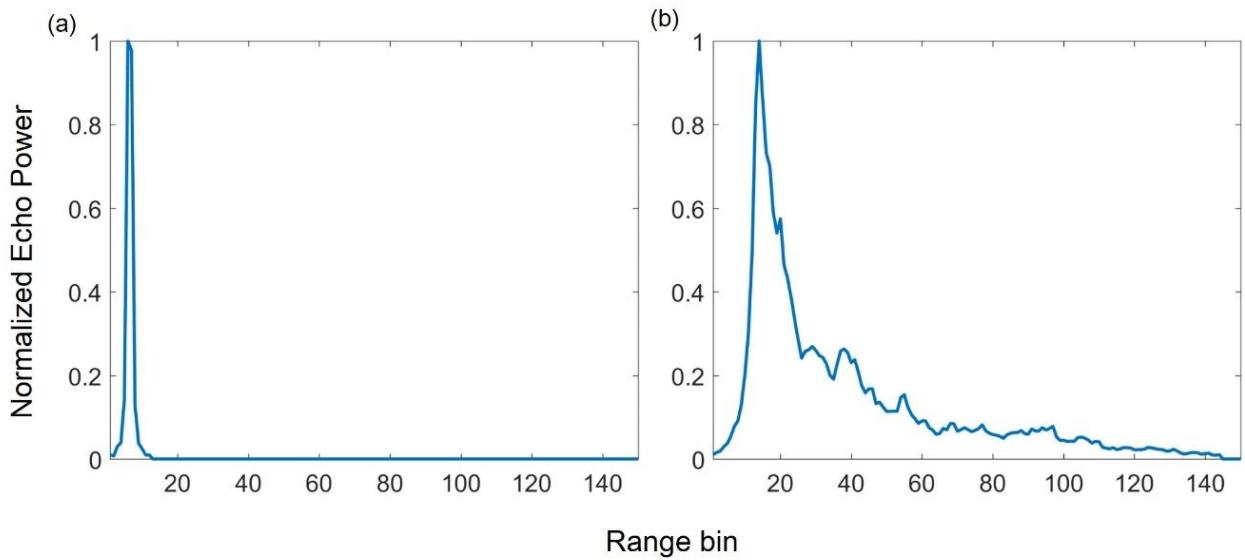
Spectral mixture analysis based on the assumption that the spectra measured by sensors for a pixel are a linear combination of the spectra for all components within the pixel (Keshava and Mustard, 2002) was first applied to the altimetry research field in the Polar Region by Chase and Hollyer (1990). They estimated sea ice type and concentration 20 using spectral mixture analysis based on Geosat waveforms. However, Geosat with a relatively small number of range bins and coarser spatial resolution is not sufficient to detect small leads in the winter (DJF) and spring seasons (MAM) in the Arctic. In this study, we adopted the linear mixture algorithm concept to waveforms from Synthetic Aperture Interferometric Radar Altimeter (SIRAL), CryoSat-2, to identify leads and produce monthly pan-Arctic lead fractions from January to May and October to December between 2011 and 2016. Waveform endmembers are crucial to implement spectral mixture 25 algorithm (Fig. 1). The N-FINDR (N-finder) algorithm was used to select waveform endmembers from extracted waveforms by Decision tree (DT) from Lee et al. (2016), which avoids the subjective selection of endmembers. The detected leads were visually evaluated with MODIS images (at 250 m resolution) and compared with other thresholds based lead detection methods. The lead detection using the proposed waveform mixture algorithm does not need to change any parameters used in the algorithm to detect leads when the CryoSat-2 baseline is updated, which is a significant advantage compared to the 30 existing threshold-based lead detection methods. The main objectives of this study are to 1) develop a novel lead detection method based on waveform mixture algorithm, 2) compute recent pan-Arctic lead fractions, and 3) examine the spatiotemporal distribution of lead fractions.

## 2. Data

### 2. 1 CryoSat-2

CryoSat-2, the carrying Synthetic Aperture Interferometric Radar Altimeter (SIRAL) was launched in April 2010 by the European Space Agency (ESA). CryoSat-2 is a satellite dedicated to Polar research. SIRAL is a radar altimeter with a central frequency of 13.575 GHz (K<sub>u</sub>-band) and a bandwidth of 320 MHz. CryoSat-2 takes an advantage of SIRAL to detect smaller leads with an efficient use of the instrument's energy compared to the previous radar altimeter missions such as GeoSat and Jason (Wingham et al., 2006). In this study, we used Synthetic Aperture Radar (SAR) mode, mainly operating on sea ice regions; and SAR Interferometric (SIN) mode, mainly operating on steep regions such as on the margin of an ice shelf and ice sheet of level 1b baseline C data. The SAR and SIN modes have 256 and 1024 range bins, respectively (Scagliola, 2014). The period of CryoSat-2 level 1b baseline C data in this study is in Jan. – May, Oct. – Dec. 2011-2016.

CryoSat-2 transmits bursts of radar pulses (i.e., 64) with high Pulse Repetition Frequency (PRF, 18.181kHz), which forms so-called Doppler beams because of the along-track movement of the satellite (Wingham et al., 2006). With the help of the high PRF, each Doppler beam is coherently correlated and pointed at the same location on the Earth surface. This is called beam stacking. Multi-looking is conducting by averaging the stacking beams to reduce speckles and thermal noises (Salvatore. 2013). Exemplary results waveforms in the L1b SAR data are shown in Fig. 1. Such waveforms represent the temporal distribution of reflected power when the radar pulses reach the surface, describing a flat or rough surface. In this case, since the leading edge of each waveform starts from a different range bin, the beginning of the waveform was set at 1% of the maximum echo power (Fig. 1). For a more detailed explanation about the processes to develop L1b waveform data, refer to Salvatore (2013).



**Figure 1:** Representative waveforms of (a) leads and (b) sea ice over the Arctic Ocean selected by N-FINDR algorithm during January to May and October to December between 2011 and 2016. Refer to the methods section for N-FINDR algorithm.

## 5 2.2 Sea ice edge data

The European organization for the exploitation of METeorological SATellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI SAF) provides multiple sea ice products such as sea ice concentration, sea ice edge, sea ice type, sea ice emissivity, and sea ice drift. The sea ice edge product was developed using the polarization ratio of 19 GHz and 91 GHz, the spectral gradient ratio of 37 GHz and 19 GHz from Special Sensor Microwave Imager/Souder (SSMIS), 10 and anisFMB from The Advanced Scatterometer (ASCAT) with Bayesian approach (Aaboe et al., 2016). In this study, monthly averaged sea ice edge data was used to mask out monthly lead fraction maps. The open ice cover in the sea ice edge product was regarded as an open ocean.

## 2.3 Monthly lead fraction maps

Lead fraction maps produced from previous studies (Röhrs and Kaleschke, 2012; Wernecke and Kaleschke, 2015; 15 Willmes and Heinemann, 2016) were used to compare to the lead fraction maps generated using the proposed waveform mixture algorithm in this study. Röhrs and Kaleschke (2012) produced daily thin ice concentration maps using AMSR-E data with a 6.25 km grid, which can detect leads that are wider (i.e., width) than 3 km. The daily thin ice concentration that was over 0.5 (i.e., 50%) was considered to be a lead and binary daily lead maps were averaged to properly compare other monthly lead fraction maps. A threshold optimization based lead detection method with the CryoSat-2 was used in Wernecke 20 and Kaleschke (2015) and monthly lead fraction maps were calculated with the grids of 99.5 km. The thin ice concentration

maps (Röhrs and Kaleschke, 2012) and the lead fraction maps using CryoSat-2 (Wernecke and Kaleschke, 2015) are available on their website (<http://icdc.cen.uni-hamburg.de/1/daten/cryosphere.html>). Willmes and Heinemann (2016) also produced daily lead maps over the entire Arctic Region, classifying land, cloud, sea ice, lead-artefact, and lead with the spatial resolution less than 2 km. The lead class was only considered to calculate daily binary lead fraction maps. The sum of 5 the lead pixels was divided by days in a month (i.e., 28, 30, or 31) to make monthly lead fraction maps. This data is available on their website (<http://dx.doi.org/10.1594/PANGAEA.854411>). In this study, we compared the monthly lead fraction maps from January to March 2011 as AMSR-E based lead fraction maps were only available until 2011.

### 3. Methods

#### 10 3.1 Waveform mixture algorithm

An endmember in remote sensing data represents a spectrally pure ground component in a single pixel. For example, it could be pure water, vegetation, bare ground or a soil crust pixel in remote sensing data. Endmembers play the most important role in conducting spectral mixture analysis. Spectral mixture analysis assumes that the spectra measured by sensors for a pixel is a linear combination of the spectra of all components within the pixel (Keshava and Mustard. 2002). 15 This technique is widely used to resolve spectral mixture problems in image analysis (Foody and Cox, 1994; Lu et al., 2003; Wu. 2004; Iordache et al., 2011). Spectral mixture analysis determines the fractions of the components (i.e., classes) found in mixed pixels by producing abundances of the components based on endmembers. The proposed waveform mixture algorithm adopts the concept of spectral mixture analysis. Since the waveform of altimetry within a footprint could be considered to be a mixture of leads and various types of sea ice, spectral mixture analysis can be applied in this framework. In this study, 20 waveforms of CryoSat-2 L1b data were used as endmembers such as the waveform of pure lead and first-year ice (FYI) (Fig. 1). The lead and ice endmembers are used as reference data for separating leads and ice. In order to successfully implement waveform mixture algorithm, the proper selection of lead and ice endmembers is essential.

The basic waveform mixture model is defined as follows in equation 1.

$$Y_k = \sum_{k=1}^K a_{ik} E_k + r_k \quad (1)$$

25 where  $Y_p = \{Y_1, Y_2, Y_3, \dots, Y_k\}$  represents waveform vectors and  $k$  means a range bin in the waveform.  $a_{ik}$  is an abundance fraction, which provides lead and ice proportion in terms of lead and ice endmember.  $E_k$  is the endmember vector. The  $r_k$  represents un-modeled residual. The equation 1 is constrained under  $\sum_{k=1}^K a_{ik} = 1$  and  $a_{ik} \geq 0$ . The abundance can be derived by using a least square method to minimize the un-modeled residual ( $r_k$ ).

