

# Author Response

Thank you for pointing out sections which had to be clarified as well as proposing changes of the used retracker and in the SSH visualization. We feel that we have been able to improve the understandability and reliability of the manuscript by revising this points.

## Anonymous Referee #3:

*I just have one technical correction which I suggest be implemented if the authors wish, it will likely not change the results much but would make the SSH results more accurate:*

*Section 3.2: Note that the use of the TFMRA tracker at the 40% level will lead to a biased estimate of SSH for lead returns. This is because the waveform shape is closer to that of a  $\text{sinc}^2$  function which should be tracked closer to the peak of the return as shown in Kurtz et al., [2014]. This probably will not have a large effect on the variance estimation, but I do not know this for sure. I suggest tracking by fitting a  $\text{sinc}^2$  function as in Kurtz et al., [2014] or a Gaussian plus exponential as described in Laxon et al., [2013].*

We changed the tracking procedure and now fit a  $\text{sinc}^2$  function to the bins around maximum of the waveform. As you mentioned, Kurtz et al., [2014] have shown that the tracking point should be close to the echo peak and that the error introduced by using the peak as tracking point is below 3cm (for leads with reasonable values of  $\alpha$ ). We used the center of the  $\text{sinc}^2$  fit as tracking point, which is comparable to the specular retracker of Armitage and Davidson, [2014] (but they used a Gaussian) and as mentioned in agreement with the results of Kurtz et al., [2014] (but they did not fit a simple function to the waveforms but the results of a physical model). The results do not differ in the general pattern.

## Dr. Christian Haas (Editor):

*I ask you to address [...] the better presentation of the retrieval of SSH proposed by myself earlier.*

A new figure showing the SSH anomaly along a single track has been added. Thereby the differences in the stability of the SSH are illustrated in a more intuitive way.

*L55: of MEASURING relatively thin ice*  
Corrected

*L69: SARM: Here and elsewhere, please make sure to spell out and possibly define acronyms before first using them*

Corrected

*L103: consists OF 722...*

Corrected

*L107 and elsewhere: Beaufort Sea (capital Seal); see also fig 2*

Corrected

*L110: THESE two SAR mode versions...*

Corrected

*L113-114: check formulation and improve*

Done

*L138-141: improve clarity of paragraph*

Done

*L165-166: Please state what the rejection rate is of waveforms not used for further analysis.*

7.6% of the ground truth measurements have a first peak with a power between 50% and 80% of the maximum value. This is now stated. The rejected measurements are nearly exclusively (to 98%) ice measurements which agrees well with the presence of narrow unnoticed leads in the ice. The rejection of them might therefore support the significance of the ROC analysis.

*L180: unclear*

Revised

*L218: has MAX0.5 been properly introduced?*

It has been introduced together with all other one dimensional classifiers by "We use the parameter acronym with the weight as index to point at the corresponding one dimensional classifier."

*L 260-61: check formulation*

Done

*L272: remove north polar stereographic grid statement. Not necessary and not quite correct?*  
All SSH estimates within a grid cell (with a nominal size of 10X10 km) are used to calculate the variance of this cell. Figure 6 shows histograms of the variances of all grid cells. This grid is in a north polar stereographic projection, but there is no figure showing it (the SSH or SSH variance in this grid as a map).

We mentioned the projection as it can introduce some errors in the grid cell size (i.e. the surface area contributing to one grid cells at the pole is actually a bit larger than it is further south in a polar stereographic projection). As this distortion is less than 6 percent (in our case) and as the variance is not sensitive to small changes in the grid cell size, we removed the statement.

*L274-279: improve paragraph.*

Done

*The variance is not only representing noise, but also natural variability?*

We agree but the differences between different classifiers can not be explained by natural variability as it influences them in the same way. The different components of the SSH-estimate variance are now listed.

*Values of what converge?*

The right hand side of the histograms, i.e. the amount of grid cells with high variances.

*L337: strongly parallel development?*

Corrected

*Table 1 caption: unclear what "lower three classifiers" are?*

Corrected

*Figure 1 caption: replace Example by Typical;*

Corrected

*Clarify what you mean by normalized waveforms. They don't look normalized in figure (i.e. they have different amplitudes; do they have same area/power or what?)*

You are right, the shown waveforms are not normalized in any way. The statement is of hypothetical kind: If they would be normalized (by amplitude) the area beneath the waveform would be inversely proportional to the PP (as  $P_{max}=1$  and integral equivalent sum). The normalized waveforms are easy to imagine as the peak values of the waveforms reach the same height in the figure i.e. only the numbers on the y-axis have to be ignored. This is the only way to illustrate the MAX and the PP definition in the same figure. The hypothetical kind of the statement is now clarified.

*Figure 4&5, and other figures: please assure that annotations and labels are large enough to be readable in final size.*

Done

*Figure 7&8 caption: be explicit about what that different mode is.*

Corrected

# Lead detection in Arctic sea ice from CryoSat-2: quality assessment, lead area fraction and width distribution

A. Wernecke and L. Kaleschke

Institute of Oceanography, University of Hamburg, Bundesstrasse 53,  
20146 Hamburg, Germany

*Correspondence to:* A. Wernecke (andreas.wernecke@uni-hamburg.de)

**Abstract.** Leads cover only a small fraction of the Arctic sea ice but they have a dominant effect on the turbulent exchange between the ocean and the atmosphere. A supervised classification of CryoSat-2 measurements is performed by a comparison with visual MODIS scenes. For several parameters thresholds are optimized and tested in order to reproduce this prior classification. The maximum power of the waveform shows the best classification properties amongst them, including the Pulse Peakiness. The sea surface height is derived and its spread is clearly reduced for a classifier based on the maximum power compared to published ones. Lead area fraction estimates based on CryoSat-2 show a major fracturing event in the Beaufort Sea in 2013. The resulting Arctic wide lead width distribution follows a power law with an exponent of  $2.47 \pm 0.04$  for the winter seasons from 2011 to 2014, confirming and complementing a regional study based on a high resolution SPOT image.

## 1 Introduction

Sea ice affects all interaction between ocean and atmosphere, namely heat, mass and momentum transports in ice covered regions. It strongly reduces most of these transports and thereby leaving these processes basically to openings in the ice. These openings, called leads, appear even in regions which are typically covered by thick ice, like the central Arctic. Shear and divergence in the ice cover create new leads (Miles and Barry, 1998). Those areas can exhibit huge temperature differences between cold air and relative warm water. The resulting heat loss causes fast formation of new ice. Even leads covered by thin ice show much higher heat fluxes than the surrounding thick ice (Maykut, 1978). The low albedo of leads promotes an energy flow in the opposite direction which increases

the amount of absorbed insolation, resulting in a warming of the underlying water. Leads reduce the internal strength of the sea ice, enabling higher drifting velocities (Rampal et al., 2009) and are expected to influence the atmospheric boundary layer chemistry (e.g. Moore et al., 2014).

25 Large scale satellite remote sensing studies of lead occurrences have been done based on visual and thermal imagers (e.g. Lindsay and Rothrock, 1995; Willmes and Heinemann, 2015). They are generally limited by the resolution of thermal infrared measurements of about one kilometer and by the influence of clouds. By using passive microwave data, Röhrs et al. (2012) avoided the requirement of free sky conditions but reduced the resolution even further to 6.25 km. Despite this resolution a good agreement with CryoSat-2 (CS-2) and ASAR—the Advanced Synthetic Aperture  
30 Radar (ASAR) based estimates of the lead occurrence for leads wider than 3 km has been reported in Röhrs et al. (2012). CryoSat-2 based lead detection is expected to be a good complement to previous estimates as it combines an increased resolution of some hundred meters with a strong atmospheric independence. The quality of this approach has been assessed by Zygmontowska et al. (2013) for airborne surveys and is topic of this study for CS-2 measurements.

35 Apart from the lead area, also the width distribution is important for the turbulent heat transport in ice covered regions. A convective boundary layer evolves over leads which increases in thickness towards the downwind side of the lead (Andreas et al., 1979). This boundary layer dampens the heat flux per lead area which is therefore higher for narrow leads than for wider ones. This has led to different lead-width dependent heat transfer formulations (e.g. Andreas and Murphy, 1986). Marcq  
40 and Weiss (2012) show that the turbulent heat flux over leads is up to 55 % higher if using a power-law distribution down to a lead width of 10 m instead of considering all leads as one large area of open water.

The Arctic sea ice extent declined substantially over the last decades (Serreze et al., 2007), while comparable studies for the ice thickness are rare and struggle with uncertainties (Lindsay  
45 and Schweiger, 2015). Ice thickness estimates based on upward looking sonars on submarines (e.g. Rothrock et al., 2008) or moorings (Proshutinsky et al., 2009) have a relatively sparse temporal and spatial coverage. Airborne and helicopter based thickness measurements utilize the strong difference between the electromagnetic inductances of seawater and ice. They are of great value for regional studies and validation, but are restricted by the limited number of conducted surveys (Haas et al.,  
50 2010; Renner et al., 2013, 2014; Maaß et al., 2015).

Sea ice thickness is retrieved from satellites by radiometry, i.e. the influence of the ice thickness, salinity and temperature on the emissivity and transmittance. Various passive thermal to microwave sensors have been used (AVHRR, MODIS, SSM/I, AMSR-E, MIRAS) (Yu and Rothrock, 1996; Singh et al., 2011; Martin et al., 2005; Kaleschke et al., 2012; Tian-Kunze et al., 2014). As the ice  
55 thickness information saturates for all these sensors at a certain level, this approach is only capable of measuring relatively thin ice, typically well below one meter (e.g. Kaleschke et al., 2010).

