

Accepted Manuscript

Lead Detection using Cryosat-2 Delay-Doppler Processing and Sentinel-1 SAR images

Marcello Passaro, Felix L. Müller, Denise Dettmering

PII: S0273-1177(17)30505-7

DOI: <http://dx.doi.org/10.1016/j.asr.2017.07.011>

Reference: JASR 13321

To appear in: *Advances in Space Research*

Accepted Date: 7 July 2017



Please cite this article as: Passaro, M., Müller, F.L., Dettmering, D., Lead Detection using Cryosat-2 Delay-Doppler Processing and Sentinel-1 SAR images, *Advances in Space Research* (2017), doi: <http://dx.doi.org/10.1016/j.asr.2017.07.011>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Lead Detection using Cryosat-2 Delay-Doppler Processing and Sentinel-1 SAR images

Marcello Passaro*

*Deutsches Geodätisches Forschungsinstitut der Technischen Universität München,
Arcisstraße 21, 80333 Munich, Germany*

Felix L. Müller and Denise Dettmering

*Deutsches Geodätisches Forschungsinstitut der Technischen Universität München,
Arcisstraße 21, 80333 Munich, Germany*

Abstract

In the Arctic and Antarctic Ocean, where part of the sea surface is seasonally or continuously covered by sea ice, the sea level monitoring from satellite altimetry relies on the localisation of open water areas, especially on the detection of leads: long and narrow fractures in the sea ice, which dominate the radar echoes even if hundreds of meters away from nadir.

The Cryosat-2 altimetry mission is based on the Delay-Doppler processing, in which the averaged waveform is formed by summing up several looks acquired at different look angles and stacked together. This imaging technique and the resulting improved along-track resolution are here exploited to evaluate different lead identification schemes.

In particular, stack and power statistics of Cryosat-2 waveforms are used to classify leads on a subset of 12 tracks in which the altimetry-based classification is compared to a classification based on Sentinel-1A SAR images. For this scope, a dedicated SAR-image automated processing is proposed to avoid the manual classification.

Results show that the adoption of a single new stack parameter (the Stack Peakiness) can perform equally well as the use of multiple stack parameters currently available. Moreover, a multi-waveform analysis of the Stack Peak-

*Corresponding author

Email address: marcello.passaro@tum.de (Marcello Passaro)

iness helps to isolate the point where narrow leads cross the tracks at nadir.

For all the tested strategies, the number of altimetry-detected leads that are unidentified by SAR is comparable to the number of detections from both sensors. This could be due to presence of narrow leads, not detected by SAR due to resolution limits, but still dominant in the radar altimeter return due to the high backscatter .

Keywords: Leads Detection; Delay-Doppler Altimetry; Cryosat-2 Stack Data; Sentinel-1; SAR Image Processing;

1. Introduction

The measurement of sea level variability in the global ocean is considered among the most important climatic indices. It relies on in-situ observations provided by a wide but unevenly distributed set of tide gauges and, since more than 20 years, on measurements collected by the radar altimeters on board of several satellite missions.

The coverage of satellite altimetry over the ocean cannot completely be defined as global, since a large part of the Arctic and Antarctic oceans is excluded. On one side this is due to the limited latitude extent of most of the altimetric missions due to their orbit configuration. On the other side the ocean in the northernmost latitudes is partially covered by sea ice, which reflects the radar signal before it hits the sea surface, preventing the possibility to measure sea level. The estimation in the sea-ice covered regions is limited to the leads, narrow cracks in the sea ice that can be several tens of kms long. Since these ocean patches are very smooth and do not have a developed wave field, the signal returned to the satellite is much stronger than the one reflected from the surrounding ice and can dominate the registered waveforms even if the lead is not located at nadir. In particular, Armitage and Davidson (2014) have shown that a lead can be the dominant return in the waveform up to about 1.5 km away from the sub-satellite point (nadir). Such off-nadir returns, if not properly spotted, result in erroneous estimations of the sea level.

Cryosat-2 (CS-2) offers ways for improving the sea level records in these regions. With its orbit configuration, it provides coverage up to 88° in latitude. Thanks to the Delay-Doppler processing of its echoes (when operating in the so-called "SAR mode" over sea ice, not to be confused with SAR imaging from Sentinel-1 used in this study), it stores the signal registered

by the satellite looking at the same resolution cell on the ground from different look angles. In particular, the beam-limited along-track footprint size (305 m, (Scagliola, 2013)) should guarantee a more precise determination of the lead position. Nevertheless, due to the size of the pulse-limited across-track footprint (1.65 km), the distinction of a lead return at nadir from an off-nadir reflection is still challenging. Most of the leads have width of less than a km (Lindsay and Rothrock, 1995; Kwok et al., 2009), while Cryosat-2 has a sampling interval of roughly 300 m (using the 20-Hz rate): in most of the cases, only one range measurement per lead will correspond to the distance at nadir. Being able to correctly identify the nadir echoes of these narrow, but numerous open water openings can increase the amount of sea level measurements and therefore improve the records.

Previous studies on past altimetry missions have used lead-detection algorithms that distinguish leads from sea ice based on the shape of the received signal: Empirical thresholds were assigned in order to classify the waveforms based on the "pulse peakiness" (Peacock and Laxon, 2004). Laxon et al. (2013); Ricker et al. (2014); Rinne and Similä (2014) have classified CS-2 signals using a combination of different waveform parameters (including the pulse peakiness) available in the European Space Agency (ESA) Baseline C Product files or computable from the waveforms. Recently, Wernecke and Kaleschke (2015) argued that it is possible to obtain an efficient lead classification only based on the absolute value of the maximum waveform power.

Leads can be also determined using thermal infrared sensors (Willmes and Heinemann, 2015), microwave radiometers (Röhrs and Kaleschke, 2012) and SAR images (Ivanova et al., 2016). SAR images have the advantage of being independent from weather conditions, while providing a good resolution (40m for Sentinel 1A). They can be therefore used for comparison with the altimetry-based lead classification, but the time difference between the acquisition of the two different data sources needs to be taken in consideration, since sea ice moves on average from 4 km/day in winter up to over 9 km/day in summer (as measured by buoys in Rampal et al. (2009)) and leads can quickly refreeze and close (Weeks, 2010).

The objective of this study is to provide a first assessment of the lead-classification methodologies based on the Delay-Doppler processing of Cryosat-2 echoes in comparison to SAR images from Sentinel-1A. Our classification, based on a new parameter computed using the Delay-Doppler processing of CS-2 (in particular from the full stack information) and on a multi-waveform analysis to isolate the nadir return, is compared with the methodologies de-

rived from the recent literature. A SAR-image processing chain is proposed to provide a reference for validation and, for the first time, is used to provide an objective comparison that is not based on a visual recognition of lead-like features.

A description of the dataset and the area of study is provided in Section 2. Section 3 describes the methodology used to analyse the altimetry and SAR dataset and to classify the leads. In Section 4 the results of the comparison are presented and discussed. Section 5 draws the conclusions and the outlook for future research.

2. Dataset

2.1. Cryosat-2 L1B-S data

By exploiting the Doppler frequency and the coherence of consecutive pulses, Delay-Doppler altimeters are able to perform multi-looked acquisitions, i.e. to associate to a resolution cell a certain number of looks (variable depending on the processing settings) acquired at different look angles as the satellite moves over the imaged area (Raney, 1998).

Using processing techniques inherited from the SAR processing, such as Range compression and Range migration correction, all the returns corresponding to the resolution cell (a 20-Hz sampling of the illuminated surface, i.e. one measurement every 300 m roughly) are aligned in a stack diagram (Figures 1a and 2a). The Cryosat-2 multilooked radar waveforms, such as the one in Figure 1c and 2c, are obtained by the incoherent sum of all the echoes in the stack. By summing up the returns in the across-track (Range) dimension (Figure 1b and 2b), the so-called stack waveform, or Range Integrated Power (RIP) waveform, can be generated. It contains information concerning the backscattering properties of the illuminated surface, but it also reveals details of the distribution of the scatterers as the satellite spans different look angles passing over the nadir position (Wingham et al., 2006).

When the satellite moves over a very smooth surface, such as for small lakes or leads (Figure 1), the signal will be specularly reflected back and the RIP will be peaky. On the opposite, when flying over areas containing scatterers with different orientation, such as for wavy seas or ice, the backscattered power will be more normally distributed (Figure 2).

L1B products provide statistical parameters that describe the RIP behaviour, but do not provide the full stack waveform, limiting therefore the

possibilities of analysis. The ESA Grid Processing on Demand (G-POD) service (see acknowledgements) is currently the only freely and easily available source to access these data. The ESA Baseline C version of this L1B stack (L1B-S) data set is used in this study (Scagliola and Fornari, 2015).

2.2. Sentinel-1A SAR images

One of the ways to verify the classification of altimetric echoes is the comparison with satellite SAR images. These have been used for lead detection, since very smooth water areas reflect electromagnetic waves like a mirror: consequently the slant incident radar waves reflect away from the spacecraft and flat water areas appear dark (Dierking, 2013). Sentinel-1A SAR images are provided with two polarization modes HH and HV (where "H" indicates horizontal and "V" vertical). HH- and HV-polarization are particularly suitable for ice versus open water discrimination because of decreasing ocean clutter and smaller sensitivity to wind and wave scattering.

In the present investigation Level-1 dual-polarized SAR Sentinel-1A extra-wide-swath mode data at medium resolution (S1A-EW-GRDM-1SDH) are employed. The images are ground-range detected showing a 40-meter spatial resolution and a 400-km swath width, which allows a wide spatial coverage and a short revisit time. The images were pre-processed using the following standard procedures: (1) Thermal noise removal, (2) Radiometric calibration, (3) Range Doppler terrain correction (Veci (2016)). The latter includes a coordinate transformation into an azimuthal equal area map projection to provide the same coordinate background as the sea ice motion vectors used to relate the images to the time of the altimetry overflight (see 2.3). In a last step, a type conversion to *uint8* is performed in order to get grayscaled values and to reduce disk space.

In this study, two sets of control data of Sentinel-1A HH-polarised images are used (see Figure 3): Set 1 comprises six images from September 2015 from the Arctic Ocean north of the Fram Strait. They are taken as a reference for comparison between the classification proposed in this study and the one described in Ricker et al. (2014). Set 2 includes six additional images taken between the north-east coast of Greenland and the Fram Strait. They are exploited as a further comparison using a different area at various times of the year (November 2014 to January 2015).