Chase and Holyer (1990) were concerned by two problems with the application of spectral mixture analysis to the waveform of altimeter data. First, the waveform within a footprint may not be linearly mixed between leads and sea ice. CryoSat-2 is more sensitive to the specular reflection of leads than the diffuse reflection of sea ice when both leads and sea ice exist within the same footprint, which implies the waveform may tend to be similar to the endmember of leads (Chase and Holyer. 5 1990). Since CryoSat-2 data have a large number of range bins, indicating higher vertical resolution than the range bins from Geosat, they could be used to reduce the overestimation of leads. Secondly, the waveform of the altimeter (i.e., Geosat) is somewhat weighted on the centre of a footprint rather than representing an entire footprint. This could be an error source when applying spectral mixture analysis to waveform data (Chase and Holyer. 1990). However, the CryoSat-2 L1b waveform is produced by averaging more than 200 weighted waveforms with various incidence angles, which can alleviate 10 such a problem.

### 3.2 Endmember selection

The selection of endmembers is essential in the framework of waveform mixture algorithm. Among CryoSat-2 orbit files in Jan. to May and Oct. to Dec. between 2011 and 2016, a total of 48 orbit files were selected to extract endmember samples by month (15<sup>th</sup> from Jan. to May and from Oct. to Dec.), which fully transverse the broad Arctic Ocean (Fig. 2). The lead 15 and ice waveforms are extracted by using the decision trees (DT) algorithm developed for lead detection by Lee et al. (2016). The DT has proven to be very effective in various remote sensing classification tasks (Kim et al., 2015; Torbick and Corbiere, 2015; Amani et al., 2017; Tadesse et al., 2017; Hisabayashi et al., 2018). The lead and sea ice endmembers (i.e., the most representative waveforms) are a key factor in the successful implementation of the waveform mixture algorithm. In order to avoid the subjective selection of endmembers, a number of endmember candidates were extracted by the DT 20 algorithm (Lee et al., 2016) and the N-FINDR algorithm determined the optimum lead and ice endmembers. The N-FINDR algorithm basically uses the fact that the N spectral dimension and the N-volume (V), defined by a simplex with pure pixels, are always greater than any other combinations (Winter 1999). It operates by inflating a simplex inside of the data 25 (endmembers), starting with any pixel set. The endmember is replaced with another endmember, and the volume is recalculated. The endmember is replaced with the spectrum of the new pixel if the volume increases. This process repeats until the volume does not increase (i.e., until there is no replacement).

$$\mathbf{E} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \overrightarrow{e_1} & \overrightarrow{e_2} & \cdots & \overrightarrow{e_i} \end{bmatrix} \quad (2)$$

Where  $\overrightarrow{e_i}$  represents a column vector of the endmember i.

$$V(\mathbf{E}) = \left| \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \overrightarrow{e_1} & \overrightarrow{e_2} & \cdots & \overrightarrow{e_i} \end{pmatrix} \right| / (i-1)!$$
(3)

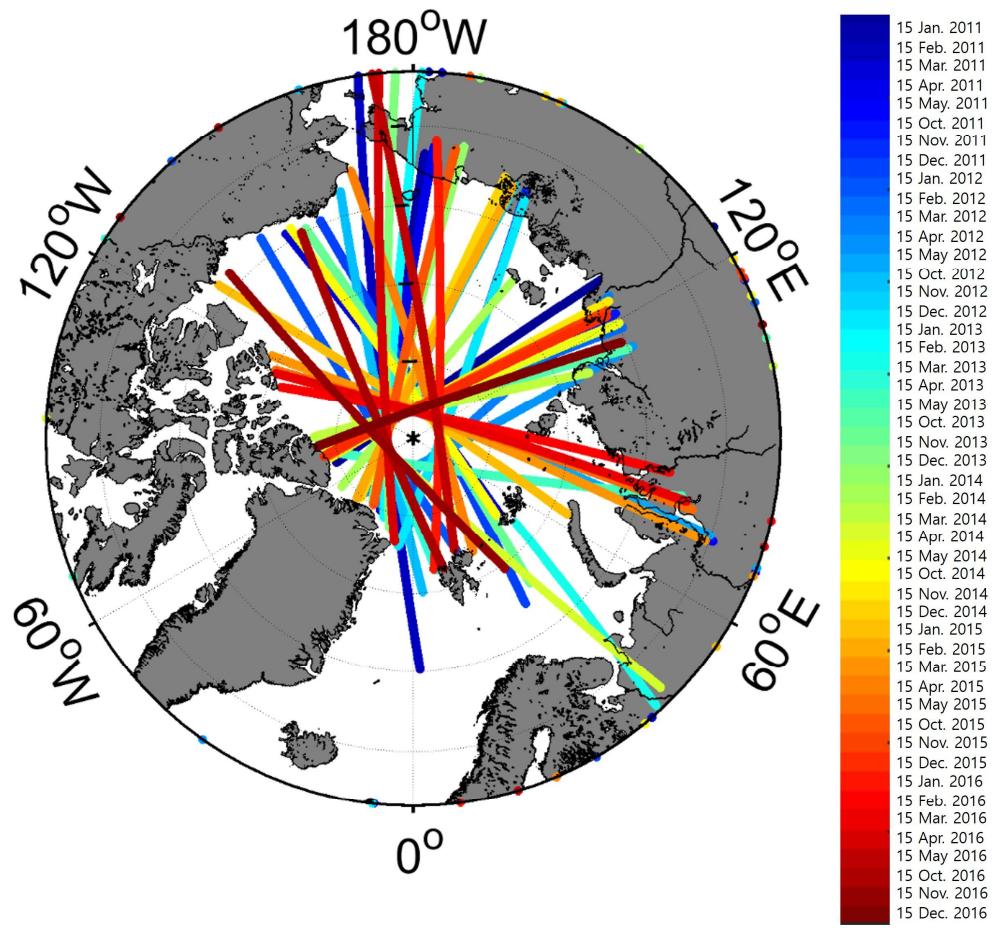
The Volume (V) of the simplex containing synthetic endmember sets is proportional to the determinant. This algorithm has been widely used for automatically selecting representative endmembers (Winter, 1999; Zortea and Plaza, 2009; Ertürk and plaza, 2015; Ji et al., 2015; Chi et al., 2016).

5 The DT model from Lee et al. (2016) was developed using data (i.e., stack standard deviation, stack skewness, stack kurtosis, pulse peakiness, and backscatter sigma-0) collected in March and April 2011-2014. Thus, the waveforms in other months and years should be compared with the waveforms in March and April 2011-2014 through visual analysis to identify whether the waveforms derived by the DT model during the study period are appropriate to implement the waveform mixture algorithm. Waveforms from March to April between 2011 and 2014 were compared to those from January to May, 10 and October to December between 2011 and 2016 (not shown), resulting in little difference between them. This justified the use of the DT algorithm proposed by Lee et al. (2016) to extract waveform samples of leads and sea ice. The total number of sea ice and lead waveforms is 420,858 and 8,501, respectively. However, such visual analysis cannot guarantee how the waveforms are quantitatively different by month and year.

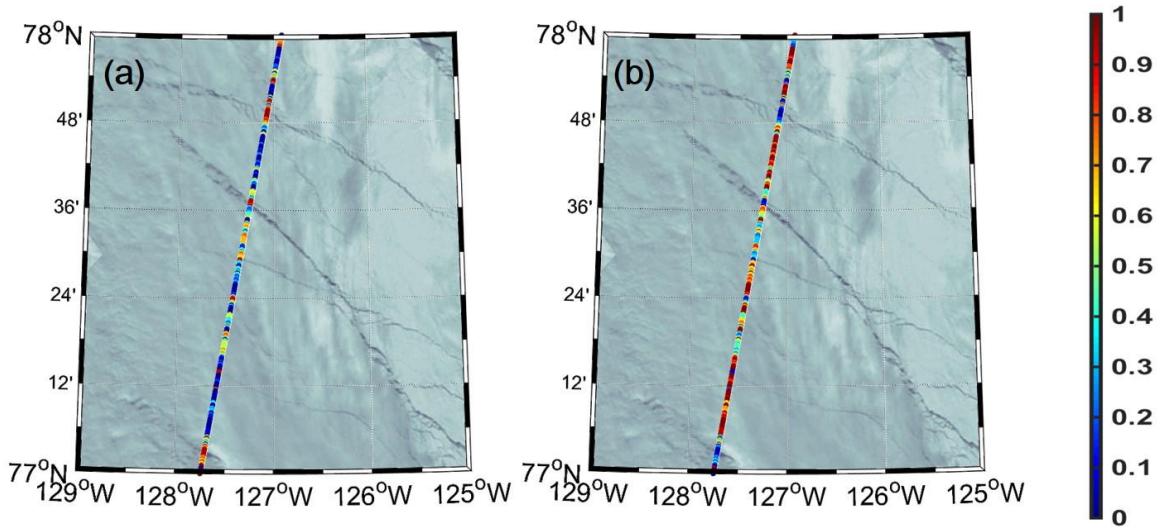
15 The lead classification based on waveform mixture algorithm was evaluated with 250 m MODIS images collected from March to May and in October. We used Earth View 250m Reflective Solar Bands Scaled Integers in MOD02QKM and adjusted the contrast to emphasize leads from sea ice in the images. It should be noted that since MODIS images with spatial resolution of 250 m were not available in January, February, November, and December due to polar nights, the evaluation with MODIS images and lead classification results based on CryoSat-2 could not be used. To secure the reliability of the comparison, the temporal difference between the MODIS images and CryoSat-2 data was always under 30 minutes.