Another approach utilizes altimetry in order to derive the snow or ice freeboard, i.e. the elevation difference between the Sea Surface Height (SSH) and snow or ice surface, respectively. Laser signals only reach the snow surface, while radar altimeters are basically showing the snow-ice interface elevation. By considering the relevant densities and the snow thickness those freeboards can be converted into ice thickness by assuming hydrostatic equilibrium. Sea ice thickness has been derived from  $K_u$  band radar altimetry from the European Remote Sensing Satellites ERS-1 and ERS-2 as well as Envisat and CS-2 (Laxon et al., 2003; Giles et al., 2008; Laxon et al., 2013; Ricker et al., 2014). These radars are not restricted to clear sky conditions, but limited knowledge of the snow loading and the radar interaction with the snow layer currently limits the accuracy of altimeter derived sea ice thickness estimates (Willatt et al., 2011; Kwok, 2014). Advantages of the radar on CS-2, over earlier  $K_u$  band altimeters are the reduced footprint size and noise due to the synthesis of overlapping measurements, its orbit which allows a coverage up to 88° North and South and the potential of interferometric measurements (Wingham et al., 2006). In most parts of the Arctic ~~the SARM-Ocean the Synthetic Aperture Radar (SAR) mode is used except for most-coastal-areas-and many coastal areas where the SAR Interferometric (SARIn) mode finds application. Until July 2014 the so called 'Wingham Box' (80-85° N and 100-140° W) where the SAR Interferometric mode finds application was another area of SARIn mode measurements.~~

The SSH is crucial for altimeter based ice thickness retrievals. For this reason the altimeter measurements are separated into those from ice and those from leads (see Fig. 1 for examples from CS-2). The lead measurements are used to derive the SSH, which acts as reference for the freeboard. Leads covered by thin ice and falsely detected leads (i.e. thick ice) result in an overestimation of the SSH and therefore in a negative bias in the derived freeboard and thickness. If considering only a very few, assured lead measurements the statistical error increases (Armitage and Davidson, 2014). It is therefore of high interest to find a lead detection method which is very trustworthy and detects as many leads as possible.

In this study the quality of CS-2 based lead detection procedures is assessed by a comparison with MODIS measurements. Previously published classifiers are implemented and compared with newly derived ones in a Receiver Operating Characteristics (ROC) graph. The most promising one is subsequently used to derive the lead area fraction and the lead width distribution. Thereby this study attempts to close a gap of knowledge about the differences of lead detection procedures from CS-2 and makes suggestions for improvements, which has direct implications for sea ice thickness estimates.

## 2 Methods

### 90 2.1 The ground truth

In order to optimize and compare the performance of different classification routines, we choose a supervised classification approach. Visual Moderate Resolution Imaging Spectroradiometer (MODIS) measurements can be used to distinguish between sea ice and water (Su et al., 2012). Two MODIS instruments are in operation on the NASA satellites Terra and Aqua. They cover the earth surface every 95 1 to 2 days and measure in 36 spectral bands from visual (used here) to infrared (Barnes et al., 1998). We identify land and cloud influences manually and are therefore able to rely only on the MODIS band 2 (around 857 nm wavelength) level 1B reflectance as reference data. It has a resolution of 250 m and seems to be even more suited to identify leads than band 1 (not shown). Dark areas with sharp edges and linear shapes in the MODIS images are interpreted as leads. CS-2 measurements 100 from these areas, recorded less than one hour before or after the MODIS acquisition, are manually labeled as lead. In the same way we identify CS-2 measurements of ice while all measurements with a mixture of both classes within the footprint are excluded from this study (see also Fig. 2a). The CS-2 footprint is assumed to be 300 m in and 1500 m across flight direction. This ice/lead information is in the following considered as ground truth, regardless of possible mislabeling for example 105 caused by unexpected high ice velocities.

The ground truth consists of 722 lead and 5768 ice measurements. Note that this method is limited by the resolution of MODIS. CryoSat-2 measurements which look like originating from ice in MODIS scenes can actually contain small amounts of leads. See section 4.2 for a discussion of this circumstance. The ground truth is acquired from February to the beginning of May in 2012 and 110 2013 from seven MODIS granules in the eastern Beaufort ~~sea~~Sea and north of the Canadian Arctic Archipelago. For this time of the year optical MODIS scenes are available and surface melting can be ruled out. Within this study we use CryoSat-2 Level 1b data with processor versions “SIR1SAR/4.0” and “SIR1SAR/4.1” (Baseline B). ~~This~~These two SAR mode versions are equivalent.

### 2.2 Relation to physical properties

115 Large scale roughness results in a spread in time of the received CS-2 signal as exposed parts of the surface are reached earlier than low lying parts. ~~Additionally small scale roughness is promoting surface scattering instead of reflection, meaning that more energy is scattered in oblique directions~~Roughness with a scale smaller than the wavelength ( $\sim 2.2$  cm for  $K_u$  band) reduces the specularity of the surface. Therefore measurements of the same position from altering ~~look~~incidence 120 angles are more similar for rough surfaces (Wingham et al., 2006). ~~Additionally~~In addition areas further away from the nadir point have a stronger contribution, leading to an emphasized signal following the first (nadir) peak ~~for rough surfaces~~ (Laxon, 1994a). Energy conservation conditions



a reduced maximal receivable signal if the emitted power is scattered in all directions by a rough surface.

125 The characteristic impedance of the surface layer might also influence the signal amplitude (Laxon, 1994a). If the difference in impedance at 13.5 GHz of the uppermost layer and the air is small, there is less reflection and more transmission into the ice/snow. Within the medium it is partly absorbed and scattered by inhomogeneities, again leading to a spread of the signal with lower maximum values and a more homogeneous angular distribution. This process could for example be  
130 favored by a layer of snow with moderate temperature.

As leads are locally bound, the fetch is too small for bigger waves to evolve in the water. The thin ice cover, if present, is yet neither physically deformed nor covered with snow. Furthermore the microstructure of young ice is more compact than of older ice as most brine pockets are filled and less connections have evolved. Therefore leads can be characterized by their commonly flat surface  
135 with relatively high impedance difference to the air. The returns originating from leads are expected to be compressed in time with higher maximum values and stronger incidence angle dependency (specular returns).

~~In the SAR Mode, The Doppler Shift is used in the CS-2 returns from the same position on the ground are received from different look angles. After allocating them by using the Doppler Shift, they are~~  
140 ~~SAR mode to split each returning echo into 64 beams with different along track incidence angles. For each processed point on the ground all beams targeting this point from altering satellite positions are~~ combined to one waveform (Wingham et al., 2006) i.e. the returned power as function of time (see Fig. 1 for ~~examples~~typical ice and lead waveforms). The following waveform based parameters are used: Maximum Power, Pulse Peakiness, Leading Edge Width and Trailing Edge  
145 Width.

~~In addition to stacking the beams to one waveform, they are also~~While in the process of waveform formation the information of the angular dependency is disregarded, the beams are additionally  
integrated over time (summed) ~~separately. A Gaussian distribution curve is used to approximate the returned~~individually. Thereby the incidence angle information is maintained in exchange for  
150 the temporal development. The returning energy as function of ~~beams~~beam number (i.e. incidence angle) is approximated by a fitted Gaussian distribution curve. We use the Stack Standard Deviation and the Stack Excess Kurtosis parameters which are based on this curve.

### 2.3 Parameter definition

– The Maximum Power (MAX) is the highest recorded power of the calibrated waveform in  
155 Watts.

- The Pulse Peakiness (PP) has been established by Laxon (1994b) and is defined as the MAX divided by the accumulated power ( $P^{WF}$ ) of all bins constituting the waveform:

$$PP = \frac{\max(P^{WF})}{\sum_{i=1}^{128} P_i^{WF}}, \quad (1)$$

160 which is the same definition as used by Armitage and Davidson (2014), while the values of Laxon et al. (2013) are divided by 100 and those of Ricker et al. (2014) by 128 for consistency.

- The Left and Right Pulse Peakiness (PPL and PPR) from Ricker et al. (2014) for Baseline B data are defined as (Robert Ricker, personal communication, January 2015):

$$PPL = \frac{15 \cdot \max(P^{WF})}{\sum_{i=i_{\max}-6}^{i_{\max}-2} P_i^{WF}}, \quad (2)$$

$$165 \quad PPR = \frac{15 \cdot \max(P^{WF})}{\sum_{i=i_{\max}+2}^{i_{\max}+6} P_i^{WF}}, \quad (3)$$

170 where  $i_{\max}$  is the index of the maximal value of the waveform. The PPL and PPR have been proposed in order to reject off-nadir leads, the influence of which can not be quantified based on our methodology (see Sect. 2.1). Therefore the PPL and PPR are not fully included in this study. However, they are defined as we use the classifier of Ricker et al. (2014) for comparisons.

- The Leading Edge Width (LEW) is defined as the width between 1 and 99 % of the amplitude of a Gaussian fit to the leading edge of the waveform. The fitted area starts at the first bin reaching one percent of the Maximum Power and ends at the second bin, following the first peak. The first peak is the first local maximum reaching at least 50 % of the Maximum Power. To avoid bimodal waveforms, we ~~completely~~ exclude measurements with a ~~first peak~~ smaller than 80% of the Maximum Power from ~~this study~~the ground truth. About 7.6% of the waveforms are discarded in this way. Similar fits and constrains are used by Kurtz et al. (2014).

- The Trailing Edge Width (TEW) is defined as the width between 99 and 1 % of the amplitude of an exponentially decaying fit to the trailing edge of the waveform. The fitted area starts at the position of the Maximum Power and ends at the last bin (e.g. Legresy et al., 2005).

- The Stack Standard Deviation (SSD) is the standard deviation (SD) ~~a random variable would have, if of~~ the mentioned Gaussian ~~approximation distribution~~ of the energy as function of ~~beams would be its Probability Density Function. It is therefore describing its width and beam~~ number (i.e. incidence angle). The SSD is describing the width of the Gaussian, it is not the SD

of the energy values ~~itself. The used unit is themselves. We use the SSD in units of “Beams”~~ which ~~but it~~ can also be expressed in ~~terms of angles~~ degrees. Due to the more specular characteristics of leads, the spread of power with incidence angle is expected to ~~narrow~~ be smaller and so is the SSD ~~for leads~~ (Wingham et al., 2006).

