Arctic sea ice radar freeboard retrieval from ERS-2 using altimetry: Toward sea ice thickness observation from 1995 to 2021

Marion Bocquet[1,2], Sara Fleury[1], Fanny Piras[2], Eero Rinne[3,4], Heidi Sallila[3], Florent Garnier[1], and Frédérique Rémy[1]

[1]LEGOS, Université de Toulouse, CNES, CNRS, IRD, UPS (Toulouse), France
[2]Collecte Localisation Satellites (CLS), Toulouse, France
[3]Marine Research, Finnish Meteorological Institute, Helsinki, Finland
[4]University Centre in Svalbard (UNIS), PO Box 156, N-9171 Longyearbyen, Norway

Correspondence: Marion Bocquet (marion.bocquet@legos.obs-mip.fr)

Abstract.

Sea ice volume's significant interannual variability requires long-term series of observations to identify trends in its evolution. Despite improvements in sea ice thickness estimations from altimetry during the past few years thanks to CryoSat-2 and ICESat-2, former ESA radar altimetry missions such as Envisat and especially ERS-1 and ERS-2 have remained under-exploited so far. Although solutions have already been proposed to ensure continuity of measurements between CryoSat-2 and Envisat, there is no time series integrating ERS. The purpose of this study is to extend the Arctic radar freeboard time series back to 1995. The difficulty to handle ERS measurements comes from a technical issue known as the pulse-blurring effect, altering the radar echos over sea ice and the resulting surface height estimates. Here we present and apply a correction for this pulse-blurring effect. To ensure consistency of the CryoSat-2/Envisat/ERS-2 time series, a multi-parameters neural network-based method to calibrate Envisat against CryoSat-2 and ERS-2 against Envisat is presented. The calibration is trained on the discrepancies observed between the altimeter measurements during the missions-overlap periods and a set of parameters characterizing the sea ice state. Monthly radar freeboards are provided with uncertainty estimations based on a Monte Carlo approach to propagate the uncertainties all along the processing chain, including the neural network. Comparisons of corrected radar freeboards during overlap periods reveal good consistencies between missions, with a mean bias of ~~3~~ 0.30 cm and a standard deviation of 9.7 cm for Envisat/CryoSat-2 and ~~2~~ 0.20 cm bias and a standard deviation of 3.8 cm for ERS-2/Envisat. The monthly maps obtained from Envisat and ERS-2 are then validated by comparison with several independent data such as airborne, moorings, direct measurements and other altimeter products. Except for two data sets, comparisons lead to correlation ranging from ~~0.42~~ 0.41 to 0.94 for Envisat, and ~~0.6 to 0.76~~ 0.60 to 0.74 for ERS-2. The study finally provides radar freeboard estimation for winters from 1995 to 2021 (from ERS-2 mission to CryoSat-2).

1 Introduction

Several indicators illustrate the evolution of sea ice in response to climate change. Arctic sea ice extent has strongly decreased since the beginning of satellite observation era by radiometry (Stroeve et al., 2012; Meier et al., 2014; Stroeve and Notz, 2018).

The perennial ice proportion has declined significantly since 1984, the amount has halved in April between 1984 and 2018 (Stroeve and Notz, 2018). The end of summer 2021 became the second lowest amount of multi-year ice since 1985 (Meier

25 et al., 2021). To improve our knowledge and forecast its evolution, an additional dimension becomes crucial: the thickness. Thick and old ice is disappearing and being replaced by younger, thin ice that has a higher mechanical sensitivity. Thin ice is ~~indeed more sensitive to climatic hazards than thicker ice but it especially enables to compute the volume~~more prone to deformation (Stroeve and Notz, 2018) that induce area changes, and is more sensitive to climate hazards such as cyclones or strong winds (Rheinlænder et al., 2022). Thickness is a key parameter for sea ice study, it varies a lot according to the regions

30 and it modulates the sea ice volume evolution in the Arctic ocean (Landy et al., 2022). Various campaigns have been carried out in the Arctic since the middle of the 20th century to measure sea ice thickness (Lindsay and Schweiger, 2013; Krishfield et al., 2014). However, these space and time-limited measurements do not allow conclusions to be drawn at basin-scale sea ice volume variations. A ~~global~~ quasi-global approach is possible through satellite altimetry, especially with radar altimetry, which is not impacted by the cloud cover and whose missions are continuous since 1991.

35 Sea ice thickness estimation by ~~Ku-band radar altimetry is~~ spatial altimetry was first introduced by Laxon (1994) and Peacock and Laxon (2004) ~~. From the difference of altimeter measured surface heights on the ice floes and on the leads, we can deduce the thickness of the emerged part of the floe, named the radar freeboard . It is commonly accepted that the Ku frequency penetrates the snow layer when it is sufficiently cold, in other situations this assumption can be questioned (Ricker et al., 2014; Nandan et al., 2017). Nevertheless, it is necessary to correct the estimated freeboard due to the~~ based on

40 the freeboard methodology. The radar freeboard is obtained by taking the difference between the height measured over floes and the height over leads interpolated under the floes. The radar freeboard has to be corrected of the radar signal slowdown ~~when penetrating~~ within the snow layer to retrieve the ice freeboard which is the thickness of the emerged part of the floe. Given the snow depth, the density of water, ice and snow, it is possible to derive the thickness of the ice assuming hydrostatic equilibrium. Lead and floe heights can be estimated using a heuristic retracker or, more recently developed, physical retrackers

45 (Kurtz et al., 2014; Landy et al., 2019; Laforge et al., 2020). Implementing this method requires the assumption that the Ku-band radar wave completely penetrates the snow layer, which is still widely discussed and is not the subject of a definitive consensus (Ricker et al., 2014; Nandan et al., 2017).

The launch of the CryoSat-2 (CS-2) mission and featuring a high-resolution Synthetic Aperture Radar (SAR) mode has allowed important advances in the estimation of sea ice thicknesses. The benefits are many, especially when compared with

50 altimeters from past missions (ERS-1, ERS-2 and Envisat) operating with older technology, the Low-Resolution Mode (LRM) which has a larger surface footprint, making thickness estimation more difficult. To reconstitute Envisat sea ice thickness estimation, Guerreiro et al. (2017), Paul et al. (2018) and Tilling et al. (2019) rely on the differences between Envisat and CS-2 during their common flight period to be able to calibrate the Envisat freeboard. To go back to the ERS missions, an additional problem coming from the instrument appears, the 'pulse-blurring' described in Peacock (1998) and in Peacock and Laxon

55 (2004). Laxon et al. (2003) and Giles et al. (2008) published thereafter the first and last ERS thickness estimations for the Arctic and Antarctic sea ice so far (as a map averaging all the estimates of the different winters over the whole flight period of ERS-1 and ERS-2).

This study presents a method to recover a continuous time series of the Arctic sea ice freeboard back to ERS-2. To minimize intermission bias along with the series, ERS-2 freeboard estimates are adjusted on Envisat freeboard estimates which in turn
60 have been previously adjusted on CryoSat-2, taking advantage of respective common flight periods. The consistency between missions is preserved by using the same processing chain regardless of the mission ~~, with the heuristic retracker TFMRA50 (Helm et al., 2014)~~ (before calibration), starting by the chosen retracking algorithm : the empirical threshold first-maximum retracker algorithm (Helm et al., 2014) with a threshold of 50 % (TFMRA50). We also present the method used to correct the ERS-2 measurements for the effect of pulse-blurring, which is a prerequisite for using ERS measurements over sea ice. The
65 adjustment of LRM measurements on CryoSat-2 is done using machine learning based on the surface state of the ice in Sect. 3. The associated uncertainties are derived using a Monte Carlo approach. Sect. 4 compares the monthly Envisat and ERS-2 radar freeboard data with various in-situ, space-borne data sets or other altimetry products available during this period and presents the 27 year-long time series.

2 Data

70 ### 2.1 Satellite Altimetry Data

2.1.1 Cryosat-2

CryoSat-2 is an ESA altimeter mission launched in 2010. With a nearly polar and geodetic orbit, it enables observations up to 88 °N that makes it particularly adapted for cryosphere observations. Additionally, CryoSat-2 incorporates nadir SAR and SARin technologies (Wingham et al., 2006). The two altimetry approaches exploit the Doppler capabilities of the instrument
75 to reduce the along-track footprint from several kilometers to approximately 300 m compared to LRM (corresponding to a reduction of the footprint area from 5 km^2 to ~~150~~ 180 km^2 (Stammer, 2018)). Increasing the along-track resolution of the aperture radar has led to considerable advances in ~~the measurement of~~ sea ice thickness estimation. For this study, we use SAR and SARin data at 20 Hz, from the ESA baseline-D L1b product~~(SAR mode : , SARin mode :)~~. We derive the radar freeboard for the 7 coldest Arctic months (from October to April) of each year, from November 2010 to present.

80 #### 2.1.2 Envisat

ESA's Envisat mission, was launched in 2002, reaching latitudes of 81.5° N/S. The satellite carried the radar altimeter RA-2 (operating in Low-Resolution Mode, LRM), with a high Pulse Repetition Frequency (PRF) of 1795 Hz allowing a large number of measurements per second to be performed, resulting in a better accuracy. The return pulses are averaged in batches of 100 to constitute each waveform (Roca et al., 2009). RA-2 L1b ~~v3~~ version 3 products including waveforms provided by ESA ~~()~~ are
85 used in this analysis. Sea ice freeboard is computed for the coldest 7 months of each year from October 2002 to ~~April~~ March 2012.

Contrary to SARM, LRM altimetry measurements are strongly impacted by the surface roughness of the surface illuminated by the radar, also affecting the freeboard ~~estimate~~measurement (Raney, 1995) (see Sect 3.4 for more information). Our

approach is to make use of these processing mode differences to derive an LRM corrected freeboard. To that end, we compare

90 Envisat and CryoSat-2 datasets during missions-overlap period which runs from November 2010 to ~~April~~ March 2012 (see Sect. 3.4).

2.1.3 ERS-2

In the 1990s, ESA has launched two European Remote Sensing (ERS) satellites, ERS-1 in July 1991 and ERS-2 in April 1995. ERS-1 was able to perform nominally until June 1996 and ERS-2 until November 2003. To extend the time series while

95 ensuring the continuity with Envisat mission, ERS-2 products from the ESA Reaper project (Brockley et al., 2017) ~~()~~ were used until July 2003. The ERS RA altimeter operated at a lower PRF than Envisat RA-2 altimeter (1020 Hz against 1795 Hz respectively). Thus, the 20 Hz waveforms are made up of 50 elementary echoes instead of 100 for RA-2. That leads to a higher speckle noise for ERS missions than for Envisat (see Sect. 3.5 for further details).

Another significant difference between RA and RA-2 comes from the tracker-board control loop that aims at centering the

100 expected echo in the altimeter acquisition window. The delay between the transmission and the reception of the radar waveform depends on the vertical distance between the altimeter and the Earth's surface. This distance varies along with the satellite orbit and the ground topography. The time between the transmission of the radar wave and the opening of the acquisition window must therefore be constantly adapted (called window-delay (s) or tracker-range (m) if it is whether a time or a distance). The distance, or range, between the altimeter and the measured surface, is then equal to the sum of the tracker range and the epoch,

105 i.e. the position of the waveform in the window. Since a waveform is an average of 50 individual pulses, it is important that each pulse is correctly centered in the window (to be aligned with each others). Otherwise, the resulting averaged 20 Hz waveforms will be blurred. This is unfortunately what ~~happens~~ happened to ERS altimeters over sea ice-covered surfaces (Peacock and Laxon, 2004).