20 The waveform mixture model produces abundance data (i.e., lead and sea ice abundance) at along-track points with respect to each endmember of the leads and sea ice (Fig. 3). While the lead abundances are high on the leads, the ice abundances are low on the leads, and vice versa (Fig. 3). Thresholds have to be determined to make a binary classification between leads and sea ice. Optimum thresholds to produce binary lead classification from lead and sea ice abundances were identified through an automated calibration. To implement the automated calibration, reference point data of leads and sea 25 ice were determined by visual inspection of four MODIS images collected on 17 April 2014, 25 May 2015, 10 October 2015, and 27 March 2016. While the calibration was conducted using half of the reference data randomly selected, the validation was performed using the remaining data. The size of the leads detected by the proposed waveform mixture algorithm is at least 250m or greater because the calibration and validation processes were conducted using MODIS images with 250m spatial resolution. It should be noted that leads smaller than 250m are hardly seen in MODIS images, which implies that 30 there is some uncertainty in the comparison of the lead detection methods for small leads. Threshold combinations from 0.2

to 0.9 with a step size of 0.01, for both lead and sea ice abundances, were tested and the one resulting in the highest accuracy was determined to be an optimum threshold combination.



**Figure 2:** The 48 CryoSat-2 orbit files from Jan. 2011 to Dec. 2016 used for extraction endmember waveforms. The CryoSat-2 orbit files relatively cover the entire Arctic Ocean.



**Figure 3:** Lead and ice abundance derived by waveform mixture analysis on 10 Oct. 2015. (a) Lead abundance, (b) Ice abundance. The colour bar expresses abundances from 0 to 1.

Lead detection results were evaluated using three accuracy metrics—producer's accuracy, user's accuracy, and overall accuracy (Tab. 1). Producer's accuracy (i.e.,  $a/(a+c)$  in the table), which is associated with omission errors, is calculated as the percentage of correctly classified pixels in terms of all reference samples for each class. User's accuracy (i.e.,  $a/(a+b)$  in the table), which is related to commission errors, is calculated as the fraction of correctly classified pixels with regards to the pixels classified to a class. Overall accuracy (i.e.,  $(a+d)/(a+b+c+d)$  in the table) is calculated as the total number of correctly classified samples divided by the total number of validation sample data. The lead and ice reference data using MODIS images and CryoSat-2 tracks were labeled through visual interpretation.

**Table 1:** Error matrix for calculation of user's, producer's and overall accuracy in terms of lead and ice classification.

		MODIS references		
		Lead	Ice	Sum
CryoSat-2 based classification	Lead	a	b	(a+b)
	Ice	c	d	(c+d)
	Sum	(a+c)	(b+d)	(a+b+c+d)

A monthly lead fraction was derived by dividing the number of lead observations by the number of total observations within a 10 km grid in a month. It is noted that while there are more than 30 CryoSat-2 observations in the 10 km grid around the centre of the Arctic, less than 5 observations are generally found in each 10 km grid in the marginal zones of 5 Arctic Ocean. This will be dealt with in the results section with more details. It also should be noted that it is hard for the altimeter-based lead detection methods used in such as Wernecke and Kaleschke (2015) and this study to identify the propagating, opening, closing of leads because sea ice and leads generally move when the altimeters revisit a certain grid.

### 3.3 Calculation of sensitivity in a 10x10 km grid

Since each grid has a different number of CryoSat-2 observations, a sensitivity analysis was conducted in terms of the 10 number of observations by grid. We tested various percentage values to identify which percentage is appropriate to represent a grid sensitivity. As the percentage increased, the grid sensitivity (i.e., standard deviation) also increased but the spatial difference was not significant, hence a 30 % was chosen. Thirty (30) percent of the lead and ice observations in 10x10 km 15 grids was randomly permuted 50 times, and the standard deviation of the resultant lead fractions through the 50 iterations were calculated by grid. The higher the standard deviation in a grid, the more sensitive the observed lead fraction is to the number of available observations. It should be noted that the standard deviation is zero when no lead observation is found, which means lead fraction is also zero. Sensitivities were calculated from January to April 2011 because these months were used to compare the lead fractions from the proposed waveform mixture analysis to those in the existing literature.

## 4 Results

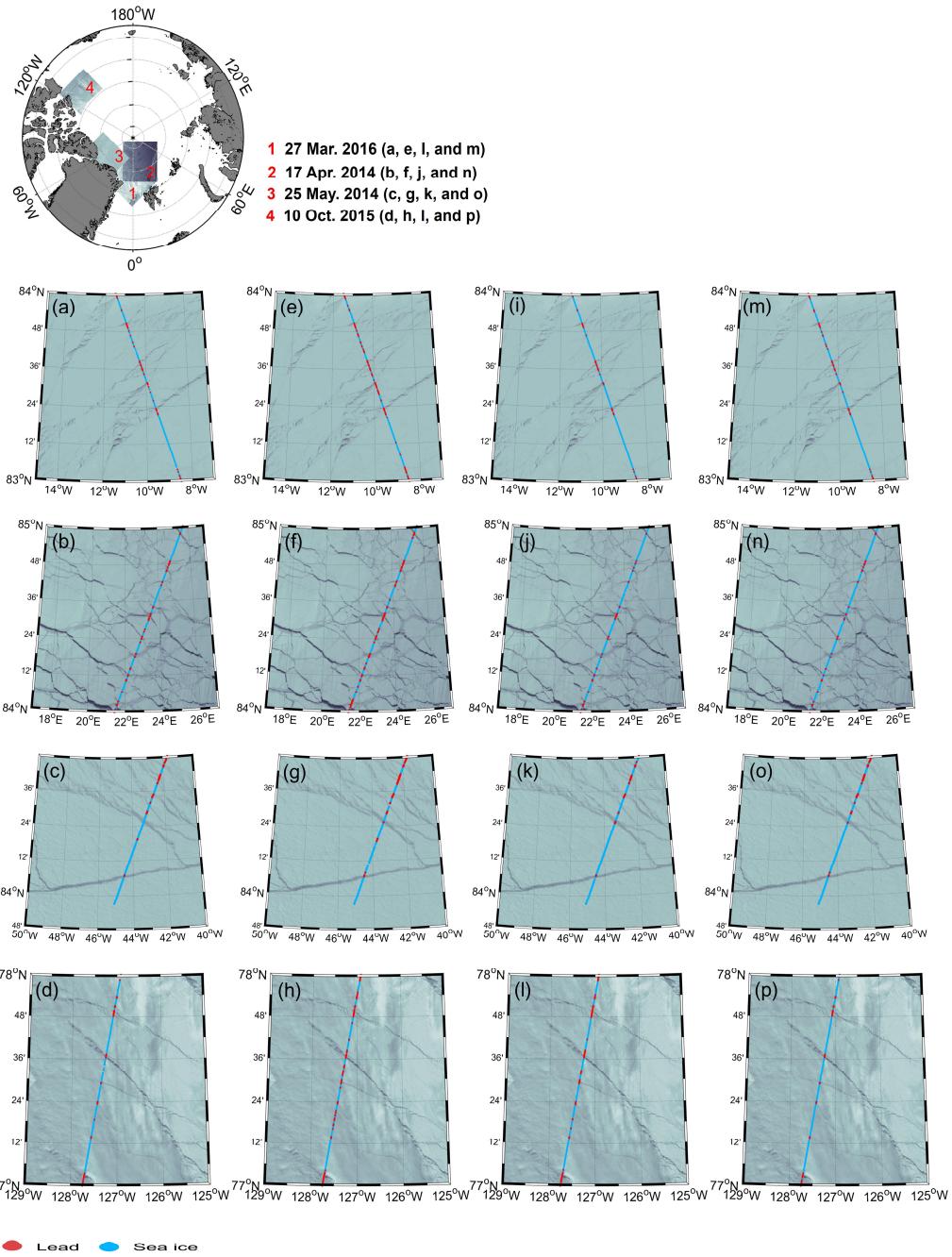
### 4.1 Performance of lead classification

Fig. 1 shows representative waveforms of leads and sea ice extracted by the N-FINDR algorithm as endmembers. The waveform of leads is dominated by specular reflection, resulting in a narrow peak curve. The representative waveform of sea