- 190 – The Stack ~~Excess~~ Kurtosis (SK) is also obtained from the Gaussian approximation of the energy as function of ~~beams. In contrast to the SSD it is not derived as a fitting variable but~~ from the function, ~~evaluated at the positions of the beams. While all~~ beam number. Continuous Gaussian functions have in general ~~a kurtosis of three, discrete points of them do not have if~~ the Gaussian is ~~badly represented. This is an excess~~ kurtosis of zero, so how can the SK reach  
195 ~~other values? This is attained by evaluating the Gaussian at the beam numbers. The Excess~~ Kurtosis of these discrete values is the SK (Veit Helm, personal communication, June 2014; Wingham et al., 2006). ~~The fitting of the case for very narrow (badly resolved) and extremely~~ wide functions ~~(only partly covered)~~ Gaussian to the measured beam energies and subsequent  
200 ~~evaluation of it at the very same positions can be understood as a smoothing procedure. It is~~ worth mentioning that this procedure might also limit the information the SK provides. The Kurtosis is a measure of the peakedness which is expected to be ~~high~~ higher for leads.

## 2.4 Threshold optimization

Threshold based classifications are widely used to identify leads from  $K_u$  band altimeters. We use a repeated random Cross-Validation technique to derive and test thresholds ( $\Theta$ )(interested readers  
205 are referred to chapter 9 in Duda et al. (2001)).  $\Theta$  consists of one threshold for each parameter used for the respective classifier. The Cross-Validation involves a random separation of the ~~ground~~ truth samples into a training and a testing subset, each of which consist of 50 % of all samples. From the training subset we derive  $\Theta$  by using Eq. 4 and apply it to the testing set to investigate its performance. The random assignment into subsets and the testing of the newly derived  $\Theta$  is  
210 repeated 200 times for each classifier to get an overall performance and an estimation of its spread. These steps are illustrated in Fig. 3.

As mentioned in Sect. 1 there are different applications for lead detection algorithms also resulting in different demands on its characteristics. One plausible aim is to reduce the total amount of false detections to a minimum. But one might also be interested in a more conservative lead detection by  
215 reducing the amount of ice being detected as lead (False Leads) at the cost of less correctly detected leads (True Leads). A more conservative detection might be used for a freeboard retrieval as False Leads might result in a bias while high True Lead Rates are not always of high importance.

To take these different demands into account we include a weighting factor  $w$  in the cost function.

$$\text{cost}(\Theta) = w \cdot \text{False\_Ice}(\Theta) + \text{False\_Leads}(\Theta), \quad (4)$$

220 where False\_Ice represents the amount of lead samples classified as ice.  $\Theta$  is derived by minimizing the cost function on the training subset using the Nelder–Mead simplex algorithm (Nelder and Mead, 1965) with up to 400 initial guesses to find the global minimum. The Nelder–Mead method is an unconstrained direct search algorithm for multidimensional minimization. This optimization reduces primarily False Leads for  $0 < w < 1$ , while for  $w = 1$  the total amount of false classifications (False  
225 Ice + False Leads) is minimized. We use the parameter acronym with the weight as index to point at the corresponding one dimensional classifier.

This methodology is applied to all single parameters and all possible pairs of them. In the latter case  $\Theta$  is derived as the combination of both thresholds with the smallest value of the cost function.

### 3 Results

#### 230 3.1 Classification performance

In Fig. 2 the CS-2 track is essentially crossing three wider leads, two of which are brighter at the northern side. This indicates that they are covered by ice on this side, while the southern side might exhibit open water. The third wider lead around  $71.2^\circ$  N and a thinner one at  $71.75^\circ$  N seem both to be completely covered by thin ice. The manual classification in Fig. 2a only visualizes the method-  
235 ology as the time difference is larger than one hour and this scene is therefore not part of the ground truth. Gaps in the track occur when the MODIS information of CS-2 footprints cannot be assigned unambiguously to leads or ice. ~~In Fig. 2b the  $MAX_{0.5}$  detects only some leads and those only partly. It shows an accentuation on newly formed parts of the leads. The  $PP_1$ ,  $MAX_1$  and the classifier developed by Ricker et al. (2014) (from now on hereinafter called RI14) show strong similarities as~~  
240 ~~both they~~ detect all relevant leads while lead detections are very rare where the MODIS scene shows ice. However ~~for wider leads they can show all of them show in some cases~~ a mixture of ice and lead detections ~~which is in some cases stronger for the  $MAX_1$  and in other cases for the RI14 within wide leads~~ (not shown). The classifier used by Laxon et al. (2013) (from now on hereinafter called LX13) detects all visible leads without a significant amount of missing lead detections, but it also  
245 detects leads where no or only weak indications for them can be found in the MODIS scene.

Figure 4 shows a Receiver Operating Characteristics (ROC) graph of all tested classifiers. Each classifier is represented by one point in the graph, the position of which is defined by its True Lead Rate (TLR, the amount of correctly detected leads divided by the amount of tested lead samples) and False Lead Rate (FLR, the amount of ice measurements in the ground truth detected as lead  
250 divided by the amount of tested ice samples). The upper left corner corresponds to ideal classifiers and the principle diagonal represents random assignments. For each parameter and pairs of them, we use different weights resulting in different  $\Theta$  and corresponding performances. For single parameter classifiers 15 different weights (0.001, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1, 2, 5, 10, 30, 100) are applied to capture the development of the performance (from the lower left corner to the upper right

255 in Fig. 4) while  $w = 0.001$ ,  $w = 0.5$  and  $w = 1$  are used in the two dimensional case. To follow the performance of e.g. MAX based classifiers one can start with small  $w$ , implying high values of  $\Theta$  which is detecting only a few leads (lower left corner in Fig. 4). With increasing  $w$  and decreasing  $\Theta$  the TLR increases in the beginning much faster than the FLR. At some point the amount of correct lead detections is mostly constant while a further lowering of  $\Theta$  is mainly increasing the amount  
260 of ice measurements which are detected as lead. As relative performances are shown the classifier closest to the upper left corner is not necessarily the “best” one but if one classifier is on the upper left side of another it can be considered as superior. Further remarks on ROC graphs are given by Fawcett (2006).

It is at this point not important how the thresholds ( $\Theta$ ) are derived but only the combination of its  
265 value and performance. This allows us to compare the classifiers found here with those other authors have developed, independent from the optimization routine.

Classifiers based on the Maximum Power (MAX) appear on the upper left side of all others on the whole range of Fig. 4. Only classifiers using two parameters including the MAX (black marker) reach the single feature MAX classifications and are in all cases very close to them. All other classifiers  
270 based on pairs also show very similar results to that classifier based on the single, more suited parameter within its pair.

The PP and LEW based classifiers show strong similarities and have the second-best combinations of True and False Lead Rates.

Figure 5a illustrates the spread within the runs in terms of the SD of the True and False Lead  
275 Rates. The differences between all shown one dimensional classifiers and the corresponding two dimensional ones are clearly smaller than the inherent fluctuations and are therefore considered as not significant. The classifiers based on the MAX are separated from the others by more than their SDs for small weights, while they are not for higher weights. However, ~~Fig. 5b shows that the fluctuations~~ the fluctuation in classifier performance of individual runs with the same weight occur  
280 mostly in the direction of ~~neighboring weights, which shows that the performance of MAX based classifiers is even stronger separated from the others~~ the mean performances of neighboring weights on the same parameter (i.e. along the lines) as shown in Fig. 5b.

### 3.2 Sea Surface Height

The SSH is calculated as a second stage of accessing the quality of classifiers. To derive the SSH  
285 from leads is a popular application, to test the classifier behavior in this context is therefore a very practical approach. This is done statistically by investigating the stability of SSH estimates from different classifiers.

~~We use the threshold first maximum re-tracker (TFMRA,  $TL=0.4$ ) to calculate~~ The function  
 $A \text{sinc}^2(\pi B_w(\tau - \tau_0))$  is fitted to the waveform from  $P_{\text{imax}-2}^{\text{WF}}$  to  $P_{\text{imax}+2}^{\text{WF}}$ . Where  $A$  is the amplitude,  
290  $B_w = 320$  MHz is the received bandwidth and  $\tau$  the delay time.  $\tau_0$  is the center of the fit and is used

as tracking point, i.e. the delay time which is assumed to correspond the return from the main scattering surface. Kurtz et al. (2014) have shown that specular returns are well approximated by a sinc<sup>2</sup> function and that the tracking point should be defined close to the maximum of the waveform. The range is corrected for atmospheric influences (ionosphere, wet and dry troposphere, dynamic atmosphere and the inverse barometric effect) and tides (namely: ocean, long period, solid earth, polar and ocean loading tides) as provided in the CS-2 L1B data. The surface elevation of lead measurements is considered as SSH and the variance is calculated on a 10 km × 10 km north polar stereographic grid. All SARM. All SAR mode measurements from January to March of the years 2011 to 2014 are used brought to a 10 km × 10 km grid.

300 The mean field represents the stable SSH signal which Figure 6 shows the SSH anomaly, i.e. the difference of individual measurements along a CS-2 track from the multi-year mean SSH field. The LX13 shows the largest amount of lead detections and the strongest SSH anomalies. The other three classifiers show a more similar behavior but with the MAX<sub>1</sub> having a notably reduced amount of large (outside of ±0.2 m) SSH anomalies.

305 The mean SSH field could be used as reference for ice thickness estimates. The variance around it is acts as an indicator for its reliability. Differences and is caused by SSH variability, noise and the lead detection behavior. We expect differences of the variance between the classifiers are expected to be caused by the only by the detection behavior, namely the inclusion/exclusion of ice measurements and/or off-nadir leads lead measurements. Figure 7 shows the variance distribution of selected classifiers. The MAX<sub>1</sub> based on the gridded SSH estimates. The MAX<sub>1</sub> shows in general the smallest variances, while the amount of grid cells with high values converges to zero histograms converge to zero with increasing variances for all classifiers in a similar way.

### 3.3 Spatial distribution

In the following sections we are using the MAX<sub>1</sub> which has been derived by minimizing the total amount of false classifications and its results are therefore taken as the best representation of the overall lead occurrence.

Figure 8 shows the lead fraction in the Arctic region as derived from CS-2 by dividing the amount of detected lead measurements by the total amount of measurements from January to March 2011. The AMSR-E Arctic lead area fraction (Röhrs and Kaleschke, 2012; Röhrs et al., 2012) (downloaded in September 2014) is also shown, combined over the same period and brought to the same spatial resolution.