2.2 Ancillary Data

110 Whether it is for the calculation of the radar freeboard itself, the LRM calibration or the comparison of our results to in situ data, we use various additional data sets. We present here additional data sets that have been used for that purpose.

The sea ice concentration field is needed to restrain the freeboard computation over sea ice covered area. The product used is the NSIDC 0051 product based on Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data (Cavalieri et al., 1996). This product is also used to compute radar freeboard volume in Sect. 4.3 ~~.~~

115 ~~This product is also used~~ as well as for the LRM/SARM calibration~~training and correction, as well as a Multi-Year Ice (MYI) fraction product. This information comes~~ . The study also requires a sea ice type product, this information is derived from the NSIDC 0061 sea ice age product (Tschudi et al., 2019) that is aggregated into two classes (MYI and FYI) according to the age of the ice (FYI : ice age between 0 and 1 year, MYI : ice age of at least one-year). Data are respectively available as daily and weekly map with a 12,5 km grid resolution. The fraction of MYI is derived from the ice type information during the gridding

120 processing step.

4
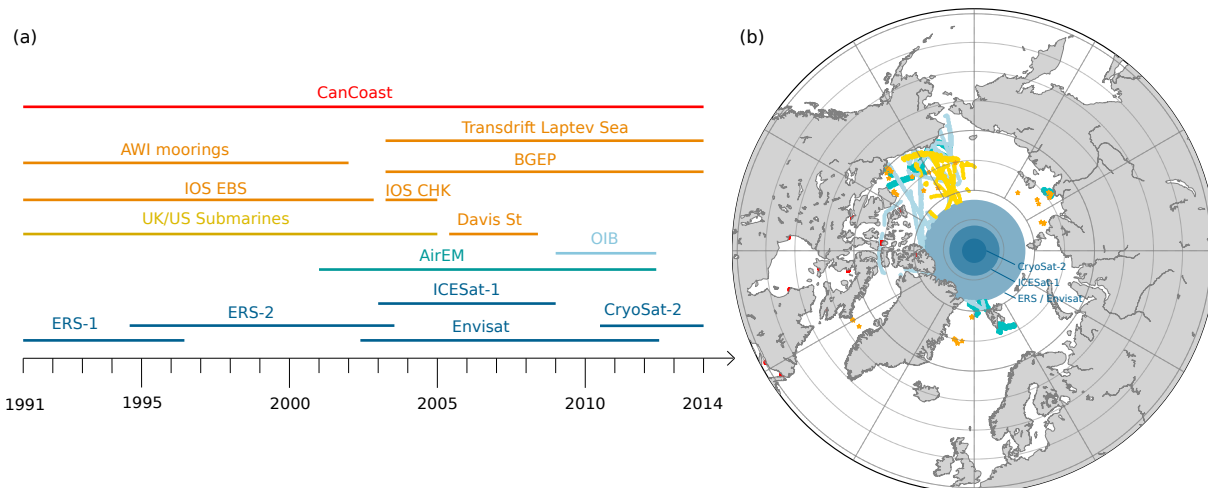
Figure 1. Summary of various available dataset for Envisat and ERS validation. Colors distinguish the different types of data. Dark blue for satellite products, light blue for airborne data, yellow for submarines, orange for anchored moorings, green for buoys and red for direct measurements. (a) Temporal availability (b) Spatial availability and extent of missions data gaps. Blue rounds represent altimeters coverage limitation due to their orbit inclination (81.5°N for Envisat, 86°N for ICESat-1 and 88°N for CryoSat-2)

SnowModel-LG (Liston et al., 2020a; Stroeve et al., 2020) is a snow depth product from a snow evolution model forced by different reanalysis, we use the version forced by ReAnalysis-5th Generation (ERA5) (Liston et al., 2020b). The data set is available from August 1, 1980, and July 30, 2018, in a 25 km resolution. Although the most commonly used product is still W99 climatology (Warren et al., 1999), it is no longer consistent for the recent period, as well as an altimetry-based
125 product such as ASD climatology (Garnier et al., 2021) would not be a relevant choice before 2010s, we then justify the use of SnowModel-LG to ensure the relevance of comparisons presented in Sect. 4.

2.3 Validation data

The results obtained in this study are compared with different independent data sets presented in this section. Comparisons are detailed in Sect. 4. Most of the following data sets are included in Lindsay and Schweiger (2013) data set and further described
130 in the corresponding publication Lindsay and Schweiger (2015). Data availability is summarized in Fig. 1.

2.3.1 Airborne

Operation Ice Bridge (OIB) was a mission led by NASA. It consisted of airborne measurement campaigns using scanning lidar altimeter and snow radar to measure both snow depth and ice thickness (Kurtz et al., 2013). The data we use are from the Unified Sea Ice Thickness Climate Data Record of ~~(Lindsay and Schweiger, 2013)~~Lindsay and Schweiger (2013), Operation
135 Ice Bridge Version 2 processed by ~~(Kurtz et al., 2013)~~Kurtz et al. (2013). These measurements were carried out between 2009 and 2013, during each early spring or early autumn near the coasts of the Canadian Archipelago and Alaska.

5

Airborne Electromagnetic Induction (AirEM) can measure total thickness (snow plus sea ice), the methodology is described in Haas et al. (2009). Air-EM data that are used in this study is provided by (Lindsay and Schweiger, 2013) and are available from 2001 until 2013 from to 22 campaigns in the Arctic Ocean and Fram Strait.

2.3.2 Moorings & Submarines

The following data are all measured with upward looking instruments that are installed either on anchored moorings or on board of submarines. These instruments measure the sea ice draft i.e. the height of the immersed part, from which can be derived the sea ice thickness to be compared with altimetry data.

The Beaufort Gyre Expedition Project (BGEP) is composed of a network of 4 moorings located in the Beaufort Sea (Krishfield et al., 2014). The moorings, equipped with ULS record drafts every 2s with a precision evaluated to ± 0.3 cm. Data are currently available from August 2003 to September 2018. The data were collected and made available by the Beaufort Gyre Exploration Project based at the Woods Hole Oceanographic Institution().

Belter et al. (2020) perform and diffuse a daily sea ice draft data set based on Upward-Looking Acoustic Doppler Current Profilers (ADCPs). These data are located in the Laptev Sea and are available from August 2003 to September 2016 located in the Laptev Sea().

The Institute of Ocean Sciences (IOS) provides two ULS draft measurement data sets named IOS-Eastern Beaufort Sea (IOS-EBS) and IOS-Chukchi Sea (IOS-CHK) (Melling, 2008). For IOS-EBS, data are available from April 1990 to September 2003 on a network of 9 sites in the Beaufort Sea. IOS-CHK is composed of data from a single site located in the Chukchi Sea between August 2003 and August 2005. The sea ice draft product for IOS-EBS comes from Melling (2008), and IOS-CHK from Lindsay and Schweiger (2013). Draft can be measured with precision about 0.05 m for young ice and can be overestimated up to 0.3 m for older and rougher ice.

Davis Strait sea ice draft (Davis St) product from anchored moorings has been detailed in Drucker et al. (2003). Data used in the study comes from the Unified Sea Ice Thickness Climate Data Record (Lindsay and Schweiger, 2013) and are available from 2005 to 2008.

Alfred Wegener Institute for Polar and Marine Research (AWI) have processed and distributed a sea ice draft data set consisting of data from 11 moorings in the Greenland Sea and Fram Strait (Witte, 2005). The data span from 1991 to 2002, and the draft is recorded with a 5 minutes frequency with an accuracy of ± 0.2m.

The last sea ice draft data set presented in this section is derived from data collected by both U.S. Navy and Royal Navy submarines in the Arctic Ocean from 1975 to 2005 (National Snow and Ice Data Center, 2006; Wadhams and Horne, 1980; Wadhams, 1984; Wensnahan, 2005). It gathers data from 39 cruises. According to Rothrock and Wensnahan (2007), sea ice drafts are estimated to have an overall bias of 29 cm and a standard deviation of 25 cm from the actual draft.

2.3.3 ~~Buoys and coastal~~ Coastal stations

Environment and Climate Change Canada compiled weekly measurements from 27 monitoring stations along the coasts of the Canadian Archipelago in one product named CanCoast. Measurement methods can vary from a station to another (boreholes,

170 hot-wire thickness gauges, etc.) but all stations provide at least sea ice thickness and snow depth estimation with an accuracy of less than a centimeter. ~~The data set is available at .~~

~~The CRREL-Dartmouth Mass Balance Buoy Program (IMB) gathers sea ice thickness and snow depth (on sea ice) data from more than 100 ice mass balance buoys released between 2000 and 2018 in the Central Arctic, the Beaufort Sea, the Canadian Archipelago and the Baffin Bay (Perovich et al., 2022). Sea ice thickness and snow depth are provided at a 2 hours frequency.~~

175 **2.3.4   Satellite Altimetry products**

Two satellite altimeter sea ice freeboard products have also been used for comparisons.

One of these is the product generated by the ESA Climate Change Initiative Program (CCI) for sea ice thickness estimation (Hendricks et al., 2018), which includes the whole Arctic sea ice covered region for all winters (October-April) of Envisat mission (2002-2012). It provides monthly grids of sea ice thickness, radar and ice freeboard combined with the related uncer-

180 tainties. The corresponding methodology is described in Paul et al. (2018).

ICESat-1 was a mission operated by NASA, was launched in 2003 and ceased operating in 2009. It was composed of a laser altimeter that allows retrieving the total sea ice freeboard (snow depth plus sea ice freeboard). ICESat-1 product provides estimations for 15 periods of about 30 days between February 2002 and November 2008 in a 25 km grid resolu-tion~~(Zwally et al., 2008; Yi and Zwally, 2009)~~, the version used is the NASA-Goddard one processed by Zwally et al. (2008); Yi and Zwall

185 . The ICESat-1 total sea ice freeboard measurement accuracy is estimated to be about 0.05 m.