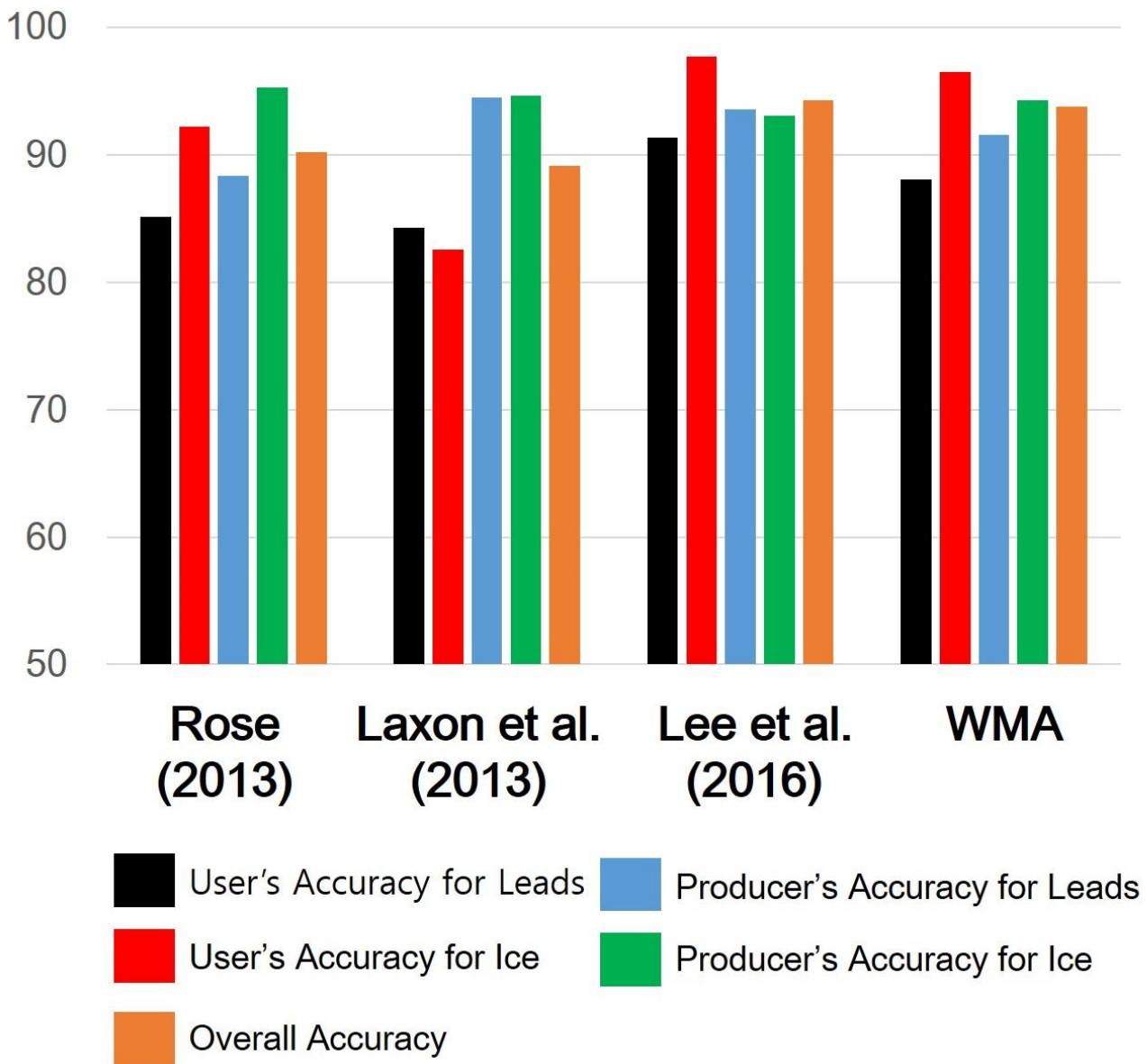
ice has a wider distribution due to its rough surface when compared to that of leads. Considering different types of sea ice such as young ice, FYI, and Multi-Year Ice (MYI), the representative waveform of sea ice is not significantly different from that of FYI based on visual inspection (Zygmuntowska et al., 2013; Ricker et al., 2015; Lee et al., 2016).

The optimum thresholds for the lead and sea ice abundances were determined to be 0.84 and 0.57 through the automated calibration, respectively. According to the thresholds, leads were identified with the conditions of lead abundance  $> 0.84$  and sea ice abundance  $< 0.57$ . Selected examples of lead detection results based on waveform mixture algorithm are presented in Fig. 4 with threshold-based lead detection results from the existing literature (Rose, 2013; Laxon et al., 2013; and Lee et al., 2016). Simple thresholding approaches based on two waveform parameters, pulse peakiness (PP) and stack standard deviation (SSD) were used in Rose (2013), Laxon et al. (2013), and Lee et al. (2016), respectively. It should be noted that since the existing methods were developed using parameters such as beam behaviour parameters and backscatter sigma-0 extracted from baseline B data, rescaling was conducted on the parameters extracted from a newly updated baseline C data for reasonable comparison. Since the contrast between the parameters of baselines B and C data is not linear, we rescaled the parameters by adding the difference of the parameters between the two baseline data to baseline C data.

Multiple lead classification methods based on CryoSat-2 data were evaluated by visual inspection with high resolution (250m) MODIS images. Leads (i.e., red dot) and sea ice (i.e., light blue dot) are distinguished, depending on the surface condition of lead and sea ice (Fig. 4). For better comparisons, a quantitative assessment is required (Fig. 4). DT from Lee et al. (2016) produced the highest overall accuracy (95.19%), followed by the waveform mixture algorithm (95%), Rose (2013) (93.26%), and Laxon et al. (2013) (91.70%). DT from Lee et al. (2016) produced the highest user's accuracy for leads, while the proposed approach produced the highest producer's accuracy for leads, which implies a slight over-detection of leads by the proposed waveform mixture algorithm. The user's accuracy for leads of Laxon et al. (2013) is the lowest, resulting in much over-detection of leads (i.e., many leads on sea ice; Fig. 4). Similarly, the user's accuracy for ice of Rose (2013) is lower than that of the proposed waveform mixture algorithm, indicating the detection of leads on sea ice, which is shown in Figs. 4b and c. While the performance of the waveform mixture analysis was comparable to the DT algorithm from Lee et al. (2016), the waveform mixture algorithm slightly over-estimated leads resulting in a lower user's accuracy for leads than that by DT (Figs. 4 and 5). These are inevitable results because waveforms used in the waveform mixture algorithm are basically extracted by DT from Lee et al. (2016). The lead classification results should be assessed during all the months (i.e., January to May, and October to December) and years (i.e., 2011 to 2016) using MODIS images to thoroughly evaluate the proposed waveform-based algorithm for lead detection. However, the lead classification results in January, February, November, and December were not assessed using MODIS images due to polar nights. Thus, the lead classification results in these months could possibly have uncertainties. It should be also noted that the validation was limited as the MODIS images did not fully cover the entire Arctic region (top in Fig. 4).



**Figure 4:** Visual comparison of lead classifications: (a) – (d) lead classifications based on Rose (2013), (e) – (h) lead classifications based on Laxon et al. (2013), (i) – (l) lead classifications based on decision trees from Lee et al. (2016), and (m) – (p) lead classifications based on the proposed waveform mixture analysis. The MODIS data were collected on 27 March 2016 (a, e, i, and m), 17 April 2014 (b, f, j, and n), 25 May 2015 (c, g, k, and o), and 10 October 2015 (d, h, l, and p). An overview map of the location of cropped MODIS images is in top of the figure.



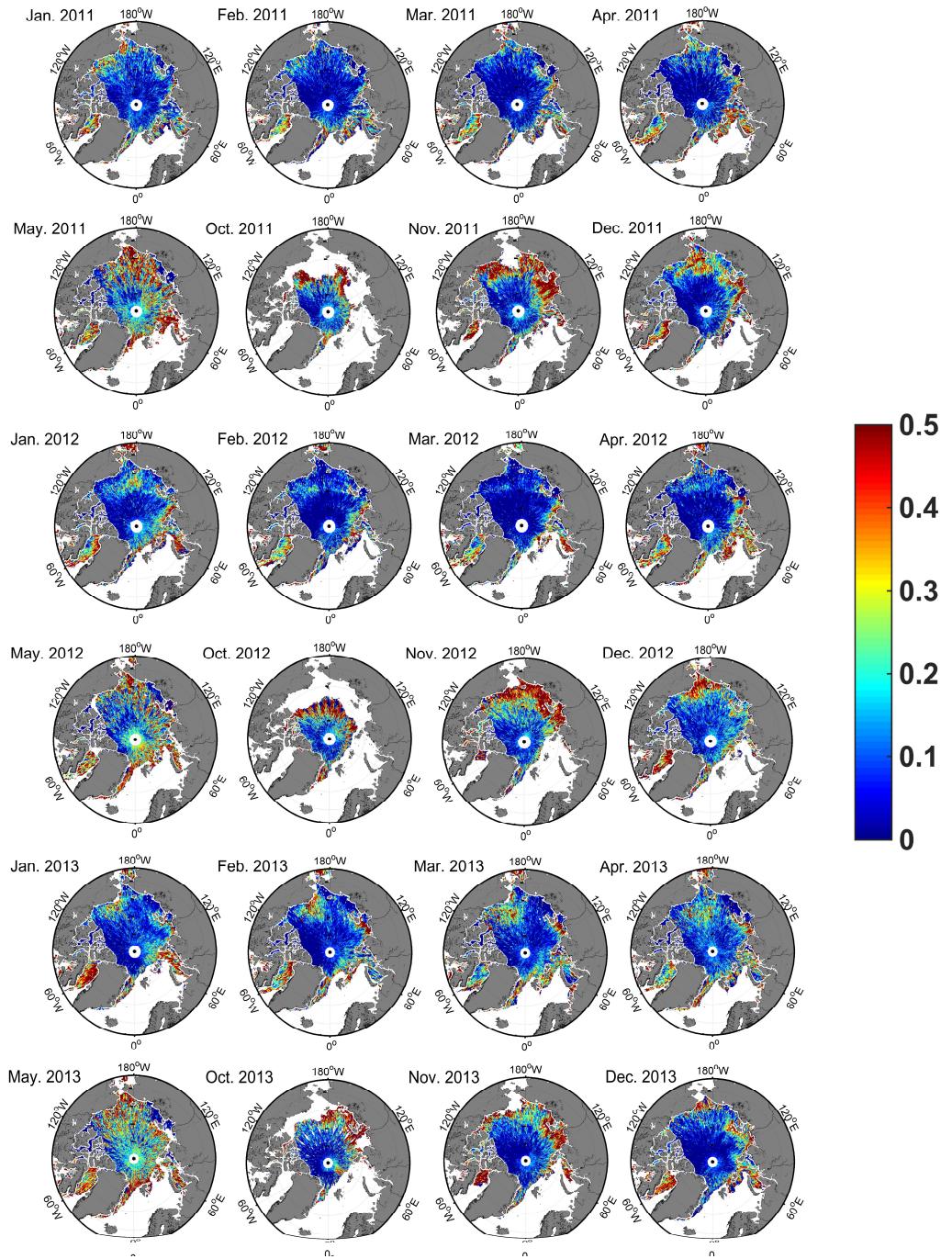
\*Unit: Percentage (%)

**Figure 5:** Accuracy assessment results for lead detection by method—three existing methods and the proposed waveform mixture analysis (WMA).