Lead detections from CS-2 are most common in the Baffin Bay, the Fram Strait region, the northern Barents Sea and the Kara Sea, as well as in the western Laptev and the Chukchi Sea, all with lead fractions up to around 15 % (Fig. 8a). The central Arctic including the area north of the Canadian Arctic Archipelago and the northern Canada Basin show low lead fractions of around 0–1.5 %. In

the southern Beaufort sea-Sea and especially its shear zone next to the coast line, lead fraction values of up to 6 % occur.

A somewhat different picture of the lead fraction pattern emerges by using the AMSR-E Arctic lead area fraction from Röhrs et al. (2012) (Fig. 8b). In areas covered by both estimates the CS-2 based one mostly appears to be higher than the AMSR-E based estimate. This is not the case in the south eastern Beaufort sea-Sea where the AMSR-E product shows values of 15 % and more while they reach from 1.5 to 5 % for the CS-2 based estimate. We observe reasonable agreements in the Fram Strait region, the East Siberian Sea and the Chukchi Sea. Increased values occur for both estimates near islands like Svalbard, Franz Josef Land, Severnaya Zemlya and the Wrangel Island. However there are big differences between the datasets in the Baffin Bay, the Fram Strait regions close to the ice edge, the northern Barents Sea and the Kara Sea where CS-2 consistently detects more leads than the AMSR-E lead area fraction indicates.

While a daily open ocean mask is provided for the AMSR-E product, we consider all areas north of 65° N for the CS-2 based estimates. The ice edge on the Atlantic side, as indicated by the AMSR-E mask agrees well with the transition of CS-2 lead fractions from zero to higher values.

By the end of February 2013 the whole Beaufort Sea was pervaded by leads. Favored by storms the ice started in mid February to move into the direction of the Bering Strait, causing a divergence in the pack ice. This is the reason for the opening of leads, beginning in the western part and propagating to the east. This process accelerated around 27 February after which all but the fast ice at the Canadian coast and the sea ice at the Canadian Arctic Archipelago was fractured. See also Beitsch et al. (2014) for further descriptions.

By comparing the CS-2 lead fractions from February and March 2013 (Fig. 9) the pattern of this fracture event are reproduced with a proper shape and amplitude. Most lead pattern can be observed in both months, in many cases slightly decreasing in amplitude towards March. However, while in February only in the western part of the Beaufort Sea noticeable amounts of leads are detected, the complete region shows 8 to 15 % lead coverage in March.

### 3.4 Apparent lead width

To investigate the lead width distribution we use a proxy which we call apparent lead width. The apparent lead width is the amount of consecutive  $MAX_1$  lead detections multiplied by the approximate distance between two positions of 300 m. It can be seen as a measure of the CS-2 track interval over a crossed lead or as the width of a lead how it appears in the one dimensional domain of the CS-2 track. If the lead orientation is orthogonal to the CS-2 track, the apparent lead width is our best estimate of the actual lead width. We do not allow any ice detection within a lead which will in case of false detections split a lead into smaller ones.

The apparent lead width distribution follows a power-law in winter months with an exponent  $a$  of 2.47 for values of 600 m and more (Fig. 10). A quantity  $z$  is called power-law distributed if its

Probability Density Function  $p(z)$  satisfies:

$$p(z) \propto z^{-a}, \quad (5)$$

where  $a$  is the power-law exponent. It is derived following the approximation of Clauset et al. (2009)

for discrete distributions with a simple adjustment for a step size of 300 m as shown in Eq. (6).

$$a \approx 1 + N_Z \left( \sum_{i=1}^{N_Z} \ln \frac{z_i}{z_{\min} - \frac{1}{2} \cdot 300 \text{m}} \right)^{-1} \quad (6)$$

For the calculation of the power-law exponent only apparent lead widths  $z_i$  with a width equal or higher than  $z_{\min} = 900 \text{m}$  are considered with  $N_Z$  being the amount of them. A line representing a power-law with the calculated exponent is displayed in Fig. 10. It shows the validity of this approximation down to 600 m, as both lines show a strongly parallel development as the slopes of both lines agree very well.

The interannual variability is small with exponents between 2.42 in 2013 and 2.52 in 2011 with a SD of 0.04 amongst all four years. Differences between January, February and March of the same year are even smaller while the exponent decreases towards spring and autumn. All calculated distributions are following a power-law for apparent lead widths of 600 m and more.

## 4 Discussion

### 4.1 Classification performances

Classifiers based on the MAX parameter show generally the best ratio between True and False Lead Rate. A classifier using  $\text{MAX} > 2.58 \times 10^{-11} \text{ W}$  as threshold ( $\text{MAX}_1$ ) detects 68.18 % of all leads correctly while only 3.41 % of the tested ice measurements in the ground truth are detected as leads. The  $\text{PP}_1$  using 0.35 as threshold has a TLR of 64.66 % (instead of ~~68.14~~ 68.18 %) and a FLR of 4.09 % (instead of 3.41 %). The differences are even stronger for higher thresholds of  $1.22 \times 10^{-10} \text{ W}$  and 0.425, respectively ( $\text{MAX}_{0.5}$  and  $\text{PP}_{0.5}$ , Table 1).

The performances of individual runs overlap only slightly for  $w = 1$  and are well separated for  $w = 0.5$ . This shows that the performance improvement is significant. The increased fluctuation in the direction of neighboring weights in Fig. 5b is likely to be caused by a variability of the thresholds caused by the repeated optimization.

For airborne surveys with a device very similar to SIRAL on CS-2, Zyguntowska et al. (2013) also found the MAX parameter to have less false lead classifications than all other parameters. The best combination of parameters (MAX & TEW) with a Bayesian classifier improves its rate only little from 6.5 to 6.2 %. Zyguntowska et al. (2013) define the False Lead Classification Rate (FLCR) as percentage of all lead detections originating from sea ice. This is different to our False Lead Rate as we use the amount of true ice measurements as base. The FLCR calculated from the absolute values



in Table 1 are 28.6 and 12.5 % for  $MAX_1$  and  $MAX_{0.5}$ , respectively. One reason for higher error rates of CS-2 is the reduced resolution of  $300\text{ m} \times 1500\text{ m}$  in contrast to around  $10\text{ m} \times 50\text{ m}$  for the airborne device. Thereby it becomes much more likely that different surface types occur within one footprint. Further we have to allow for some temporal differences in the data acquisition and have to collocate the datasets, while for the airborne surveys optical images are taken simultaneously. Deficiencies of the ground truth which might be caused by ice drift and opening/closing of leads between the data acquisition, collocation and unnoticed narrow leads increase the error rates which might therefore be overestimated.

Compared to their MAX classifier, the PP classifier of Zygmuntowska et al. (2013) is detecting more leads from both, ice and lead measurements. This is directly connected to applied thresholds and is not a parameter property. For a solid decision which parameter is suited best for lead detection it is necessary to vary the thresholds.

~~We are able to propose the use of the MAX parameter instead of the PP.~~

Three classifiers developed by other authors are included in this study. With the same amount of False Leads, the True Lead Rates can be increased for our dataset from 9 to ~~~13~~ ~13 % (Röhrs et al., 2012), from 83 to ~~~89~~ ~89 % (LX13) or from 61 to ~~~79~~ ~79 % (RI14) if a MAX based classifier is used instead (Fig. 4).

The shown classifiers using two parameters detect a lead if both thresholds are reached. This logical 'and' criterion is now replaced by an 'or'. A classifier based on the MAX and the PP could for example define a measurement as a lead if its MAX value is above ~~10=11~~ 10<sup>-11</sup> W or if its PP value is above 0.3 (one of those is now sufficient). This is influencing the amount of False Ice and False Lead detections (i.e. the cost function). As a result our example has higher thresholds than it would have for the same weight and parameters using the 'and' criterion. Performing the same test as before but now using the 'or' criterion for all pairs of parameters brought no improvement of the classification (not shown).

The fact that combining two parameters seems to have no benefit at all indicates that the parameters are basically all utilizing the same physical information and that the instrument and fading noises have either an correlated influence on all parameters or the influence is not significant at all. As some of the parameters are derived in a very different way (e.g. waveform and stack based ones) we do not expect the noise to affect them equally. We conclude that noise plays probably only a minor role for the classification errors.

## 4.2 Narrow leads and Sea Surface Height

It has been shown that leads which cover only a small fraction of a radar altimeter footprint can dominate the signal, due to the high amplitude of specular returns (Drinkwater, 1991). Therefore CS-2 detects leads which are simply not visible for MODIS despite its higher resolution. The fraction of this leads in the ice class of the ground truth cannot be quantified by our approach. Those narrow

430 leads either cover the nadir point or not, while leads covering the whole footprint ('True Leads') do for sure. Therefore one could expect True Lead measurements to ensure a higher quality (see section 4.3) for the derivation of the SSH.

This expectation is supported by the smaller spread of the SSH estimate based on the  $MAX_T$   $MAX_I$  compared to the  $PP_T$   $PP_I$  with nearly the same amount of lead detections (True + False  
435 Leads, Table 1). This advantage is on the other hand certifying that narrow, unnoticed leads in the ice class do not reverse the ROC analysis.

### 4.3 Off-nadir leads

As mentioned before, leads which are not directly in nadir direction can dominate the signal. As this can cause a bias in elevation estimates Ricker et al. (2014) introduced the Left and Right Pulse Peak-  
440 iness to avoid off-nadir leads. It has further been shown that it is, to some extent, possible to reduce the influence of off-nadir leads by increasing the Pulse Peakiness threshold of a single parameter classifier (Armitage and Davidson, 2014). This is done at the cost of discarding  $\sim 60$  up to  $60$  % of the lead detections and thereby increasing the statistical error. The underlying process allowing for this reduction is the influence of the surface orientation towards the sensor on the maximum return.  
445 The relative orientation, favoring high maximum values the most, is expected to be found close to the nadir point. The further away from this point the main scattering surface (i.e. the lead) is, the more power is reflected away from the sensor instead of back towards it. This process is influencing the MAX value in the first place which has then implications for the PP (Armitage and Davidson, 2014). Therefore it is reasonable to assume that the influence of off-nadir leads is also reduced for  
450 high MAX thresholds, potentially even stronger than for the PP as the process causing this reduction has a more direct impact on it. This assumption is supported by the reduced SSH variance of the  $MAX_T$   $MAX_I$  even though we can not say whether this reduction is caused by the elimination of off-nadir leads or incorrectly classified ice measurements (or a combination of both).