~~Summary of various available dataset for Envisat and ERS validation. Colors distinguish the different types of data. Dark blue for satellite products, light blue for airborne data, yellow for submarines, orange for anchored moorings, green for buoys and red for direct measurements~~

3   Methods

190 **3.1   ERS pulse-blurring correction**

The values of the on-board tracker heights (h_{trk}) while ERS-2 flights over sea ice reveal instabilities of several meters that can not be explained by the sea surface topography. This phenomenon is better observed by computing the Surface Height Anomaly (h_{rtrk}) over all type of surfaces. This height anomaly is calculated according to Eq.(1), where alt is the altitude of the satellite, $range$ is the on-board tracker range, $epoch$ is obtained after the waveform retracking using TFMRA50, MSS

195 is the DTU15 Mean Sea Surface, (Andersen et al., 2016) and $geophysical\_corrs$ is the sum of all geophysical corrections. Figure 2, shows the TFMRA50 retracked height estimation for Envisat and ERS-2 along the collocated pass 25 respectively for cycles 12 and 80 (beginning of January 2003). Regarding the behavior of Envisat measurements, ERS-2 heights anomalies show instabilities of about 1 m that make the measurements unusable.

$$h_{rtrk} = alt - range - epoch - MSS - geophysical\_corrs \qquad (1)$$
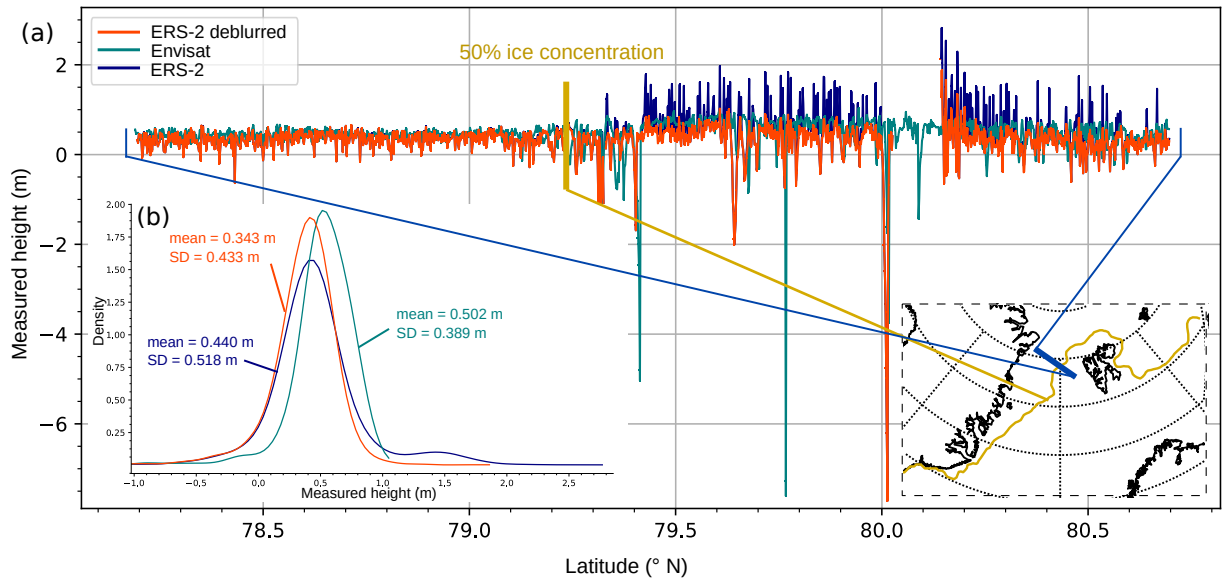
Figure 2. Profiles of surface height anomaly over sea ice and ocean for pass 25 between 78° N and 81° N for Envisat in blue-green (cycle 12), ERS-2 in blue and ERS-2 deblurred in orange (cycle 80). The red line represents the limit of 50 % concentration of sea ice, so as the limit between open ocean and an ice-covered area. The dark blue line shows the location of the pass between Svalbard Island and Greenland. (a) The surface height along the latitude and (b) the probability density function of surface height for the three passes with the associated statistics, the average and the standard deviation (SD). The color legend is identical for both sub-figures.

The instabilities of the height anomalies (mainly over sea ice) are known as "pulse blurring" and are a consequence of the on-board tracker settings. This phenomenon occurs for both ERS-1 and ERS-2 missions. A simplified version of the ocean mode tracking system is represented in Peacock (1998, p.71). The tracking system is composed of three tracking loops to maintain echoes within the radar acquisition window: the Height Tracking Loop (HTL), the Slope Tracking Loop (STL), and the Automatic Gain Control (AGC). The role of the HTL is to maintain the successive waveforms in the middle of the acquisition window. For this purpose, the tracking system is able to estimate the position of the tracker for each individual echo that composed the average sequence. The tracker position is therefore adapted at 1020Hz with a low-pass $\alpha\beta$ filter described by Eq.(2) and Eq.(3) according to the Height Tracking Loop error ε. T is the interval for the low-pass filter to be updated, $T = \dfrac{1}{PRF}$ s for the HTL and h_n the tracker height for the nth echo, with n between 0 and 49. In ocean mode, the algorithm used to estimate ε is a Sub-optimal Maximum Likelihood Estimator (SMLE). Nevertheless, the SMLE has been developed for Brown-like waveforms, that can be found over the open ocean, but is not suitable for specular waveforms found over sea ice.

$$h_n = h_{n-1} + \alpha\varepsilon + T\dot{h}_n \tag{2}$$

$$\dot{h}_n = \dot{h}_{n-1} + \frac{\beta\varepsilon}{T} \tag{3}$$
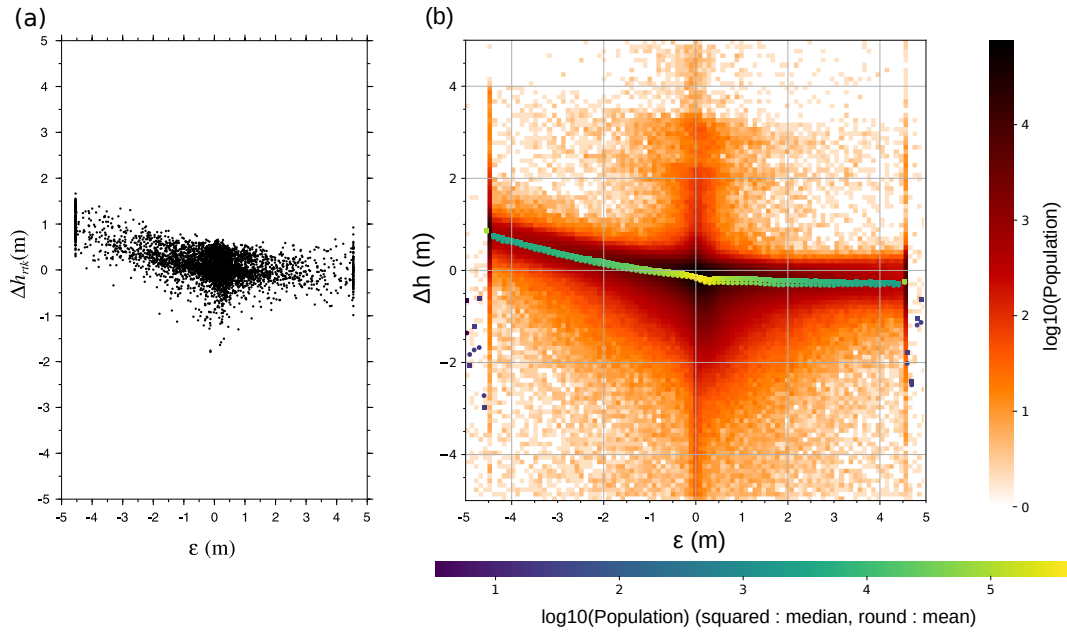
Figure 3. Difference between the raw surface measurements and the interpolated ocean level measurements as a function of the height error parameter ε. (a) is taking from Peacock and Laxon (2004) and (b) a reproduction for the Cycle 83 of ERS-2. On (b) squares and circles are respectively the median and the mean of Δh for each value of ε (on the x-axis).

Both Eq. (2) and Eq. (3) can be combined to give Eq. (4) which shows that the range window correction increases with a n^2 factor.

$$h_n = h_0 + nT\dot{h}_0 + n\alpha\varepsilon + \frac{n}{2}(n+1)\beta\varepsilon \qquad (4)$$

A high ε (due to an inappropriate sea ice height error estimation algorithm (Roca et al., 2009)) coupled with this low-pass filter could drive to the large variation of range window inside the same averaging sequence, which will "blur" the final averaged waveform echo due "the bad overlay between the individual echoes" (Peacock and Laxon, 2004). Because the height error is estimated at the end of each averaging sequence, a sudden change in the range window is then possible, especially since the error estimate is also affected by the pulse-blurring and can explain tracker height oscillations. The problem mainly comes from the choice of the SMLE to estimate the HTL error. Indeed, it has been elaborated for ocean-like waveforms with a long trailing edge contrary to peaky waveforms whose power decreases suddenly after the max power peak.

A methodology has been developed by Peacock and Laxon (2004) to deal with this issue. This method consists of finding a relation between the height error parameter ε and the difference between the measured surface (over an area covered by ice) and the "same area if it was not covered by ice" (Peacock, 1998). We interpreted it as the difference between the raw surface measurements and the interpolated ocean level measurements Δh.

Figure 3 illustrates that our interpretation of Peacock (1998) theory (on the right) fits with his results (on the left). Our results reproduce the linear relation between Δh and ε found by Peacock (1998) and Peacock and Laxon (2004) with a slope equal to $\dfrac{-1}{5}$ when ε is negative (Eq.(5)).

$$\begin{cases} h_{corr} = & h - \dfrac{\varepsilon}{5}, \quad \varepsilon \quad \leq 0 \\ h_{corr} = & h, \quad\quad \varepsilon \quad > 0 \end{cases} \tag{5}$$

This correction is applied and presented in Fig.2. The correction of the pulse-blurring effectively reduces the instabilities of measurements. The correction is similarly asymmetric so that the variations towards the positive height anomaly are more corrected than the others. The deblurred surface height anomaly of ERS-2 now appear more similar to Envisat in terms of noise and amplitude of variation. For this particular pass, the standard deviation has been reduced by 16 % and get closer to Envisat's one. Figure A1 shows more results on the impact of deblurring on ASA noise reduction comparing to Envisat during a whole cycle (80 for ERS and 12 for Envisat).

3.2 Along track radar freeboard retrievals

This section aims to describe the radar freeboard (FBr) processing chain. This procedure is common to all missions to preserve homogeneity and continuity.

The FBr is the difference between the sea ice surface height measured over floes and the sea surface height measured over leads. The use of satellite altimetry for these estimates was introduced by Laxon (1994); Laxon et al. (2003). Sea ice freeboard (FBi) can be derived from FBr after a correction of the lower wave-propagation speed into the snow layer (Kwok, 2014). In this study, we focus on the radar freeboard (without the speed propagation correction) to avoid the introduction of errors related to snow depth estimation.

It is therefore necessary to discriminate leads and sea ice floes. This essential step is based on the peakiness of the radar waveform (Eq. (6)) that quantifies the specularity of the surface (Laxon et al., 2003; Peacock and Laxon, 2004), we use the definition of Guerreiro et al. (2017) (Eq.(6)). Pulse peakiness thresholds can depend on the radar altimeter. To ensure the continuity between the missions Envisat and ERS-2, these thresholds have been adapted to keep the same lead/floe proportion during their common flight periods.period, see Tab.A1.

$$PP = \dfrac{max(WF)}{\sum_{i=0}^{Nb_{WF_{bins}}} WF_i} \tag{6}$$

Range are estimated using the TFMRA50 retracker for all surfaces and all missions mentioned here to maintain inter-missions continuity. Height anomaly measurements are then expressed relatively to the DTU15 Mean Sea Surface (Andersen et al., 2016) and corrected for the common geophysical corrections (oceanic, polar, solid earth and load tides as well as tropospheric and ionospheric corrections) according to Eq. (1). The tide model used is the FES14 from Carrere et al. (2015). ERS-2 the pulse-blurring correction (detailed in Sect. 3.1) must be applied at this stage of the freeboard computation.