All images were selected because of the time proximity with collocated Cryosat-2 tracks (never more than four hours of time difference) to reduce

the influence of sea ice motion between the acquisition date of the imaging SAR and the altimetry record.

2.3. Sea ice velocities

Since sea ice can move significantly even in short time periods, it is desirable to take the ice velocity into account when comparing altimetry results with SAR images. For this purpose daily sea ice motion vectors provided by the National Snow and Ice Data Center (NSIDC) are used. At the time of writing, daily NSIDC ice motion vector fields are only available until 31 May 2015: therefore, the sea ice motion vectors are not applied to Set 1. The "Polar Pathfinder Daily 25km EASE-Grid Sea Ice Motion Vectors, Version 3" are based on data derived from different active and passive satellite sensors as well as in-situ data. The data are provided at 25-km spatial resolution (Tschudi et al., 2016). The ice motion information is provided as zonal and meridional velocity grids.

Based on these data a mean velocity (magnitude and direction) is estimated by averaging all points in a ± 35 km rectangle box around each Cryosat-2 track. The SAR image, i.e., each of its pixel coordinates, is then shifted taking the acquisition time difference between SAR and CS-2 into account.

3. Methodology

3.1. Processing of Delay-Doppler altimetry data

3.1.1. Application of the Hamming window

When using L1B-S data in G-POD, it is possible to exploit different Delay-Doppler processing configurations. In order to apply the lead classification derived in this study, the Hamming-windowing step before the along-track Fast Fourier Transform is selected, which is currently the baseline of the distributed Cryosat-2 product (Bouzinac, 2012). Although this slightly lowers the along-track resolution and therefore creates some degree of dependence between consecutive echoes, it is needed in order to cut the energy coming from the sidelobes of the antenna.

Figure 4 shows the so-called "radargrams" of a CS-2 track in a sea-ice covered regions for the multilooked waveforms and for the RIP, without (a and c) and with (b and d) the Hamming window application. Each column corresponds to a 20-Hz RIP (a and b) or multilooked waveform (c and d). Before and after the peaky echoes typical of lead-like backscatter, high-power

features are seen preceding the leading edge in the multilooked waveform radargram and in the non-zero look angles of the stack radargram. In fact, the return coming from a sidelobe that sees a lead at nadir when the main lobe is side-looking has a shorter range and therefore is registered before the leading edge corresponding to the resolution cell.

The sidelobe effects create false leading edges, influence the statistical analysis of the Range Integrated Power and add backscattering of the same order of magnitude of the nadir return in the look angles closer to zero. These features mostly disappear after the application of the Hamming window, although residual signatures are visible, as highlighted by the arrows in Figure 4b.

3.1.2. Definition of Stack Peakiness

In order to characterise the RIP shape, the Stack Standard Deviation (SSD) and the Stack Kurtosis (SK) are already given in the ESA Baseline C product. These two indices, although useful to classify the kind of waveform, are not sufficient to isolate the nadir return of a group of waveforms influenced by a lead backscatter. The SSD is based on a gaussian fitting of the RIP, which is a poor approximation for peaky returns such as in Figure 1b. The SK is highly influenced by remaining sidelobes effects in the looks that are close to the zero look, which can have a similar power. In order to compare the power at the zero look angle with the backscatter registered in the other looks, a new parameter called Stack Peakiness (SP) is defined in this study from the RIP normalised by its maximum value in the following way:

$$SP = \frac{1}{\overline{RIP}_{l,r}} \quad (1)$$

with

$$\overline{P}_{l,r} = \frac{\sum_{i=1}^N RIP(i)_{l,r}}{N} \quad (2)$$

where N is the number of looks excluding the nadir look and $RIP(i)_{l,r}$ is the power from the look angle i , excluding the nadir look (i.e., at its right or left). A similar index of peakiness of the main return in comparison with the rest of echo was already defined in Ricker et al. (2014). Nevertheless this statistics was computed on the multilooked waveform, which is actually a time series of the received signal, while SP is able to compare the power reflected from the same resolution cell at different view angles.

The expected behaviour of SP in the case of a narrow lead crossing the CS-2 track is sketched in Figure 5 and a verification with real data is provided in section 4.1. When a lead enters the across-track pulse-limited footprint, the SP will be higher than a purely diffusive backscatter event, since the lead will scatter more energy back to the satellite. Nevertheless, the lead will still be slightly off-nadir in the across-track direction: Part of the incoming power will be specularly reflected away. Off-nadir leads are usually characterised by lower levels of backscatter power compared to leads at nadir (Wernecké and Kaleschke, 2015). Consequently, a lead located off-nadir in the across-track direction will scatter less power back to the satellite, if compared with the same lead illuminated at nadir. The maximum SP, i.e. the time when the power at the zero look angle is strongest in comparison with the backscatter received at the other look angles, is therefore expected to correspond to the position in which the lead is at nadir.

3.2. Lead classification based on Cryosat-2

3.2.1. Use of Stack Peakiness

In this study, a local maxima of SP, such as the one in Figure 5, is considered a potential lead waveform. The SP is almost constant over sea ice, but peaks in presence of a lead, as shown in section 4.1. Two additional criteria are used to identify the nadir leads:

- Median SP: The analysis of SP over all CS-2 tracks used for the validation with Set 1 and Set 2 (considering the whole length of each track over the Delay-Doppler geographical box) shows a median value of $SP_{median} = 8.67$ with a Median Absolute Deviation (MAD) of $SP_{MAD} = 8.72$. In order for a local maxima of SP to be classified as lead, the SP of the RIP before and after the local maxima has to be higher than $SP_{median} + SP_{MAD} \simeq 17$. This is done in order to avoid false detections, by utilising the fact that a lead that crosses the altimeter track influences the SP value also when not yet at nadir.
- Minimum SP: An empirical threshold $SP_{minlead} = 40$ is identified as the minimum SP of the RIP local maxima to be classified as lead. The threshold has been set by empirical observations of the locations of CS-2 points characterised by low SP values on the corresponding SAR images. This additional criteria is added in order to limit the recognition of leads that enter the field of view of the satellite, but

never cross the nadir position. Their SP maxima is therefore expected to be lower than the SP maxima of nadir-crossing leads.

In the following sections, SP will be also used as acronym of the corresponding classification method.

3.2.2. Use of Received Power

Each Delay-Doppler waveform is characterised by a received power. The received power depends on the backscattering characteristic of the surface: flat surfaces such as the still water of small leads or melting ponds specularly scatter most of the incoming radar signal back in the same direction, while ice surfaces are characterised by diffuse scattering, which decreases the amount of power reflected back to the altimeter. In this study (section 4.3) the adoption of an absolute threshold on received power to identify leads is verified and discussed. The applied threshold is 2.58×10^{-11} W, as proposed by Wernecke and Kaleschke (2015).

3.2.3. External lead classification

As previously mentioned, the classification proposed in Ricker et al. (2014) is used for comparison. The results of the classification were provided by the Alfred Wegener Institut (see Acknowledgements) for Set 1. The method consists on the use of thresholds set on three waveform-derived parameters (pulse peakiness, peakiness right of the power maximum, peakiness left of the power maximum), two RIP-derived parameters (SK and SSD) and a sea-ice concentration index. The peakiness right and left of the power maximum was directly computed from the waveform, while all the other parameters are provided in the raw data.

3.3. Lead classification based on Sentinel-1 SAR images

A set of image processing algorithms is applied to the Sentinel-1A scenes that should be used as a reference for the lead identification based on CS-2, aimed at extracting open water areas, i.e. black or near black surface areas, by converting the SAR images into binary format. Leads or polynyas are represented by ones, while ice is coded with zeros. In order to replace the common visual classification an automated SAR image processing is proposed that enables quantitative comparisons with altimetry classification results. The following steps are applied to the SAR images (previously shifted considering the ice motion as described in section 2.3):

- Noise reduction: To reduce noise in the image due to interfering scattering, a median filter is applied to the grayscaled image. For this purpose a window size of 5 x 5 pixels (equal to 200 m x 200 m spatial scale) has been chosen to emphasize the transition between ice and water pixels and to minimize a false detection of open water pixels. The filter size is a compromise between noise reduction and compliance with the original image and was experimentally determined.
- Dark pixel emphasizing: After median filtering the grayscale image undergoes a minimum, non-linear filtering emphasizing dark pixel values. This is necessary to compensate uncertainties of the image shifting due to the ice motion (Sect. 2.3). To control the effect of the minimum filtering a convolution matrix or kernel is needed. Considering the linear and circular shape of openings in the ice, reliable results are reached by using an octagon kernel with a radius of 3 pixels around the center pixel. In Figure 6, a SAR image before (a) and after median and minimum filtering (b) as well as the used kernel (c) are shown.
- Conversion to binary map: To convert the filtered grayscale image into binary values, a segmentation based on thresholding is applied. For this purpose an adaptive threshold algorithm is employed to compensate spatial variations in contrast and illumination. We follow the approach of Bradley and Roth (2007) that divides the SAR image in foreground and background pixels. In a first step the integral image, a summation of pixel values from top left to bottom right, is computed. The next step computes the average of every pixel in a given neighborhood. The last processing step separates the SAR image in background and foreground by comparing the averaged pixel to the integral image.
- Interconnection of lead fragments: The spatial extent and the shape of a lead can vary very quickly from one meter to over 500 m due to persistent ice motion and refreezing (Onstott and Shuchman (2004)). In SAR images leads can show different pattern and pixel values. For example if the leads are refrozen or covered by frost flowers, the pixel values brighten up. Furthermore open water SAR signatures are sensitive to wind conditions. If there are calm conditions, leads appear small and disconnected and could be obscured by surrounding ice (Onstott and Shuchman (2004)). Additionally leads can be segmented due to limited resolution of the SAR image and inaccuracies of the thresholding.

