Empirical Parametrization of Envisat Freeboard Retrieval of Arctic and Antarctic Sea Ice Based on CryoSat-2: Progress in the ESA Climate Change Initiative

Stephan Paul[1,2], Stefan Hendricks[1], Robert Ricker[1], Stefan Kern[3], and Eero Rinne[4]

[1]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany
[2]Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany
[3]Integrated Climate Data Center, Hamburg, Germany
[4]Finnish Meteorological Institute , Helsinki, Finland

Correspondence to: Stephan Paul (stephan.paul@awi.de)

Abstract. In order to derive long-term changes in sea-ice volume, a multi-decadal sea-ice thickness record is required. CryoSat-2 has showcased the potential of radar altimetry for sea-ice mass-balance estimation over the recent years. However, precursor altimetry missions such as Envisat have not been exploited to the same extent so far. Combining both missions to acquire a decadal sea-ice volume data set requires a method to overcome the discrepancies due to different foot-print sizes from either pulse-limited or beam-sharpened radar echoes. In this study, we implemented an inter-mission consistent surface-type classification scheme for both hemispheres, based on the waveform pulse peakiness, leading-edge width, and surface backscatter. In order to achieve a consistent retracking procedure, we adapted the Threshold First Maximum Retracker Algorithm, previously used only for CryoSat-2, to develop an adaptive retracker threshold that depends on waveform characteristics. With our method, we produce a global and consistent freeboard data set for CryoSat-2 and Envisat. This novel data set features a maximum monthly difference in the mission-overlap period of $2.2\,\mathrm{cm}$ ($2.7\,\mathrm{cm}$) for the Arctic (Antarctic) based on all gridded values with spatial resolution of $25\,\mathrm{km} \times 25\,\mathrm{km}$ and $50\,\mathrm{km} \times 50\,\mathrm{km}$ for the Arctic and Antarctic, respectively.

1 Introduction

The Arctic sea-ice extent has reduced over the last decades (e.g., Stroeve et al., 2012; Meier et al., 2014), while the Antarctic sea-ice extent has been slightly increasing (Parkinson and Cavalieri, 2012; Parkinson and DiGirolamo, 2016) but is subject to substantial inter-annual variation (e.g., Turner et al., 2017). Arctic sea ice is also thinning as observed by a variety of sensors such as upward-looking sonar measurements from submarines, aircraft measurements, as well as autonomous measurements (e.g., Rothrock et al., 1999; Meier et al., 2014; Lindsay and Schweiger, 2015). For the Antarctic, however, our knowledge about changes in the sea-ice thickness is much more limited than for the Arctic. Only few localized measurements are available from upward-looking sonars (e.g., Behrendt et al., 2013), drillings (e.g., Ozsoy-Cicek et al., 2013), as well as ship- and airborne measurements (e.g., Haas, 1998; Haas et al., 2008; Leuschen et al., 2008; Worby et al., 2008a). A different approach is the use of satellite laser altimetry utilizing the Ice, Cloud and land Elevation Satellite (ICESat; e.g., Farrell et al., 2009; Kwok and Rothrock, 2009). While this approach benefits from a very small sensor footprint, ICESat data is limited temporarily to

autumn and spring acquisition seasons as well as spatially through present cloud cover. It is widely accepted however that that measurement of sea ice thickness at circumpolar scales in both polar regions can be achieved only with satellite-altimetry (Laxon et al., 2003; Giles et al., 2007; Kwok et al., 2009; Kurtz and Markus, 2012; Laxon et al., 2013; Kern et al., 2016; Schwegmann et al., 2016; Ricker et al., 2017; Tilling et al., 2017).

5   The general methodology of retrieving sea-ice freeboard and sea-ice thickness using satellite radar altimetry is based on the pioneering work of Laxon (1994); Laxon et al. (2003) as well as Peacock and Laxon (2004). In a first step, the echo power waveforms are classified as returns from either sea-ice floes or returns from the sea surface of leads between sea-ice floes. These measurements are then converted into distance measurements that let one calculate the elevation difference of the snow surface or the sea-ice surface relative to the sea surface in the leads. Here, one can differentiate between the height
10 difference between the top of the snow surface and the sea surface (i.e., the total freeboard) and the height difference between the sea-ice surface and the sea surface (i.e., the sea-ice freeboard). When estimating sea-ice freeboard from radar altimeters, it is often assumed that the retrieved distance over sea ice using Ku-band radar always coincides with the snow/ice interface. However, this assumption is not true, especially for a highly stratified sea-ice snow cover and/or for multi-year sea-ice regimes (e.g., Armitage and Ridout, 2015). Therefore, the retrieved freeboard from the altimeter is often referred to as radar freeboard.
15 Additionally, a correction for the lower wave propagation speed in the sea-ice snow cover needs to be applied. The total sea-ice thickness can then be calculated from the sea-ice freeboard by assuming hydrostatic equilibrium (e.g., Ricker et al., 2014).

    The objective of the European Space Agency's (ESA) Climate Change Initiative for sea ice (SICCI) is to achieve a consistent sea-ice freeboard and sea-ice thickness climate data record (CDR) for both polar regions by combining radar altimetry data from all available missions with full error characterization and procedures based on existing algorithms. For our study we
20 use data from the Environmental Satellite (Envisat) as well as CryoSat-2. Envisat carries the pulse-limited Radar Altimeter 2 (RA-2) whereas CryoSat-2 on the other hand utilizes the along-track beam-sharpened Synthetic Aperture Interferometric Radar Altimeter (SIRAL).

    During the first phase of SICCI (SICCI-1), the focus was set on creating a processing scheme for Envisat data with the possibility to derive Arctic and Antarctic sea-ice freeboard and sea-ice thickness (Schwegmann et al., 2016). Here, the surface-
25 type classification was based solely on the use of a single classifier parameter to positively identify waveforms as either sea ice or leads from otherwise mixed waveform records. In general, this resulted in very few classified sea-ice-type waveforms and in turn comparably high lead fractions for the Antarctic (Schwegmann et al., 2016), but also for the Arctic. As a consequence of the very low amount of sea-ice-type classifications, only a very coarse resolution of $100\,\text{km} \times 100\,\text{km}$ could be realized for the gridded final data product due to otherwise insufficient coverage.

30 Furthermore, two different existing retracking schemes for Envisat were employed for lead-type and sea-ice-type waveforms. For lead-type waveforms, a retracker based on multiple fitting functions was used (Giles et al., 2007), whereas sea-ice-type waveforms were retracked by utilizing the standard offset-center-of-gravity (OCOG) retracker (Wingham et al., 1986). On the other side, studies such as Ricker et al. (2014) utilize multi-parameter threshold approaches for CryoSat-2 data to differentiate between lead-type and sea-ice-type waveforms and employ a threshold first maximum retracker algorithm (TFMRA; Helm
35 et al., 2014; Ricker et al., 2014) to both. Inconsistencies were also present in the use of differing auxiliary data sets for sea-ice

concentration, as well as snow and sea-ice type information. Additionally, different sensor configurations result in varying instrument footprints with associated discrepancies in the degree of surface-type mixing.

In this study, we focus on deriving an inter-mission consistent waveform interpretation scheme over sea-ice areas for Envisat and CryoSat-2 in the framework of the second phase of SICCI (SICCI-2). Therefore, the focus of this study lies not in a further optimization of the CryoSat-2 freeboard retrieval, but in the application of an evaluated methodology as is (Ricker et al., 2014). Based on this approach, we want to find an optimal way to match the freeboard retrieval of Envisat to that of CryoSat-2 and build a consistent sea-ice freeboard data record that takes the different sensor configurations and differing footprints between both sensors into account. We have developed an empirical approach to minimize inter-mission biases in the surface-type classification as well as in the range retracking and subsequent freeboard retrieval based on CryoSat-2 reference data for the mission-overlap period (MOP) from November 2010 to March 2012. The resulting parametrization takes into account differences between sea-ice surface properties in both hemisphere as well as the seasonal cycle. In this study we focus on the derivation of freeboard since the conversion from freeboard to thickness is identical for both missions and relies on additional auxiliary data sets.

In the following sections we describe the derivation of a mutual threshold-based surface-type classification from a mix of unsupervised clustering and supervised classification. Additionally, the derivation and application of a waveform-parameter dependent adaptive threshold retracker scheme for the Envisat freeboard retrieval is presented. Resulting data sets and key benchmarks from Envisat and CryoSat-2 for the MOP are then presented and discussed.

2   Data and Methods

This section gives an overview about the input data used and necessary pre-processing and filtering steps. Moreover, we describe the inter-mission consistent surface-type classification and range retracking scheme.

2.1   Input data

2.1.1   Altimetry data

For our study, we use geolocated level 1b (L1b) data for both CryoSat-2 and Envisat. In case of CryoSat-2, we make use of all available SIRAL Baseline-C data acquired in synthetic aperture radar mode (SAR) as well as in the SAR interferometric (SIN) mode. However, the specific interferometric information is not used during the processing. No further filtering based on quality control is conducted.

For Envisat, we use version 2.1 of the sensor geophysical data record (SGDR). All data is provided by the ESA. Here, we investigated the measurement-confidence data flags in the SGDR for problematic records. All data with 'Packet Length Error' (Flag 0), invalid OnBoard Data Handling (Flag 1), an Automatic Gain Control fault (Flag 4), a Rx Delay Fault (Flag 5) or an Waveform Fault (Flag 6) raised are removed from processing.

In the second step, all data is filtered regionally for both hemisphere by latitudinal boundaries to areas where sea ice is present. Data are only considered if located north of 60 °N for the Arctic and south of 50 °S for the Antarctic.

Finally, all processing for both sensors is limited to waveforms flagged as ocean.

2.1.2 Auxiliary data

5   For our surface-type classification, as well as for the conversion of elevations to sea-ice freeboard, we utilize a range of different auxiliary data sets. Our objective is to consequently maintain methodological as well as auxiliary data consistency. This is especially important for a multi-mission climate data record.

In this study, for both hemispheres we use the sea-ice concentration data obtained from the Ocean and Sea Ice Satellite Application Facility (OSISAF; ftp://osisaf.met.no/reprocessed/ice/conc/v1p2) as well as the mean sea-surface height product

10   provided by the Danish Technical University (DTU; Andersen et al., 2016; ftp://ftp.spacecenter.dk/pub/DTU15/) in its 2015 version. Sea-ice concentration data is used mainly to discard waveforms based on a minimum required sea-ice concentration threshold of 5 %, whereas the mean sea-surface height data is utilized to eliminate undulations due to the geoid before retrieving sea-ice freeboard (Ricker et al., 2014). We use the same sea-ice concentration and mean sea-surface height data for both hemispheres.

15   Additionally, information about the sea-ice snow cover are required. These are necessary for the range correction due to the lower wave-propagation speed in the snow pack. For the Arctic, we use the Warren snow climatology (Warren et al., 1999). As the Warren climatology is based on data sets obtained from Arctic drift stations primarily on multi-year sea ice (MYI), snow-depth values are suspected to be biased high in first-year sea-ice (FYI) regime. Therefore, we apply a correction to the Warren climatology over FYI (Kurtz and Farrell, 2011). As a consequence, the correction is a linear proportional reduction of the origi-

20   nal snow depth with the present FYI fraction down to 50 % of its original value over sole FYI. In order to discriminate between FYI and MYI in the Arctic, we use a MYI fraction data set based on the Special Sensor Microwave Imager (SSM/I)/Special Sensor Microwave Imager Sounder (SSMIS) sensors on-board of the Defense Meteorological Satellite Program (DMSP) satellites provided by the Integrated Climate Data Center (ICDC). This MYI fraction data set is tailored to be consistent for the entire ERS 1/2 - Envisat - CryoSat-2 period. It is based on NSIDC daily gridded 25 km grid resolution brightness temperatures

25   (Maslanik and Stroeve, 2004, updated 2017) of DMSP-f11, DMSP-f13 and DMSP-f17, inter-sensor calibrated to the level of DMSP-f17 SSMIS measurements. The MYI fraction is computed using the NASA-Team algorithm (Cavalieri et al., 1999) with monthly MYI and FYI tie points computed from inter-sensor calibrated brightness temperatures using a gradient-ratio (at 37 and 19 GHz vertical polarization) threshold approach following an idea formulated in Comiso (2012); monthly open water tie points are computed from the same data set from grid cells with the 2 % lowest brightness temperatures over open water.

30   The resulting MYI area computed from the obtained MYI fraction data set agrees well with the results of Comiso (2012) and Kwok and Cunningham (2015). More information is given in Kern (technical manual, 2016).