Surface height anomaly is then split into two variables using the Pulse-Peakiness classification explained above, the "ice level anomaly" (ILA), over floes and the "sea level anomaly" (SLA), over leads. ~~Outliers are filtered out with a three standard deviation threshold along 25~~ ILA and SLA outliers are removed by filtering data that are outside the interval : rolling mean ± 3 rolling Standard Deviation, with a 60 km ~~sliding windows.~~ large sliding window. After filtering, ILA and SLA are ~~interpolated respectively over leads and floes and smoothed with a 25 rolling mean . The difference between the measured height over floes and over leads is finally made to retrieve the radar freeboard. For the remains of the~~ smoothed using a rolling mean at 12.5 km, then SLA and ILA are linearly interpolated (including bellow floes for SLA and above leads for ILA) and are again smoothed using a rolling mean at 12.5 km. No limit of distance is used to discard radar freeboard, but the interpolation as well as smoothing and filtering is not done between values separated by land. Indeed, the processing is done within ocean segments, separated by land, in order to isolate statistics between segments.

In this study, we will only use the FBr measurements ~~made above the floesbecause the characteristics of the waveforms are used.~~ that are made over floes, indeed, the LRM data calibration, explained in section 3.4, is based on floes characteristics.

3.3 Data gridding

The calibration (see Sect. 3.4) between the different missions is performed using monthly maps in EASE2 (Brodzik et al., 2012) with a 12.5 km resolution. The FBr gridding is done by averaging the values within a 25 km radius from each pixel weighted by the inverse of the uncertainty, for other variables not weighting is applied. Only sea ice concentrations above 50% are considered. To limit the outliers of LRM FBr at the sea ice-ocean boundary, data with ice concentration lower than 85% is removed when the waveforms have a LEW higher than 2.5 gates.

3.4 ~~Correction~~ Calibration of LRM ~~Freeboards~~ radar freeboards against ~~SAR/SARIn~~ SARM freeboards using neural networks

The radar freeboard maps obtained from the process presented in sect. 3.2 for Envisat and CryoSat-2 are shown in Fig. 4.

Important differences between Envisat and CryoSat-2 can be noticed, both in terms of patterns and in mean values. Negative radar freeboards are mainly due to the ~~choice of the~~ retracker choice. Indeed, a TFMRA50 ~~retracker for~~ is used to retrack heights on both leads and floes~~which can result in surface heights that are higher for leads than for floes. In LRM, most of this error comes from a constant bias of the Sea Level Anomaly (SLA) over leads~~, this introduces a bias on the height over leads. The TFMRA threshold to retrack heights over leads should be closer to 80% and the use of a 50% threshold corresponds to the position of the retrack point for ocean surfaces, not specular ones Poisson et al. (2018), the surface over leads is measured to be higher than it is and even higher than the surface over floes. The SLA bias (over leads) is evaluated constant for SARM altimeter in the study of (Laforge et al., 2020), this conclusion is also relevant for LRM altimeters as waveforms over leads are peaky and similar from a lead to an other. This positive constant bias over leads results to a negative bias on the radar freeboard. To avoid this bias, the retracker threshold could be adapted for leads or the SLA could be calibrated on CryoSat-2 one. Nevertheless, a threshold of 50% ensure the stability of the range (Poisson et al., 2018)[fig.9]) contrary to higher thresholds (80%-95%) that could lead up to 47 cm of random error on the SLA. A TFMRA at 50 % for both leads and floes is preferred in this study as a
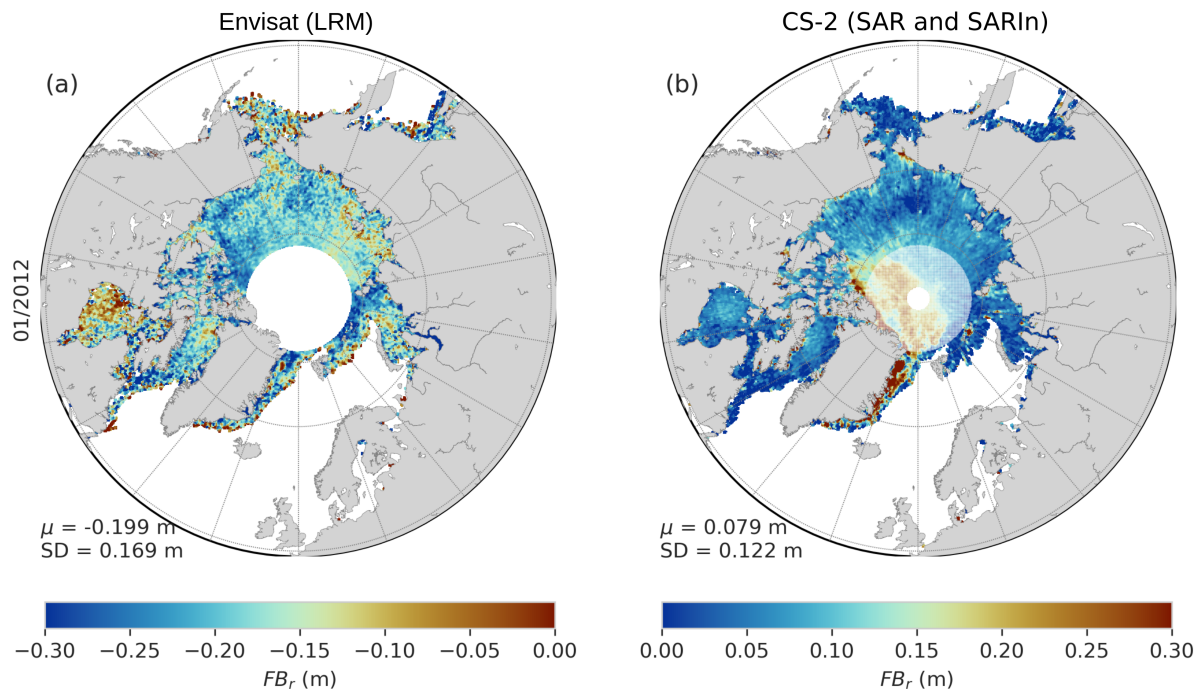
Figure 4. Pan-Artic radar freeboard maps for January 2012 for (a) Envisat uncorrected and (b) ~~Cryosat2~~ CryoSat-2

constant bias is easier to correct than an undetermined random error. Nevertheless, beyond this bias, a lack of representative sea
290 ice patterns can be observed, for instance, thick ice regions do not appear for the Envisat mission (cf Fig.4). This phenomenon
also appears in the ERS-2 FBr since it is also in LRM with the TFMRA50 retracker. This inconsistency comes from the LRM
itself (plus the retracker choice), which has a larger footprint than the ~~actual SAR processing~~resulting one in SARM. The larger
footprint leads to a large impact of the surface roughness on the reflected echo. It is therefore impossible to distinguish the
contribution of heights and roughness without a physical model of the effect of ~~ice roughness (such as the sea state bias on the~~
295 ~~open ocean). Indeed,~~ sea ice roughness on the reflected echo.

Paul et al. (2018) and Guerreiro et al. (2017) show that there is a significant correlation between the patterns of the parameters
characterizing the surface roughness and the differences between the FBr of CS-2 and Envisat. These observations led to the
development of calibration technics of Envisat relatively to CS-2 taking advantage of the missions-overlap period. In the same
way, we propose to calibrate ERS-2 against Envisat, once ~~corrected~~calibrated, using the common flight period between Envisat
300 and ERS-2.

So far, several empirical methods of Envisat freeboard correction using Cryosat-2 have been developed (Guerreiro et al.,
2017; Paul et al., 2018; Tilling et al., 2019). In Guerreiro et al. (2017), the ~~correction~~ calibration consists of finding a link
between Envisat and CS-2 freeboard differences and the PP of Envisat's waveforms. The correction proposed by Paul et al.
(2018), adapts the TFMRA threshold of Envisat waveform retracking according to the Leading Edge Width (LEW) of the
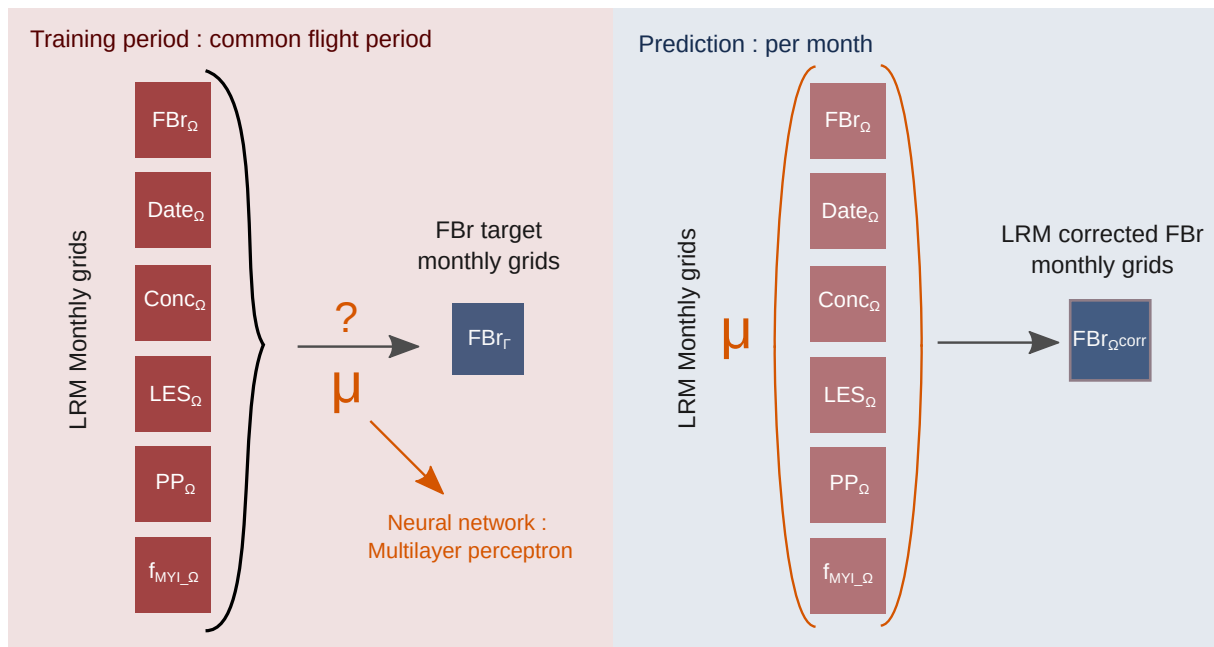
Figure 5. Diagram illustrating the principle of freeboard ~~correction~~ calibration by neural network with the two main steps: on the left panel, the neural network training phase, and on the right panel, the prediction (correction) phase. Ω corresponds to the inputs and Γ the output of the Neural Network. FB: radar freeboard, Conc : Sea ice concentration, LES: Leading Edge Slope, PP, Pulse Peakiness, f_{MYI}, MYI fraction.