### 4.4 Spatial distribution

455 The CS-2 lead fraction shows a reasonable spatial distribution. It is small in the central Arctic and north of the Canadian Arctic Archipelago which are typical regions of thick multi-year ice. It shows high values in regions of high drifting velocities or known to favor the development of polynyas like the Fram Strait, the western Laptev Sea and the Chukchi Sea. The lead fractions are also increasing around most islands and coasts which introduce shear between the land fast ice and the drifting pack  
460 ice. Small lead fractions in the eastern Laptev sea  $Sea$  and the western parts of the East Siberian Sea could indicate the presence of large amounts of land fast ice. The absolute lead fraction values tend to be higher but are mostly in agreement with those of Lindsay and Rothrock (1995). They found lead fractions of 2 to 3 % for the central arctic and 6 to 9 % in the peripheral seas in the winter using the Advanced Very High Resolution Radiometer (AVHRR).

465 In nearly all regions the CS-2 lead fraction exceeds the AMSR-E Arctic lead area fraction from Röhrs et al. (2012) (Fig. 8). While the AMSR-E product only detects most leads with a width of 3 km and more, a width of at least some hundred meters is sufficient for detection by CS-2. As shown in Sect. 3.4 the apparent lead width is following a power law on the scale of kilometers, implying that measurements from narrow leads are largely outnumbering those from wider leads.  
470 In contrast to the CS-2 lead fraction, the AMSR-E product is additionally not including very large regions of thin ice like huge polynyas as a spatial high-pass filter is used.

The ice edge towards the North Atlantic is captured by both approaches quite similar. In Fig. 8a we expect the ice edge to be at the interface between areas of no lead detections around the Norwegian and central Barents Sea and neighboring areas of higher lead fractions. This allows the inference that  
475 the  $MAX_1$  detects no leads over the open ocean. For this reason the lead fraction of grid cells at the very ice edge ~~are likely to underestimate the lead fraction~~ is likely to be underestimated relative to the ice covered part of the cell.

While the AMSR-E lead fraction drops relative consistently down to values around 2–3 % within a belt of around 200 km from the ice edge, CS-2 based estimates show much higher values of around  
480 12 % in this areas. The high values in the marginal ice zone are reasonable as this area is likely to be fractured due to the influence of ocean waves. Especially in the Baffin Bay, the northern Barents Sea and the Kara Sea high rates of new ice formation can occur in winter which is in good agreement with high CS-2 lead fractions of these regions. The general reasonable distribution and its alternation ~~increase~~ enhance our confidence in ~~our~~ the CS-2 lead detection algorithm.

#### 485 4.5 Apparent lead width

Compared to the power-law, the found number of apparent lead width of 300 m is smaller than expected. This is a typical feature on the lower bound of the resolution as leads of this size are not always covered by a single measurement but partially by more, not necessarily leading to a detection. This is intensified by the elongated footprint of CS-2 as small leads may only be detected if they  
490 cover most of the width of the footprint. The ~~MAX<sub>T</sub>~~ MAX<sub>L</sub> is optimized mainly on leads wider than a single measurement which could also cause the relative small number of apparent lead width of 300 m. Therefore it is likely that the bend on the lower bound of the distribution in Fig. 10 is an artifact and not a valid part of the lead distribution.

Marcq and Weiss (2012) have found a power-law exponent similar to our between 2.1 and 2.6  
495 for scales from 20 m to 2 km by analyzing a single SPOT image with a resolution of 10 m. In two submarine based surveys, power-laws with exponents of 2 and 2.29 were found for the regions from the Fram Strait to the North Pole and the Davis Strait, respectively (Wadhams, 1981; Wadhams et al., 1985). In both cases resolutions of about 5 m are present and the power-law holds for the range from 50 to 1000 m. The examination of submarine and mooring data by Kwok et al. (2009) also indicates  
500 a strong accumulation of lead widths down to 5 m but the distribution has not been analyzed. For the

central Arctic, a study of Lindsay and Rothrock (1995) also states a power-law distribution, but with a mean exponent of 1.6 for scales from one to around 50 km. It is based on thermal to near visible infrared measurements from the AVHRR, which is despite its resolution of one kilometer expected to detect leads with a minimum width below this size. It has been discussed whether the lead width distribution might be scale dependent (Lindsay and Rothrock, 1995; Marcq and Weiss, 2012) which seems not to be the case, as we found a stable power-law behavior on scales partly covering those of all other studies.

The results of Lindsay and Rothrock (1995) are contradictory to ours as we found a higher power-law exponent, implying a higher fraction of narrow leads. One explanation would be the relative coarse resolution of the AVHRR in combination with its high sensitivity to leads. This could cause leads to appear wider than they are, as well as several narrow ones to appear as one wide lead, resulting in a less steep apparent lead width distribution. Comparisons with MODIS images indicate that the classifier used in this study switches in some cases between lead and ice detections over refrozen leads. This could result in an overestimation of the power-law exponent. The estimates might also have a different tolerance of refrozen leads, while both include at least the early stages of freezing. The size of leads is often growing with time as the surrounding ice floes keep drifting apart, meaning that estimates which include older leads are also likely to show less steep apparent lead width distributions.

Another reason could be an actual shift in the distribution between the periods from 1989 to 1995 and 2011 to 2014. This would be consistent with the results of Marcq and Weiss (2012) but would not explain the differences to those studies by Wadhams (1981) and Wadhams et al. (1985). However, this shift could be driven by observed changes in the amount of perennial ice, the ice thickness and drifting velocities (Nghiem et al., 2007; Haas et al., 2008; Rampal et al., 2009). Rampal et al. are further linking a found increase in winter strain rates between 1978 and 2007 to a weakening in mechanical strength of the ice and increased fracturing. We found no sign for a trend of the power-law exponent within the four years of CS-2 data.

#### 4.6 Implications of apparent lead width distribution

As most leads are not crossed orthogonally, the apparent lead width is typically larger than the actual width of the lead. A transformation to the latter is not possible without profound knowledge of the sensitivity of lead detections and requires assumptions about the shape and orientation of leads. This is impeded by a nonuniform distribution of lead orientation (Bröhan and Kaleschke, 2014). For most applications it is not necessary to perform this transformation as this is the way leads appear to anything moving along sea ice, including the wind acting on the ocean surface.

The apparent lead width distribution is showing a strong intensification towards smaller lead widths. The area contribution of leads having the width  $z$  is  $z \cdot p(z) \propto z^{-2.47+1}$ , which is still relatively fast decreasing with increasing width. This indicates that every lead area estimate which

is not capable of detecting narrow leads is very likely to underestimate the total lead area. For a parametrization of lead area estimates it is of high interest to know down to which bound the power law behavior holds. This defines not only the mean lead width but also the fraction of lead area which is not captured by the estimate.

## 5 Conclusions

This study presented the potentials of several parameters and combinations of them to distinguish between CryoSat-2 measurements from leads and those from ice. They have been tested by deriving thresholds and analyzing their capabilities of reproducing a prior classification. The combination of parameters, even though common practice, has not shown any advantage for threshold based classifications. Using the maximum value of the waveform has in all cases shown better results than any other tested parameter, including the Pulse Peakiness. Compared to the classifier used by Laxon et al. (2013) a threshold of  $2.58 \times 10^{-11}$  W on the MAX detected only 68 instead of 83 % of ensured lead measurements but showed a much more stable SSH estimate by reducing the amount of ice being detected as lead and/or off-nadir leads. A solid lead detection, which ensures that nearly all lead classifications are actually originating from leads is facilitating a precise, unbiased freeboard retrieval. It thereby helps to improve ice thickness estimates, which is one of the major aims of the CryoSat-2 mission.

The threshold of  $2.58 \times 10^{-11}$  W was further used as the best representation of the overall lead occurrence. It showed reasonable spatial distributions with relatively high lead fractions of around 12 % in the marginal ice zone. The apparent lead width was derived from the amount of consecutive lead detections. Its distributions is following a power-law with exponent of  $2.47 \pm 0.04$  which implies a concentration of both, amount and area contribution at small lead widths. Embedding this work into those of others, a scale independent lead width distribution from 20 m to 50 km is likely. The implications for a parametrization of low resolution lead area estimates were addressed and its dependency on the lower bound of the found distribution emphasized. The turbulent heat transport over ice covered regions is known to be strongly lead width dependent on small scales. The found distribution is suggesting that the work of Marcq and Weiss (2012), based on a single SPOT scene, can be generalized. This implies a much higher heat transport per lead area than it would be obtained by wide leads. In this manner the presented findings can help to improve the parameterization of this fundamental process in coupled ocean–ice–atmosphere models.

**The Supplement related to this article is available online at  
doi:10.5194/tc-0-1-2015-supplement.**

*Acknowledgements.* This work is supported by the German Federal Ministry of Education and Research (FKZ: 570 01LP1126A) and in parts through the Cluster of Excellence “CliSAP” (EXC177), University of Hamburg, funded through the German Science Foundation (DFG). CryoSat-2 data are provided by the European Space Agency (ESA). Additionally we would like to acknowledge the use of Rapid Response imagery from the Land Atmosphere Near-real time Capability for EOS (LANCE) system operated by the NASA/GSFC/Earth Science Data and Information System (ESDIS) with funding provided by NASA/HQ. The AMSR-E Arctic lead area 575 fraction is provided by the Integrated Climate Data Center (ICDC, <http://icdc.zmaw.de/>), University of Hamburg, Hamburg, Germany. We would like to thank Veit Helm, Stefan Hendricks and Robert Ricker for the kind and fruitful discussions.