In order to reconnect these leads, a morphological closing operation is applied to the binary images. The closing operator is a consecutive execution of a morphological dilatation followed by a morphological erosion. It enlarges pixel areas by mainly keeping the original boundary shape, thus it fills gaps and connects objects in a specific range. More details regarding the effect of closing operation on binary images can be found in Gonzalez and Woods (2008). As a closing operator, an octagonal kernel with radius of 12 pixels around the center pixel has been chosen based on empirical tests. Figure 6 visualizes the effect of the closing operation on the segmented binary image (from c to d). The thin adjacent lead fragments are connected by mainly preserving their natural linear spatial extent. As a downside, independent openings that are closely located can be linked, resulting in one widespread open water area.

For the statistical comparison between CS-2 and SAR lead classifications, the binary pixel values of the processed SAR image are interpolated to the altimetry track coordinates, using nearest-neighbor interpolation.

4. Results and discussion

4.1. Analysis of stack parameters

Figure 7 shows the evolution of the SP over a sea ice covered area in comparison with the Kurtosis and SSD stored in the ESA Baseline C product. Points that are identified as leads by the SP classification described in section 3.2.1 are highlighted by circles. For comparison, points that are identified as leads using the classification from Ricker et al. (2014) (Ricker from now on) are highlighted by squares.

The evolution of SP in the lead areas closely resembles the scheme of Figure 5: a peak, which corresponds to the strongest return from the zero look angle compared to the other looks, is easily identifiable, but the lead returns influence also the measurements nearby. The lead areas are also characterized by high Kurtosis and low SSD, but these two indices are often not able to univocally show a local maximum or minimum: in some cases, the Kurtosis shows multiple peaks in a same sequence of points influenced by a lead, which may be attributed to high power in non-zero look angles due to residual sidelobe effects; the SSD, being based on a Gaussian fitting,

is not able to distinguish subtle differences in the power distribution of the very peaky RIP waveforms in the lead areas.

Classifications that are based on these two stack parameters, such as Ricker, tend to classify as leads more points related to the same feature than the SP classification, which performs a sort of multi-waveform approach by looking for local maxima, rather than only considering thresholds on single measurements. In the case of narrow leads, a single-waveform approach can imply that off-nadir returns are being considered as leads. An example is found in Figure 8, in which the CS-2 track crosses a narrow lead: SP is able to detect the return in which the lead is at nadir, while Ricker classifies as lead also the neighbouring point, in which the lead is seen off-nadir.

4.2. Quantitative comparison with SAR images

The altimetry-based classification is rated considering the following parameters:

- Fraction of False Detections (FFD), i.e. the fraction of points along the CS-2 track that are identified as leads by the altimeter-based method, but identified as ice on the SAR image;
- Fraction of Correctly Classified Leads (FCCL), i.e. the fraction of leads on the SAR image (along the CS-2 track) that are also seen by the CS-2 classification. Note that this statistics concerns the number of leads, therefore two consecutive lead detections are considered as part of the same lead.

The coincidence between CS-2 and SAR lead detection is verified by simply interpolating the SAR binary image generated as in section 3.3 onto the altimeter track. An along-track tolerance of 400 m is applied in the comparison, corresponding to the along-track resolution of CS-2 after the Hamming window application (Scagliola, 2013). The results are summarised in Table 1. The FCCL and the FFD are computed for each SAR image - CS-2 track couple and are shown taking their mean and standard deviation. The total number of leads detected by the SAR images is 232 for September 2015 (Set 1) and 275 for Set 2, while the table reports the total number of detections from the different altimetry-based methods. As a final score for each method, the ratio FCCL/FFD is adopted.

While the Relative Power Threshold and its results will be discussed in the next section, we firstly compare SP with Ricker. The two methods have

the same score of 1.2 and this result is also similar for SP in the control dataset, where the ice motion is applied. The only difference between the two methods is that SP finds 9% more of the leads recognised in the SAR images, but scores 9% worse in the FFD statistics. Essentially, the adoption of the SP index as a criteria is almost equivalent to the use of the 6 indices tuned in Ricker for the purpose of lead detection.

In both methods, the high standard deviation of FCCL and FFD attests the variability of the results depending on the different SAR scenes. We have not found a relationship with the seasonality, since the values are similar in both control data sets. The influence of seasonality cannot be excluded, but to assess it a larger amount of CS-2/SAR combinations are needed. This is not trivial, since an archive of these combinations does not exist and since the Stack data are not distributed in the ESA Baseline C CS-2 product and have to be acquired from an external source (G-POD).

Ricker was also validated against MODIS images by Wernecke and Kaleschke (2015). Their True Lead Rate (against MODIS taken as ground truth), essentially equivalent to the FCCL in this research, has a mean value of 60%, in line with our result considering the standard deviation. Despite the different ground truth, the statistics is therefore robust, but the validation here presented can be easily replicable since the lead extraction from the ground truth does not rely on visual criteria such as in the previous studies.

We argue that the reason for which the score of Ricker and SP is equivalent lies in the filtering of the SAR images needed for noise reduction. Figure 9 shows two examples of the comparison between leads seen through SAR image processing and altimetry-based classification with SP and Ricker. The figure shows that the filtering widens the borders of the lead, which has the effect of considering as a correct nadir detection the off-nadir CS-2 points that are classified as leads in Ricker. Nevertheless, the filtering is a necessary step to achieve an automatic and objective procedure for validation, as shown in Figure 10: a binary map formed without the noise reduction described in section 3.3 would result in several incorrect lead-like isolated features and as a consequence would strongly decrease the reliability of the comparison. Even with the filtering, the SAR processing is not always able to extract leads that are very thin, as for example in Figure 9b. The binary map recognises the wider part of the lead that crosses the CS-2 track, but misses the elongated feature, due to the insufficient contrast between the ice edge and thin lead.

A validation experiment was also undertaken to observe the effect of the closing operation. The results are listed for each altimetry classification

method in Table 1. By comparing altimetry with SAR images without the closing, a general increase of the mean FFD by 3-4% is observed as well as a decrease of the FFCL, which result in a worse score. The reason is observed in Figure 10 c: the closing connects a lead that is seen fragmented in the original SAR image, which can be due to partial refreezing or noise. Without the closing, the lead identified by the altimeter would be interpolated on a black (ice) SAR binary pixel, resulting in an apparent false detection.

For the purpose of comparing the lead detection from satellite altimetry with the one applied on SAR images, the latter is considered as a ground truth. This assumption is only meant to provide a common ground for the comparison of different altimetry-based strategy, but it is an approximation of the reality: Despite the high resolution, Sentinel-1 is not able to distinguish leads that are narrower than 40 m, while such cracks in the ice could still be the dominant return in the altimetric waveforms. This difference is even more stringent after the application of the 200 m x 200 m median filter for noise reduction in the SAR image. Moreover, the ability of both the automated SAR technique and the altimetry methods to distinguish between leads and melt ponds, which according to recent studies can occupy as much as 70% of the first-year ice area in the melting period (Divine et al., 2016), remains hard to verify and further research is needed in this matter.

Finally, the impact of the time difference between S1 and CS-2 and of the application of the sea ice motion vectors in Set 2 is displayed in Figure 11, considering the SP classification statistics. Regardless of the sea ice motion vectors, in both Sets the best results in terms of FCCL are obtained for S1-CS2 coupling close in time: in particular, only CS-2/SAR comparisons within 90 minutes show $FCCL > 0.4$ and the closest CS-2/SAR comparison has the best FCCL score. Moreover, in 4 out of 5 comparisons within 90 minutes, FCCL is higher than FFD, while the opposite happens in 6 out of 7 comparisons over 90 minutes, regardless of the sea ice motion vectors application. The latter has nevertheless a limited, but constant positive effect, reducing FFD and improving FCCL.

4.3. Received power as lead classifier

Using the distinction between true leads and false leads on the base of the comparison with SAR images, figure 12 shows, for every SAR image of the two datasets, the mean and standard deviation of the received waveform power from the points classified as leads by CS-2. The figure also shows as a dotted line the value of 2.58×10^{-11} , which has been proposed by Wernecke

and Kaleschke (2015) as the threshold for the best representation of lead occurrence. This threshold is not entirely comparable with the values here presented, since it has been computed using the ESA Baseline B release of the CS-2 data, while at the time of writing this has been substituted with ESA Baseline C: small differences in the Delay-Doppler processing used to build the waveforms can lead to different power output associated to each echo.

Considering the results, the implementation of an absolute threshold to classify the leads does not look feasible. The proposed threshold, although it avoids false detections, does not find leads in three of the tracks from September 2015 and misses almost any lead in Set 2. The return power distribution of false leads and true leads is not constant: for example, the power distribution of false leads in images 1 and 2 from September 2015 is almost coincident with the power distribution of true leads in images 1 and 2 from Set 2. The power backscattered by a lead in fact does not depend only on the off-nadir or nadir view, but also on the presence of sea ice in the illuminated area, on possible refrozen areas of the lead and on its width. In general, the higher is the selected absolute power threshold, the higher is the confidence that the selected points are real leads, but the lower is the number of detections.

Nevertheless, the mean power of the true leads distribution is constantly higher than the one of the false leads distribution. We can use this information in a relative sense by computing the ratio between the power of the CS-2 returns classified as leads and the median of the power in the segment of the track considered (Power Ratio). As a tentative approach in this study, we considered the whole length of the CS-2 segment downloaded for each corresponding Sentinel-1 image date. These vary between roughly 200 and 800 km). A histogram of the Power Ratio in the two datasets is shown in Figure 13. Although they refer to different areas at different times, both datasets show that when the Power Ratio is below 10, the False Leads are predominant. Since the results are consistent, we keep the previously defined approach to define the Power Ratio. This relative threshold is therefore used as a "Relative Power Threshold" classification and the results are displayed in Table 1. Although the FCCL is lower than for the other methods, the false lead detections are also considerably less. This brings to an overall score FCCL/FFD of 1.7 in Set 1 and 1.5 in Set 2. The Relative Power Threshold scores therefore best, but it is here derived *a posteriori* and further studies in different locations and at different times of the year are needed in order to un-

derstand whether the same threshold can be used systematically. Compared to the use of an absolute threshold, the performances of this method are not dependent on time and location of the track, but, as seen in Table 1, also the Relative Power Threshold brings as a drawback a significant reduction of the number of detected leads, in particular w.r.t. SP.