305 waveforms and the surface backscatter. Tilling et al. (2019) suggests correcting the Envisat sea ice thickness (SIT) utilizing the distance between leads and floes. All methods are based on the comparison of monthly gridded data. The first two studies share the common approach of trying to correct the Envisat FBr using the surface roughness (characterized by one or more parameters as proxy). They both propose a 3rd degree polynomial function to link Envisat and CS-2 FBr differences with surface properties. Paul et al. (2018) was the first to propose to use two distinct parameters that characterize two roughness

310 scales to correct LRM measurements that can impact differently the waveform shape. Our method follows the same approach (Pulse Peakiness, Leading Edge Slope) with other additional parameters that define sea ice state such as ice concentration, ice type, season, etc. The Leading Edge Slope refers to the LE-height divided by the LE-width computed between 30% and 70% of the max power of the first waveform peak. The LRM correction model procedure is based on a neural network in order to manage strong non-linearities. The procedure is illustrated on Fig. 5.

315   For each mission to be calibrated, the neural network (NN) is trained on the common flight period between the mission considered as a reference and the one to be ~~corrected~~calibrated. It takes as inputs monthly grids of the following parameters: LRM FBr (to be corrected), pulse peakiness, leading-edge slope, ice concentration, MYI fraction, and the period and as a target, the FBr of reference (SARM FBr or ~~LRM-corrected~~ LRM-calibrated FBr). Inputs and targets are standardized before the training step.

320 The neural network used is a multilayer perceptron ~~regressor~~ (MLP)~~composed by~~. Both calibrations have been processed with Scikit learn (Pedregosa et al., 2011). The MLP is composed of 5 hidden layers, each composed of 100 neurons. The ~~activation function used is a sigmoid.~~

~~Hyper-parameters have been tuned by dichotomy by choosing at each step the hyper-parameter combination with the~~ ~~highest mean score (average score made on 5 models) on the test sample. The score used for this regression is the Pearson~~
325 ~~correlation coefficient. To determine the most suitable hyper-parameter combination, the dataset is randomly split into a~~ ~~training and a testing dataset, corresponding respectively to~~ choice of hyperparameters : number of neurons, the learning rate, the regularization term, batch size, activation functions, solver for the weights optimization, have been done using a grid search method. The evaluation criterion, called the score, is chosen as the determination coefficient. Models are trained on 90~~% and~~ ~~10~~% of the ~~initial dataset. To avoid overfitting, we use early stopping to interrupt the training when~~ data set and tested on the
330 remaining 10%, the splitting in random. During the tuning step, models are cross validated, it means that they are each trained 5 times with the same combination of hyperparameters but without the same train/test data set, the 5 scores are then analyzed to determine the best combination. Cross validation give a better idea of the model performance as the dependence to the training data set is limited. The activation function for the hidden layers neurons is a sigmoid, motivated by possible negative radar freeboard values, and the optimizer is and ADAM (Kingma and Ba, 2014). Moreover, in order to avoid over-fitting, an early
335 stopping criterion is used to stop the model training as soon as the score is not ~~improving anymore. Once the hyper-parameter~~ improved during 10 consecutive iterations, with a defined tolerance.

Finally, once the hyperparameters combination is set, the MLP is trained ~~with the whole dataset. The NN trained~~ on the whole data set to provide the calibration function. The trained model is then applied to the LRM monthly grids to obtain a monthly LRM-corrected radar freeboard.

340 ## 3.5 ~~Freeboard~~ Radar freeboard uncertainties quantification

This section aims to estimate the uncertainties for Envisat and ERS-2 FBr estimations. The uncertainty budget is split into two steps corresponding to the two main parts of the freeboard processing chain. The first step covers the along-track processing up to the gridding which is common to all missions, and the second step concerns the correction of the LRM freeboard which consists of predicting the FBr-corr with the neural network.

345 The uncertainty budget methodology concerning the first part is taken from Landy et al. (2020) and Ricker et al. (2014). We assume that for this step there are 3 sources of uncertainty. Two of them are random uncertainties: the speckle noise, largely discussed in Wingham et al. (2006) and the accuracy of the Sea Level Anomaly (SLA) measurement. The last one is linked to both retracker choices, surface roughness and snow radar signal penetration (Ricker et al., 2014).

According to Wingham et al. (2006), the speckle noise generates an ~~error~~ "error" (σ_{l1}) from 7 to 14 cm (depending on
350 the acquisition mode) on the range measurement for CryoSat-2 altimeter. Estimations of speckle noise ~~error~~ "error" on range for other missions can be ~~rely~~ relied on the individual echoes number used to compute the averaged waveforms (for LRM sensors). Indeed, the speckle noise generates an ~~error~~ "error" on the range that function of $\sqrt{n}$, where n is the ratio between the number of individual pulses used for each averaging sequence for CS-2 and for the mission we want to estimate the

14

Table 1. Summary of the "range error" for various missions governed by the speckle noise (Wingham et al., 2006)

Mission (RA mode)	σ_{l1} (m)
Cryosat-2 (SAR)	0.10*
Cryosat-2 (SARIn)	0.14*
Cryosat-2 (LRM)	0.07*
Envisat	0.068
ERS-2	0.096

* Wingham et al. (2006)

~~error (Calafat et al., 2017).~~ "error" (Calafat et al., 2017) (see Tab. A2 for LRM RA characteristics). All final σ_{l1} values are summarized in Tab. 1. Note that Wingham et al. (2006) uses the terminology "error" but it refers to uncertainty.

SLA uncertainty (σ_{SLA}) estimation ~~depends on the surface type (leads or floes). For leads, we take~~ is taken from (Ricker et al., 2014). The uncertainty on the SLA is estimated to be the standard deviation of the ~~measured height~~ SLA within a sliding window of 25 km ~~window. Concerning floes, the uncertainty is estimated~~ if there are some leads within this window. If not, the SLA uncertainty is taken as the difference between the ~~height measurement and the mean elevation along the track (Ricker et al., 2014)~~ interpolated and smoothed SLA and the mean SLA computed as the mean of raw SLA measurement at leads within a segment of ocean (if the pass is over land, statistics are made segment of ocean by segment of ocean). Finally, we consider that σ_{SLA} and σ_{l1} are not correlated and can be combined to give the random part of radar freeboard uncertainties (σ_R) following Eq.(7).

$$\sigma_R^2 = \sigma_{SLA}^2 + \sigma_{l1}^2 \tag{7}$$

~~Summary diagram of the uncertainty budget from along track to the propagation by the neural network~~

The radar freeboard (including uncertainties) gridding methodology is taken from Ricker et al. (2014, Sect. 2.4) in order to take into account the random uncertainties in the radar freeboard gridding process.

~~Landy et al. (2020) decomposed it in two~~ In Landy et al. (2020), the FBr systematic uncertainty budget is decomposed in two parts, on the one hand, the uncertainties due to the penetration of the signal in the snow (depending on its salinity or if it is composed of metamorphic snow, according to the type of ice) and in the other hand, the surface roughness. ~~Surface roughness is identified as the largest source of uncertainty and we, therefore , choose to consider only this source in our systematic uncertainty evaluation. Roughness is~~ We assume, as in Ricker et al. (2014), that the comparison of the freeboard from different retrackers does not enable to separate the contribution of the roughness from the signal partial penetration. We therefore assume to consider both sources as one mixed contribution, estimated to be respectively about 20 % and 30% of the sea ice thickness for FYI and MYI (Landy et al., 2020). The systematic uncertainties can be underestimated as the penetration of the radar waves in the snow uncertainty may be poorly handled. Note that this systematic uncertainty budget only ~~concern~~ concerns CS-2 mission which ~~are~~ is afterward propagated to Envisat and ERS-2, indeed other ~~mission~~ missions will be "~~corrected~~calibrated" from surface roughness effect during the calibration procedure.

Therefore, we combined random and systematic uncertainties with a quadratic sum to have the total radar freeboard uncertainty in a grid as well as for the related radar freeboard estimation. The uncertainty of the other inputs (LES, PP, sea ice concentration, MYI fraction), is considered to be, for each grid cell, two times the standard deviation of the measurements used to calculate the average value (grid cell value) divided by the number of tracks passing through the corresponding grid cell.

As explained in sect. 3.4, the LRM radar freeboard correction is predicted by a neural network. The uncertainty propagation through the neural network is not straightforward since the inputs and outputs of the model are not linked by a mathematical relation. We have chosen to estimate the propagation of uncertainties through a Monte Carlo approach. This method allows to propagate uncertainties through non-analytically represented systems. Nevertheless, the Monte-Carlo method requires a representation of the distribution of the input and target parameters. As discussed above, there is no thorough knowledge of the distribution of the input data for the NN.

The method consists ~~of~~ in training a number M of NN with noisy inputs~~. The~~ noise has been added to all inputs (for each grid cell and each month) according to a Gaussian distribution ~~), and then to analyze~~ centered on the estimated value and the corresponding uncertainty as standard deviation. The calibration processing (training and prediction) is done for all the noisy inputs/output, then the distribution of MxN radar freeboard predictions (from the M noisy NN ~~applied on noisy N inputs for each considered grids~~models applied on N noisy inputs) has been analyzed of each grid cell and each month. The whole uncertainty budget process is summarized in Fig. 6.

~~This processing provides the distribution of the M*N predicted FBr for each grid cell for each month, from which it is possible to get information on the uncertainties.~~

Unfortunately, we have not been able to identify known distributions such as normal, log-Normal, gamma, etc., for ~~FBr output~~ predicted FBr distributions. However, we can still derive various statistics, such as the quantiles at 2.5% and 97.5%, which represent 95% of the values or the standard deviation to describe the FBr distribution ~~of~~ for each pixel of all monthly grids.

4 Results

This section first presents the calibration performance when applied to Envisat with respect to CS-2 and thereafter to ERS-2 with respect to Envisat. After this we present the Envisat and ERS-2 freeboard comparisons against independent validation data sets.

4.1 Calibration performances

The calibration method presented in Sect. 3.4 was successively applied to Envisat (Env) with respect to CS-2, and then to ERS-2 with respect to Envisat (Env) ~~corrected~~ calibrated estimation. Figure 7 compares Envisat and CS-2 radar freeboard during December 2010 and April 2011. Figure 8 presents the same feature for Envisat and ERS-2 radar freeboard ~~.~~during December 2002 and April 2003.

As can be noticed on these maps, Envisat calibration allows recovering typical patterns of the Arctic sea ice with thick ice near the coasts of the Canadian Archipelago and Greenland, and thinner ice on the eastern part of the basin.

Nevertheless, compared to CS-2, the radar freeboard of thick ice is ~~lightly~~ slightly underestimated, and thin ice is ~~lightly~~ slightly overestimated. However, the mean difference between the two products is close to zero (1 or 2 mm) for both months. We also note that the calibration results in an asymmetric distribution, with a tendency towards a log-normal distribution at the basin scale, whereas the CS-2 distribution for the same mask is centered and appears ~~more Gaussian~~Gaussian-like.

Radar freeboards of ERS-2 and Envisat (see Fig.8) are even more similar. Again, the bias is negligible, and the standard deviation (SD) of their difference is two to three times lower than for Env/CS-2, with a SD of 0.03 m and 0.05 m.