For the Antarctic, we assume only a single sea-ice type being present. As the Warren climatology is only available for the Arctic, we use a snow-depth climatology derived from data acquired by the Advanced Microwave Scanning Radiometer-EOS (AMSR-E) and AMSR-2 aboard GCOM-W1 for the Antarctic (Kern et al., 2015; data access via: http://icdc.cen.unihamburg.de/

projekte/esa-cci-sea-ice-ecv0.html). This data set is based on a revised version of the approach described by Markus and Cavalieri (1998) and Markus et al. (2011). Daily snow depths of 13 full seasonal cycles (August through July of years 2002/03 through 2010/11 and years 2012/13 through 2015/16) are used to compute a daily Antarctic snow depth on sea ice climatology. Note that even though this climatology is based on snow depth data derived with a version of the original empirical algorithm which is now developed directly from AMSR-E brightness temperatures (Frost et al., 2015), the limitations of the algorithm in terms of snow depth under-estimation over deformed sea ice and snow depth sensitivity to snow properties such as wetness (e.g., Worby et al., 2008b; Kern and Ozsoy-Çiçek, 2016) essentially remain the same.

2.2 Surface-type classification

2.2.1 Importance and general issues

The surface-type classification is a crucial part in the processing chain, because the detection of leads is essential for determining the instantaneous sea-surface height anomaly with respect to the mean sea-surface height at the ice-floe location. The resulting sea-surface height at the ice-floe location in turn is used as the reference from which the sea-ice freeboard is calculated. Moreover, a clear distinction between leads and sea ice improves the quality and accuracy of resulting sea-ice freeboard estimates. Ambiguous signals are excluded from the freeboard retrieval.

In general, leads feature a specular reflection due to their rather smooth surface, whereas sea ice features a diffuse reflection due to a higher surface roughness. With smaller instrument footprint sizes, less surface-type mixing occurs and the return signal is easier to classify. However, leads often dominate acquired waveforms due to their specular reflection. Off-nadir leads still represent sources of strong backscatter and therefore result in false range estimates. In case of Envisat, the nominal circular footprint is 2 km in diameter (Connor et al., 2009). Despite its much smaller footprint (1.65 km×0.30 km), CryoSat-2 can also be affected by off-nadir leads, which will result in erroneous freeboard estimates (Armitage and Davidson, 2014).

In contrast to the work conducted during SICCI-1, where a single threshold classification scheme for Envisat was used alongside a multi-parameter classification scheme for CryoSat-2, we aim for a inter-mission consistent surface-type classification scheme for Envisat and CryoSat-2. Therefore, a set of classifier parameters that is available for both sensors is necessary. Here, we use the surface backscatter, the leading-edge width, and the pulse peakiness as classifier parameters to identify lead-type and sea-ice-type waveforms from mixed- or ambiguous-type waveforms.

We define pulse peakiness (pp) slightly different as compared to the one used by Laxon et al. (2003). Ours follows the definition of Ricker et al. (2014), where N_{wf} is the number of range bins, wf_i is the echo power at range bin i of the waveform, and $max(wf)$ is the maximum echo power in the given waveform:

$$pp = \sum_{i=1}^{N_{wf}} \frac{max(wf)}{wf_i} \times N_{wf} \tag{1}$$

The leading-edge width is defined as the width in range bins along the power rise to the first local maximum between 5 % and 95 % of the first-maximum peak power while using a ten-times oversampled waveform.

The choice for using three classifier parameters in SICCI-2 also allows for less strict thresholds compared to the previously used single threshold parameter classification for Envisat during SICCI-1.

Over the course of a winter season, ice conditions can change substantially. Similar to leads, young- and thin-ice areas cause specular reflections compared to other ice types (Zygmuntowska et al., 2013). Furthermore, the amount of leads varies both seasonally and regionally. Based on fixed thresholds for a whole winter season, these changes are difficult to capture and the rejection rate is increased unnecessarily. Hence, we decided on using monthly thresholds to improve the overall results and data quality.

There is a general lack of ground-truth data as collocated measurements of the same sea-ice situation are very difficult due to sea-ice drift. However, received waveforms have very distinct characteristics and are well described in literature (e.g., Ricker et al., 2014; Schwegmann et al., 2016). These characteristics can also be deduced from the chosen set of classifier parameters. In order to overcome the lack of ground-truth, we decided to use a combination of unsupervised clustering and supervised classification.

Based on this combination, we are able to determine suitable thresholds for data acquired by Envisat as well as CryoSat-2. The work-flow of how we derived the surface-type thresholds is summarized in Figure 1 and described thoroughly in the following subsections.

2.2.2 Monthly classifier parameters and k-means clustering

In a first step, the three classifier parameters of surface backscatter, pulse peakiness, and leading-edge width are computed for all available L1b data per sensor and month in the mission overlap period (MOP) from November 2010 to March 2012.

We only use waveforms that are located between $70\,°N$ and $81.5\,°N$ for the Arctic and feature a minimum sea-ice concentration of $70\,\%$. The northern limit of 81.5°N was chosen to assure a maximum of consistency between Envisat and CryoSat-2 with their differing orbital parameters. Until an update to the geographic mode mask in July 2014, CryoSat-2 operated in SIN mode in an area between 80–85 °N and 100–140 °W (referred to as "Wingham Box"). For this area, as well as all other Arctic areas that are covered while CryoSat-2 operates in SIN mode, we use all waveforms acquired north of $70\,°N$. For the Antarctic the same restrictions apply, but waveforms are geographically limited to an area south of $65\,°S$ to exclude the majority of the marginal-ice zone to reduce the impact of ocean swell.

Next, a subset of $1\,\%$ is sampled at random without replacement (i.e., each original waveform with corresponding surface backscatter, pulse peakiness, and leading-edge width can only appear once) for each month in the MOP and for each sensor independently. This data sample is then separated into three clusters using an unsupervised methodology named k-means clustering (MacQueen, 1967; Hartigan and Wong, 1979). This unsupervised method (i.e., without any a-priori information about the data) is widely used to separate input data of N observations into K clusters of equal variance. In our case, based on the input classifier parameters of surface backscatter, pulse peakiness, and leading-edge width, whereby the within-cluster sum-of-squares are iteratively minimized (MacQueen, 1967; Hartigan and Wong, 1979). The result is a 'labeled' data set where each input waveform with corresponding surface backscatter, pulse peakiness, and leading-edge width is labeled as an either sea-ice-type, lead-type, or ambiguous-type waveform.
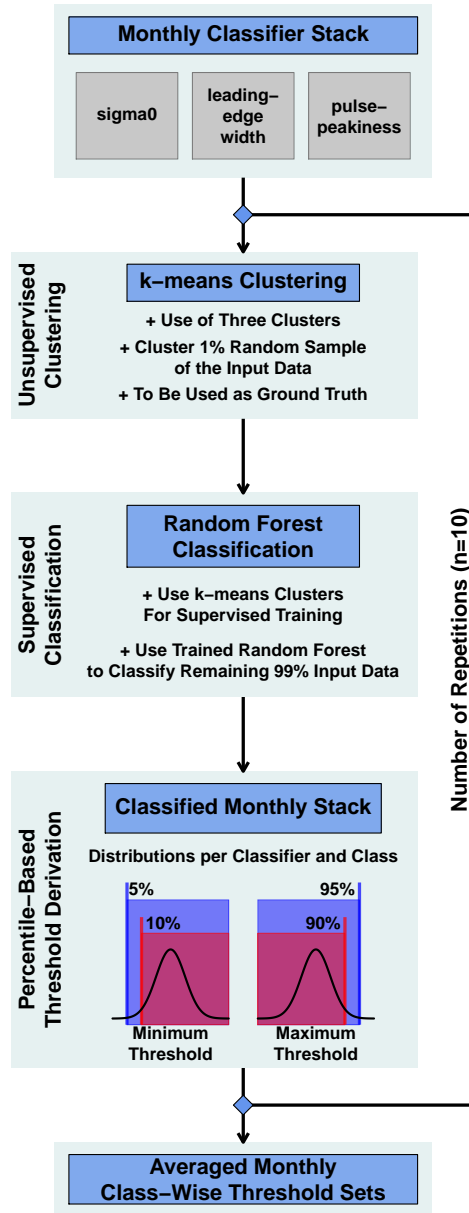
Figure 1. Flowchart visualizing the important sub steps of unsupervised clustering and supervised classification in order to derive the new surface-type thresholds from monthly stacks of surface backscatter, leading-edge width and pulse peakiness.

Generally, the preselection of the number of clusters can be a problem when utilizing k-means clustering. However, while we tested a higher number of initial clusters with perspective of later reunion of similar clusters, a separation into just three clusters turned out to be sufficient. Overall, lead waveforms account for a smaller fraction of the total measurements than

sea-ice waveforms. Because of this and the fact that k-means clustering tends towards generating equal-size clusters (this is a presumption of k-means clustering algorithms), sole use of k-means clustering for the complete data set was not feasible.

2.2.3 Random forest classification

As k-means clustering can not be used for classification of the complete data set due to its unevenly distributed nature, the initially clustered 1% data sample is instead used as a-priori information (i.e., a training data set of classified waveforms as either sea ice, leads, or ambiguous) for a supervised classification. In our case, we use an ensemble supervised machine-learning method called random forest (Breiman, 2001).

Random forests are based on multiple decision trees. A decision tree is a rather simple statistical tool to predict data categories based on thresholds. Over several steps, the input data set is split at each step (called a 'node') based on a threshold of a given parameter until all input data is categorized. When visualized, a decision tree resembles a tree with an increasing numbers of branches, leading to the final categories. (Breiman, 2001).

The procedure of fitting all single decision trees to the random sub-samples is called training. During this training, each decision tree in the random forest is grown following certain rules (Breiman, 2001):

- First, from the training data of size N, N cases are sampled randomly with replacement as a 'new' and specific training data set for each single tree. This means that the resulting 'new' training data set for each tree has the same size as the input training data, but any single waveform with corresponding surface backscatter, pulse peakiness, and leading-edge width can appear multiple times (i.e., "with replacement").

- Second, for M input parameters (in our case surface backscatter, pulse peakiness, and leading-edge width), Breiman (2001) states that ideally a fixed number $m \ll M$ of the given input parameters is specified and randomly selected out of M. The best split on these selected parameters m is then used to split the node. Throughout the growing of the forest, the value of m is held constant.

- Third, each tree is grown out fully, i.e., to its largest possible extent. No pruning is applied. In contrast to single decision trees that tend to overfit (i.e., match data too precisely and therefore fail for any additional data), random forests do not overfit and are also capable of dealing with unbalanced data sets (Breiman, 2001).

For our purpose, we always grow a total number of 500 decision trees per training in each month. Due to the small number of input parameters ($M = 3$), we set m to one, following the suggestion by Breiman (2001) to approximate m by $\sqrt{M}$.

The result is an ensemble of 500 single uncorrelated decision-tree classifiers called a random forest (Breiman, 2001). After initial training, the random forest can be used for classification of the remaining 99% of the initial monthly data. During this classification, each decision tree in the now trained random forest categorizes each waveform based on surface backscatter, pulse peakiness, and leading-edge width into either a sea-ice type, a lead-type, or an ambiguous-type waveform. In the end, the majority of all decision trees in the random forest decides the resulting class.

Available data from months that are covered twice during the mission-overlap period are merged for the random-forest training. The trained random forest for each month is then used to classify the remaining 99 % of the corresponding monthly data. From this classified data set, distributions for each of the three classifier parameters for each month in the mission-overlap period are obtained. These distributions feature clear distinctions for each surface-type class (leads, sea ice, and ambiguous).

5 For example, leads feature in general high values in surface backscatter and pulse peakiness as well as shorter leading-edge widths. The opposite can be seen for sea ice. The class of ambiguous signals is placed in between.

2.2.4 Percentile-based averaged thresholds

Thresholds are then obtained from the resulting classifier-parameter distributions by using either the 5 % or 10 % percentile for a minimum threshold, or the 90 % or 95 % percentile in case of a maximum threshold (Figure 1). The exact numbers

10 were chosen arbitrarily after visual screening of all resulting classifier-parameter distributions to eliminate outliers. The choice of using the more strict (10 %/90 %) or less strict (5 %/95 %) percentile thresholds depends on the sensor. Due to its larger footprint and therefore an expected higher degree of surface-type mixing, we chose the more strict thresholds for Envisat, and the less strict thresholds for CryoSat-2 due to its smaller footprint.

The decision, whether to derive a minimum or maximum threshold depends on the surface-type class. For example, lead-type

15 waveforms are generally characterized by high values in pulse-peakiness as well as surface backscatter due to their specular reflection. Lead-type waveforms therefore feature a very steep increase in echo power which results in short leading-edge widths. Hence, the 5 %/10 % percentiles of the surface backscatter and pulse-peakiness distributions would be used alongside the 90 %/95 % percentile of the leading-edge-width distribution. Sea-ice type waveforms on the other hand should have smaller values in pulse peakiness. Due to their rather diffuse reflection, sea-ice type waveforms also feature low backscatter and a less

20 steep increase in echo power, which results in longer leading-edge widths. As a result, the 90 %/95 % percentiles are used for the surface backscatter and pulse-peakiness distributions. For the leading-edge width, the 5 %/10 % percentile of its distribution is used. As we positively identify sea-ice and lead waveforms from all available measurements using these thresholds, all remaining waveforms are classified as ambiguous.