## References

- Andreas, E. L., Paulson, C. A., William, R. M., Lindsay, R. W., and Businger, J. A.: The turbulent heat flux  
580 from Arctic leads, *Boundary-Layer Meteorology*, 17, 57–91, doi:10.1007/BF00121937, 1979.
- Andreas, E. L. and Murphy, B.: Bulk transfer coefficients for heat and momentum over leads and polynyas, *J. Phys. Oceanogr.*, 16, 1875–1883, 1986.
- 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 T. Geosci. Remote*, 52, 529–536, 2014.
- 585 Barnes, W. L., Pagano, T. S., and Salomonson, V. V.: Prelaunch characteristics of the moderate resolution imaging spectroradiometer (MODIS) on EOS-AM1, *Geoscience and Remote Sensing, IEEE Transactions on*, 36, 1088–1100, doi:10.1109/36.700993, 1998.
- Beitsch, A., Kaleschke, L., and Kern, S.: Investigating high-resolution AMSR2 sea ice concentrations during the February 2013 fracture event in the Beaufort Sea, *Remote Sens.*, 6, 3841–3856, 2014.
- 590 Bröhan, D. and Kaleschke, L.: A nine-year climatology of Arctic sea ice lead orientation and frequency from AMSR-E, *Remote Sens.*, 6, 1451–1475, 2014.
- Clauset, A., Shalizi, C. R., and Newman, M. E.: Power-law distributions in empirical data, *SIAM Rev.*, 51, 661–703, 2009.
- Duda, R. O., Hart, P. E., and Stork, D.G.: *Pattern classification*, John Wiley & Sons, second edition., 2001.
- 595 Drinkwater, M. R.: Ku band airborne radar altimeter observations of marginal sea ice during the 1984 Marginal Ice Zone Experiment, *J. Geophys. Res.-Oceans (1978–2012)*, 96, 4555–4572, 1991.
- Fawcett, T.: An introduction to ROC analysis, *Pattern Recogn. Lett.*, 27, 861–874, 2006.
- Giles, K. A., Laxon, S. W., and Ridout, A. L.: Circumpolar thinning of Arctic sea ice following the 2007 record ice extent minimum, *Geophys. Res. Lett.*, 35, L22502, doi:10.1029/2008GL035710, 2008.
- 600 Haas, C., Pfaffling, A., Hendricks, S., Rabenstein, L., Etienne, J.-L., and Rigor, I.: Reduced ice thickness in Arctic Transpolar Drift favors rapid ice retreat, *Geophys. Res. Lett.*, 35, L17501, doi:10.1029/2008GL034457, 2008.
- Haas, C., Hendricks, S., Eicken, H., and Herber, A.: Synoptic airborne thickness surveys reveal state of Arctic sea ice cover, *Geophys. Res. Lett.*, 37, L09501, doi:10.1029/2010GL042652, 2010.
- 605 ~~Helm, V., Humbert, A., and Miller, H.: Elevation and elevation change of Greenland and Antarctica derived from CryoSat-2, *The Cryosphere*, 8, 1539–1559, doi:, 2014.~~
- Kaleschke, L., Maaß, N., Haas, C., Hendricks, S., Heygster, G., and Tonboe, R. T.: A sea-ice thickness retrieval model for 1.4 GHz radiometry and application to airborne measurements over low salinity sea-ice, *The Cryosphere*, 4, 583–592, doi:10.5194/tc-4-583-2010, 2010.
- 610 Kaleschke, L., Tian-Kunze, X., Maaß, N., Mäkynen, M., and Drusch, M.: Sea ice thickness retrieval from SMOS brightness temperatures during the Arctic freeze-up period, *Geophys. Res. Lett.*, 39, L05501, doi:10.1029/2012GL050916, 2012.
- Kurtz, N. T., Galin, N., and Studinger, M.: An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting, *The Cryosphere*, 8, 1217–1237, doi:10.5194/tc-8-1217-2014, 2014.
- 615 Kwok, R.: Simulated effects of a snow layer on retrieval of CryoSat-2 sea ice freeboard, *Geophys. Res. Lett.*, 41, 5014–5020, 2014.

- 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, C07005, doi:10.1029/2009JC005312, 2009.
- 620 Laxon, S. W.: Sea ice altimeter processing scheme at the EODC, *International Journal of Remote Sensing*, 15, 915–924, doi:10.1080/01431169408954124, 1994a.
- Laxon, S. W.: Sea ice extent mapping using the ERS-1 radar altimeter, *EARSel Adv. Remote Sens.*, 3, 112–116, 1994b.
- Laxon, S. W., Peacock, N., and Smith, D.: High interannual variability of sea ice thickness in the Arctic region, 625 *Nature*, 425, 947–950, 2003.
- 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, *Geophys. Res. Lett.*, 40, 732–737, 2013.
- Legresy, B., Papa, F., Remy, F., Vinay, G., van den Bosch, M., and Zanife, O.-Z.: ENVISAT radar altimeter 630 measurements over continental surfaces and ice caps using the ICE-2 retracking algorithm, *Remote Sens. Environ.*, 95, 150–163, 2005.
- Lindsay, R. and Rothrock, D.: Arctic sea ice leads from advanced very high resolution radiometer images, *J. Geophys. Res.-Oceans*, 100, 4533–4544, 1995.
- Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite 635 observations, *The Cryosphere*, 9, 269–283, doi:10.5194/tc-9-269-2015, 2015.
- Maaß, N., Kaleschke, L., Tian-Kunze, X., Mäkynen, M., Drusch, M., Krumpen, T., Hendricks, S., Lensu, M., Haapala, J., and Haas, C.: Validation of SMOS sea ice thickness retrieval in the northern Baltic Sea, *Tellus A*, 67, 24617, doi:10.3402/tellusa.v67.24617, 2015.
- Marcq, S. and Weiss, J.: Influence of sea ice lead-width distribution on turbulent heat transfer between the ocean 640 and the atmosphere, *The Cryosphere*, 6, 143–156, doi:10.5194/tc-6-143-2012, 2012.
- Martin, S., Drucker, R., Kwok, R., and Holt, B.: Improvements in the estimates of ice thickness and production in the Chukchi Sea polynyas derived from AMSR-E, *Geophys. Res. Lett.*, 32, L05505, doi:10.1029/2004GL022013, 2005.
- Maykut, G. A.: Energy exchange over young sea ice in the central Arctic, *J. Geophys. Res.-Oceans*, 83, 3646– 645 3658, 1978.
- Miles, M. W. and Barry, R. G.: A 5-year satellite climatology of winter sea ice leads in the western Arctic, *J. Geophys. Res.-Oceans*, 103, 21723–21734, 1998.
- Moore, C. W., Obrist, D., Steffen, A., Staebler, R. M., Douglas, T. A., Richter, A., and Nghiem, S. V.: Convective forcing of mercury and ozone in the Arctic boundary layer induced by leads in sea ice, *Nature*, 506, 81–84, 650 doi:10.1038/nature12924, 2014.
- Nelder, J. A. and Mead, R.: A simplex method for function minimization, *The computer journal*, 7, 308–313, doi:10.1093/comjnl/7.4.308, 1965.
- Nghiem, S., Rigor, I., Perovich, D., Clemente-Colon, P., Weatherly, J., and Neumann, G.: Rapid reduction of Arctic perennial sea ice, *Geophys. Res. Lett.*, 34, L19504, doi:10.1029/2007GL031138, 2007.