5. Conclusions

This study aimed at testing different Cryosat-based methods for lead classification taking the SAR images from Sentinel-1 as reference. SAR images have been processed in order to provide an automatic distinction between leads and ice.

A new parameter based on stack data, the Stack Peakiness, has been proposed. A visual analysis of the CS-2 tracks and SAR images has shown how SP can be used to isolate the nadir return from narrow leads crossing the track. The statistics show that in terms of correctly identified leads and false detections, the SP method has comparable results to the method proposed by Ricker et al. (2014), which combines six different waveform indices derivable from the CS-2 returns. The automated SAR processing is not able to highlight significant differences between SP and Ricker, due to the necessary filtering, nevertheless it constitutes a reliable, objective and easily replicable validation method.

In order to quantitatively understand whether SP avoids the off-nadir returns coming from the leads when approaching the altimetry track, an easier and systematic access to CS-2 stack data is needed. In this case, further research could be planned to produce maps of sea surface height in the Arctic region using the identified leads and comparing the sea level variability with previous estimates.

Given the substantial differences of the power distribution of the leads observed in the datasets, the use of an absolute threshold on the return power to classify leads and avoid off-nadir returns, as proposed in previous studies, is not considered reliable. The best performances are nevertheless obtained by using the return power in a relative sense in comparison with the average power of the CS-2 returns in the area considered.

The percentage of false lead detections in comparison with the fraction of correctly classified leads is high for all the tested strategy. If the classification based on SAR images could be taken as the ground truth, this result would undermine the reliability of sea surface estimates in the Arctic Ocean, since

it would imply that several reflections from the sea ice are considered for sea level measurements. Nevertheless, given that previous studies on-ground have shown that the width of a lead can be well below 1 km, it is likely that the altimeters spot leads whose width is below the SAR resolution. The time difference between CS-2 and SAR acquisitions plays also a key role and the comparability between the two sources clearly decreases after 60-90 minutes, despite the attempt of taking in consideration the sea ice velocity.

Although the SAR images do not represent the ground truth in terms of lead classification, given the resolution limits, they represent a well established comparison. The strength of this work is that the automatised validation can be easily reproduced in other areas, if an archive of CS-2 passes and Sentinel-1A images that are sufficiently close in time is provided. Indeed, further validation is needed in different sea-ice covered regions (multi-year and first year ice) and at different times of the year, since the presence of melt ponds and the refreezing of the leads could considerably alter the performances of any classification algorithm.

Future research should also address the exploitation of the SP for classification of the waveforms from the SAR-Interferometric mode of CS-2 (SARIn). In SARIn, the use of a second across-track antenna allows the localisation of different returns using the phase difference between the echoes reaching the two antennas. While the availability of SARIn in the sea ice region is limited to few patches, Armitage and Davidson (2014) have shown that, combining the classification with an off-nadir ranging correction to characterise the off-nadir leads, the accuracy and the precision of sea ice freeboard measurements can be improved compared to the SAR mode.

Acknowledgements

The authors are thankful to Eero Rinne (Finnish Meteorological Institute, Helsinki) for providing a list of available Sentinel-1 data and the corresponding Cryosat-2 passes in September 2015, and to Stephan Hendricks and Robert Ricker (Alfred Wegener Institute) for providing the corresponding lead classification. To all of them and to Andreas Wernecke (University of Hamburg) goes our gratitude for the useful discussions.

The Cryosat-2 data were obtained using the ESA G-POD Service (<https://gpod.eo.esa.int/>). All Sentinel 1-A images are provided by ESA and downloaded at "Sentinels Scientific Data Hub". The pre-processing was done by applying the "Sentinel Application Platform (SNAP)" in particular the

Sentinel-1 toolbox (S1TBX), version 3.0.0 provided by ESA under the terms of the GNU General Public License. The Sea Ice velocity data were obtained using the Polar Pathfinder EASE-grids (<http://nsidc.org/data/nsidc-0116/versions/3>)

This work was supported by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program. The study was partly funded by the Deutsche Forschungsgemeinschaft (DFG) through grant BO1128/13-1. The publication is funded by the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program.

Citations

References

- Armitage, T. W. and Davidson, M. W.: Using the interferometric capabilities of the ESA Cryosat-2 mission to improve the accuracy of sea ice freeboard retrievals, *IEEE Trans. Geosci. Remote Sens.*, 52, 529–536, 2014.
- Bouzinac, C.: *CryoSat Product Handbook*, European Space Agency, 2012.
- Bradley, D. and Roth, G.: Adaptive thresholding using the integral image, *Journal of Graphics, GPU, and game tools*, 12, 13–21, 2007.
- Dierking, W.: Sea ice monitoring by synthetic aperture radar, *Oceanography*, 26, 2013.
- Divine, D. V., Pedersen, C. A., Karlsen, T. I., Aas, H. F., Granskog, M. A., Hudson, S. R., and Gerland, S.: Photogrammetric retrieval and analysis of small scale sea ice topography during summer melt, *Cold. Reg. Sci. Technol.*, 129, 77–84, 2016.
- Gonzalez, R. C. and Woods, R. E.: *Digital image processing*, Prentice Hall, Upper Saddle River, New Jersey, USA, 3 edn., 2008.
- Ivanova, N., Rampal, P., and Bouillon, S.: Error assessment of satellite-derived lead fraction in the Arctic, *The Cryosphere*, 10, 585–595, 2016.
- Kwok, R., Cunningham, G., Wensnahan, M., Rigor, I., Zwally, H., and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008, *J. Geophys. Res.: Oceans*, 114, 2009.

- Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., et al.: CryoSat-2 estimates of Arctic sea ice thickness and volume, *Geophys. Res. Lett.*, 40, 732–737, 2013.
- Lindsay, R. and Rothrock, D.: Arctic sea ice leads from advanced very high resolution radiometer images, *J. Geophys. Res.: Oceans*, 100, 4533–4544, 1995.
- Onstott, R.G. and Shuchman, R.A.: SAR Measurement of Sea Ice. In Jackson and Apel [eds.]: *Synthetic Aperture Radar: Marine User's Manual*, 81–115. US Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, Office of Research and Applications, 2004.
- Peacock, N. R. and Laxon, S. W.: Sea surface height determination in the Arctic Ocean from ERS altimetry, *J. Geophys. Res.: Oceans*, 109, 2004.
- 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.: Oceans*, 114, 2009.
- Raney, R. K.: The delay/Doppler radar altimeter, *IEEE Trans. Geosci. Remote Sensing*, 36, 1578–1588, 1998.
- Ricker, R., Hendricks, S., Helm, V., Skourup, H., and Davidson, M.: Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation, *The Cryosphere*, 8, 1607–1622, 2014.
- Rinne, E. and Similä, M.: Utilisation of CryoSat-2 SAR altimeter in operational ice charting, *The Cryosphere*, 10, 121–131, 2014.
- Röhrs, J. and Kaleschke, L.: An algorithm to detect sea ice leads by using AMSR-E passive microwave imagery, *The Cryosphere*, 6, 343–352, 2012.
- Scagliola, M.: CryoSat footprints, Aresys technical note, SAR-CRY2-TEN-6331, Tech. rep., Aresys/ESA, Italy, 2013.
- Scagliola, M. and Fornari, M.: Main evolutions and expected quality improvements in BaselineC Level1b products, Aresys technical note, C2-TN-ARS-GS-5154, Tech. rep., Aresys/ESA, Italy, 2015.

- Serreze M.C. and Barry R.G.: The Arctic Climate System. Cambridge Atmospheric and Space Science Series. Cambridge University Press, Cambridge, UK, 2014.
- Tschudi, M., Fowler, C., Maslanik, J., Stewart, J., and Meier, M.: Polar Pathfinder daily 25 km EASE-grid sea ice motion vectors (Last Access: 08/08/2016), Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, 2016.
- Weeks, W.: On sea ice, University of Alaska Press, Fairbanks, USA, 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, 2015.
- Willmes, S. and Heinemann, G.: Pan-Arctic lead detection from MODIS thermal infrared imagery, Ann. Glaciol., 56, 29–37, 2015.
- Wingham, D., Francis, C., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., de Chateau-Thierry, P., Laxon, S., Mallow, U., Mavrocordatos, C., et al.: CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields, Adv. Space Res., 37, 841–871, 2006.
- Veci L. SENTINEL-1 Toolbox SAR Basics Tutorial, Array Systems Computing Inc., <http://step.esa.int/main/>, 2016.

Figures

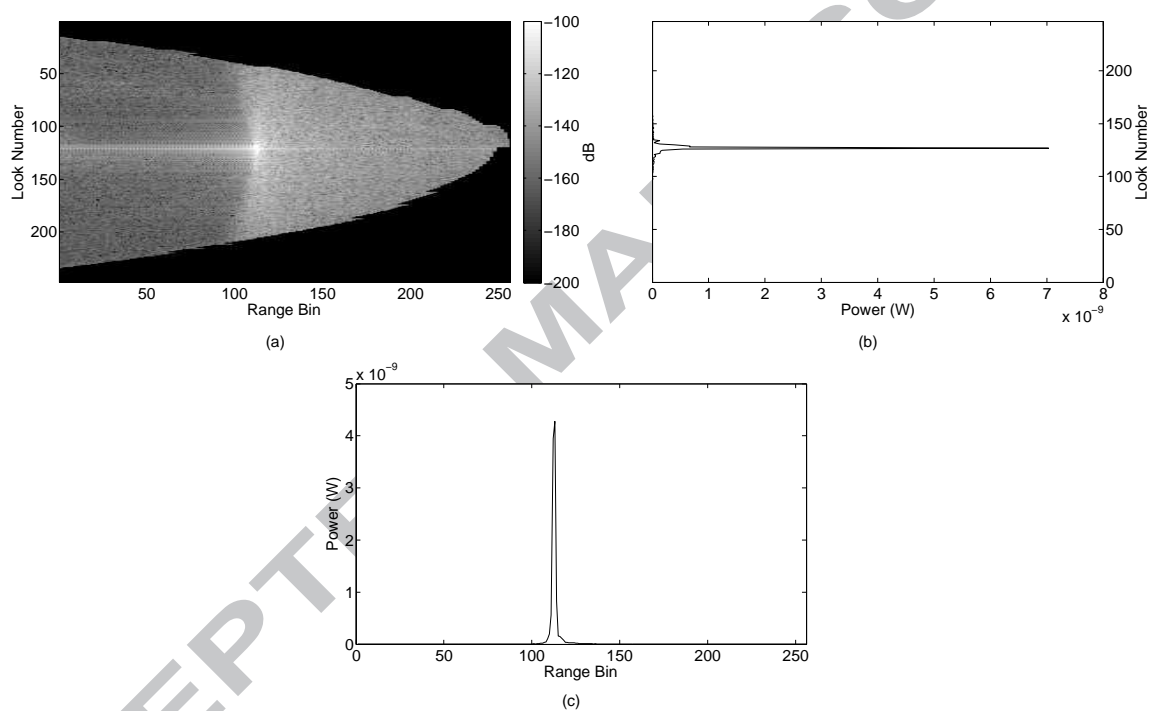


Figure 1: Example of a stack diagram (a), a RIP (b) and a multilooked waveform (c) acquired by Cryosat-2 over a lead.