Additionally, for all classifications of leads as well as sea ice in both hemispheres we set a minimum requirement of 70 %

25 sea-ice concentration.

The whole procedure, starting with randomly sampling 1 % from the initial monthly stack, is then repeated ten times. As the whole procedure is initially based on random sampling, this repetition is done to compensate for the odd case of an insufficient representation of lead or sea-ice waveforms in the sampled data. In a last step, the average minimum/maximum thresholds for each classifier parameter, surface-type class, and month in the MOP are estimated for each sensor. These thresholds are

30 summarized in Tables A1 through A6 in the appendix.

2.3 Range retracking and freeboard retrieval

The range-retracking algorithm for Envisat and CryoSat-2 waveforms is identical for sea-ice-type and lead-type waveforms. The used Threshold First Maximum Retracker Algorithm (TFMRA, Helm et al., 2014; Ricker et al., 2014) is based on the following steps:

- Either estimating the noise level as the average of the first five bins of the waveform (CryoSat-2) or discarding all counts in the first five bins of the waveform as these just contain artifacts of the fast Fourier transformation (Envisat);

- Oversampling of the waveforms by a factor of 10 using linear interpolation;

- Smoothing of the oversampled waveforms with a running-mean window-filter size of 11 (Envisat, CryoSat-2 SAR) or 21 (CryoSat-2 SIN) range bins respectively;

- Locating the first local maximum of the waveform. This maximum has to be higher than the noise level by 15 % of the absolute peak power; and finally,

- Obtaining the range value (i.e., the elevation) at a specified percentage threshold of the power at the detected first maximum, by linear interpolation of the smoothed and oversampled waveform.

The conversion from range estimates into sea-ice freeboard follows by subtracting the interpolated sea-surface height (the

15 sum of mean sea-surface height taken from the DTU2015 product and the instantaneous sea-surface height anomaly estimated from the interpolated elevation between present leads) at the floe location from the elevation of the sea-ice floe. Given a wave-propagation speed correction based on the auxiliary snow depth data, sea-ice freeboard can be calculated. A thorough description on the calculation of sea-ice freeboard (and also sea-ice thickness which is not part of this study) is described in Ricker et al. (2014).

20 Continuing on the last point of the general TFMRA retracking procedure, the choice of retracker threshold is pivotal for the range estimation. Following the AWI's implementation for CryoSat-2 (Ricker et al., 2014), we use a threshold of 50 % from the first maximum peak power both for lead-type and sea-ice-type waveforms. For pulse-limited altimetry such as Envisat, retracking near the maximum power for leads proved to be essential to retrieve reasonable sea-ice freeboard estimates (e.g., Giles et al., 2007). Hence, we chose a threshold of 95 % for leads from Envisat waveforms. In a very recent study by Guerreiro

25 et al. (2017), the use of a 50 % for lead-type wave forms resulted in initial average conditions where the lead surface elevation was detected above that of the surrounding sea-ice floes. When we used a single fixed threshold for the range retrieval over sea ice similar to that for CryoSat-2 (i.e., a 50 % threshold from the first local maximum peak power), our Envisat sea-ice-freeboard estimates featured an overall smaller variation and range compared to CryoSat-2 estimates. We relate this behavior to the much larger footprint and the therefore increased mixing of surface types of different surface-roughness scales in every obtained

30 Envisat waveform.

We used our methodology as illustrated in the previous paragraphs and in Figure 1 to compute the sea-ice freeboard for every month of the MOP separately for Envisat and CryoSat-2. Subsequently, we computed the sea-ice freeboard difference
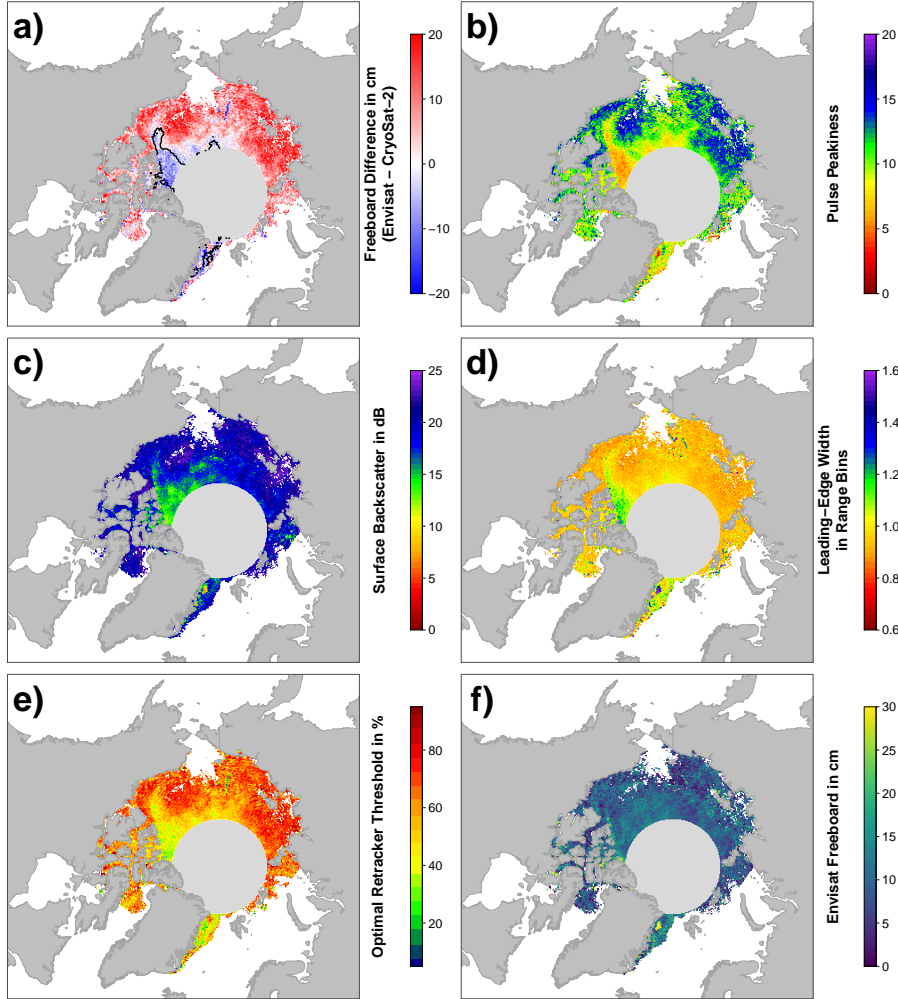
Figure 2. Exemplary visualizations of freeboard differences between Envisat and CryoSat-2 (a; in cm; the black isoline resembles the boundary of 75 % multi-year-ice fraction), Envisat pulse peakiness (b; unitless), Envisat surface backscatter (c; in dB), Envisat leading-edge width (d; in range bins), optimal retracker threshold (e; in %), and the resulting Envisat freeboard using our adaptive retracker procedure (f; in cm) for the Arctic in November 2011.

Envisat minus CrysoSat-2, which is shown together with parameters of the retrieval in Figures 2 and 3. From Figures 2a and 3a it appears that there are substantial differences in the resulting sea-ice freeboard between both sensors. However, these patterns of sea-ice freeboard differences are related to differences in the Envisat waveform parameters of pulse peakiness, surface backscatter, and leading-edge width (Figures 2b-d/3b-d). These waveform parameter variations in turn reflect changes in the surface properties. Areas of MYI near the Canadian Archipelago and areas influenced by MYI export are in general substantially thinner for Envisat that CryoSat-2 (e.g., about 20 cm and more in March, Figure 3a). On the other side, areas
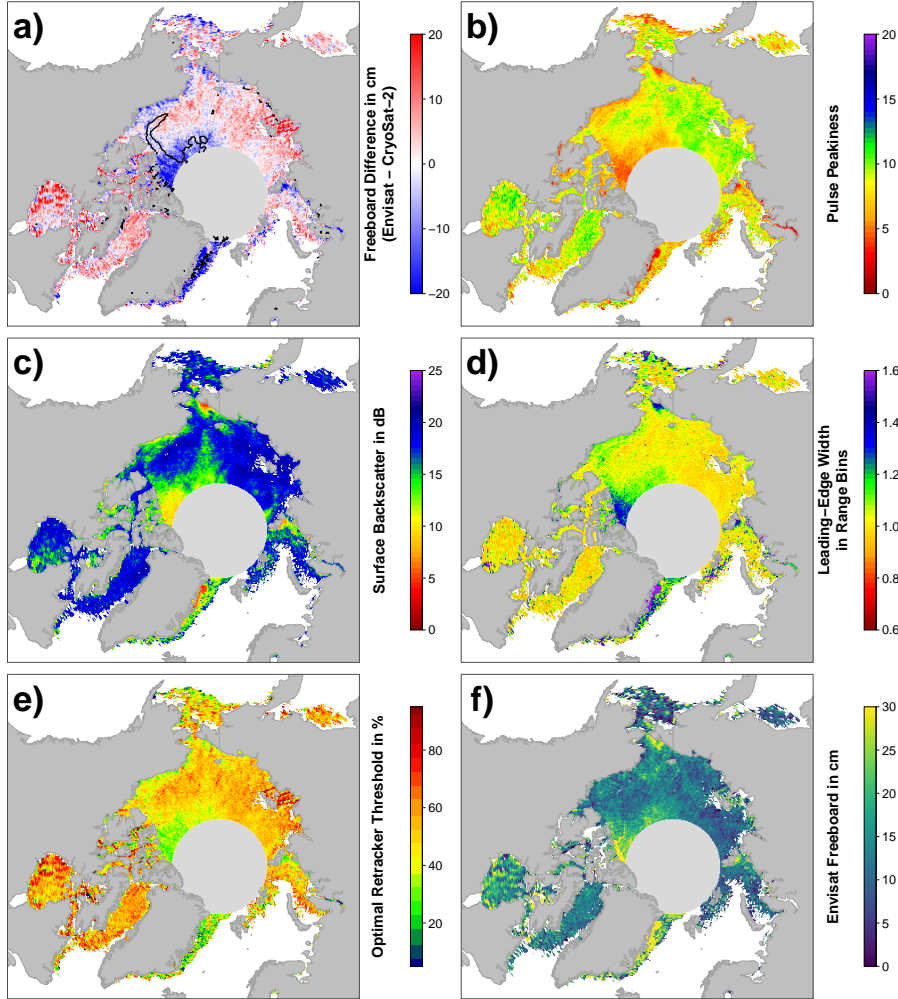
Figure 3. Exemplary visualizations for the Arctic in March 2012 in the same setup as Figure 2.

of predominantly FYI are in general thicker in the Envisat data (Figure 2a). However, the level of freeboard difference is not constant throughout a winter season but rather appears to be seasonal, where Envisat appears to be unable to keep track of these seasonal changes.

As these differences in sea-ice freeboard between CryoSat-2 and Envisat appear to be indeed strongly correlated to pat-
5 terns in the surface backscatter and the leading-edge width of Envisat waveforms (Figures 2c-d/3c-d), we decided to apply a novel empirical tuning scheme by computing an adaptive range-retracker threshold as a function of surface backscatter and the leading-edge width to mitigate the differences. Due to the already mentioned larger footprint of Envisat and hence increased mixing of different surface types, it appears to be necessary to treat waveforms differently according to the wave-form shape (and hence surface properties) by means of retracking the main scattering horizon. Guerreiro et al. (2017) proposed

in their study a correction scheme deriving a relationship between monthly-gridded pulse peakiness and the monthly-gridded freeboard differences between CryoSat-2 and Envisat based on a third order polynomial fit. In contrast to applying a similar post-retracking correction to the resulting freeboard estimates, we apply our correction already during waveform retracking.

In order to derive a functional relationship between retracker threshold and surface backscatter/leading-edge width, we first processed all available Envisat data for the complete MOP. This processing was done using the TFMRA with a fixed threshold for leads of 95 % and a threshold for sea-ice-type waveforms that was changed in each run. This sea-ice threshold ranged between 5 % and 95 % in steps of 5 %. For example, in the first run the complete data set was processed using a retracker threshold of 5 % for sea-ice-type waveforms and the resulting sea-ice freeboard was calculated. In the next run, a fixed threshold of 10 % was used for all sea-ice-type waveforms and so on. This continued until the last run was computed with a retracker threshold of 95 % for sea-ice-type waveforms and the resulting sea-ice freeboard was calculated.

From this data set, the optimal threshold, i.e., the threshold that yields the smallest absolute difference in sea-ice freeboard between Envisat and CryoSat-2, was iteratively derived. Exemplary results are shown in Figures 2e and 3e. Again, the seasonal change observed in the waveform parameters is also reflected in the resulting optimal threshold values. These show a varying range of optimal-threshold values that are in general higher for the early winter compared to the late winter.