The better performance of the calibration for ERS-2 can be explained ~~by the fact that~~ because ERS-2 and Envisat carried a similar LRM altimeter, and they flight on a same orbit with 28 minutes delay between the two missions which allow them to observe nearly the same surface. On the other hand, CryoSat-2 operates in SARM and flies on a quasi-polar orbit with cycles and sub-cycles very different from Envisat ones. These comparisons show that the neural network calibration gives satisfying results, at least during the common flight periods. The calibration function is then applied to the 10 years of Envisat mission and 8 years of ERS-2 one.

Table A3 and Table A4 ~~represent respectively~~ respectively represent statistics for Envisat/CryoSat-2 and ERS-2/Envisat radar freeboard for each month of missions-overlap ~~period~~periods. For both calibrations, averaged radar freeboard are ~~closed. The higher mean difference is~~ close. The highest mean difference reaches 7 mm ~~and concerns~~ in February 2011 ~~, the Envisat calibration. For the~~ for Envisat calibration, i.e. 9.5% of the mean Envisat radar freeboard. For ERS-2 calibration, the mean freeboard difference ~~with~~between ERS-2 and Envisat does not exceed 3 mm, 3.3% of ERS-2 mean radar freeboard. Concerning all the overlap times, the mean difference is 3 mm for Env/CS-2 calibration, 4.1% of Envisat mean radar freeboard and -2 mm for ERS-2/Env one, about 2.2% of ERS-2 mean radar freeboard.

In both Fig.7 and Fig.8, the uncertainties presented for Envisat and ERS2 are in fact two times the standard deviation of the respective ~~corrected~~ calibrated radar freeboard distributions as output of Monte Carlo simulations for each pixel grid. Detailed statistics for uncertainties are also provided in Table A3 and Table A4.

The median uncertainty on the radar freeboard for the period 2010/2012 is ~~0.063~~6.3 m for Envisat and ~~0.020~~2 cm for CS-2, regardless of the month, the mean and median uncertainties of Envisat are always larger than those of CS-2. ~~Similarly, for~~ Concerning the period 2002/2003, the ~~mean and median uncertainties of ERS-2 are always larger than those of Envisat by about 6~~ median uncertainty is 8 cm ~~over the median radar freeboard~~ for ERS-2 radar freeboard and 7.3cm for Envisat, similarly, statistics on uncertainties are globally higher for ERS-2 estimates (see Tab.A3 and Tab.A4 for ~~more~~ detailed statistics).

4.2 Validation

In this section, the calibrated Envisat and ERS-2 freeboard are evaluated against a large set of independent data. These data are presented in Sect. 2.3 and include in situ, airborne and space-based measurements providing sea ice freeboard, draft or thickness. Data have been converted to sea ice thickness to make it comparable, except for SI-CCI Envisat product which also

provides FBr. The conversions are based on the hydrostatic balance assumption of sea ice covered by snow in sea water that is described in Eq.(8).

$$SIT = \frac{\rho_w}{\rho_w - \rho_i} \cdot FB_i + \frac{\rho_s}{\rho_w - \rho_i} \cdot h_s \tag{8}$$

445 The snow depth (h_s) is taken from the data set itself if given (e.g. OIB, CanCoast~~, IMB) . Otherwise~~) with a constant snow density, $rho_s = 300\mathrm{kg \cdot m^{-3}}$. Otherwise, the SnowModel-LG, ERA5 version is used for the snow load. The water density value used is $\rho_w = 1024\mathrm{kg \cdot m^{-3}}$, ~~$\rho_s = 300\mathrm{kg \cdot m^{-3}}$,~~ $\rho_i^{FYI} = 917\mathrm{kg \cdot m^{-3}}$ and $\rho_i^{MYI} = 882\mathrm{kg \cdot m^{-3}}$ (Alexandrov et al., 2010). The sea ice density depends on the MYI-fraction within the grid cell, such as described by Eq.(9).

$$\rho_i = (1 - f_{MYI})\rho_i^{FYI} + f_{MYI}\rho_i^{MYI} \tag{9}$$

450 For Ku-band measurements, the speed reduction of the wave in the snow layer is taken into account to obtain the ice freeboard $FB_i = FB_r + h_s(\frac{c}{c_s} - 1)$ (Mallett et al., 2020) with c the speed of light in vacuum and c_s the speed of light in the snow estimated to be $c_s = c(1 + 0.00051\rho_s)^{-1.5}$ determined by Ulaby et al. (1986)~~, as we consider ρ_s constant~~.

The different data sets are then gridded into monthly EASE2 grids of a 12.5 km resolution, as to facilitate the comparison to Envisat and ERS-2 monthly grids. ~~Static data are monthly~~ Concerning moorings or coastal measurement stations, data are
455 averaged to get one value per month.

4.2.1 Envisat

Although numerous data sets are available during Envisat flight period, the spatio-temporal coverage of the Arctic basin remains very patchy. The following comparisons are presented with several types of sensors to reinforce the relevance of the validation. The consistency and discrepancies are discussed in the following.
460 Figure 9. gathers comparisons between Envisat and different data sets coming from airborne, space borne, submarines, drifting buoys and fixed coastal stations, Fig.10 presents comparisons with fixed moorings.

Comparisons presented in Fig. 9 provide satisfying statistics, with correlation values between ~~0.40~~ 0.41 (CanCoast) and 0.71 (Envisat SI-CCI). Correlation between Envisat and OIB is in good agreement with Kurtz et al. (2014) that showed that the auto-correlation of OIB varies from 0.46 to 0.60. Nevertheless, ~~IMB and AirEM~~ Air EM data are poorly correlated with
465 Envisat and statistics reveal high bias and RMSE (up to 1 m). Disregarding CanCoast and space borne estimations, comparisons reveal negative bias with Envisat from ~~-0.37 to -0.25~~ -3.9 cm to -2.9 cm, that could suggest an underestimation of Envisat sea ice thickness. The relevant statistics with CanCoast, with a bias of ~~8~~ 8.9 cm and a RMSE of ~~0.62~~ 6.4 cm suggest that this underestimation could not only be attributed to Envisat but maybe also to snow depth or other parameters. However, bias remains within the estimated range of uncertainty. The bias between OIB and Envisat estimation ~~can~~ could also be attributed
470 to the OIB snow depth ~~estimation that remains slightly different from one algorithm to another (Kwok et al., 2017)~~ which estimation seems sensitive to the algorithm used (Kwok and Haas, 2015; Kwok et al., 2017).

The comparison with ICESat-1 data reveals a strong dispersion with a low bias of ~~21~~ 20 cm and a correlation of 0.50. Envisat radar freeboard product established in the framework of the CCI and the version presented in this study are coherent with a bias close to zero, a standard deviation of the difference of ~~5~~ 5.2 cm and a fairly high correlation of ~~0.74~~ 0.71. Both Envisat data sets are consistent.

The new Envisat solution is also compared ~~with several anchored~~ to several moorings (BGEP, IOS, Davis strait and Transdrift Laptev sea) data sets (Fig. 10). The four campaigns yield fairly high correlations with Envisat estimates, greater than or equal to ~~0.63~~ 0.62 and reasonable standard deviations of ~~40-50~~ 50-60 cm, ~~except~~ down to 32 cm for the Laptev Sea estimates from Transdrift~~, which fall to 26~~. While Envisat has ~~a positive bias of 4~~ low negative bias of -9.2 cm with respect to BGEP, other campaigns biases are largely negative, with values from ~~-37~~ -53 cm to ~~-63~~ -76 cm.

Figure 10 (b) also provides a comparison between Envisat SI-CCI version and BGEP similarly to Fig. 10 (a). ~~All~~ With respect to BGEP estimations, the statistics are slightly ~~worse for SI-CCI than for~~ better for the product presented in this study ~~with respect to BGEP estimations~~ than for SI-CCI's. Nevertheless, the two products seem to be relatively consistent with each other.

4.2.2 ERS-2

The ERS-2 freeboard are also compared with measurements from AirEM, US and UK submarines and CanCoast in Fig.11.

The results of ERS-2 validation are close to those of Envisat. Figure 11 shows similar strong discrepancies with ~~AirEM~~ Air EM, but even better results for CanCoast and submarines comparison in terms of correlations (respectively ~~0.6 and 0.72 instead of 0.4 and 0.57~~ 0.60 and 0.74 instead of 0.41 and 0.55 for Envisat) and standard deviations of the differences are equivalent to those found for Envisat (62 and ~~58~~ 57 cm compared to ~~62 and 51~~ 63 and 53 cm). The comparisons between Envisat and the mooring illustrated in Fig.12 are also relevant, with correlations of 0.67 and ~~0.76~~ 0.66. Similarly to Envisat comparisons with ULS data reveal non-negligible negative biases. As a comparison, the bias between CryoSat-2 and OIB between 2010 and 2019 is about 16 cm and the RMSE about 77 cm. Concerning CS-2/BGEP comparisons, the bias is 21 cm with the same overestimation of FYI thickness for CS-2 and with Transdrift Laptev Sea comparisons shows a negative bias of -38 cm.

Although draft measurements by moorings are among the most accurate measurements (they measure 90% of the total thickness), they will tend to overestimate the true thickness when the ice bottom surface is rough, which is inherent to the method. Indeed, the chaotic aspect of the lower surface of sea ice can impact the ULS returned echoes. This strong deformation concerns mainly the thick and ~~old~~ rough ice, which can explain the tendency of ULS measurements to overestimate thick ice relatively to Envisat and ERS-2. Eventually, the methods used for above comparisons can be questioned insofar as monthly averages are compared with punctual measurements (spatially and temporally), which may indeed induce biases (e.g. OIB, AirEM, etc.). Conversions from radar freeboard to sea ice thickness can also be suspected to bias comparisons, as it depends on the snow depth product~~and the assumption of constant snow density can be questioned~~. SnowModel-LG has been chosen because it is the only continuous data set at the Arctic-bassin scale product that covers more than the 27-year required in this study.

4.3 Radar freeboard volume time series from 1995 to 2021

The time series represented in Fig. 13 is derived from monthly maps processed as developed in Sect. 3 and validated in Sect. ~~2.3~~4.2. Nevertheless, even after gridding process, missing values can occur especially where tracks density becomes low (e.g. closed to ice-ocean boundary). To ensure comparable volumes from a month to another, missing data have been replaced by ~~estimated~~ interpolated values from Gauss-Seidel relaxation method implemented in pangeo-Pyinterp (https://github.com/CNES/pangeo-pyinterp), a python library developed by the CNES (French National Space Agency). It's important to note that the volumes presented in Fig. 13 are only considering value up to 81.5 °N (ERS-2 and Envisat orbit limitation).

~~Radar freeboard volume have~~ Monthly radar freeboard volume has been finally estimated ~~according to~~ with Eq. (10), ~~with $Surface$, the pixel~~ by summing radar freeboard volumes within the P cells, with S_i, the grid cell area (12500^2 m^2), ~~FB_r the~~ FB_{ri} and SIC_i, the individual radar freeboard and ~~SIC~~ sea ice concentration from NSIDC (introduced in Sect. 2.2) for each grid cell. We have decided not to convert estimations to sea ice volume to limit bias coming from snow depth estimates.