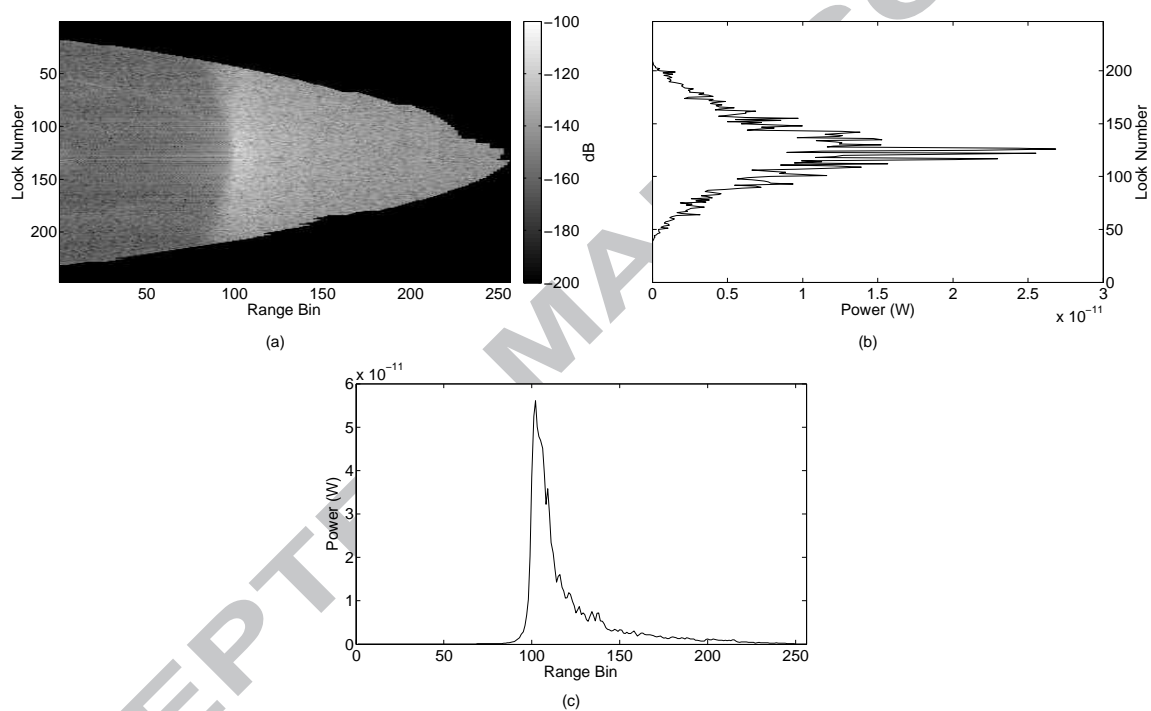


Figure 2: Example of a stack diagram (a), a RIP (b) and a multilooked waveform (c) acquired by Cryosat-2 over sea ice.

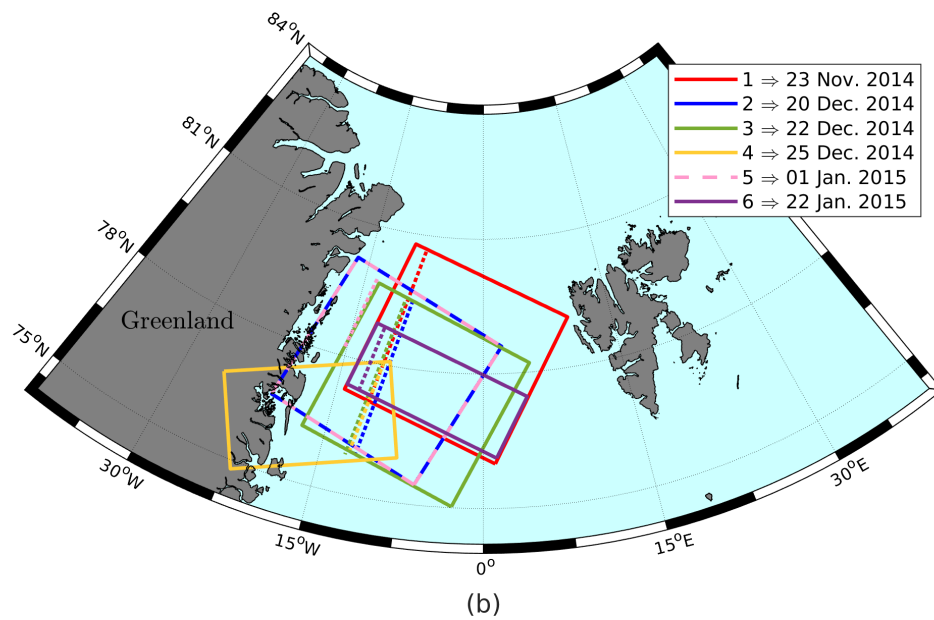
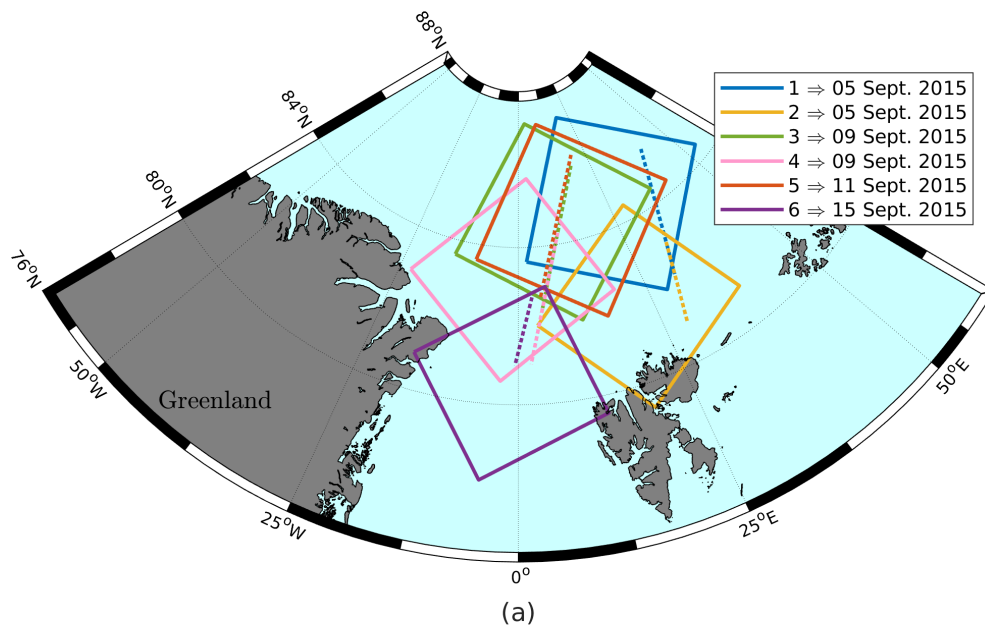


Figure 3: Location and date of the SAR-images in the two control datasets. Dashed lines are the corresponding segments of CS2 tracks on which the classification is applied.