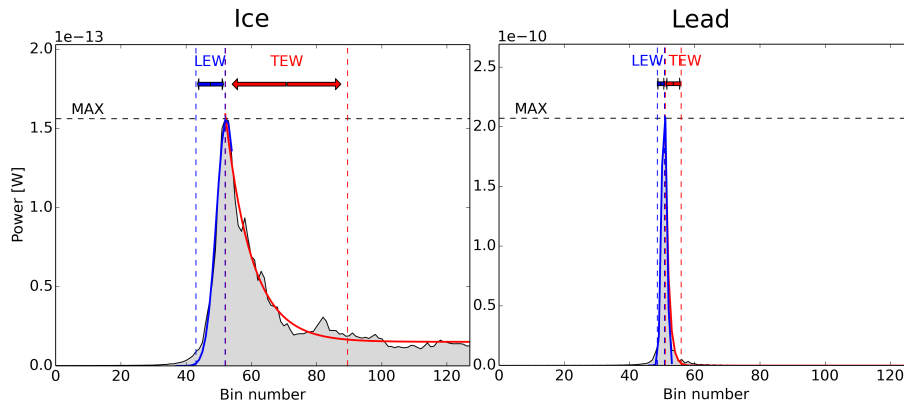


- 655 Proshutinsky, A., Krishfield, R., Timmermans, M.-L., Toole, J., Carmack, E., McLaughlin, F., Williams, W. J., Zimmermann, S., Itoh, M., and Shimada, K.: Beaufort Gyre freshwater reservoir: state and variability from observations, *J. Geophys. Res.-Oceans*, 114, C00A10, doi:10.1029/2008JC005104, 2009.
- 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, C05013, doi:10.1029/2008JC005066, 2009.
- 660 Renner, A. H. H., Hendricks, S., Gerland, S., Beckers, J., Haas, C., and Krumpen, T.: Large-scale ice thickness distribution of first-year sea ice in spring and summer north of Svalbard, *Ann. Glaciol.*, 54, 13–18, 2013.
- Renner, A. H. H., Gerland, S., Haas, C., Spreen, G., Beckers, J. F., Hansen, E., Nicolaus, M., and Goodwin, H.: Evidence of Arctic sea ice thinning from direct observations, *Geophys. Res. Lett.*, 41, 5029–5036, doi:10.1002/2014GL060369 2014.
- 665 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, doi:10.5194/tc-8-1607-2014, 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, doi:10.5194/tc-6-343-2012, 2012.
- 670 Röhrs, J., Kaleschke, L., Bröhan, D., and Siligam, P. K.: Corrigendum to “An algorithm to detect sea ice leads by using AMSR-E passive microwave imagery” published in *The Cryosphere*, 6, 343–352, 2012, *The Cryosphere*, 6, 365–365, doi:10.5194/tc-6-365-2012, 2012.
- Rothrock, D., Percival, D., and Wensnahan, M.: The decline in arctic sea-ice thickness: separating the spatial, annual, and interannual variability in a quarter century of submarine data, *J. Geophys. Res.-Oceans*, 113, C05003, doi:10.1029/2007JC004252, 2008.
- 675 Serreze, M. C., Holland, M. M., and Stroeve, J.: Perspectives on the Arctic’s shrinking sea-ice cover, *Science*, 315, 1533–1536, 2007.
- Singh, R. K., Oza, S. R., Vyas, N. K., and Sarkar, A.: Estimation of thin ice thickness from the advanced microwave scanning radiometer-EOS for coastal polynyas in the Chukchi and Beaufort Seas, *IEEE T. Geosci. Remote*, 49, 2993–2998, 2011.
- 680 Su, H., Wang, Y., and Yang, J.: Monitoring the spatiotemporal evolution of sea ice in the Bohai Sea in the 2009–2010 winter combining MODIS and meteorological data, *Estuaries and coasts*, 35, 281–291, doi:10.1007/s12237-011-9425-3, 2012.
- Tian-Kunze, X., Kaleschke, L., Maaß, N., Mäkynen, M., Serra, N., Drusch, M., and Krumpen, T.: SMOS-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification, *The Cryosphere*, 8, 997–1018, doi:10.5194/tc-8-997-2014, 2014.
- 685 Wadhams, P.: Sea-Ice Topography of the Arctic Ocean in the Region 70° W to 25° E, *Philos. T. Roy. Soc. S.-A.*, 302, 45–85, 1981.
- Wadhams, P., McLaren, A. S., and Weintraub, R.: Ice thickness distribution in Davis Strait in February from submarine sonar profiles, *J. Geophys. Res.-Oceans*, 90, 1069–1077, 1985.
- 690 Willatt, R., Laxon, S., Giles, K., Cullen, R., Haas, C., and Helm, V.: Ku-band radar penetration into snow cover on Arctic sea ice using airborne data, *Ann. Glaciol.*, 52, 197–205, 2011.
- Willmes, S. and Heinemann, G.: Pan-Arctic lead detection from MODIS thermal infrared imagery, *Ann. Glaciol.*, 56, 29–37, 2015.

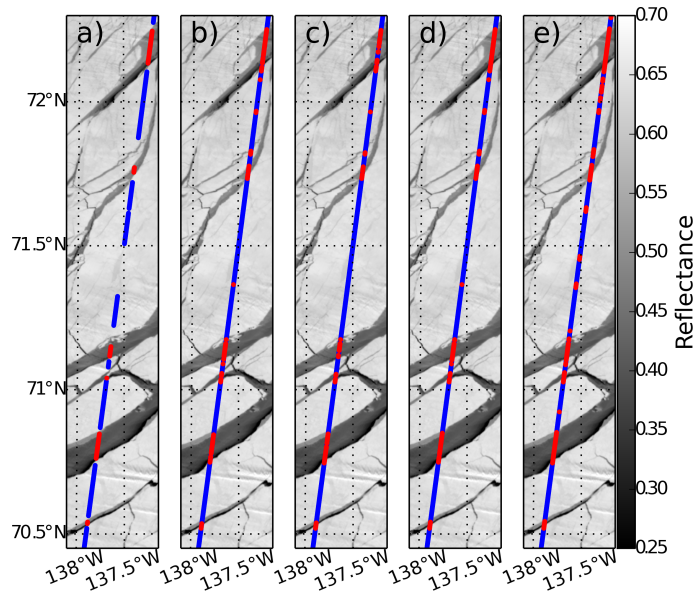
- 695 Wingham, D., Francis, C., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., de Chateau-Thierry, P., Laxon, S.,  
Mallow, U., Mavrocordatos, C., Phalippouc, L., Ratierb, G., Reyc, L., Rostand, F., Viaub, P. and Wallisa,  
D. W.: CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields, *Adv. Space  
Res.*, 37, 841–871, 2006.
- Yu, Y. and Rothrock, D.: Thin ice thickness from satellite thermal imagery, *J. Geophys. Res.-Oceans*, 101,  
700 25753–25766, 1996.
- Zygmuntowska, M., Khvorostovsky, K., Helm, V., and Sandven, S.: Waveform classification of airborne syn-  
thetic aperture radar altimeter over Arctic sea ice, *The Cryosphere*, 7, 1315–1324, doi:10.5194/tc-7-1315-  
2013, 2013.

**Table 1.** Selected classifier performance. The lower-last three classifiers are used by (from bottom to top): RI14, from Röhrs et al. (2012) and LX13. TL: True Leads; FL: False Leads; TI: True Ice; FI: False Ice; TLR and FLR: True and False Lead Rates [%]; eTLR and eFLR: SDs of TLR and FLR within runs [%]. A list of all tested classifier performances is provided as Supplement.

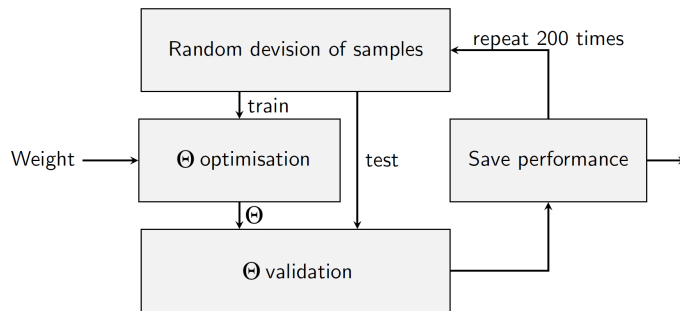
FeatPar	$w$	$\Theta$	TL	FL	TI	FI	TLR	FLR	eTLR	eFLR
MAX	$10^{-3}$	$4.28 \times 10^{-10}$	10 336	226	576 597	61 841	14.32	0.04	2.07	0.05
MAX	0.5	$1.22 \times 10^{-10}$	29 435	4220	572 634	42 711	40.80	0.73	3.26	0.24
MAX	1	$2.58 \times 10^{-11}$	49 204	19 689	557 143	22 964	68.18	3.41	5.89	0.73
MAX	1	$2.55 \times 10^{-11}$	48 808	19 580	557 305	23 307	67.68	3.39	5.61	0.72
TEW		200								
PP	0.5	0.425	22 677	7728	569 244	49 351	31.48	1.34	8.62	0.71
PP	1	0.35	46 602	23 623	553 307	25 468	64.66	4.09	4.94	0.66
PP	–	0.18	59 809	73 003	504 042	12 146	83.12	12.65	1.40	0.42
SSD		4								
MAX	–	$6 \times 10^{-10}$	6576	00	576 811	65 613	9.11	0.00	1.12	0.00
PP	–	0.3125	43 875	28 599	548 145	28 381	60.72	4.96	1.74	0.27
SSD		4								
SK		40								
PPL		40								
PPR		30								



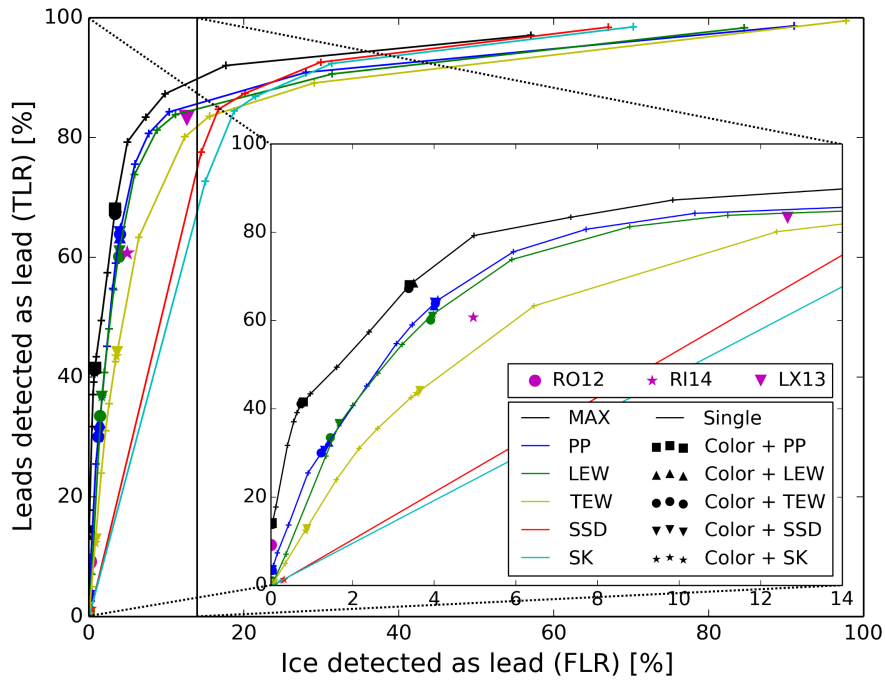
**Figure 1.** Example-Typical CryoSat-2 waveforms from ice (left panel) and a lead (right panel). The definition of the Leading Edge Width (LEW), Trailing Edge Width (TEW) and Maximum Power (MAX) are illustrated while the Pulse Peakiness (PP) is inversely proportional to the gray areas of normalized waveforms would have. The Bin number can be converted into delay time. Note the different scaling factors of the  $y$  axis ( $\times 10^{-13}$  and  $\times 10^{-10}$  for the ice and lead example waveform, respectively).