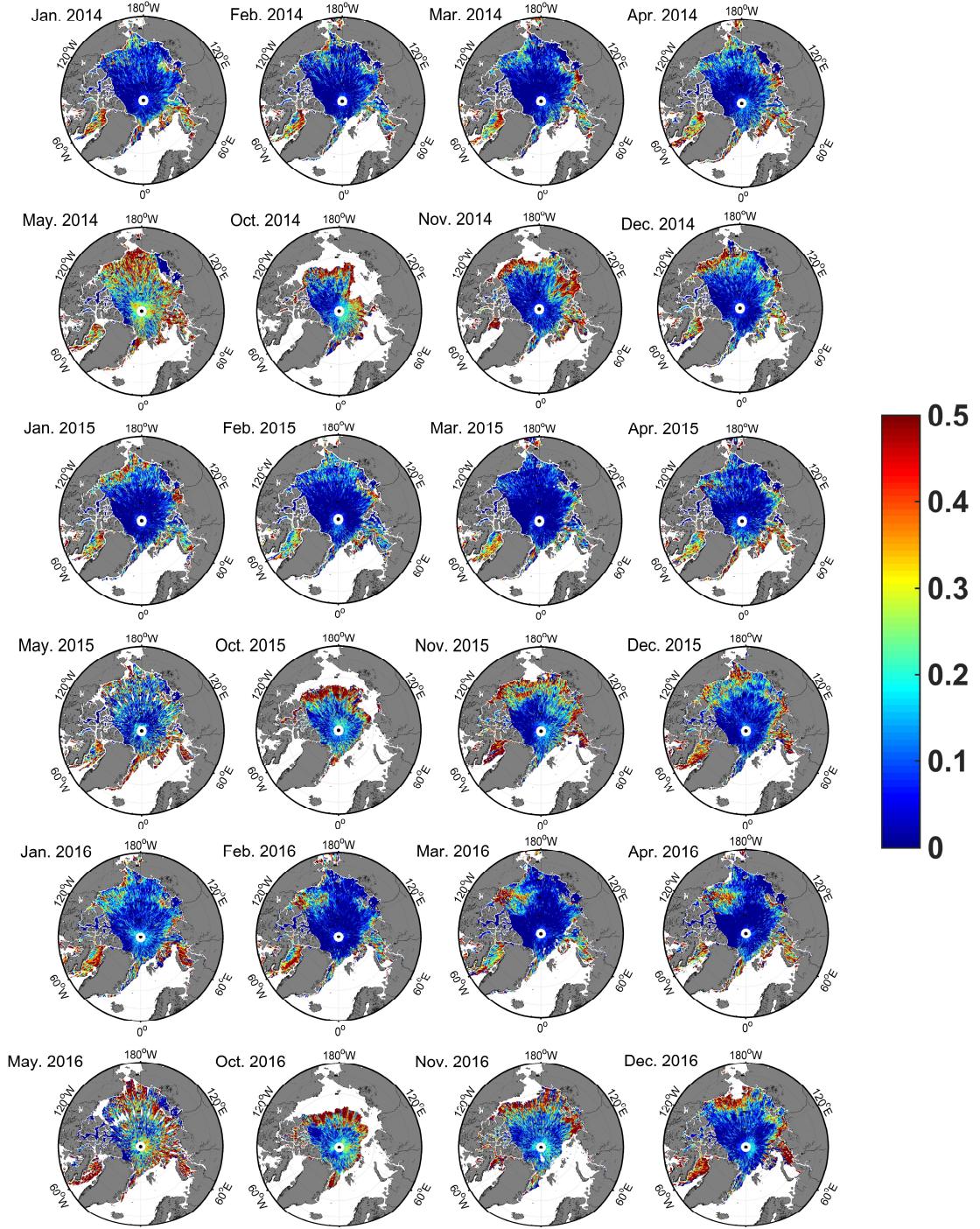
## 4.2 Spatiotemporal distribution of lead fraction maps

The monthly lead fraction maps with a 10 km grid in January to May, and October to December from 2011 to 2016 are shown in Figs. 6 and 7. The period from June to September is generally considered as the melting season. In this season, the presence of leads as well as melt pond in sea ice are dominant. It is difficult to accurately distinguish leads from sea ice due to the fact that waveform of the melt pond is quite similar to that of leads. Since the lead detection methods for the retrieval of sea ice thickness do not work well in the melting season, the sea ice thickness during the melting season is still unavailable (Tilling et al., 2017). We have compared lead fraction maps with the different spatial resolutions (i.e., 10, 50, and 100 km) to decide the proper spatial resolution. The spatial distribution of all lead fraction maps looked similar (not shown) because the ratios of lead observations to the entire CryoSat-2 observations did not significantly change among different spatial resolutions. Although the number of CryoSat-2 observations with a 10 km grid around the coastline is small (5-10), the greater number of observations in larger grids (50 and 100km) resulted in the similar distribution of lead fraction around the coastline. It is believed that the lead fraction maps with 10 km spatial resolution better represent the detailed spatial distribution of leads. The areas in the marginal ice zones line of the Arctic Ocean clearly show high lead fraction due to the shear zone (i.e., an area of deformed sea ice along the coast, and outflow of sea ice (Serreze and Barry, 2005). In particular, the high lead fraction was found around the Beaufort Sea during the spring season (MAM) because of the Beaufort Gyre, a wind-driven ocean current. It is widely known that the Chukchi Sea is the main strait through which warm Pacific water flows into the Arctic (Woodgate et al., 2006; Woodgate et al., 2010). However, the lead fraction around the Chuckchi Sea was lower than the lead fraction around the Beaufort Sea in January to April (i.e., winter season) 2011 and 2016, excluding 2015. While the lead fraction decreases from October to March (i.e., freezing season) with the minimum in March, the lead fraction starts to increase from April.

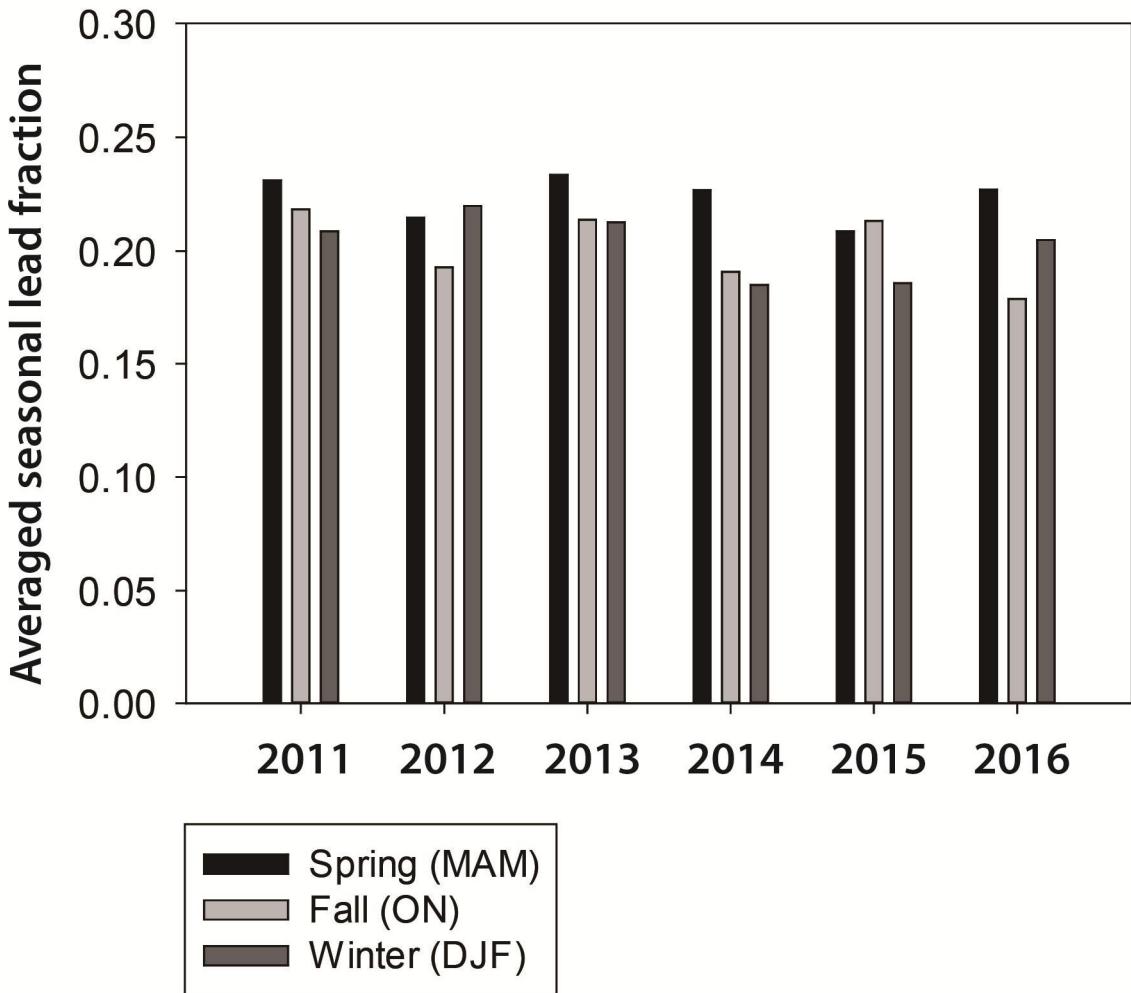
The Arctic Ocean circulations have contributed to the change in the state of sea ice. The lead fraction in Northwestern Greenland in Figs. 6 and 7 is low because of the convergence of sea ice by two major circulations, which was clearly shown in Kwok (2015). Kwok et al. (2013) revealed that the currents speed of Beaufort Gyre and Transpolar Drift increased from the years of 1982 to 2009 and this makes the fraction of multi-year ice decrease. However, the increasing lead fraction from the years of 2011 to 2016 in this study was not seen due to the high inter-annual variability of a lead fraction, particularly in the spring season (Fig. 8). High sensitivities in the marginal sea ice zone might result in not catching the increasing trend of Arctic lead fraction shown in the literature. In order to properly compare the Arctic current circulations and lead fraction, long-term lead fraction data are needed.