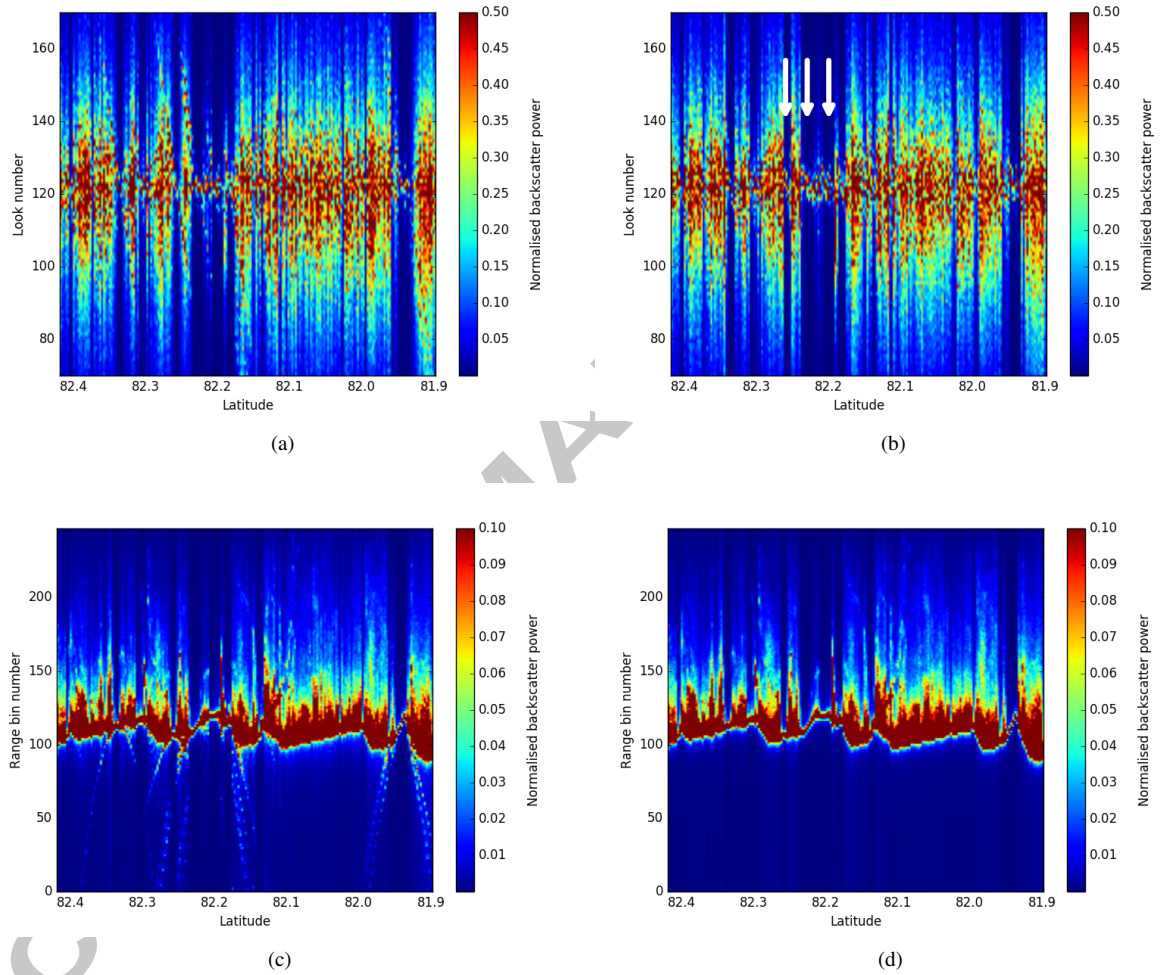


Figure 4: Waveform (c and d) and RIP (a and b) radargrams from a CS-2 pass over sea ice acquired on the 5th September 2015 (see figure 3 for track location). All plots correspond to the same pass, but Hamming windowing has been applied in b and d. The arrows highlight residual sidelobe effects also despite the Hamming windowing. The normalised power has been saturated in order to visually show the sidelobe effect.

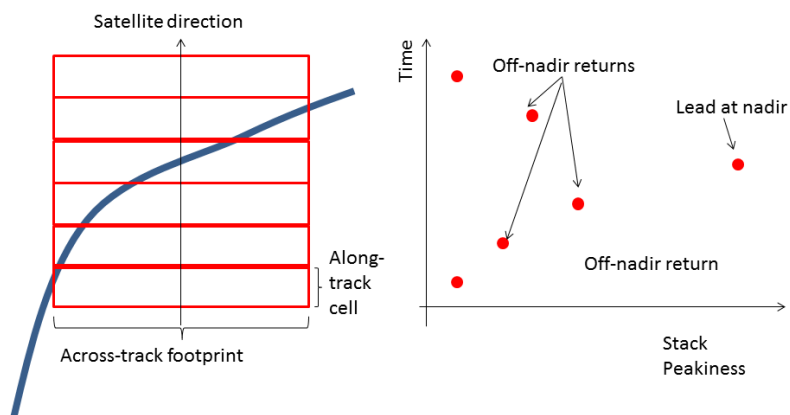
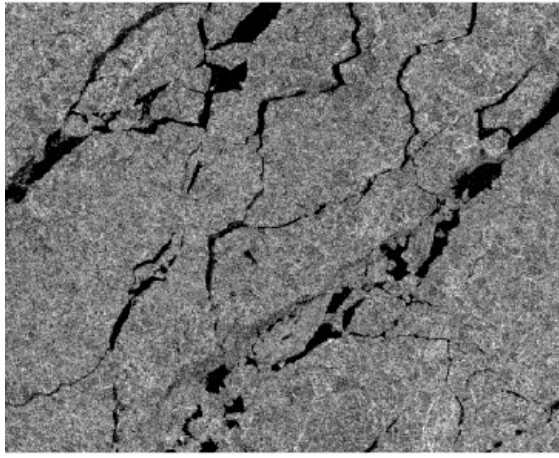
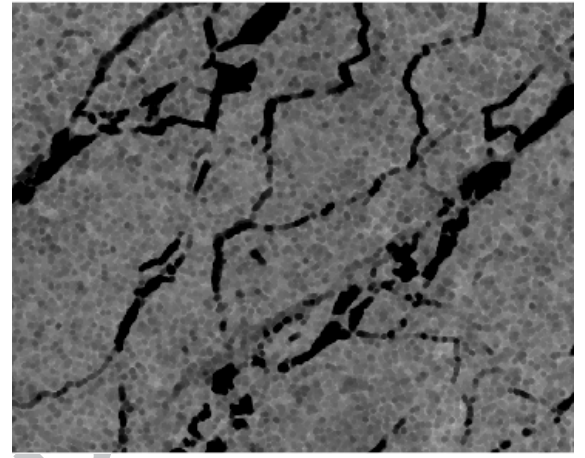


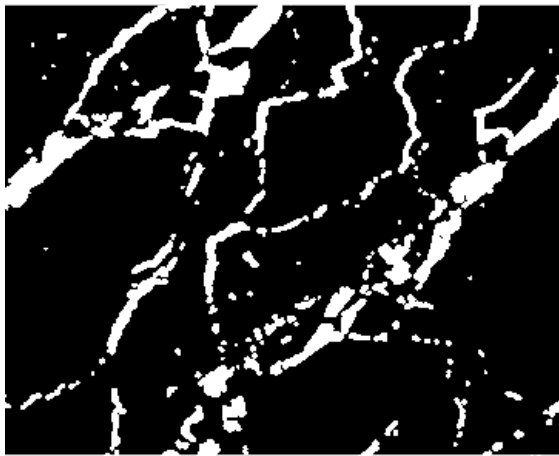
Figure 5: Modelled lead (blue line) crossing over a CS-2 track (left) and expected SP behaviour (right).



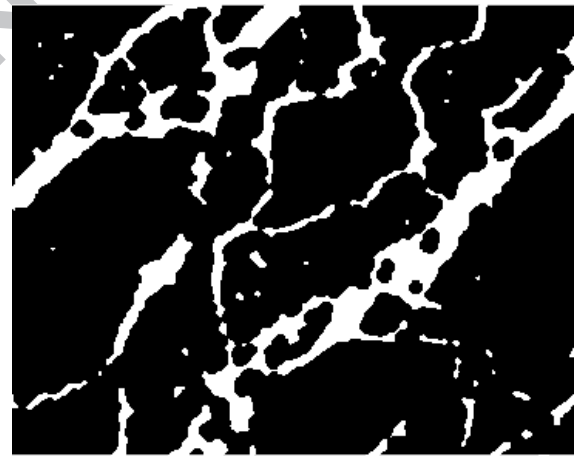
(a) Sentinel-1A image after SAR pre-processing



(b) Sentinel-1A image after median and minimum filtering



(c) Sentinel-1A image after segmentation



(d) Sentinel-1A image after morphological closing operation

0	0	1	1	1	0	0
0	1	1	1	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0

(e) Octagon kernel for minimum filtering

Figure 6: Sentinel-1A image subset of about 24x24 km: (a) the original grayscale SAR image after SAR pre-processing; (b) the same image after 5x5 median and minimum filtering; (c) the binary image after segmentation by adaptive thresholding; (d) final image after closing operation giving open water in white and sea ice areas in black. Plot (e) shows the octagon kernel with radius 3 around center pixel applied for the minimum filtering.

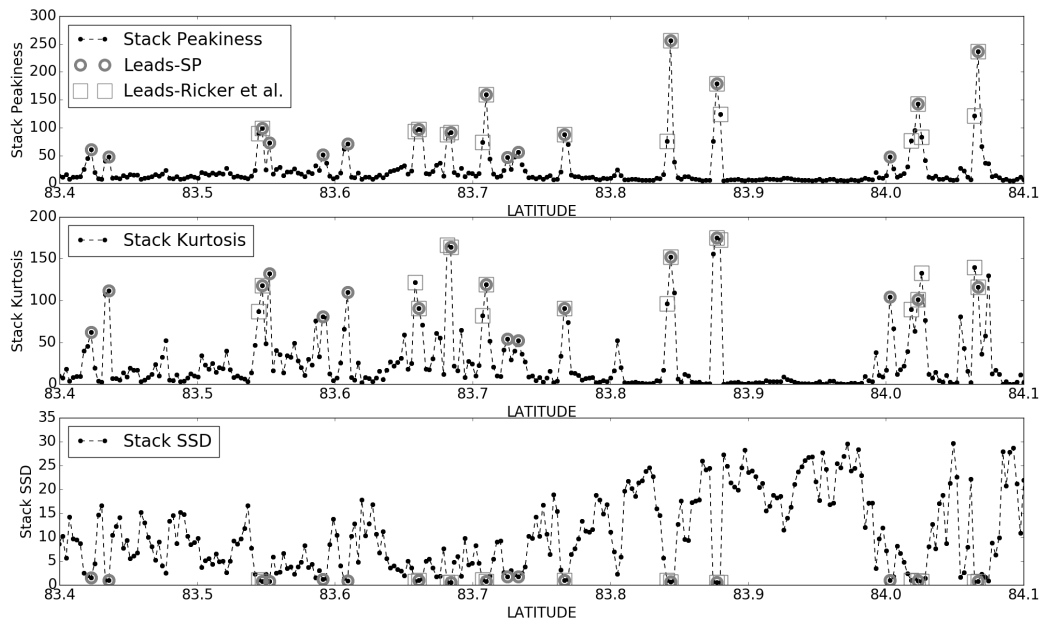


Figure 7: SP behaviour along a CS-2 track acquired over sea ice (top plot). The figure shows the points that are recognized as leads by the SP-based classification described in this study (circles) and the results of the classification based on Ricker et al. (2014)(squares). The lower plots show Stack Kurtosis and Stack SSD for comparison.