In the next step, we derive a functional relationship between optimal-threshold values and the waveform parameters of surface backscatter/leading-edge width for our adaptive threshold range-retracking. Therefore, we first average all optimal-threshold values during the mission-overlap period (MOP) for bins of 0.25 dB for the surface backscatter and 0.025 for the leading-edge width, respectively. Here, we use a three dimensional coordinate system with average optimal threshold (z-axis) against leading-edge width (x-axis) and surface backscatter (y-axis).

The months November through March are covered twice during the MOP and both occurrences were used. October and April, which were only covered once during the MOP, were each added twice to circumvent issues of under-representation in their number of data values added to the total.

Through this compilation of monthly data points, three third-order polynomial planes were fitted based on different weighting schemes in order to maximize the adjusted coefficient of determination (R^2_{adj}). R^2_{adj} is a measure for the quality of the model fit. In contrast to the normal R^2, R^2_{adj} decreases through adding useless predictors to a model and is therefore a more robust measure for model quality than the standard R^2. As weights, we either used the number of optimal-threshold values per bin in the x-y plane, the inverse standard deviation of all optimal threshold values per bin ($1/\sigma$), or no weights at all.

The optimal threshold (th_{opt}; in decimal values) to be used in the adaptive range retracking for the Arctic as a function of surface backscatter (σ^o) and leading-edge width (lew) is given by Equation 2:

$$
\begin{aligned}
th_{opt} = {} & 3.4775697362 \\
& - 5.9296875486 \times lew + 4.3516498381 \times lew^2 \\
& - 1.0933131955 \times lew^3 - 0.0914747272 \times \sigma^o \\
& + 0.0063983796 \times \sigma^{o2} - 0.0001237455 \times \sigma^{o3}
\end{aligned} \tag{2}
$$
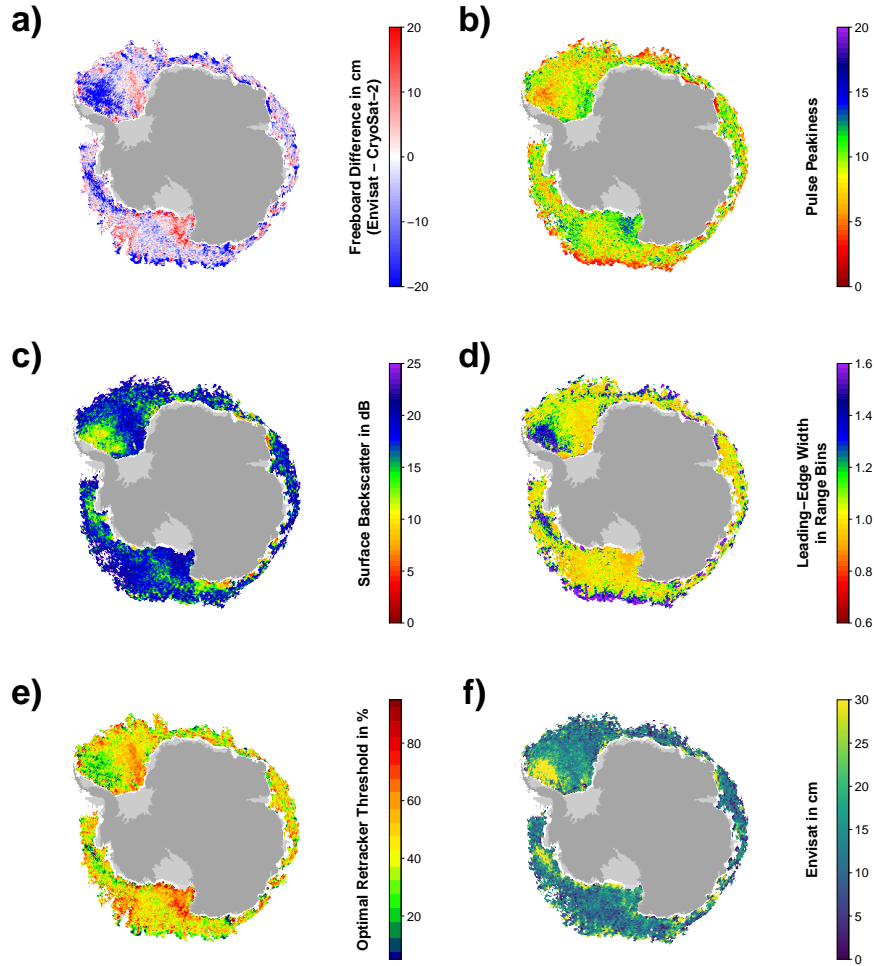
Figure 4. Exemplary visualizations for the Antarctic in May 2011 in the same setup as Figure 2.

For the Arctic, Equation 2 achieved the highest adjusted R^2_{adj} of 0.94 with the inverse standard deviation as weights. All used data points have a minimum of 50 occurrences to reduce noise and were obtained in the central Arctic only (i.e., we excluded the Canadian Arctic Archipelago and the Hudson Bay, but also extensive fast-ice areas like the Laptev Sea). The resulting monthly-gridded Envisat freeboard estimates are shown in Figures 2f and 3f.

5 In a first attempt, we applied Equation 2 also to the southern hemisphere. However, this did not result in an improvement of the freeboard differences between Envisat and CryoSat-2. The reason for that can partly be seen in Figures 4 and 5.

In contrast to the Arctic (Figures 2 and 3), there is less seasonality in the data, i.e., the differences between early and late winter are less prominent in the sea-ice freeboard differences as well as the optimal-threshold values. Overall, a wider range of
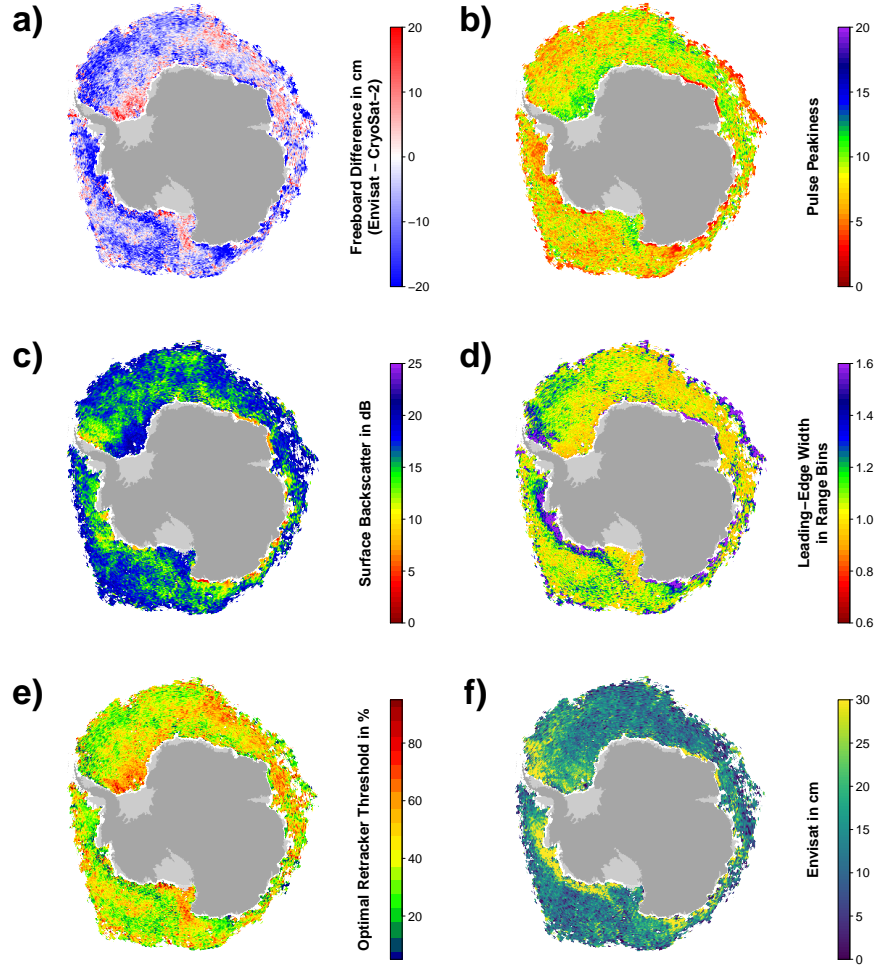
Figure 5. Exemplary visualizations for the Antarctic in September 2011 in the same setup as Figure 2.

optimal-threshold values is necessary at any given month, in order to achieve a minimum freeboard difference (Figures 4e and 5e).

Additionally, the overall range and distribution of the surface backscatter is different between the Arctic and the Antarctic (not shown) as well as patterns in surface backscatter and leading-edge width are less correlated in some areas (e.g., the MIZ but also in the central Weddell Sea; Figures 4c-d and 5c-d). This is potentially related to ice-snow interface flooding paired with subsequent refreezing and formation of snow ice, large fast ice areas with a different snow stratigraphy and depth, and a different ice-growth history than in the Arctic causing a larger fraction of rough and deformed sea ice.

For the Antarctic, a second-order polynomial fit resulted in the best statistical result (R^2_{adj} of 0.77) to describe the optimal threshold as a function of leading-edge width and surface backscatter. Equation 3 summarizes the relationship for deriving the

optimal threshold (th_{opt}; in decimal values) in the Antarctic as a function of surface backscatter (σ^o) and leading-edge width (lew):

$$
\begin{aligned}
th_{opt} = {} & 0.8147895184 \\
& - 0.5555823623 \times lew + 0.1347526920 \times lew^2 \\
& + 0.0055934198 \times \sigma^o - 0.0001431595 \times \sigma^{o2}
\end{aligned}
\tag{3}
$$

Here, the best fit is obtained using the total number of optimal-threshold values per bin as weights. All used data points also have a minimum of 50 occurrences and were obtained by excluding ice zones around the Antarctic that appear to be influenced by ocean swell (identified by surface backscatter and/or leading edge with outlier artifacts) as well as the months from December through April.

Utilizing both equations, for each range-retracking of every sea-ice waveform, the to-be-used threshold is calculated from the waveform-associated surface backscatter and leading-edge width value. This threshold is then believed to yield the mean-scattering surface in best accordance to the CryoSat-2 measurements. The resulting monthly-gridded Envisat freeboard estimates are shown in Figures 4f and 5f.

3 Results and Discussion

In this section we want to present and discuss the results obtained for the mission overlap period (MOP) between Envisat and CryoSat-2. This is presented first for the surface-type classification and then for the range retracking and the associated sea-ice freeboard retrieval.

3.1 Surface-type classification

Utilizing our surface-type classification scheme results in an overall much better agreement between CryoSat-2 and Envisat based on lead, sea-ice, and valid fractions (Figures 6 – 9). Compared to the surface-type classification used during SICCI-1 for Envisat, our approach is less strict and allows for substantially more waveforms being classified as either lead or sea-ice type that were rejected before. Additionally, a very high fraction of lead detections was present, compared to only a very small fraction of classified sea-ice type waveforms during SICCI-1 (Schwegmann et al., 2016). Furthermore, the inter-mission consistency of the surface-type classification for the Arctic as well as the Antarctic has improved substantially (Figures 8 and 9).