**Figure 6:** Monthly lead fraction maps based on waveform mixture analysis from January to May, October to December between 2011 and 2013. The range of the colour bar was set from 0 to 0.5 to emphasize lower values.



**Figure 7:** Monthly lead fraction maps based on waveform mixture analysis from January to May, October to December between 2014 and 2016. The range of the colour bar was set from 0 to 0.5 to emphasize lower values.



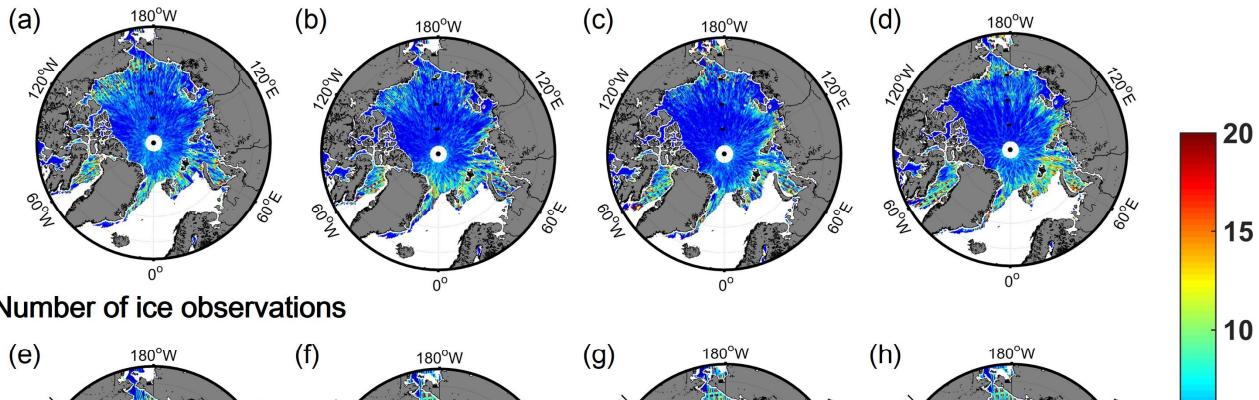
**Figure 8:** Averaged seasonal lead fraction in spring (MAM), fall (ON), and winter (DJF) between 2011 and 2016. The lead fraction from June to September was not available because leads were hard to distinguish from melt ponds using CryoSat-2 in the summer season.

## 5 4.3 Grid sensitive analysis in 10x10 km

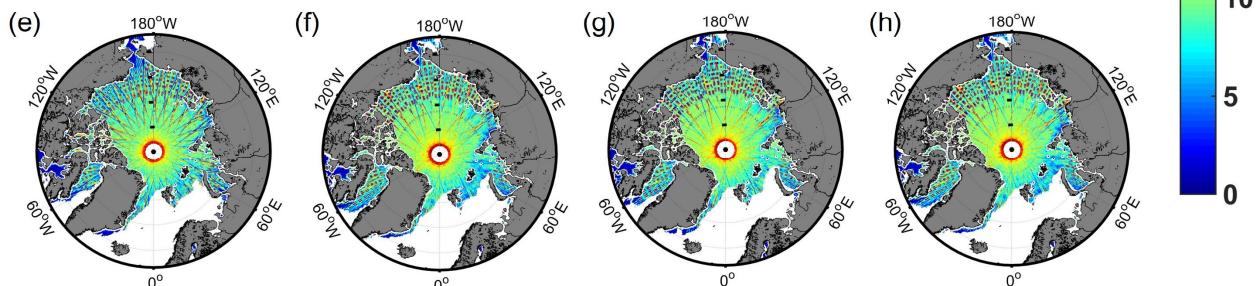
As mentioned in section 3.2, the number of CryoSat-2 observations decreases from the North Pole toward the coastline of Arctic Ocean. This results in an increase in statistical uncertainties when calculating monthly lead fraction around the coastline of Arctic Ocean based on the small number of CryoSat-2 observations. The number of lead and ice observations is

shown in Fig. 9a-h. While there are a few lead observations in the central Arctic, a large number of ice observations was found in the central Arctic. The high standard deviation values around the coastline of the Arctic Ocean imply that the reliability of lead fractions was low, while the relatively large number of CryoSat-2 observations around the North Pole produced low standard deviation indicating less sensitivity (Fig. 9i-l). There was a spatial difference of sensitivity by month (i.e., January to April) because of the different number of lead observations. Especially, since there was no lead observation in the East Siberian coast and Eastern Laptev Sea, the sensitivity (i.e., standard deviation) was also zero (Fig. 9c and d). It should be noted that the corresponding lead fraction might not represent an actual lead fraction in a 10 x 10 km grid. This is a drawback when calculating monthly lead fraction maps with satellite altimeters.

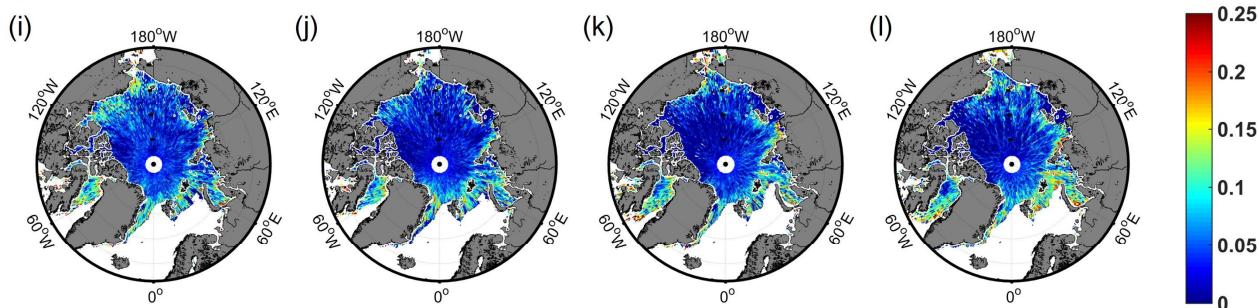
### Number of lead observations



### Number of ice observations



### Standard deviation



January 2011

February 2011

March 2011

April 2011

**Figure 9:** (a-d) the number of lead observations, (e-h) the number of ice observations, (i-l) the standard deviation of the results based on the sensitivity analysis of lead fraction from January to April 2011.

## 5. Discussion

### 5 5.1 Comparison of lead classification methods

Since the overall accuracy metrics of the proposed waveform mixture algorithm approach was comparable to those of the existing methods, especially DT, the waveform-based method can be used for estimating SSHA. Threshold-based lead detection methods have to be re-scaled whenever baseline data are updated. For example, beam behaviour parameters and backscatter sigma-0 changed slightly between when baseline B and C data were used. Thus, thresholds must also be updated 10 in order to appropriately identify leads using the threshold-based methods. However, the waveform mixture algorithm is less affected the change of baseline data because waveforms can still be used to detect leads using updated baseline data. This is the strong point of the waveform mixture algorithm when compared to the existing methods.

The use of waveform mixture algorithm might not work well to detect leads in cases of refreezing leads. In Figs. 4 c, g, k, and o, the dark area in the MODIS scenes around the latitude of  $84.26^{\circ}\text{N}$  and longitude of  $43^{\circ}\text{W}$  was determined to be a 15 lead class with visual inspection of the images and waveforms. Rose (2013) classified this region as ice. Laxon et al. (2013) and waveform mixture analysis detected one lead in that region. In Lee et al. (2016), DT detected more leads in that region than the other methods, but the validation could not entirely cover the dark area. In fact, since the leads are often refrozen, the shape of the waveforms in that region were likely more similar to the FYI waveform than the lead waveform (Zygmuntowska et al., 2013; Ricker et al., 2015; Lee et al., 2016). In the context of waveform mixture analysis, this region 20 could be classified as ice. Therefore, in order to more accurately detect leads, a surface elevation anomaly is needed as well as beam behaviour parameters, backscatter sigma-0, and the waveform mixture algorithm because the surface elevation anomaly on refreezing leads would be low, as in other leads.

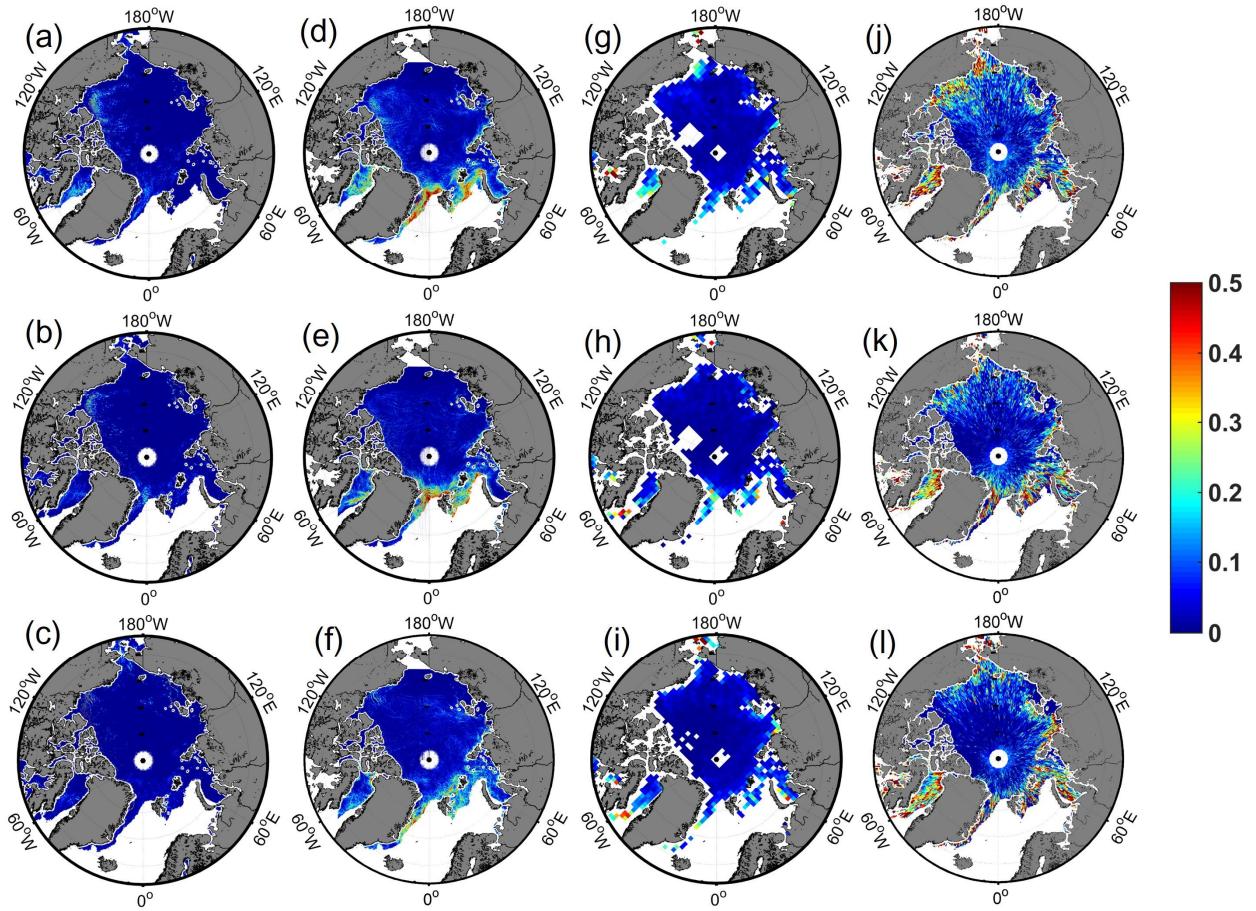
### 5.2 Comparison to other lead fraction maps