The evolution of the snow load is not taken into account in Fig. 13, which means that the evolution of the volume is not fully represented, in the same way as if the total volume were derived with a snow depth climatology. Indeed, a decrease in FBr volume may merely indicate that the snow depth is greater and the ice thickness unchanged.

$$V_{rFBr} = Surface \sum_{i=1}^{P} S_i \cdot FB_{ri} \cdot SIC_i \tag{10}$$

~~Two~~ Three trend lines fitted on winter mean volumes are represented in Fig. 13, one is computed considering Envisat and CS-2 estimates only (dashed line)~~and the other~~, a second one from all missions estimations (solid line) ~~. Both trends are negative, respectively -6.27km^3/year~~ and a third one with a FBr climatology from 1995 to 2021 (dashed-dot line). Trends are performed using the Theil-Sen estimator (Theil, 1950; Sen, 1968) using SciPy (**?**) provided with uncertainties on the regression given with a 95% confidence interval, along with a statistical significance Mann-Kendall test (Mann, 1945; Kendall, 1948). All trends are decreasing, respectively -4.57 ± 8.73km^3/year, -9.11 ± 5.16km^3/year and ~~-9.86km^3/year, the~~ -1.80 ± 0.42km^3/year. The negative trend is strengthened with the integration of ERS-2 ~~measures.~~ estimates and turns statistically significant. The uncertainties from regression remain high but are induced by the large inter annual variability. The trend in mean winter FBr volume obtained from a FBr climatology between 1995 and 2021 is lower but more reliable, suggesting that FBr variation over the last 30 years dominates FBr volume variations in contrast to sea ice concentration. (In comparison, the Ob river mean discharge is about 400km^3/year, if the radar freeboard volume is roughly converted to total volume with a factor 10, the Arctic sea ice decline rate (up to 81.5° N) is about a quarter of the Ob mean annual discharge).

5 Conclusions

This study presents a methodology to recover the freeboard of ERS-2 through the pulse-blurring effect correction and its calibration over the Envisat and CryoSat-2 missions. The pulse-blurring correction is based on Peacock (1998) and Peacock and

535 Laxon (2004) approach. The calibration function developed in this study relies on a multi-layer perceptron, which is trained during the common flight period between ERS-2 and Envisat missions, using the Envisat radar freeboard as a reference. To ensure consistency along the three altimeter time-series, Envisat radar freeboard has been preliminary calibrated against CryoSat-2 using the same neural network. The choice of the threshold retracker TFMRA to process the waveforms was motivated for continuity purpose as it can be used for all radar altimetry missions.

540 The uncertainty estimation is initially tackled, referring to previous studies of Ricker et al. (2014) and Landy et al. (2020). These uncertainties are then propagated through the neural network thanks to a Monte Carlo approach. Uncertainties of ~~LRM-corrected~~ LRM-calibrated radar freeboard run from a few millimeters up to about 15 cm depending on the ice type and the density of the along-track measurements in the grid cell. One of the limitation of the uncertainty budget is the potential underestimated impact of radar penetration within the snow layer that could lead to an underestimation of the radar freeboard 545 uncertainty.

Envisat calibration results show good consistency with CryoSat-2 estimation, with a mean bias of 3 mm for both common winters and a SD of 9.8 cm. The Envisat radar freeboard was then compared to a large sample of validation data. 9 of the ~~11~~ 10 data sets give consistent results, especially the strong correlation with the moorings (0.63 to ~~0.95~~0.94) and CanCoast stations. Apart with CanCoast, these results are nearly systematically negatively biased, suggesting an underestimation of the radar 550 freeboard. A part of the latter bias is probably due to the draft measurement's method, suggested by the increase of the bias depending on the thickness of the ice. In any case, these biases remain within the range of estimated uncertainties. The result is also consistent with the solution proposed by SI-CCI. Even if the main purpose of this study is to extend the radar freeboard time series to ERS-2, it is nevertheless fundamental to ensure the reliability of ~~the reference that Envisat represents~~Envisat as a reference.

555 ERS-2 ~~corrected~~ calibrated radar freeboard is close to Envisat ~~corrected~~ calibrated ones with a correlation of 0.88, a bias of 2 mm and 3.8 cm of standard deviation for the difference. Comparisons with the few sets of in-situ data reveal the same positive bias with CanCoast as for Envisat, and high negative biases for submerged draft measurements. Except for Air EM, the comparisons provide consistent correlation values between ~~0.60 and 0.76~~0.55 and 0.74. Indeed, these statistics are those expected for comparisons between measurements from different technologies (airborne laser, ULS moorings, etc.), recording 560 different physical quantities (draft, radar freeboard) with different spatial and temporal availability (point, monthly averages). Unless the comparison methodology is reconsidered, it seems difficult to obtain better correlations.

This work finally allows reconstructing 27 years of Arctic sea ice freeboard up to 81.5 °N and suggests a decline in sea ice radar freeboard volume of ~~almost 10km^3/year~~ 9.11 ± 5.16km^3/year (about an order of magnitude more for total volume). This decline is significantly greater than considering only Envisat/CryoSat-2 period. Radar freeboard variations have 565 a predominant influence on volume variations but also on the trend, in contrast to sea ice concentration that seems to have a moderate impact. In the near future, the methodology will be extended to ERS-1 mission as well as for Austral sea ice to recover 30 years of sea ice volume variation for both ~~hemisphere~~hemispheres. This extended data will also be freely available to the community at large.