The increased number of valid waveforms has an additional positive side effect on the overall data record: It allows for a much finer grid resolution to be used in the final Level 3 product without any concessions on overall coverage. Here, we are now able to provide a $25\,\mathrm{km} \times 25\,\mathrm{km}$ ($50\,\mathrm{km} \times 50\,\mathrm{km}$) resolution gridded data set for the Arctic (Antarctic) compared to the $100\,\mathrm{km} \times 100\,\mathrm{km}$ during SICCI-1.
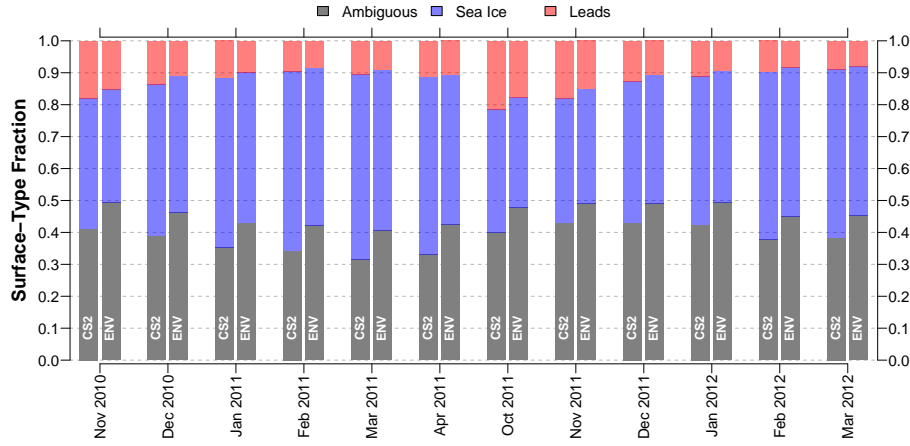
Figure 6. Time-series of surface-type fractions for the mission overlap period between CryoSat-2 (CS2) and Envisat (ENV) for the Arctic based on orbit-track (i.e., non gridded) data.
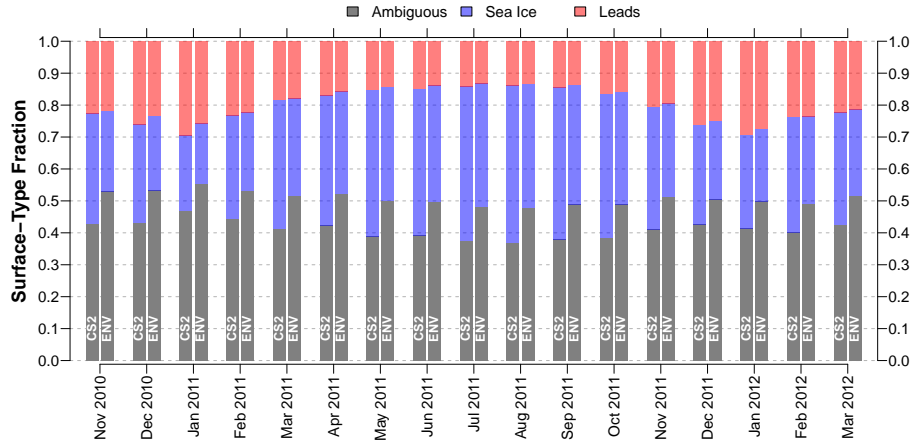
Figure 7. Time-series of surface-type fractions for the mission overlap period between CryoSat-2 (CS2) and Envisat (ENV) for the Antarctic based on orbit-track (i.e., non gridded) data.

Direct comparisons of surface-type class fractions (i.e., either ambiguous, lead, or sea-ice type) over the course of the MOP reveal an overall good agreement between CryoSat-2 and Envisat based on the non-gridded orbit data (Figures 6 and 7). While the fractions of lead and sea-ice waveforms are on average slightly smaller for Envisat compared to CryoSat-2 (about 8 % for the Arctic and 10 % for the Antarctic), both sensors show a similar seasonal development in both hemispheres. Especially, the fraction differences of detected leads are very small with a root-mean-squared difference (RMSD) of 2 % and 1 % for the Arctic and Antarctic, respectively. The discrepancy is larger for detected sea-ice waveforms with a RMSD of 6 % for the Arctic and 9 % for the Antarctic.
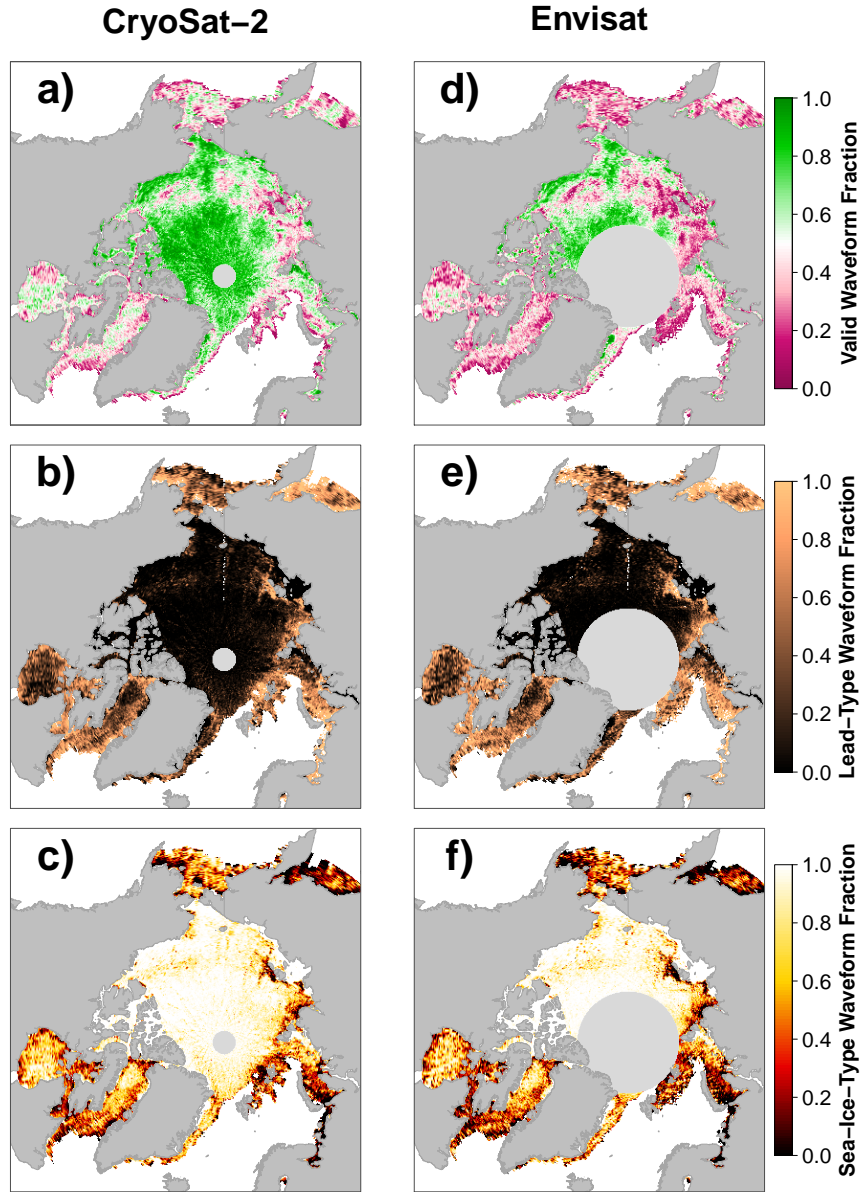
Figure 8. Comparison of gridded surface-type classification benchmarks of valid waveform fractions (a/d; ratio of either lead-type/sea-ice-type classification to the total number of waveforms per grid cell), lead-type waveform fraction (b/e; ratio of lead-type classifications to the number of valid classifications per grid cell), and the sea-ice-type waveform fraction (c/f; ratio of sea-ice-type classifications to the number of valid classifications per grid cell) between CryoSat-2 and Envisat for March 2012 in the Arctic.

Exemplary visualizations of monthly gridded inter-comparisons between Envisat and CryoSat-2 based on valid, lead, and sea-ice waveform fractions are shown in Figures 8 and 9. In these gridded data sets, the overall good agreement is confirmed.
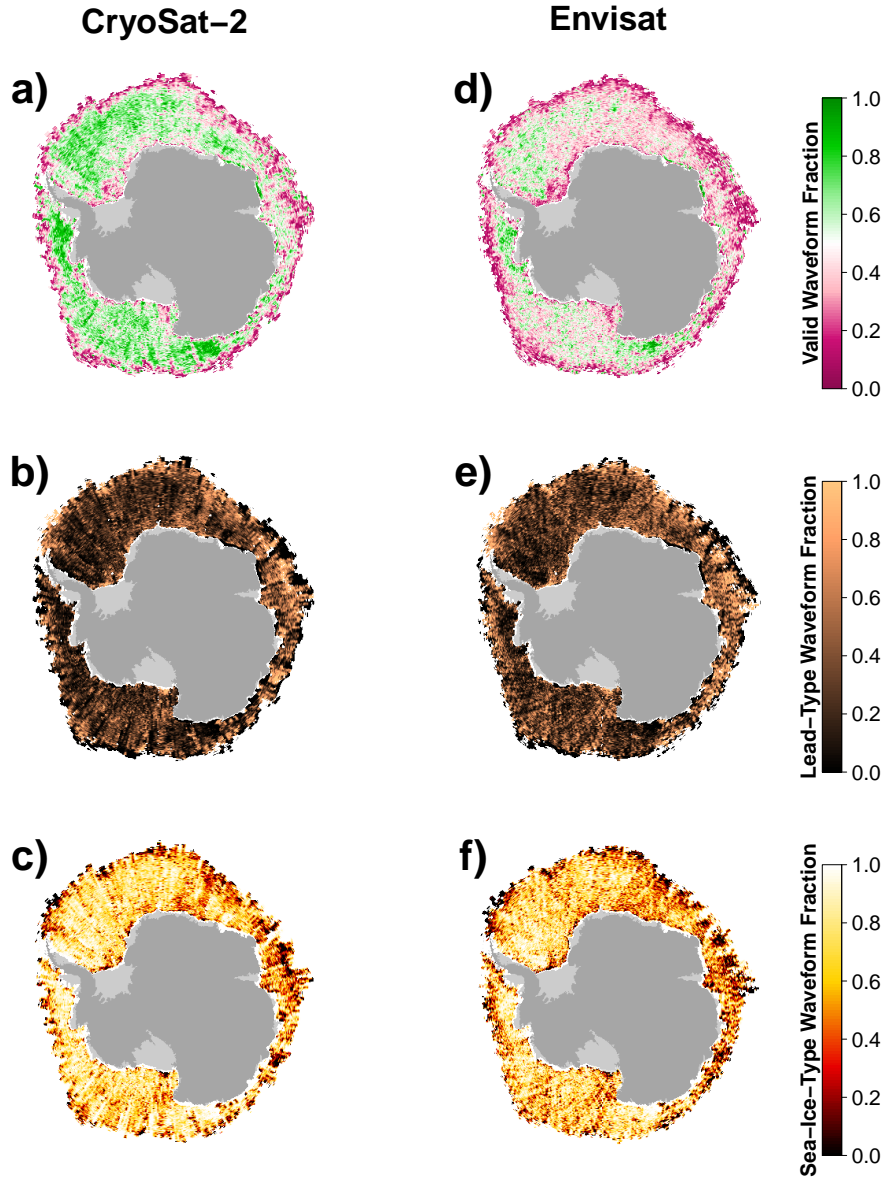
Figure 9. Comparison of gridded surface-type classification benchmarks in the same setup as Figure 8 between CryoSat-2 and Envisat for September 2011 in the Antarctic.

On average, the gridded valid waveform fraction for Envisat is 9 % (11 %) lower in the Arctic (Antarctic) than the ones achieved by CryoSat-2. This behavior is expected in regions with high rates of sea-ice dynamics such as the Beaufort Sea, where the increased surface-type mixing from the much larger footprint of Envisat likely prevents a clearer separation between waveform types. Lead and sea-ice fractions differ by 2 % and 4 % for the Arctic and Antarctic, respectively.
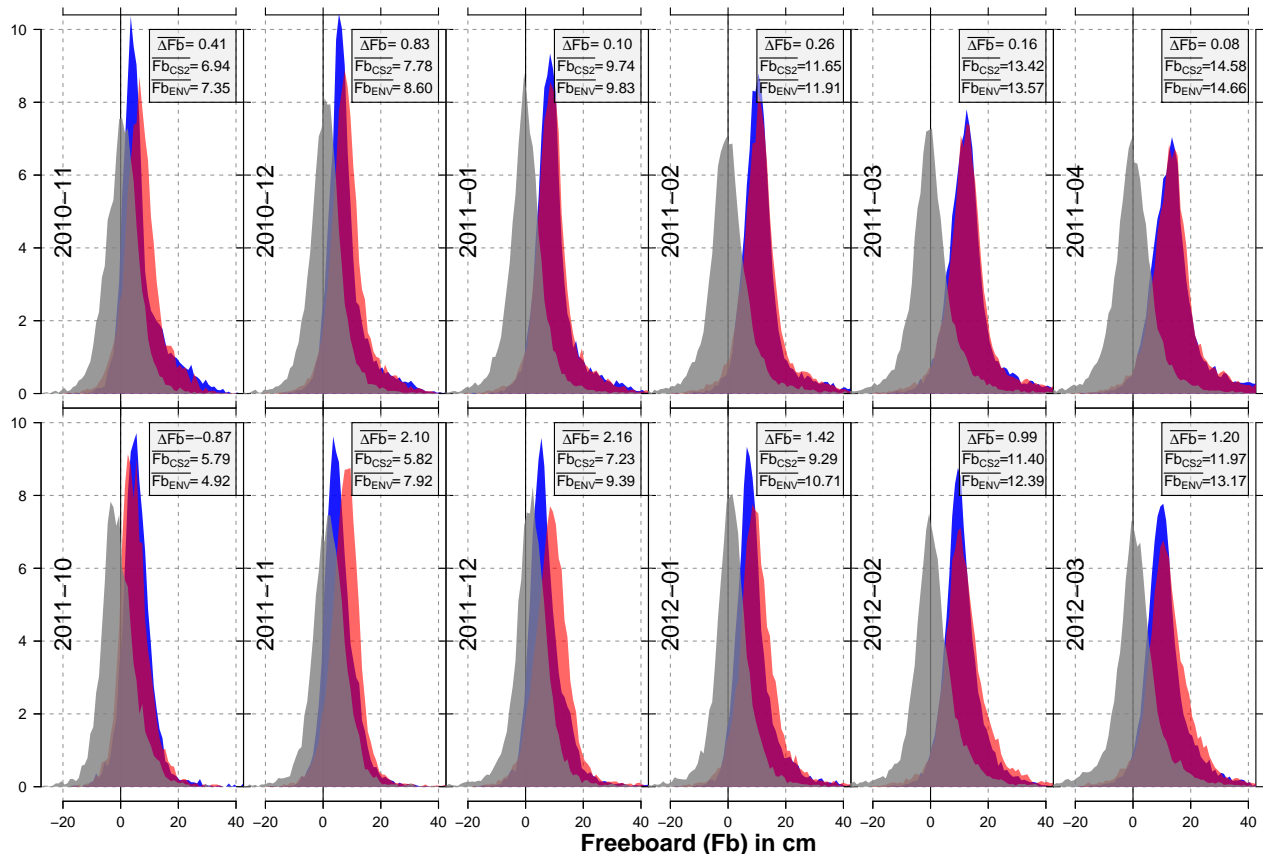
Figure 10. Histograms of freeboard for each month of the mission overlap period (in centimeters) for Envisat (red) and CryoSat-2 (blue) as well as the corresponding freeboard difference between both sensors (gray) for the Arctic. Furthermore, the average freeboard difference ($\overline{\Delta Fb}$; Envisat minus CryoSat-2), as well as the average freeboards for CryoSat-2 ($\overline{Fb_{CS2}}$) and Envisat ($\overline{Fb_{ENV}}$) are given in centimeters in the gray box for each month.