Four monthly lead fraction maps (Röhrs and Kaleschke, 2012; Wernecke and Kaleschke, 2015; Willmes and Heinemann, 25 2015) were compared to evaluate the pros and cons of each method used to produce the maps (Fig. 10). Basically, all four methods represent the spatiotemporal pattern of leads well for the freezing season from January to March. Scene-based lead fraction maps (i.e., AMSR-E in Figs. 10a, b and c, and MODIS in Figs. 10d, e, and f) and altimeter-based lead fraction maps (i.e., CryoSat-2 in Figs. 10g to l) have fundamentally different spatial characteristics as AMSR-E and MODIS are sensitive to different surface features. Scene-based lead fraction maps better represent the linear feature of leads and coastal polynya 30 than altimeter-based lead fraction maps. Since the AMSR-E-based approach only detects relatively large ( $\sim 3$  km) leads, lead fractions are generally lower than in the fraction maps using the other approaches. While altimeter-based lead fractions in

January 2011 (Figs. 10g and j) in the Chuckchi Sea were high, scene-based lead fractions (Figs. 10a to f) were low in January 2011. There are deformed and fragmented sea ices in the Chukchi Sea, which are different from the general lead shape. Altimeter-based lead detection methods identified leads between deformed and fragmented sea ices, generating a higher lead fraction in the Chukchi Sea in January 2011 (Figs. 10g and j). However, scene-based lead fraction methods did 5 not detect leads in the Chukchi Sea well, resulting in a lower lead fraction. The MODIS-based lead detection method that used ice surface temperature (IST) did not detect leads in the Chukchi Sea (Figs. 10d, e, and f). In the AMSR-E images, sea ice signals were dominant in the footprint around the Chukchi Sea and cracks between deformed and fragmented sea ices were identified as ice.

Altimeter-based monthly fraction maps might be insufficient to represent monthly lead fractions in the coastline of the 10 Arctic Ocean due to the limited number of CryoSat-2 observations in a month. Nonetheless, altimeter-based lead fraction maps documented the overall spatial distribution of leads reasonably; in particular, high lead fractions in the shear zone. Wernecke and Kaleschke (2015) used a random cross-validation technique to derive optimum thresholds based on ground references (i.e., MODIS images). They identified leads conservatively to reduce false classifications. The classification 15 results strongly depend on ground reference data. Since relatively high resolution (250m) MODIS images were used to construct reference data in this study, the waveform mixture algorithm was able to identify small leads through the calibration process of the abundance data (Fig. 4). Although the proposed waveform mixture algorithm produced lead fraction maps with a higher spatial resolution than those in Wernecke and Kaleschke (2015), the lead fractions around the coastline of the Arctic Ocean from Wernecke and Kaleschke (2015) appeared to have less sensitivity. This is because of the 20 larger number of lead observations in a much coarser grid than that from our results. The grid sensitivity analysis should be considered when interpreting the lead fraction maps around the coastline of the Arctic Ocean derived by the proposed waveform mixture analysis.

The choice of monthly lead fraction maps depends on the user's interest. Scene-based lead fraction maps better represent 25 coastal polynya and the intrinsic form of leads (Röhrs and Kaleschke, 2012; Willmes and Heinemann, 2016). CryoSat-2 based lead fraction maps might not represent the linear shape of typical leads well like cracks which include deformed and fragmented sea ices that are not in linear form. This is also a way to exchange heat and momentum transfer between the atmosphere and ocean, which can be detected as leads.



**Figure 10:** Comparison to other lead fraction maps in January to March 2011. (a-c) Monthly mean thin ice concentration maps using AMSR-E from Röhrs and Kaleschke (2011). (d-f) Monthly mean lead fraction maps using MODIS from Willmes and Heinemann (2015). (g-i) Monthly lead fraction maps using CryoSat-2 from Wernecke and Kaleschke (2015). (j-l) Monthly lead fraction maps based on waveform mixture analysis using Cryosat-2 in this study.

### 5.3 Novelty and limitations

In this study, we developed an alternative lead detection method (i.e., waveform mixture analysis) using CryoSat-2 L1b data, which can overcome the drawbacks of the previous threshold based lead detection methods. Regardless of an update in CryoSat-2 baseline data, the proposed waveform mixture algorithm can consistently identify leads without rescaling parameters such as beam behaviour parameters, pulse peakiness, and backscatter sigma-0. Such parameters must be re-scaled to implement threshold based lead detection methods when using updated CryoSat-2 baseline data. In addition, the proposed waveform mixture analysis outperformed the existing simple thresholding-based methods (Rose, 2013; Laxon et al., 2013), and was comparable to the machine learning-based thresholding method (Lee et al., 2016). **The proposed waveform mixture algorithm would be very useful in an operational system than the existing threshold-based methods including Lee et al. (2016). In addition, this study showed the high inter-annual variability of Pan-Arctic lead fractions in recent years (i.e., 2011-2016).**

On the other hand, the waveform mixture analysis depends on the quality of the endmembers. Although the use of the N-FINDR algorithm decreased the subjective selection of endmembers, waveform samples of leads and sea ice derived by DT algorithm from Lee et al. (2016) may introduce uncertainty because the algorithm was validated for March and April from 2011 to 2014. The leads that are not identifiable in the MODIS images were not considered in this study. Detecting leads smaller than the along track resolution of CryoSat-2 (~300m) with various lead detection methods should be further discussed in detail in future research using high resolution Landsat or SAR imagery. This is quite important in the retrieval of sea ice thickness using an altimeter because leads are used as the tie points for the sea surface height (SSH). For example, how the leads smaller than the along-track resolution of CryoSat-2 affect the waveform and SSH should be further investigated. The spatial resolution of monthly lead fraction maps improved up to 10 km, showing a detailed spatial distribution of leads in the Arctic. For example, 10km lead fractions showed significant variations in some regions, while 50 km or 100km lead fractions did not because lead fractions are averaged, resulting in blurred spatial patterns.

## 6. Conclusions

The waveform mixture algorithm was proposed to detect leads with CryoSat-2 L1b data. The lead and sea ice waveforms were considered as endmembers that are essential to implement waveform mixture algorithm. The endmembers (i.e., representative waveforms of leads and sea ice) were extracted by the N-FINDR algorithm among numerous waveforms (i.e., 420,858 waveforms of sea ice and 8,501 waveforms of leads). The thresholds to make a binary classification were determined by calibrating lead and sea ice abundances with reference data extracted from a high resolution (250m) MODIS images. The results show that the proposed approach robustly classified leads with comparable performance to DT from Lee et al. (2016) and slightly better than the existing simple thresholding approaches for lead detection (Rose 2013; Laxon et al., 2013). Furthermore, the lead detection of waveform mixture algorithm was comparable to the DT based lead detection

method (Lee et al., 2016), suggesting a sea ice freeboard can be retrieved with the robust lead detection method using waveform mixture analysis. Monthly lead fraction maps were produced using the proposed waveform mixture approach, showing clear inter-annual variability. The results of the lead fraction maps are consistent with the findings of recent studies (Tilling et al., 2015; Ricker et al., 2017; Kim et al., 2017).

5      Unlike the threshold-based lead detection methods, since the waveform mixture algorithm solely uses waveforms, not beam behaviour parameters, it does not need to change any parameters when the CryoSat-2 baseline version is updated, which will be useful for future altimeter missions. The recent inter-annual variability of Arctic sea ice conditions was found. In this context, this waveform mixture algorithm can be used to consistently produce monthly lead fraction maps during the extended CryoSat-2 mission for monitoring Arctic sea ice.

10

### Acknowledgments

This study was supported by the Korea Polar Research Institute (KOPRI) grant PE180120 (Research on analytical techniques for satellite observations of Arctic sea ice).

### References

15    Aaboe, S., Breivik, L., Sorensen, A., Eastwood, S., and Lavergne, T.: Global Sea Ice Edge and Type Product User's Manual, EUMETSAT OSISAF, 2016.

     Amani, M., Salehi, B., Mahdavi, S., Granger, J., Brisco, B.: Wetland classification in Newfoundland and Labrador using multi-source SAR and optical data integration. *GIScience and Remote Sensing*, in press, 2017.