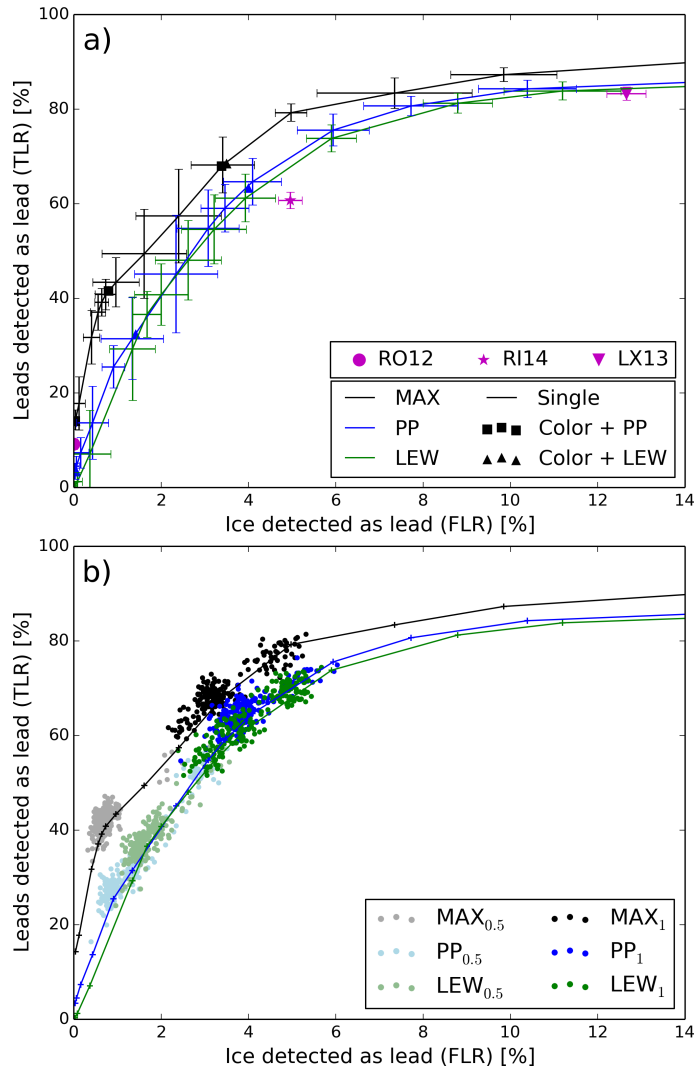
**Figure 2.** MODIS band 2 scene from 6 March 2013 in the southern Beaufort sea combined with a CS-2 track taken 83 min later on. The CS-2 samples have been classified as lead (red) and ice (blue) manually (a) or by  $MAX_{0.5} PP_1$  (b),  $MAX_1$  (c), RI14 (d) and LX13 (e). The classifier from Röhrs et al. (2012) detects no leads within this section.



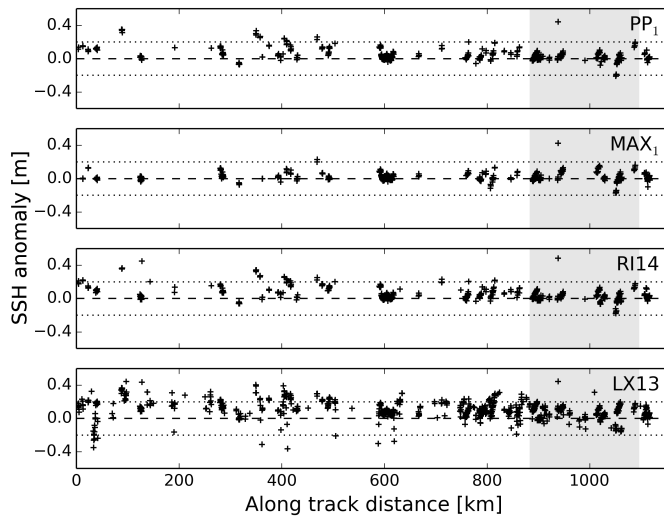
**Figure 3.** Flow chart of the used Cross-Validation scheme.



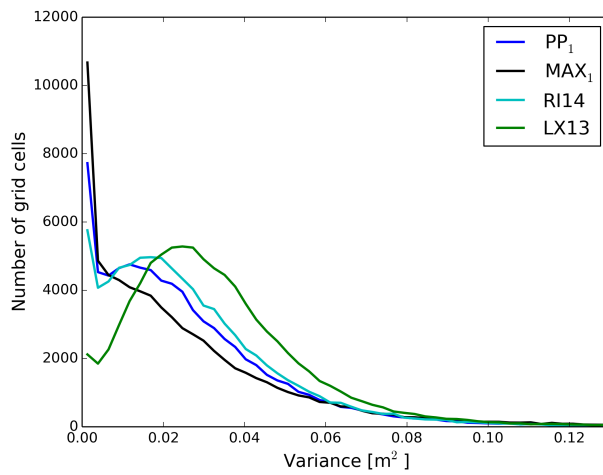
**Figure 4.** ROC graph of tested classifiers with altering thresholds ( $\Theta$ ) on one (connected by lines) and two (marker) parameters as well as predefined classifiers (magenta markers). [RO12 corresponds to the classifier used in Röhrs et al. \(2012\)](#). In the two dimensional case the color indicates one of the parameters and the shape the other one. The insertion is a zoom on small False Lead Rates.



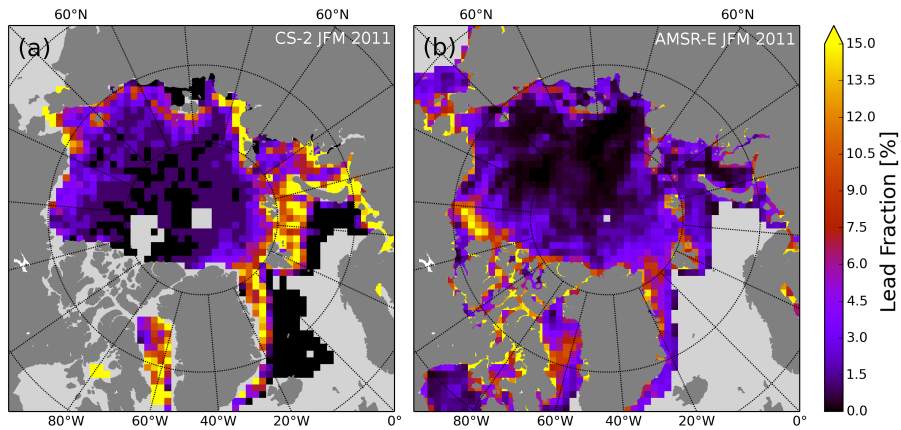
**Figure 5.** (a) ROC graph including error estimates in terms of SD of the 200 runs for each weight of the single feature classifiers using the MAX, PP and LEW as well as the performances and SDs of the predefined classifiers. For comparison the performances of selected two dimensional classifiers are included. (b) Performances of each individual run being part of the single feature classifiers using the MAX, PP and LEW with weights of 0.5 and 1 (dots) in combination with mean values for all weights (lines).



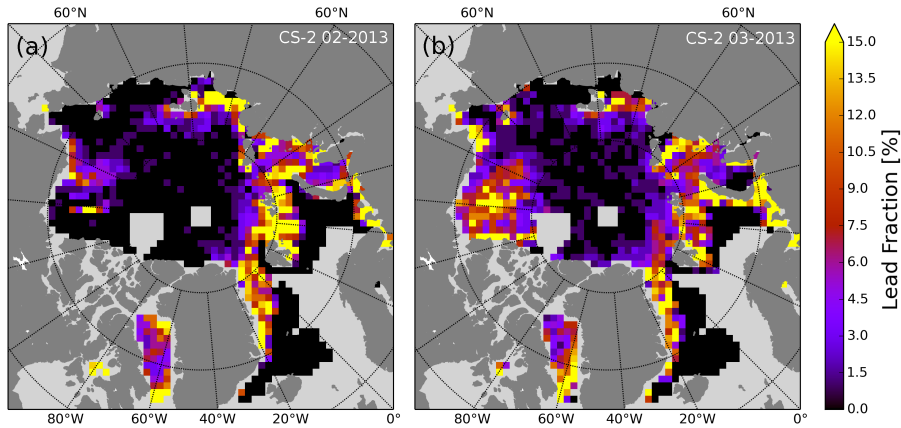
**Figure 6.** [Sea Surface Height anomaly from different classifiers along a typical descending CS-2 track from 6 March 2013. The shaded segment corresponds to the section shown in Fig. 2.](#)



**Figure 7.** Histograms of grid cell SSH variance from different classifiers. Only values based on at least three lead detections are considered.

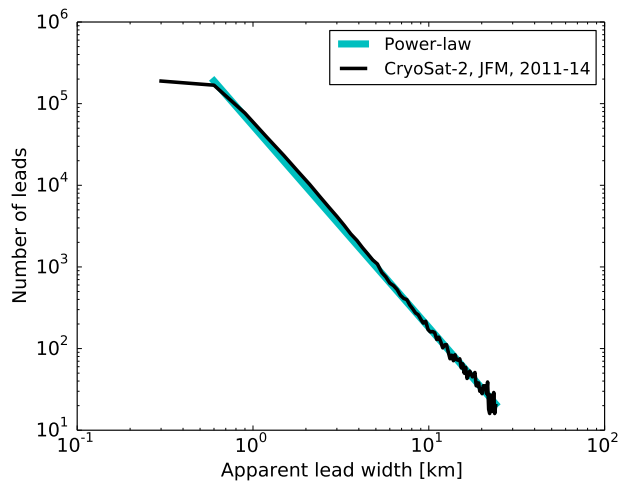


**Figure 8.** Lead fraction derived from CS-2 ~~SAR-mode~~ SAR mode (a) and from Röhrs et al. (2012) (b) on a North Polar Stereographic Grid with a resolution of  $99.5 \text{ km} \times 99.5 \text{ km}$ , merged from January to March 2011. Only values based on at least 2000 CS-2 measurements north of  $65^\circ \text{ N}$  (a) or with a grid cell data coverage of 10% or more (b) are shown. Missing CS-2 estimates north of Canada are caused by the use of ~~a different~~ the SARIn mode in the Wingham Box.



**Figure 9.** Lead fraction derived from CS-2 ~~SAR-mode~~ SAR mode on a North Polar Stereographic Grid with a resolution of  $99.5 \text{ km} \times 99.5 \text{ km}$  from February (a) and March 2013 (b). Only lead fraction values north of  $65^\circ \text{ N}$  based on at least 1000 measurements are shown. Missing estimates north of Canada are caused by the use of ~~a different~~ the SARIn mode in the Wingham Box.





**Figure 10.** Apparent lead width distribution from all CS-2 ~~SAR-mode~~ SAR mode ocean measurements north of  $65^\circ$  N in winter season (JFM) from 2011 to 2014. The distribution of a power-law with exponent of 2.47 is included for comparison, forming a straight line in a double logarithmic presentation. See text for definition of the apparent lead width.