Nevertheless, both comparisons highlight the overall good agreement that could be achieved between both sensors with this inter-mission consistent surface-type classification scheme. These results therefore lay the foundation for a proper inter-mission sea-ice freeboard and sea-ice thickness data record.

3.2 Range retracking and freeboard retrieval

5   In this subsection, we show and discuss the results of using the adaptive threshold retracker for Envisat. For the Arctic, Figure 10 shows the histograms of CryoSat-2 (blue) and Envisat (red) freeboards in centimeters as well as the histogram of the resulting freeboard differences (Envisat minus CryoSat-2; gray). Furthermore, average freeboard estimates for all distributions per month during the mission-overlap period for Envisat and CryoSat-2 are shown.
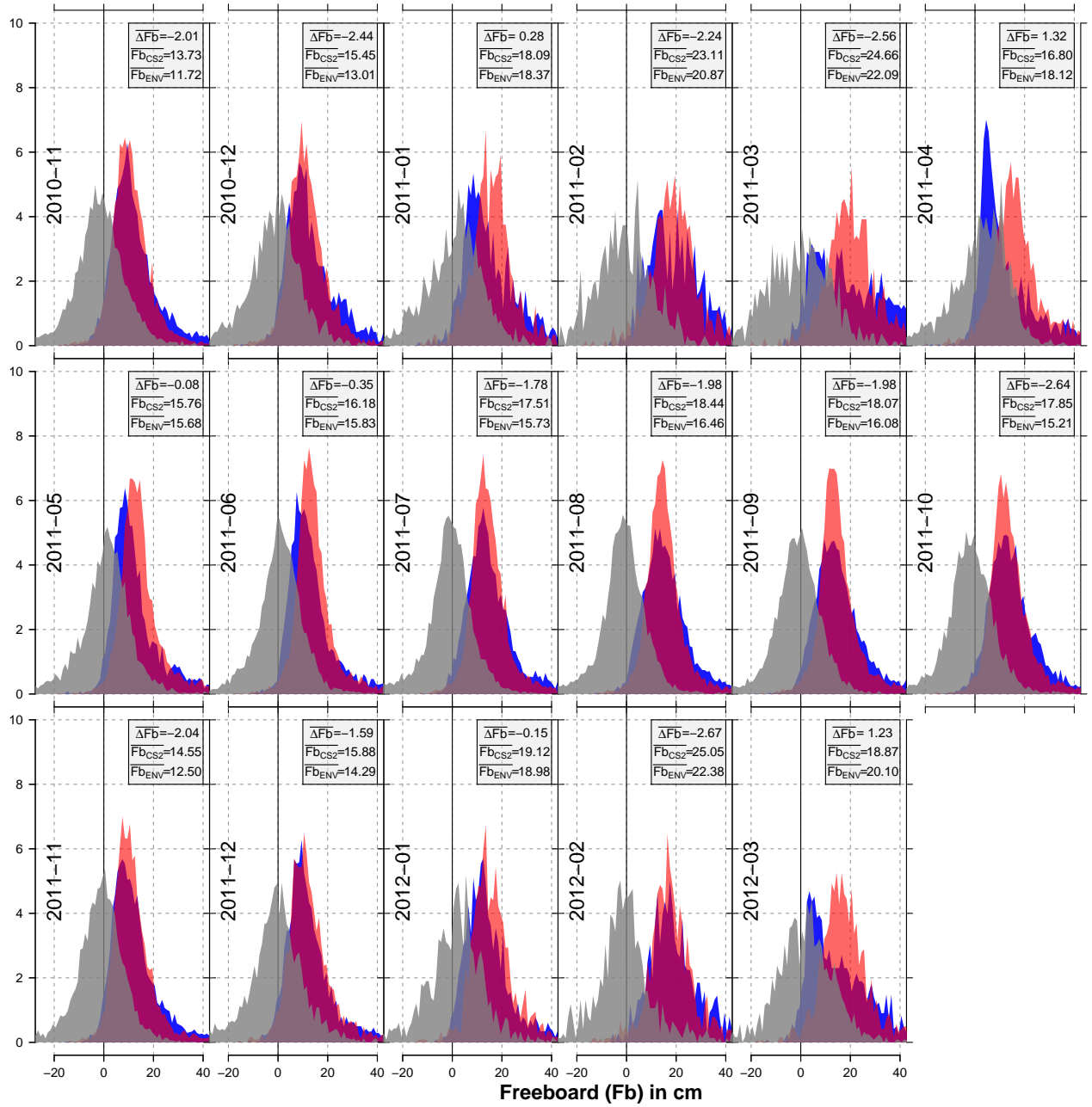
Figure 11. Histograms of freeboard for each month of the mission overlap period (in centimeters) for Envisat (red) and CryoSat-2 (blue) as well as the corresponding freeboard difference between both sensors (gray) for the Antarctic. Furthermore, the average freeboard difference ($\overline{\Delta Fb}$; Envisat minus CryoSat-2), as well as the average freeboards for CryoSat-2 ($\overline{Fb_{CS2}}$) and Envisat ($\overline{Fb_{ENV}}$) are given in centimeters in the gray box for each month.

While in the first winter season, the match is nearly perfect with absolute average freeboard differences below one centimeter, differences during the second winter season increase up to about three centimeters. Nevertheless, this is a substantial improvement over any previous comparisons conducted during SICCI-1. Especially for the Arctic spring period (March and April), differences in average freeboard are 1.2 cm or less. The overall maximum monthly average freeboard differences is 2.2 cm.

A comparison to the results of Guerreiro et al. (2017) is constrained by the different methods used. Besides the pulse-peakiness correction of the Envisat monthly freeboard, Guerreiro et al. (2017) also apply a 25 km along-track median smoothing to all freeboard estimates and afterward discard all freeboard estimates below -1 m and above 2 m, respectively. The resulting values are then used to compile the monthly gridded data set. In this study, we set the lower and upper sea-ice freeboard thresholds to -0.25 m and 2.25 m, respectively, without applying any smoothing. Values outside this range are discarded. Furthermore, we also compute freeboard results outside the central Arctic basin and take them into account for our comparison. While the differences between CryoSat-2 and Envisat freeboard appear to be comparable between both studies, the average monthly sea-ice freeboard estimates for both sensors are between about 3 cm and 10 cm larger in our study compared to the results shown in Guerreiro et al. (2017).

For the Antarctic, results are not as good as for the Arctic (Figure 11). Overall, the approach has less skill to match Envisat sea-ice freeboards to the ones of CryoSat-2. This is very likely related to other physical process such as a more complex snow stratigraphy caused by the more strongly varying weather patterns with melt-refreeze cycles even in the middle of winter. Furthermore, snow/ice-interface flooding causes a (temporarily) wet and saline basal snow layer and influences the sea-ice surface roughness. However, issues causing these sensor differences are subject to further investigation. Overall, there is a stronger seasonality in the monthly freeboard differences between summer and winter, which also leads towards a higher maximum monthly average freeboard difference of about 2.7 cm.

A different way in visualizing these results is shown in Figure 12. Here, the cumulative frequencies of absolute freeboard differences (Envisat minus CryoSat-2, i.e. the gray histograms in Figures 10 and 11) are averaged over the complete MOP for both hemispheres. For the Arctic (Figure 12a), more than 50 % of all data points are in an absolute freeboard difference range of ±4 cm. Furthermore, more than 75 %(90 %) of all data points are in a range of ±7 cm(±12 cm) absolute freeboard difference. However, for the Antarctic (12b) results are not as good. In order to achieve values of 50 %,75 %, and 90 % cumulative frequency, absolute freeboard difference increase to ±7 cm, ±12 cm,and ±20 cm, respectively.

Resulting freeboard differences using our new approach for Envisat and CryoSat-2 are shown in Figure 13 for the Arctic and in Figure 14 for the Antarctic. Shown are the same months as in Figures 2 through 5. While the overall differences are minimized, especially in the Arctic, there are still areas with rather large freeboard differences. In the Arctic, these comprise the Hudson Bay and the area east of Greenland. In the Antarctic, while the freeboard differences between both sensors are lowered through applying the here-presented methodology, the overall resulting differences remain larger than the ones estimated for the Arctic.
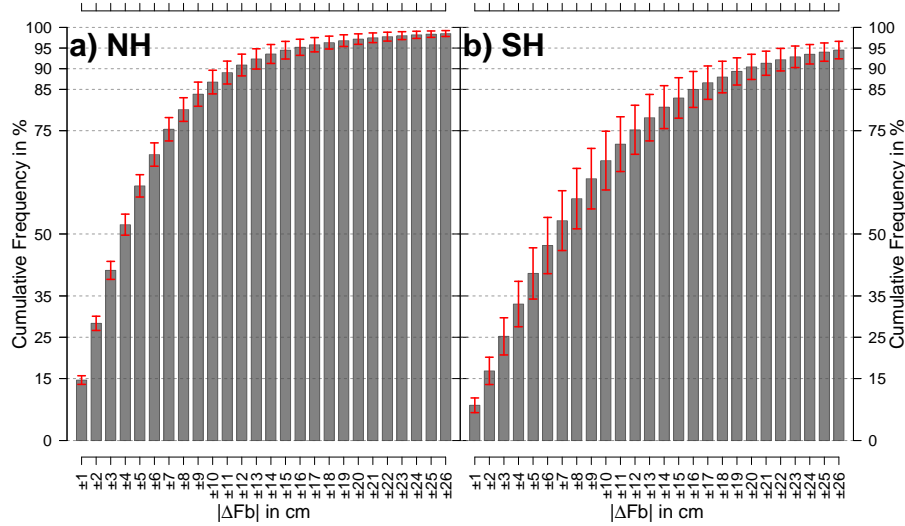
Figure 12. Cumulative frequencies of absolute freeboard differences (Envisat minus CryoSat-2) as shown in Figures 10 and 11 averaged over all months in the MOP. Data is presented for the northern hemisphere (a) as well as the southern hemisphere (b). Errorbars indicate ± 1 standard deviation. First bar covers the absolute freeboard difference range from [0cm:1cm[, second bin from [1cm:2cm[, and so on.
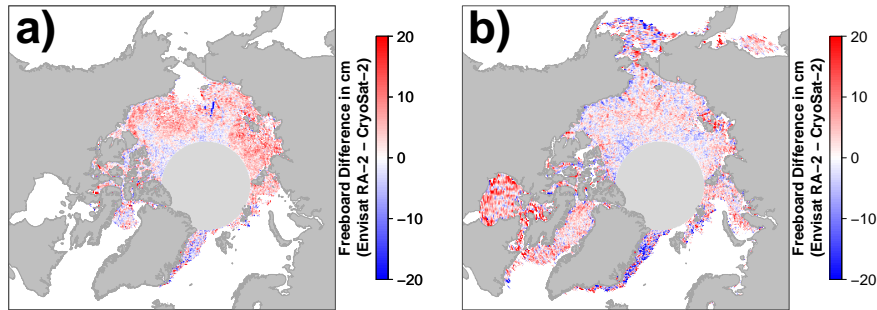
Figure 13. Visualizations of freeboard differences between Envisat using our new approach and CryoSat-2 in November 2011 (a) and March 2012 (b) for the Arctic (compare Figures 2a and 3a).