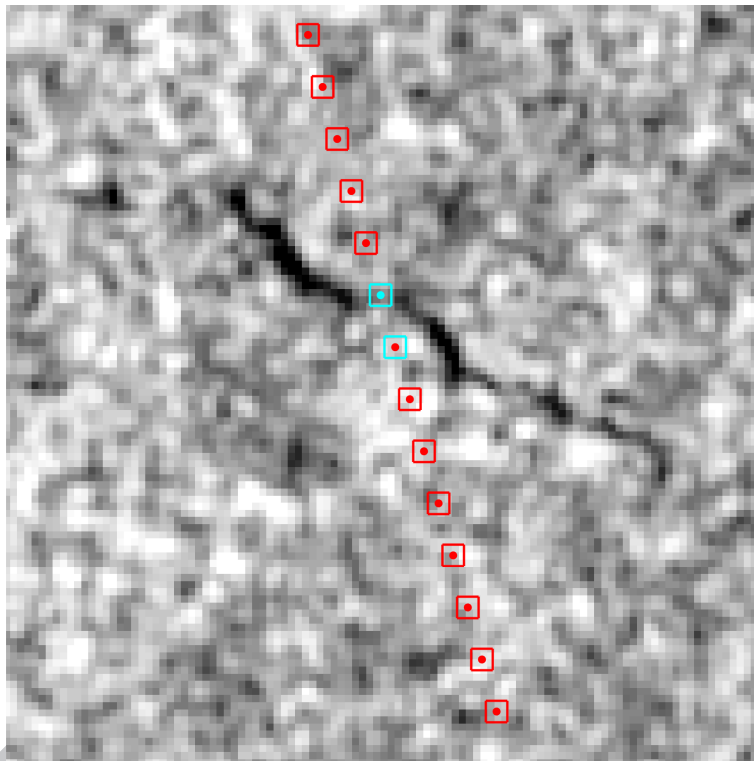


Figure 8: 4.25 by 4.25 km zoom on a lead seen by Sentinel-1 image from 05/09/2015 at 12:46 (Image 1 of Set 1, see Figure 3 for location), with Cryosat-2 crossing the area at 16.36. No ice motion correction is applied due to high variation in the ice flow direction. SP (dots) in comparison with classification results of Ricker (squares) showing in cyan lead detections.

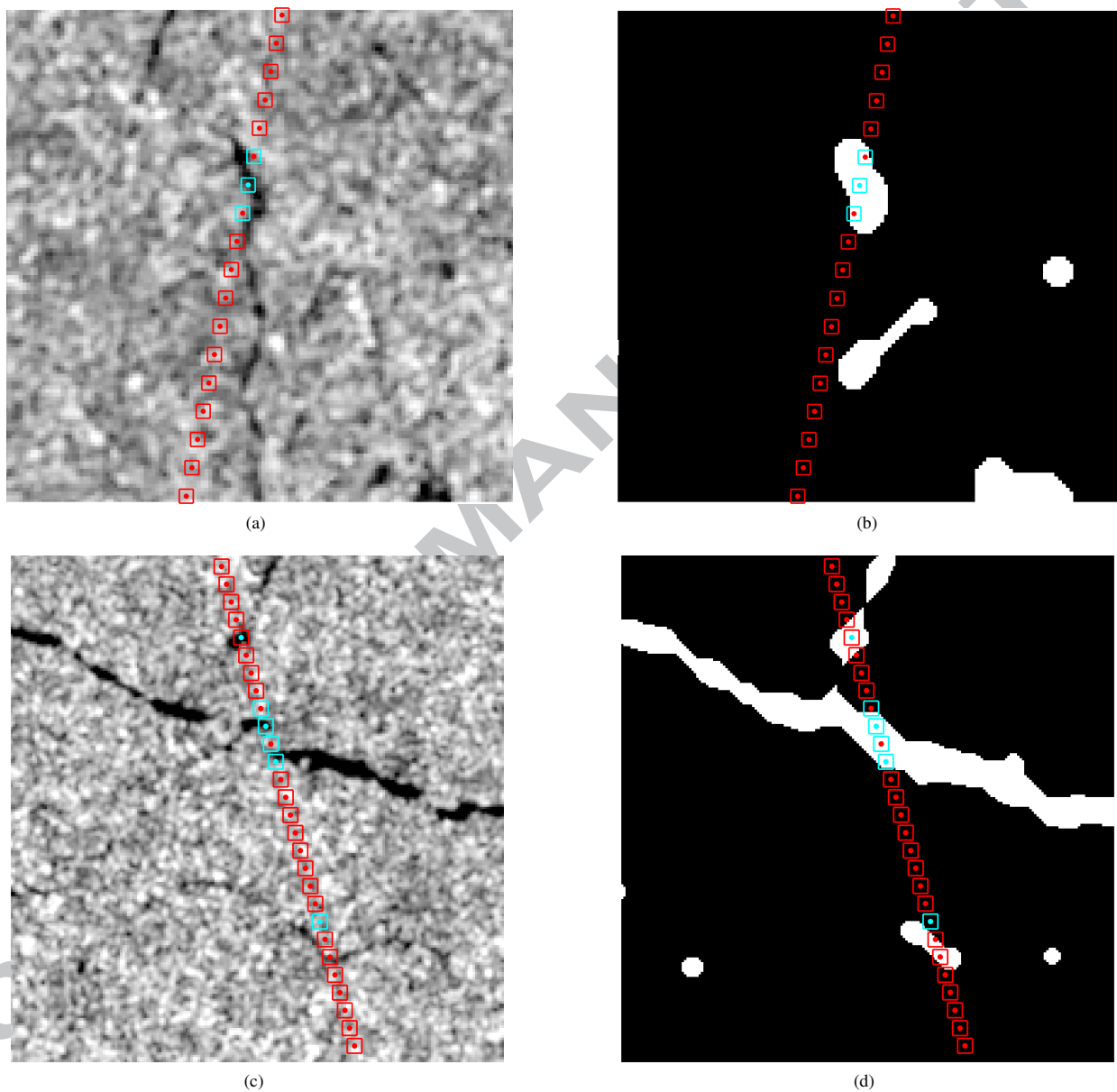


Figure 9: Examples of lead detection from altimetry against SAR images before (a and c) and after the processing (b and d). a and b show a 5.35 by 5.18 km subset from Image 4 of Set 1, c and d illustrate a 8.11 by 8.34 km zoom from Image 1 of Set 1. Red: ice detection, cyan: lead detection. Squares: Ricker classification, dots: SP classification.

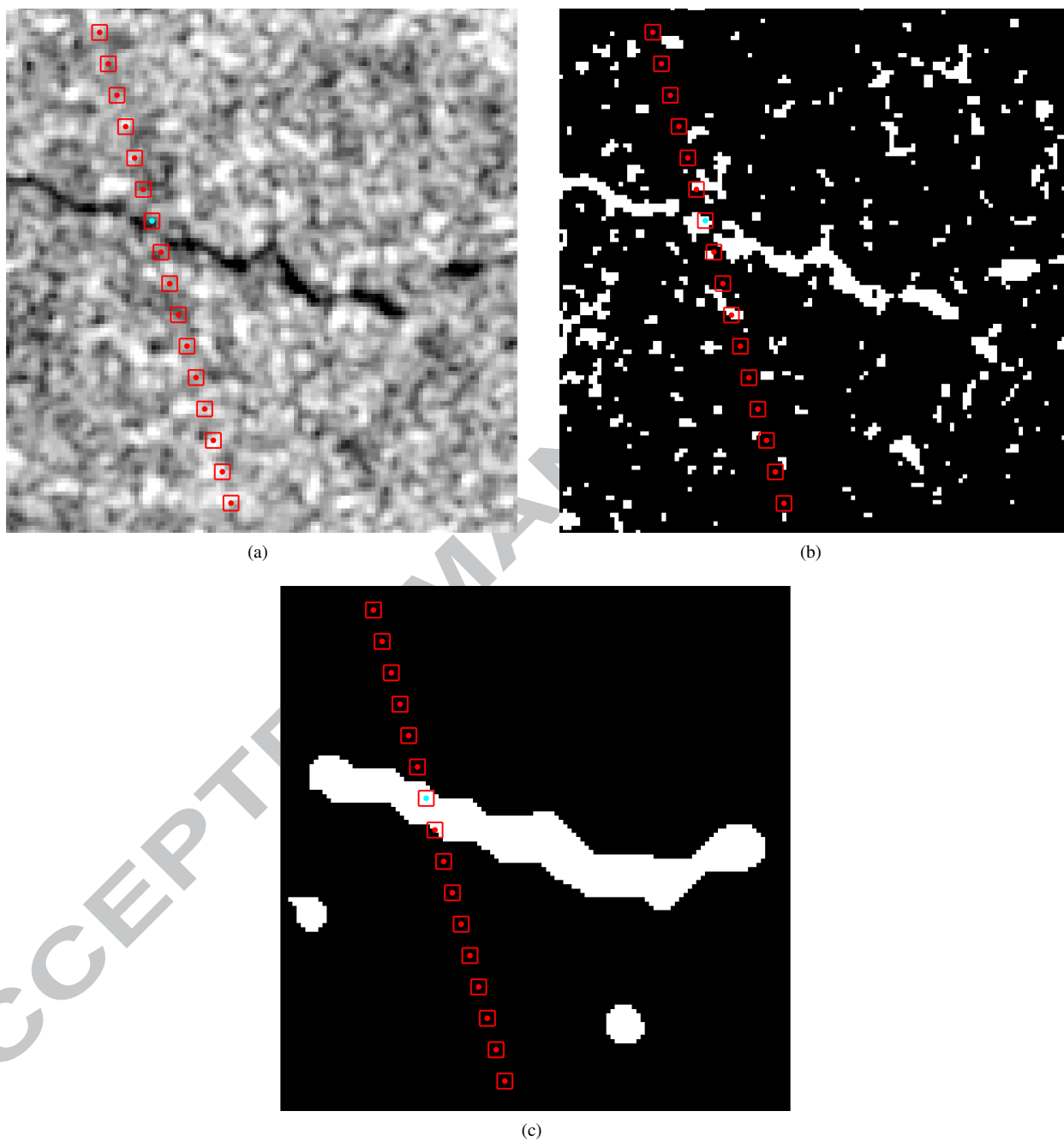


Figure 10: Examples of lead detection from altimetry against SAR images (4.75 by 4.88 km extract from Image 1 from Set 1) before the processing (a), as a binary map without the filtering and closing (b) and after the full processing (c). Red: ice detection, cyan: lead detection. Squares: Ricker classification, dots: SP classification.