     Bröhan, D., and Kaleschke, L.: A Nine-Year Climatology of Arctic Sea Ice Lead Orientation and Frequency from AMSR-E, 20    *Remote Sensing*, 6, 1451, 2014.

     Chase, J. R., and Hollyer, R. J.: Estimation of sea ice type and concentration by linear unmixing of Geosat altimeter waveforms, *Journal of Geophysical Research: Oceans*, 95, 18015-18025, 10.1029/JC095iC10p18015, 1990.

     Chi, J., Kim, H.-C., and Kang, S.-H.: Machine learning-based temporal mixture analysis of hypertemporal Antarctic sea ice data, *Remote Sensing Letters*, 7, 190-199, 10.1080/2150704X.2015.1121300, 2016.

25    Ertürk, A., and Plaza, A.: Informative Change Detection by Unmixing for Hyperspectral Images, *IEEE Geoscience and Remote Sensing Letters*, 12, 1252-1256, 10.1109/LGRS.2015.2390973, 2015.

     Foody, G. M., and Cox, D. P.: Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions, *Remote sensing*, 15(3), 619-631, 1994.

     Hisabayashi, M., Rogan, J., and Elmes, A.: Quantifying shoreline change in Funafuti Atoll, Tuvalu using a time series of 30    Quickbird, Worldview and Landsat data. *GIScience & Remote Sensing*, 1-24, 2018.

     Iordache, M. D., Bioucas-Dias, J. M., and Plaza, A.: Sparse Unmixing of Hyperspectral Data, *IEEE Transactions on*

del and observations, Journal of Geophysical Research: Oceans, 120, 4925-4944, 10.1002/2014JC010635, 2015.

Ji, L., Geng, X., Sun, K., Zhao, Y., and Gong, P.: Modified N-FINDR endmember extraction algorithm for remote-sensing imagery, International Journal of Remote Sensing, 36, 2148-2162, 10.1080/01431161.2015.1034895, 2015.

5 Keshava, N., and Mustard, J. F.: Spectral unmixing, IEEE Signal Processing Magazine, 19, 44-57, 10.1109/79.974727, 2002.

Key, J., Stone, R., Maslanik, J., and Ellefsen, E.: The detectability of sea-ice leads in satellite data as a function of atmospheric conditions and measurement scale, Annals of Glaciology, 17, 227-232, 10.3198/1993AoG17-1-227-232, 1993.

10 Kim, B.-M., Hong, J.-Y., Jun, S.-Y., Zhang, X., Kwon, H., Kim, S.-J., Kim, J.-H., Kim, S.-W., and Kim, H.-K.: Major cause of unprecedented Arctic warming in January 2016: Critical role of an Atlantic windstorm, Scientific Reports, 7, 40051, 10.1038/srep40051, <http://www.nature.com/articles/srep40051#supplementary-information>, 2017.

Kim, M., Im, J., Han, H., Kim, J., Lee, S., Shin, M., Kim, H.-C. Landfast sea ice monitoring using multisensor fusion in the antarctic. *GIScience & Remote Sensing*, 52, 239-256, 2015.

Kwok, R., and Untersteiner, N.: The thinning of Arctic sea ice, Physics today, 64(4), 36-41. 2011.

15 Kwok, R., Spreen, G., and Pang, S.: Arctic sea ice circulation and drift speed: Decadal trends and ocean currents, Journal of Geophysical Research: Oceans, 118, 2408-2425, 10.1002/jgrc.20191, 2013.

Kwok, R.: Sea ice convergence along the Arctic coasts of Greenland and the Canadian Arctic Archipelago: Variability and extremes (1992–2014), Geophysical Research Letters, 42, 7598-7605, 10.1002/2015GL065462, 2015.

20 Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., Hendricks, S., Krishfield, R., Kurtz, N., Farrell, S., and Davidson, M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, Geophysical Research Letters, 40, 732-737, 10.1002/grl.50193, 2013.

Lee, S., Im, J., Kim, J., Kim, M., Shin, M., Kim, H.-c., and Quackenbush, L.: Arctic Sea Ice Thickness Estimation from CryoSat-2 Satellite Data Using Machine Learning-Based Lead Detection, Remote Sensing, 8, 698, 2016.

25 Lindsay, R. W., and Rothrock, D. A.: Arctic sea ice leads from advanced very high resolution radiometer images, Journal of Geophysical Research: Oceans, 100, 4533-4544, 10.1029/94JC02393, 1995.

Lu, D., Moran, E., & Batistella, M.: Linear mixture model applied to Amazonian vegetation classification, Remote sensing of environment, 87(4), 456-469, 2003.

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, Geophysical Research Letters, 35, n/a-n/a, 10.1029/2007GL032461, 2008.

30 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, 10.5194/tc-6-143-2012, 2012.

Maykut, G. A.: Energy exchange over young sea ice in the central Arctic, Journal of Geophysical Research: Oceans, 83(C7), 3646-3658, 1978.

Maykut, G. A.: Large-scale heat exchange and ice production in the central Arctic, Journal of Geophysical Research: Oceans,

Miles, M. W., and Barry, R. G.: A 5-year satellite climatology of winter sea ice leads in the western Arctic, *Journal of Geophysical Research: Oceans*, 103(C10), 21723-21734, 1998

Perovich, D. K., Jones, K. F., Light, B., Eicken, H., Markus, T., Stroeve, J., and Lindsay, R.: Solar partitioning in a changing Arctic sea-ice cover, *Annals of Glaciology*, 52(57), 192-196, 2011

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

Rose, S., Forsberg, R., and Pedersen, L.: Measurements of sea ice by satellite and airborne altimetry, DMIDMI,

Ricker, R., Hendricks, S., Helm, V., and Gerdes, R.: Classification of CryoSat-2 radar echoes. In *Towards an Interdisciplinary Approach in Earth System Science* (pp. 149-158), Springer International Publishing, 2015.

Ricker, R., Hendricks, S., Girard-Ardhuin, F., Kaleschke, L., Lique, C., Tian-Kunze, X., Nicolaus, M., and Krumpen, T.: Satellite-observed drop of Arctic sea-ice growth in winter 2015-2016, *Geophysical Research Letters*, n/a-n/a, 10.1002/2016GL072244, 2017.

Salvatore D.: Guidelines for reverting Waveform Power to Sigma Nought for CryoSat-2 in SAR mode, European Space Agency, 2014.

Scagliola, M., and Fornari, M.: Main evolutions and expected quality improvements in BaselineC Level1b products (No. 1.3), C2-TN-ARS-GS-5154, 2015.

Serreze, M. and Barry, R.: The Arctic climate system, PP185-186, Cambridge, 2015.

Spall, M. A.: Circulation and water mass transformation in a model of the Chukchi Sea, *Journal of Geophysical Research: Oceans*, 112, n/a-n/a, 10.1029/2005JC003364, 2007.

Tilling, R. L., Ridout, A., Shepherd, A., and Wingham, D. J.: Increased Arctic sea ice volume after anomalously low melting in 2013, *Nature Geosci*, 8, 643-646, 10.1038/ngeo2489 <http://www.nature.com/ngeo/journal/v8/n8/abs/ngeo2489.html#supplementary-information>, 2015.

Tadesse, T., Champagne, C., Wardlow, B., Hadwen, T., Brown, J., Demisse, G., Bayissa, Y., Davidson, A.: Building the vegetation drought response index for Canada (BegDRI-Canada) to monitor agricultural drought: first results. *GIScience and Remote Sensing*, 54(2), 230-257, 2017.

Torbick, N., Corbiere, M.: Mapping urban sprawl and impervious surfaces in the northeast United States for the past four decades. *GIScience & Remote Sensing*, 52(6), 746-764, 2015.

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, 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(69), 29-37, 2015

Willmes, S., and Heinemann, G.: Sea-Ice Wintertime Lead Frequencies and Regional Characteristics in the Arctic, 2003–2015, *Remote Sensing*, 8, 4, 2016.

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., de Chateau-Thierry, P., Laxon, S. W., Mallow, U., Mavrocordatos, C., Phalippou, L., Ratier, G., Rey, L., Rostan, F., Viau, P., and Wallis, D. W.: CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields, *Advances in Space Research*, 37, 841-871, <http://dx.doi.org/10.1016/j.asr.2005.07.027>, 2006.

5 Winter, M. E.: N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data, 1999, 266-275,

Woodgate, R. A., Aagaard, K., and Weingartner, T. J.: Interannual changes in the Bering Strait fluxes of volume, heat and freshwater between 1991 and 2004, *Geophysical Research Letters*, 33, n/a-n/a, 10.1029/2006GL026931, 2006.

Woodgate, R. A., Weingartner, T., and Lindsay, R.: The 2007 Bering Strait oceanic heat flux and anomalous Arctic sea-ice  
10 retreat, *Geophysical Research Letters*, 37, n/a-n/a, 10.1029/2009GL041621, 2010.

Wu, C.: Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery, *Remote Sensing of Environment*, 93(4), 480-492, 2004

Zakharova, E. A., Fleury, S., Guerreiro, K., Willmes, S., Rémy, F., Kouraev, A. V., and Heinemann, G.: Sea Ice Leads Detection Using SARAL/AltiKa Altimeter, *Marine Geodesy*, 38, 522-533, 10.1080/01490419.2015.1019655, 2015.

15 Zortea, M., and Plaza, A.: A Quantitative and Comparative Analysis of Different Implementations of N-FINDR: A Fast Endmember Extraction Algorithm, *IEEE Geoscience and Remote Sensing Letters*, 6, 787-791, 10.1109/LGRS.2009.2025520, 2009.

Zygmuntowska, M., Khvorostovsky, K., Helm, V., and Sandven, S.: Waveform classification of airborne synthetic aperture radar altimeter over Arctic sea ice, *The Cryosphere*, 7, 1315-1324, 10.5194/tc-7-1315-2013, 2013.