4 Summary and Outlook

This study showed the potential of a combined novel surface-type classification scheme in combination with a waveform-parameter dependent adaptive threshold retracker approach in order to create a consistent data set of Envisat and CryoSat-2 sea-ice freeboard estimates. This approach is based on the observed correlation between freeboard differences between both sensors and the waveform characteristics of surface backscatter and leading-edge width. Their spatio-temporal variations in acquired Envisat waveforms reflect changes in surface properties such as the surface roughness and footprint-size dependent surface-type mixing. We applied this approach for the mission overlap period from November 2010 to March 2012 and then
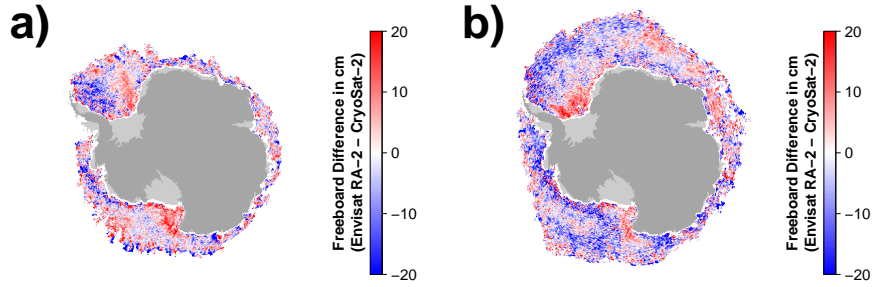
Figure 14. Visualizations of freeboard differences between Envisat using our new approach and CryoSat-2 in May 2011 (a) and September 2011 (b) for the Antarctic (compare Figures 4a and 5a).

used it to iteratively train and apply an adaptive threshold retracker to Envisat for both hemispheres. Different sea-ice conditions in both hemispheres also result in different inter-mission biases between Envisat and CryoSat-2. In contrast to previous attempts during SICCI-1, the inter-mission sea-ice freeboard biases could be minimized.

Furthermore, through the application of our inter-mission consistent surface-type classification in SICCI-2, a much higher
5 comparability between the amount and location of positively identified lead-type and sea-ice-type waveforms for Envisat and CryoSat-2 could be achieved. Additionally, due to the higher amount of identified waveforms, a resolution of 25 km × 25 km and 50 km × 50 km could be realized in the final gridded data product for the Arctic and Antarctic, respectively. This is a substantial improvement over SICCI-1, where due to the much stricter surface-type classification and associated low identification/classification rates of waveforms, a resolution of only 100 km × 100 km could be achieved without a substantial
10 drop in spatial data coverage for Envisat.

While the employed surface-type classification is truly consistent between both sensors, there still is an inconsistency with regard to the range retracking in using fixed thresholds for CryoSat-2 in contrast to an adaptive procedure for Envisat. However, the purpose of this study in the framework of SICCI-2 was to use existing data sets and algorithms wherever possible to create a cross-calibrated freeboard algorithm. In the future, we are planning to develop a truly consistent similar adaptive retracking
15 procedure for CryoSat-2 as well. As we take the CryoSat-2 data in this study as correct, the shown evaluation is not independent. A truly independent evaluation will be conducted in a future study.

The next step from this point is to start creating an inter-mission consistent and reliable climate data record on Arctic and Antarctic sea-ice thickness and volume as well as to investigate inter-annual as well as seasonal changes. Moreover, investigation of the stability of assumptions in auxiliary data sets is necessary. Here, especially the validity of the used snow-
20 depth climatologies in the freeboard to thickness conversion in a changing Arctic and Antarctic need to be investigated.

Author contributions. Stephan Paul designed the surface-type classification as well as the adaptive retracker approach, conducted the analysis and drafted the original manuscript. Stefan Hendricks developed pySIRAL, assisted with the analysis, the approach development as well as the data processing using pySIRAL. Robert Ricker, Stefan Kern, and Eero Rinne assisted in the analysis as well as in writing the manuscript.

Competing interests. The authors declare no conflict of interests.

5 *Acknowledgements.* This work was funded through the ESA Climate Change Initiative. The authors would like to thank the International Space Science Institute (ISSI Bern) for its support to an international team focused on the current scientific issues in the observation of the sea level and sea ice at polar latitudes, which fostered fruitful discussion on the ideas for this paper. Furthermore, the valuable comments of three reviewers (Thomas Armitage, Nathan Kurtz, and Sara Fleury) as well as editor Jennifer Hutchings are very much appreciated and greatly improved this manuscript.

References

Andersen, O., Knudsen, P., and Stenseng, L.: The DTU13 MSS (Mean Sea Surface) and MDT (Mean Dynamic Topography) from 20 Years of Satellite Altimetry, in: IGFS 2014, pp. 111–121, Springer International Publishing, Cham, 2016.

Armitage, T. W. K. and Davidson, M. W. J.: Using the Interferometric Capabilities of the ESA CryoSat-2 Mission to Improve the Accuracy of Sea Ice Freeboard Retrievals, IEEE Transactions on Geoscience and Remote Sensing, 52, 529–536, 2014.

Armitage, T. W. K. and Ridout, A. L.: Arctic sea ice freeboard from AltiKa and comparison with CryoSat-2 and Operation IceBridge, Geophys. Res. Lett., 42, 6724–6731, http://dx.doi.org/10.1002/2015GL064823, 2015.

Behrendt, A., Dierking, W., Fahrbach, E., and Witte, H.: Sea ice draft in the Weddell Sea, measured by upward looking sonars, Earth Syst. Sci. Data, 5, 209–226, https://www.earth-syst-sci-data.net/5/209/2013/, 2013.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Cavalieri, D. J., Parkinson, C. L., Gloersen, P., Comiso, J. C., and Zwally, H. J.: Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets, J. Geophys. Res., 104, 15 803–15 814, http://dx.doi.org/10.1029/1999JC900081, 1999.

Comiso, J. C.: Large Decadal Decline of the Arctic Multiyear Ice Cover, J. Climate, 25, 1176–1193, doi:10.1175/JCLI-D-11-00113.1, https://doi.org/10.1175/JCLI-D-11-00113.1, 2012.

Connor, L. N., Laxon, S. W., Ridout, A. L., Krabill, W. B., and McAdoo, D. C.: Comparison of Envisat radar and airborne laser altimeter measurements over Arctic sea ice, Remote Sensing of Environment, 113, 563–570, http://www.sciencedirect.com/science/article/pii/S0034425708003283, 2009.

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

Frost, T., Kern, S., and Heygster, G.: Passive microwave snow depth on Antarctic sea ice assessment, ESA CCI Sea-Ice ECV project report SICCI-ANT-PMW-SDASS-11-14, version 1.1, 2015.

Giles, K., Laxon, S., Wingham, D., Wallis, D., Krabill, W., Leuschen, C., McAdoo, D., Manizade, S., and Raney, R.: Combined airborne laser and radar altimeter measurements over the Fram Strait in May 2002, Remote Sensing of Environment, 111, 182–194, http://www.sciencedirect.com/science/article/pii/S0034425707002817, 2007.

Guerreiro, K., Fleury, S., Zakharova, E., Kouraev, A., Rémy, F., and Maisongrande, P.: Comparison of CryoSat-2 and ENVISAT radar freeboard over Arctic sea ice: toward an improved Envisat freeboard retrieval, The Cryosphere, 11, 2059–2073, https://www.the-cryosphere.net/11/2059/2017/, 2017.

Haas, C.: Evaluation of ship-based electromagnetic-inductive thickness measurements of summer sea-ice in the Bellingshausen and Amundsen Seas, Antarctica, Cold Regions Science and Technology, 27, 1–16, http://www.sciencedirect.com/science/article/pii/S0165232X97000190, 1998.

Haas, C., Nicolaus, M., Willmes, S., Worby, A., and Flinspach, D.: Sea ice and snow thickness and physical properties of an ice floe in the western Weddell Sea and their changes during spring warming, Deep Sea Research Part II: Topical Studies in Oceanography, 55, 963–974, http://www.sciencedirect.com/science/article/pii/S0967064508000386, 2008.

Hartigan, J. A. and Wong, M. A.: Algorithm AS 136: A K-Means Clustering Algorithm, 28, 100–108, http://www.jstor.org/stable/2346830, 1979.

Helm, V., Humbert, A., and Miller, H.: Elevation and elevation change of Greenland and Antarctica derived from CryoSat-2, The Cryosphere, 8, 1539–1559, https://www.the-cryosphere.net/8/1539/2014/, 2014.

Kern, S. and Ozsoy-Çiçek, B.: Satellite Remote Sensing of Snow Depth on Antarctic Sea Ice: An Inter-Comparison of Two Empirical Approaches, 2016.

Kern, S., Frost, T., and Heygster, G.: Antarctic AMSR-E snow depth product user guide, ESA CCI Sea-Ice ECV project report SICCI-ANT-SD-PUG-14-08, version 2.1, 2015.

5 Kern, S., Ozsoy-Çiçek, B., and Worby, P. A.: Antarctic Sea-Ice Thickness Retrieval from ICESat: Inter-Comparison of Different Approaches, 2016.

Kurtz, N. T. and Farrell, S. L.: Large-scale surveys of snow depth on Arctic sea ice from Operation IceBridge, Geophys. Res. Lett., 38, n/a–n/a, http://dx.doi.org/10.1029/2011GL049216, 2011.

Kurtz, N. T. and Markus, T.: Satellite observations of Antarctic sea ice thickness and volume, J. Geophys. Res., 117, n/a–n/a, http://dx.doi.
10 org/10.1029/2012JC008141, 2012.

Kwok, R. and Cunningham, G. F.: Variability of Arctic sea ice thickness and volume from CryoSat-2, Philos Transact A Math Phys Eng Sci, 373, –, http://rsta.royalsocietypublishing.org/content/373/2045/20140157.abstract, 2015.

Kwok, R. and Rothrock, D. A.: Decline in Arctic sea ice thickness from submarine and ICESat records: 1958-2008, Geophys. Res. Lett., 36, http://dx.doi.org/10.1029/2009GL039035, 2009.

15 Kwok, R., Cunningham, G. F., Wensnahan, M., Rigor, I., Zwally, H. J., and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003-2008, J. Geophys. Res., 114, n/a–n/a, http://dx.doi.org/10.1029/2009JC005312, 2009.

Laxon, S.: Sea ice altimeter processing scheme at the EODC, International Journal of Remote Sensing, 15, 915–924, doi:10.1080/01431169408954124, https://doi.org/10.1080/01431169408954124, 1994.

Laxon, S., Peacock, N., and Smith, D.: High interannual variability of sea ice thickness in the Arctic region, Nature, 425, 947–950, http:
20 //dx.doi.org/10.1038/nature02050, 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, http://dx.doi.org/10.1002/grl.50193, 2013.

Leuschen, C. J., Swift, R. N., Comiso, J. C., Raney, R. K., Chapman, R. D., Krabill, W. B., and Sonntag, J. G.: Combination of laser and
25 radar altimeter height measurements to estimate snow depth during the 2004 Antarctic AMSR-E Sea Ice field campaign, J. Geophys. Res., 113, n/a–n/a, http://dx.doi.org/10.1029/2007JC004285, 2008.

Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, The Cryosphere, 9, 269–283, https://www.the-cryosphere.net/9/269/2015/, 2015.

MacQueen, J.: Some methods for classification and analysis of multivariate observations, pp. –, 1967.

30 Markus, T. and Cavalieri, D. J.: Snow Depth Distribution Over Sea Ice in the Southern Ocean from Satellite Passive Microwave Data, in: Antarctic Sea Ice: Physical Processes, Interactions and Variability, pp. 19–39, American Geophysical Union, http://dx.doi.org/10.1029/AR074p0019, 1998.

Markus, T., Cavalieri, D. J., and Ivanoff, A.: Algorithm theoretical basis document: Sea ice products: Updated December 2011, 2011.

Maslanik, J. and Stroeve, J.: DMSP SSM/I-SSMIS daily polar gridded brightness temperatures, version 4, Boulder, Co: NASA DAAC at the
35 National Snow and Ice Data Center, pp. –, 2004.

Meier, W. N., Hovelsrud, G. K., van Oort, B. E., Key, J. R., Kovacs, K. M., Michel, C., Haas, C., Granskog, M. A., Gerland, S., Perovich, D. K., Makshtas, A., and Reist, J. D.: Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity, Rev. Geophys., 52, 185–217, http://dx.doi.org/10.1002/2013RG000431, 2014.

Ozsoy-Cicek, B., Ackley, S., Xie, H., Yi, D., and Zwally, J.: Sea ice thickness retrieval algorithms based on in situ surface elevation and thickness values for application to altimetry, J. Geophys. Res. Oceans, 118, 3807–3822, http://dx.doi.org/10.1002/jgrc.20252, 2013.

Parkinson, C. L. and Cavalieri, D. J.: Antarctic sea ice variability and trends, 1979–2010, The Cryosphere, 6, 871–880, https://www.the-cryosphere.net/6/871/2012/, 2012.