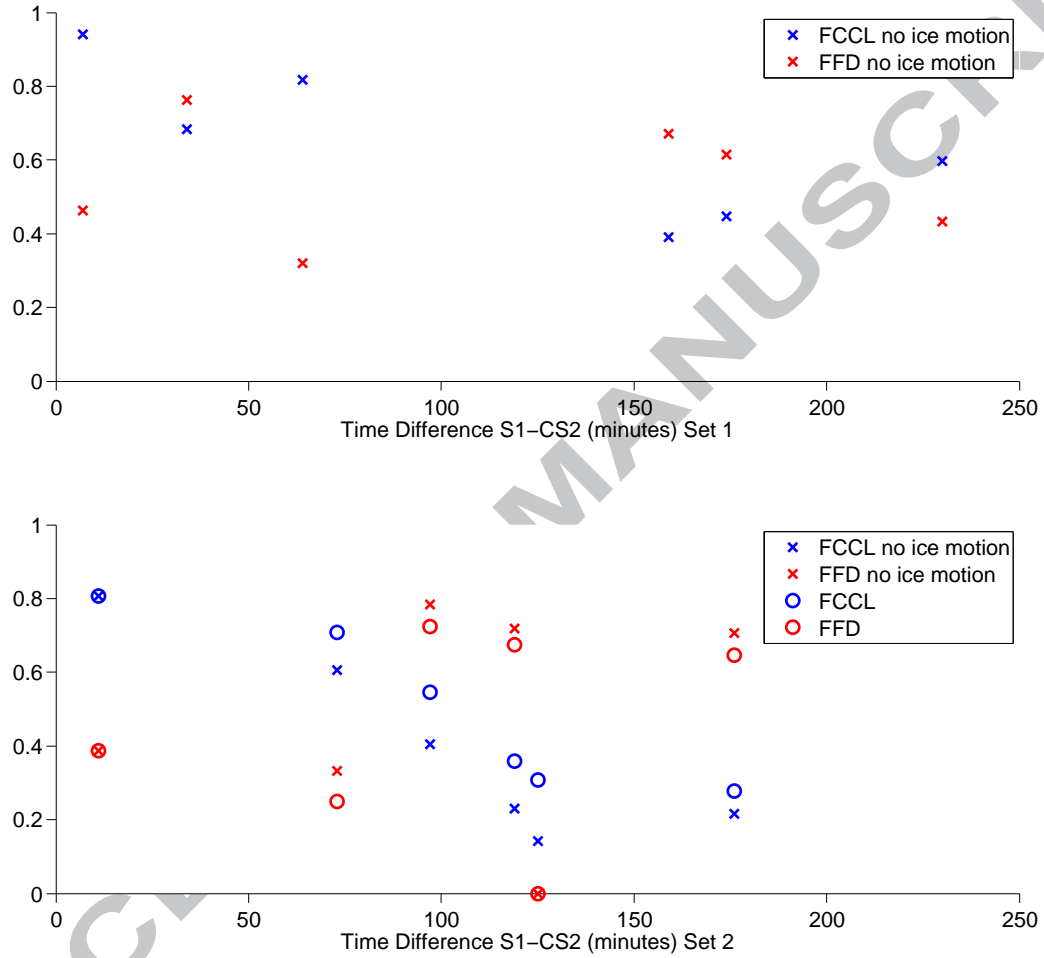


Figure 11: FCCL (blue) and FFD (red) for each SAR image of Set 1 and Set 2 according to the time difference with the corresponding CS-2 track. In Set 1, the sea ice motion vectors are not applied, as in table 1. In Set 2, statistics are shown with (circles) and without (crosses) the application of sea ice motion vectors.

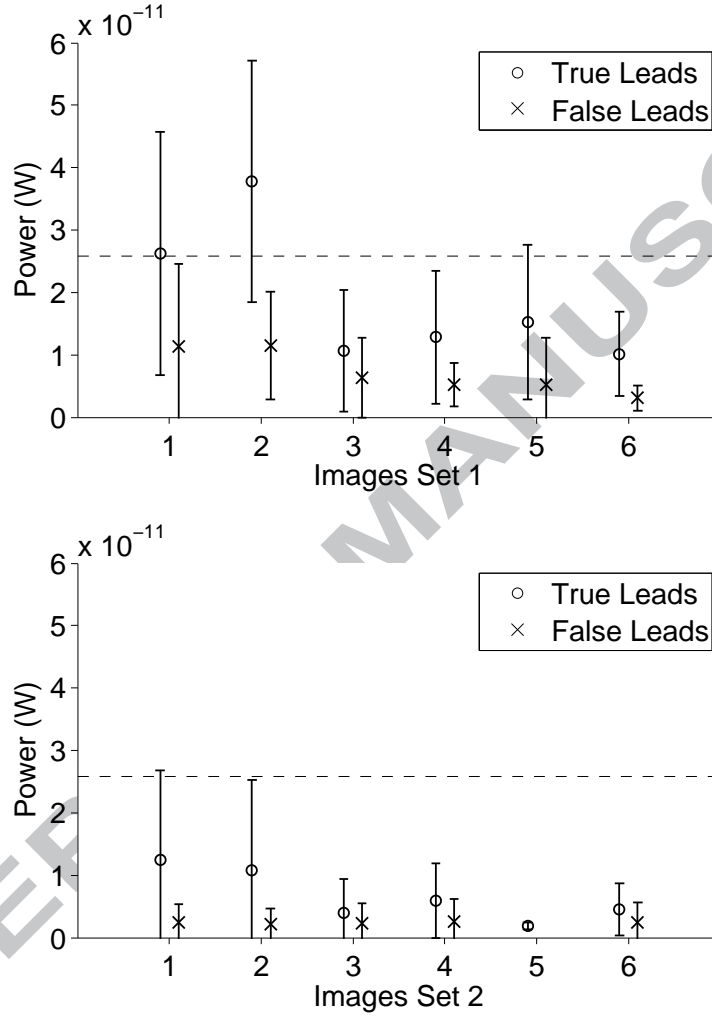


Figure 12: Mean and standard deviation of the received power of True and False Leads in each image of the two control datasets. True and False Leads are the CS-2 waveforms classified as leads with the Stack Peakiness method and validated by means of comparison with the SAR images.

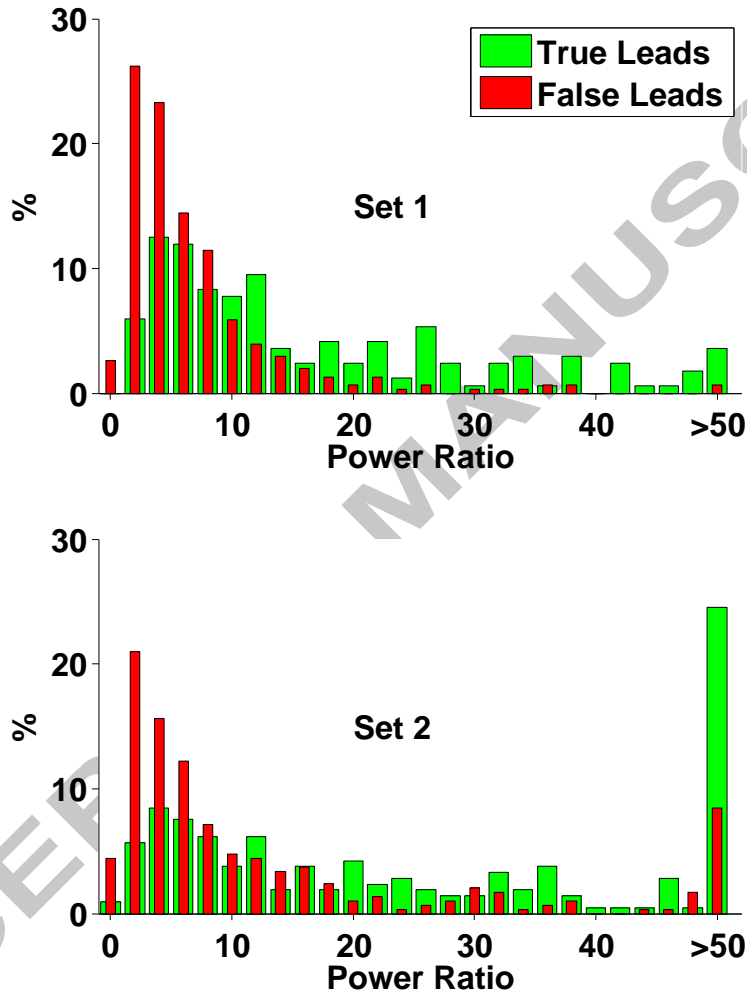


Figure 13: Histogram of the received power of True and False Leads in both control datasets. True and False Leads are the CS-2 waveforms classified as leads with the Stack Peakiness method and validated by means of comparison with the SAR images.

Table 1: Results of the CS-2-based lead classification with the methods described in the study, w.r.t. the SAR images classification. Col 1: name of the SAR images dataset; Col 2: classification method; Col 3: mean and standard deviation of the FFD (defined in 4.2); Col 4: mean and standard deviation of the FCCL (defined in 4.2); Col 5: ratio between FCCL and FFD, taken as final score of the method; Col 6: number of leads seen by each CS-2 lead classification method. For each classification method, the validation against SAR images without the morphological closing operation is identified by the suffix NC (No Closing).

Validation Dataset	Lead Detection	mean(FFD) \pm std	mean(FCCL) \pm std	FCCL/FFD	Tot lead
Set 1	Stack Peakiness	0.5444 \pm 0.1659	0.6465 \pm 0.2121	1.2	472
	Stack Peakiness NC	0.5795 \pm 0.1756	0.6065 \pm 0.2243	1.0	472
	Ricker et al. (2015)	0.4528 \pm 0.1594	0.5496 \pm 0.1982	1.2	279
	Ricker et al. (2015) NC	0.4982 \pm 0.1758	0.5427 \pm 0.2011	1.1	279
	Relative Power Threshold	0.2554 \pm 0.2199	0.4269 \pm 0.1985	1.7	144
	Relative Power Threshold NC	0.2971 \pm 0.2358	0.4276 \pm 0.2056	1.4	144
Set 2	Stack Peakiness	0.4472 \pm 0.2865	0.5007 \pm 0.2217	1.1	507
	Stack Peakiness NC	0.4869 \pm 0.2914	0.4392 \pm 0.2218	0.9	507
	Relative Power Threshold	0.2885 \pm 0.2178	0.4333 \pm 0.0728	1.5	188
	Relative Power Threshold NC	0.3349 \pm 0.2256	0.4397 \pm 0.0994	1.3	188