5 Parkinson, C. L. and DiGirolamo, N. E.: New visualizations highlight new information on the contrasting Arctic and Antarctic sea-ice trends since the late 1970s, Remote Sensing of Environment, 183, 198–204, http://www.sciencedirect.com/science/article/pii/S0034425716302218, 2016.

Peacock, N. R. and Laxon, S. W.: Sea surface height determination in the Arctic Ocean from ERS altimetry, J. Geophys. Res., 109, n/a–n/a, http://dx.doi.org/10.1029/2001JC001026, 2004.

10 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, https://www.the-cryosphere.net/8/1607/2014/, 2014.

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J., and Haas, C.: A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data, The Cryosphere, 11, 1607–1623, https://www.the-cryosphere.net/11/1607/2017/, 2017.

Rothrock, D. A., Yu, Y., and Maykut, G. A.: Thinning of the Arctic sea-ice cover, Geophys. Res. Lett., 26, 3469–3472, http://dx.doi.org/10.1029/1999GL010863, 1999.

15 Schwegmann, S., Rinne, E., Ricker, R., Hendricks, S., and Helm, V.: About the consistency between Envisat and CryoSat-2 radar freeboard retrieval over Antarctic sea ice, The Cryosphere, 10, 1415–1425, https://www.the-cryosphere.net/10/1415/2016/, 2016.

Stroeve, J. C., Serreze, M. C., Holland, M. M., Kay, J. E., Malanik, J., and Barrett, A. P.: The Arctic's rapidly shrinking sea ice cover: a research synthesis, Climatic Change, 110, 1005–1027, https://doi.org/10.1007/s10584-011-0101-1, 2012.

20 Tilling, R. L., Ridout, A., and Shepherd, A.: Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data, Advances in Space Research, pp. –, http://www.sciencedirect.com/science/article/pii/S0273117717307901, 2017.

Turner, J., Phillips, T., Marshall, G. J., Hosking, J. S., Pope, J. O., Bracegirdle, T. J., and Deb, P.: Unprecedented springtime retreat of Antarctic sea ice in 2016, Geophys. Res. Lett., 44, 6868–6875, http://dx.doi.org/10.1002/2017GL073656, 2017.

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., and Colony, R.: Snow Depth on Arctic Sea
25 Ice, J. Climate, 12, 1814–1829, doi:10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2, https://doi.org/10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2, 1999.

Wingham, D., Rapley, C., and Griffiths, H.: New techniques in satellite altimeter tracking systems, in: Proceedings of IGARSS, vol. 86, pp. 1339–1344, 1986.

Worby, A. P., Geiger, C. A., Paget, M. J., Van Woert, M. L., Ackley, S. F., and DeLiberty, T. L.: Thickness distribution of Antarctic sea ice,
30 J. Geophys. Res., 113, n/a–n/a, http://dx.doi.org/10.1029/2007JC004254, 2008a.

Worby, A. P., Markus, T., Steer, A. D., Lytle, V. I., and Massom, R. A.: Evaluation of AMSR-E snow depth product over East Antarctic sea ice using in situ measurements and aerial photography, J. Geophys. Res., 113, n/a–n/a, http://dx.doi.org/10.1029/2007JC004181, 2008b.

Zygmuntowska, M., Khvorostovsky, K., Helm, V., and Sandven, S.: Waveform classification of airborne synthetic aperture radar altimeter over Arctic sea ice, The Cryosphere, 7, 1315–1324, https://www.the-cryosphere.net/7/1315/2013/, 2013.

35 **Appendix A: Surface-Type Classification Thresholds**

Table A1. Pulse-peakiness thresholds for the surface-type classification of lead-type waveforms for Envisat, CryoSat-2 SAR mode, and CryoSat-2 SIN mode data for the Arctic and the Antarctic. Values for the pulse peakiness are unit-less.

| Month | Envisat | | | | CryoSat-2 SAR | | | | CryoSat-2 SIN | | | |
| | Arctic | | Antarctic | | Arctic | | Antarctic | | Arctic | | Antarctic | |
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Jan	46.90	–	56.60	–	67.30	–	80.70	–	264.30	–	307.40	–
Feb	46.40	–	53.20	–	66.30	–	75.10	–	257.90	–	300.70	–
Mar	46.20	–	51.90	–	66.60	–	73.20	–	253.60	–	291.70	–
Apr	48.40	–	50.70	–	69.90	–	69.50	–	264.60	–	288.50	–
May	–	–	50.10	–	–	–	69.70	–	–	–	283.70	–
Jun	–	–	49.30	–	–	–	69.30	–	–	–	284.20	–
Jul	–	–	49.50	–	–	–	69.20	–	–	–	276.90	–
Aug	–	–	49.10	–	–	–	69.50	–	–	–	284.40	–
Sep	–	–	49.30	–	–	–	69.70	–	–	–	278.90	–
Oct	52.90	–	51.60	–	76.00	–	71.70	–	291.80	–	289.40	–
Nov	51.00	–	53.90	–	73.80	–	76.00	–	288.80	–	299.40	–
Dec	47.70	–	55.10	–	68.60	–	78.10	–	272.60	–	307.70	–

Table A2. surface backscatter thresholds for the surface-type classification of lead-type waveforms for Envisat, CryoSat-2 SAR mode, and CryoSat-2 SIN mode data for the Arctic and the Antarctic. Values for surface backscatter are given in dB.

| Month | Envisat | | | | CryoSat-2 SAR | | | | CryoSat-2 SIN | | | |
| | Arctic | | Antarctic | | Arctic | | Antarctic | | Arctic | | Antarctic | |
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Jan	28.80	–	33.20	–	23.80	–	28.50	–	24.90	–	29.20	–
Feb	28.60	–	32.10	–	23.20	–	26.80	–	25.00	–	29.00	–
Mar	28.50	–	31.80	–	23.30	–	26.20	–	24.10	–	28.50	–
Apr	28.40	–	30.80	–	23.40	–	24.60	–	24.50	–	27.80	–
May	–	–	29.40	–	–	–	23.40	–	–	–	26.90	–
Jun	–	–	28.60	–	–	–	22.80	–	–	–	26.50	–
Jul	–	–	28.60	–	–	–	23.00	–	–	–	26.30	–
Aug	–	–	28.40	–	–	–	23.00	–	–	–	27.00	–
Sep	–	–	28.50	–	–	–	23.20	–	–	–	26.20	–
Oct	32.80	–	29.50	–	28.00	–	24.00	–	29.00	–	27.20	–
Nov	30.80	–	31.10	–	25.80	–	25.90	–	27.40	–	27.50	–
Dec	29.30	–	32.10	–	24.10	–	27.30	–	25.80	–	28.40	–

Table A3. Leading-edge width thresholds for the surface-type classification of lead-type waveforms for Envisat, CryoSat-2 SAR mode, and CryoSat-2 SIN mode data for the Arctic and the Antarctic. Values are in range-bin fractions for the leading-edge width.

| Month | Envisat | | | | CryoSat-2 SAR | | | | CryoSat-2 SIN | | | |
| | Arctic | | Antarctic | | Arctic | | Antarctic | | Arctic | | Antarctic | |
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Jan	–	0.82	–	0.82	–	0.77	–	0.71	–	1.10	–	1.00
Feb	–	0.82	–	0.82	–	0.78	–	0.73	–	1.11	–	1.01
Mar	–	0.82	–	0.82	–	0.78	–	0.74	–	1.13	–	1.03
Apr	–	0.82	–	0.82	–	0.76	–	0.77	–	1.09	–	1.04
May	–	–	–	0.82	–	–	–	0.77	–	–	–	1.06
Jun	–	–	–	0.82	–	–	–	0.77	–	–	–	1.05
Jul	–	–	–	0.82	–	–	–	0.78	–	–	–	1.07
Aug	–	–	–	0.82	–	–	–	0.77	–	–	–	1.05
Sep	–	–	–	0.82	–	–	–	0.77	–	–	–	1.07
Oct	–	0.82	–	0.82	–	0.72	–	0.76	–	1.02	–	1.05
Nov	–	0.82	–	0.82	–	0.73	–	0.74	–	1.03	–	1.02
Dec	–	0.82	–	0.82	–	0.76	–	0.72	–	1.07	–	1.00

Table A4. Pulse-peakiness thresholds for the surface-type classification of sea-ice-type waveforms for Envisat, CryoSat-2 SAR mode, and CryoSat-2 SIN mode data for the Arctic and the Antarctic. Values for the pulse peakiness are unit-less.

| Month | Envisat | | | | CryoSat-2 SAR | | | | CryoSat-2 SIN | | | |
| | Arctic | | Antarctic | | Arctic | | Antarctic | | Arctic | | Antarctic | |
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Jan	–	16.00	–	24.60	–	30.50	–	40.10	–	99.40	–	138.40
Feb	–	14.80	–	20.70	–	28.70	–	35.30	–	94.20	–	126.10
Mar	–	14.10	–	19.60	–	28.10	–	32.90	–	89.90	–	124.90
Apr	–	14.20	–	18.80	–	28.50	–	30.20	–	90.00	–	127.30
May	–	–	–	17.50	–	–	–	28.70	–	–	–	122.20
Jun	–	–	–	16.90	–	–	–	28.90	–	–	–	121.00
Jul	–	–	–	16.60	–	–	–	28.10	–	–	–	114.90
Aug	–	–	–	16.10	–	–	–	28.00	–	–	–	115.80
Sep	–	–	–	16.30	–	–	–	28.40	–	–	–	114.30
Oct	–	19.40	–	18.10	–	35.40	–	29.60	–	114.40	–	121.20
Nov	–	19.30	–	20.70	–	34.90	–	34.10	–	113.90	–	126.50
Dec	–	16.90	–	22.80	–	31.90	–	36.60	–	103.80	–	135.20

Table A5. surface backscatter thresholds for the surface-type classification of sea-ice-type waveforms for Envisat, CryoSat-2 SAR mode, and CryoSat-2 SIN mode data for the Arctic and the Antarctic. Values for surface backscatter are given in dB.

| Month | Envisat | | | | CryoSat-2 SAR | | | | CryoSat-2 SIN | | | |
| | Arctic | | Antarctic | | Arctic | | Antarctic | | Arctic | | Antarctic | |
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Jan	–	22.50	–	27.20	–	20.80	–	26.30	–	21.40	–	26.40
Feb	–	21.80	–	25.40	–	19.90	–	24.10	–	20.90	–	25.10
Mar	–	21.30	–	26.70	–	19.60	–	25.10	–	20.10	–	27.60
Apr	–	20.40	–	27.20	–	19.00	–	26.20	–	19.10	–	27.30
May	–	–	–	24.60	–	–	–	23.10	–	–	–	24.90
Jun	–	–	–	23.10	–	–	–	20.90	–	–	–	24.20
Jul	–	–	–	22.50	–	–	–	20.20	–	–	–	24.10
Aug	–	–	–	21.70	–	–	–	19.10	–	–	–	24.90
Sep	–	–	–	22.30	–	–	–	20.00	–	–	–	23.70
Oct	–	25.90	–	23.30	–	25.70	–	20.60	–	24.30	–	25.00
Nov	–	24.60	–	25.20	–	23.20	–	22.90	–	23.70	–	25.20
Dec	–	22.80	–	26.10	–	21.10	–	23.90	–	22.00	–	25.00

Table A6. Leading-edge width thresholds for the surface-type classification of sea-ice-type waveforms for Envisat, CryoSat-2 SAR mode, and CryoSat-2 SIN mode data for the Arctic and the Antarctic. Values are in range-bin fractions for the leading-edge width.

| Month | Envisat | | | | CryoSat-2 SAR | | | | CryoSat-2 SIN | | | |
| | Arctic | | Antarctic | | Arctic | | Antarctic | | Arctic | | Antarctic | |
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Jan	0.81	–	0.78	–	1.02	–	0.87	–	1.55	–	1.31	–
Feb	0.83	–	0.80	–	1.08	–	0.95	–	1.58	–	1.40	–
Mar	0.83	–	0.80	–	1.10	–	0.98	–	1.62	–	1.37	–
Apr	0.83	–	0.80	–	1.11	–	1.02	–	1.64	–	1.34	–
May	–	–	0.81	–	–	–	1.07	–	–	–	1.37	–
Jun	–	–	0.80	–	–	–	1.07	–	–	–	1.38	–
Jul	–	–	0.80	–	–	–	1.12	–	–	–	1.41	–
Aug	–	–	0.81	–	–	–	1.13	–	–	–	1.41	–
Sep	–	–	0.81	–	–	–	1.11	–	–	–	1.42	–
Oct	0.78	–	0.80	–	0.91	–	1.08	–	1.44	–	1.38	–
Nov	0.78	–	0.79	–	0.90	–	0.95	–	1.44	–	1.36	–
Dec	0.80	–	0.78	–	0.97	–	0.92	–	1.51	–	1.33	–