Figure 6. Summary diagram of uncertainty budget during along track, gridding and calibration steps. Top left panel corresponds to the along track to grid uncertainty budget. Top right panel defines the notations, for the Monte Carlo procedure : Ω for the Neural Network input parameters, Γ for the Neural Network output parameter (radar freeboard) with σ_Ω and σ_Γ, the corresponding uncertainties. The middle panel corresponds to the training of M models with noisy inputs and outputs. Bottom panel show the predictions of the N noisy input with the M trained neural network. γ is the predicted radar freeboard estimation for one pixel of the MxN predictions. M=100, N=200.
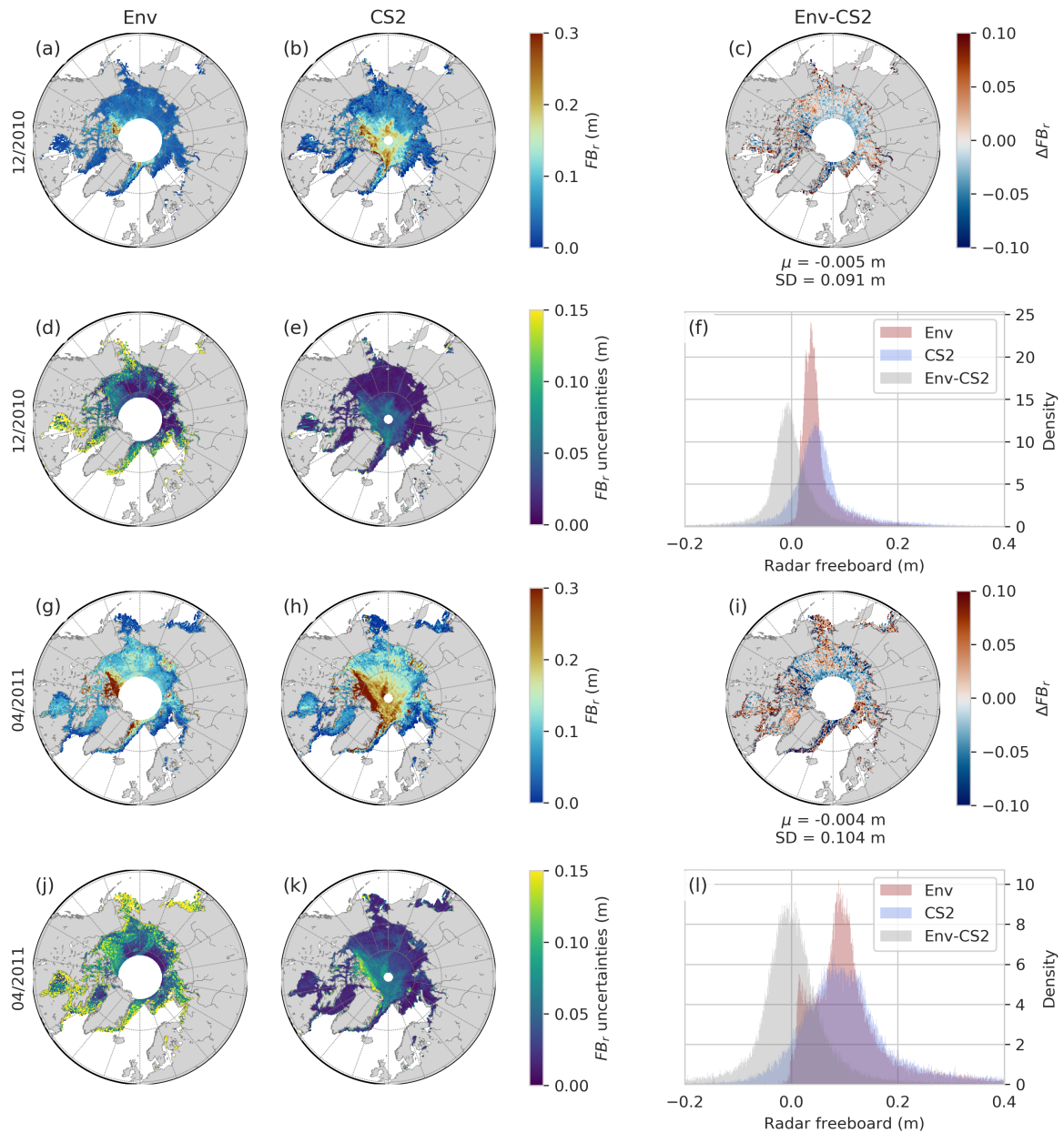
Figure 7. Comparison of Envisat calibrated radar freeboard against CryoSat-2 ~~reference~~ for December 2010 in the upper half and April 2011 in the lower half. The maps (a) and (g) ~~refers~~ refer to Envisat aside with the corresponding CryoSat-2 radar freeboard (b) and (h). Maps bellow (d), (e), (j) and (k) are the related uncertainties. The right column presents ~~differences~~ freeboard difference maps (Env-CS2) ((c) and (i)). (f) and (l) are the distribution of Envisat FBr in red, CryoSat-2 FBr in blue and ΔFBr in grey. Histograms only include common data between Envisat and CryoSat-2, data north or 81.5 °N are excluded. μ refers to the average and SD to the Standard Deviation.
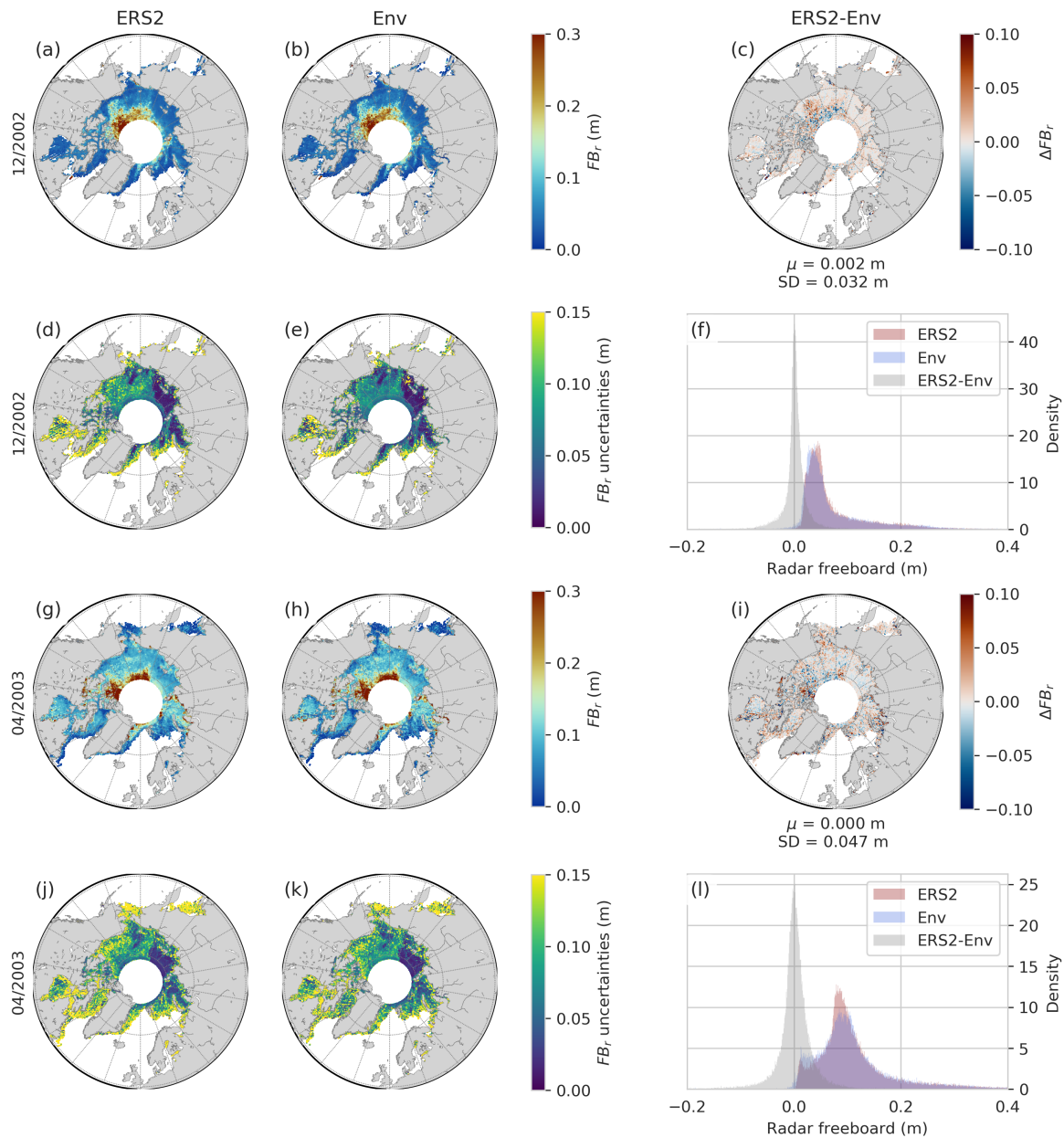
Figure 8. Same as Figure 7 but for ERS-2 and Envisat during December 2002 and April 2003. Histograms only include data for the coinciding region between ERS-2 and Envisat. μ refers to the average and SD to the Standard Deviation.
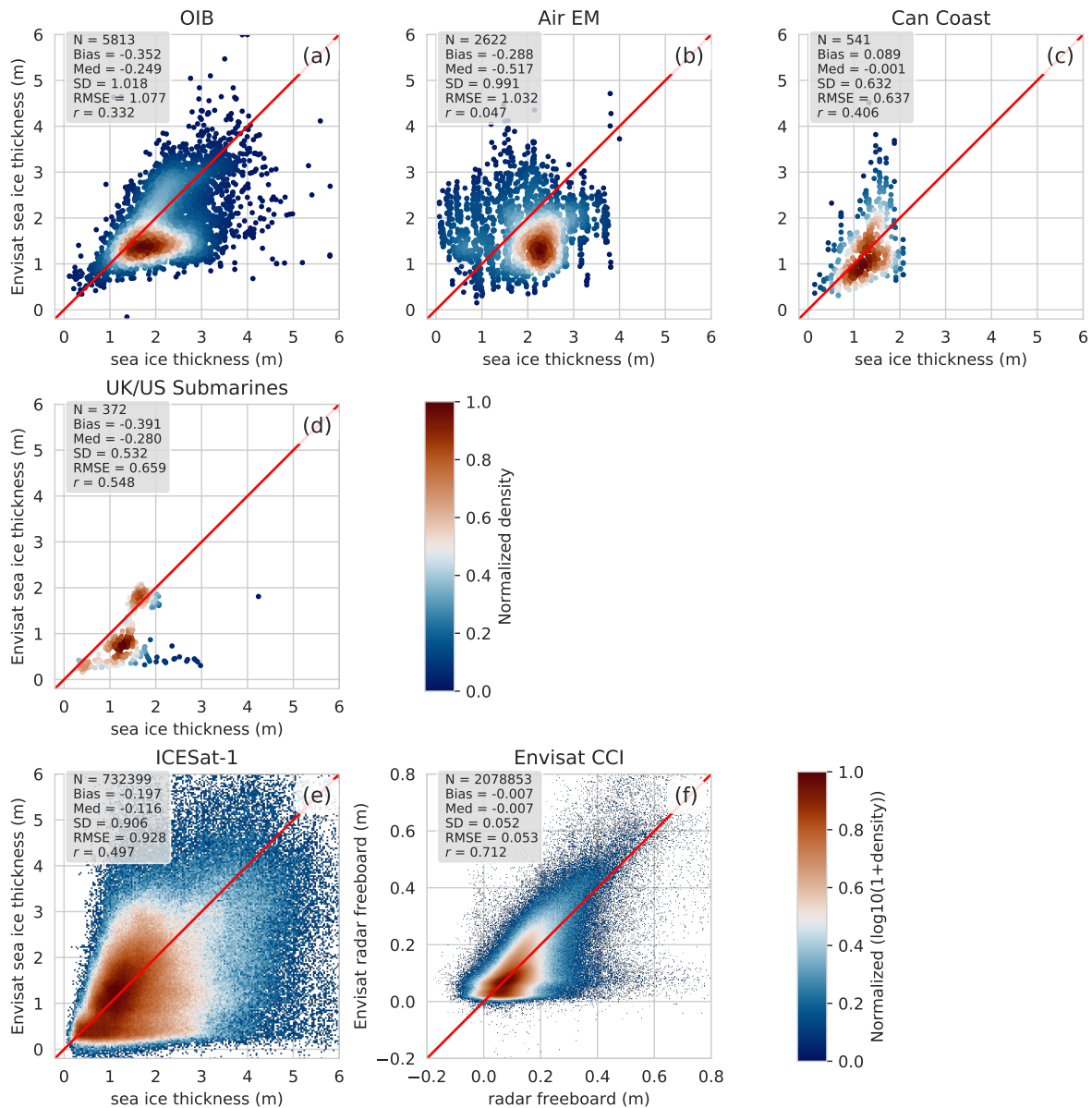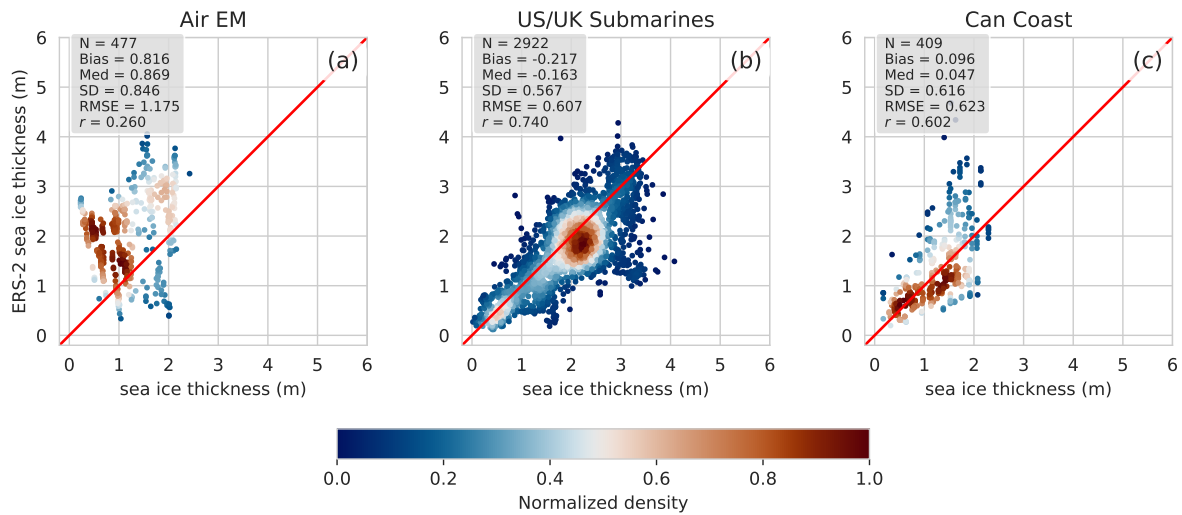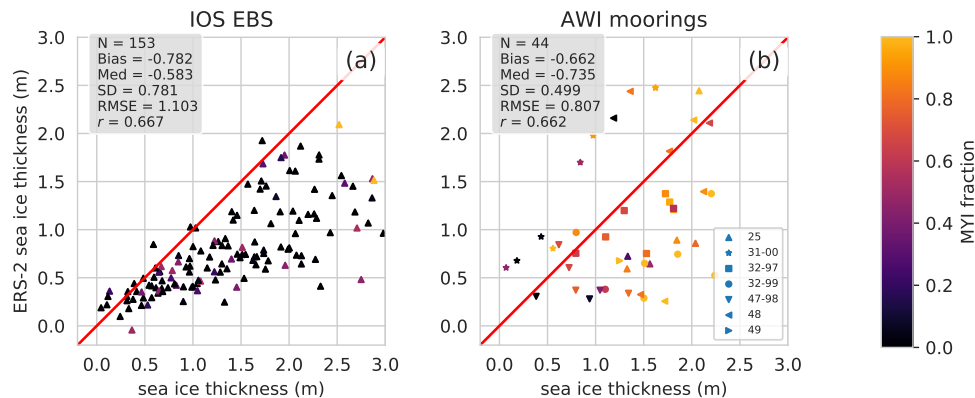
Figure 9. Comparative scatter-plots between Envisat sea ice thickness or radar freeboard estimations and other data sets. The x-axis indicates the sea ice thickness from (a) OIB total ice freeboard, (b) Air EM snow plus ice thickness, (c) Can Coast ice thickness, (d) UK/US submarines draft ~~,~~and (e) ~~IMB ice thickness and (f)~~ ICESat-1 total freeboard. (~~g~~f) compares our Envisat radar freeboard with SI-CCI Envisat solution. Colorbars represent the normalized density. A log_{10} has been applied before the normalization for (~~f~~e) and (~~g~~f) due to the large number of data. N is the number of the couple of values that are compared, Med refers to the Median, SD the Standard deviation, RMSE the Root Mean Square Error and r the correlation coefficient.
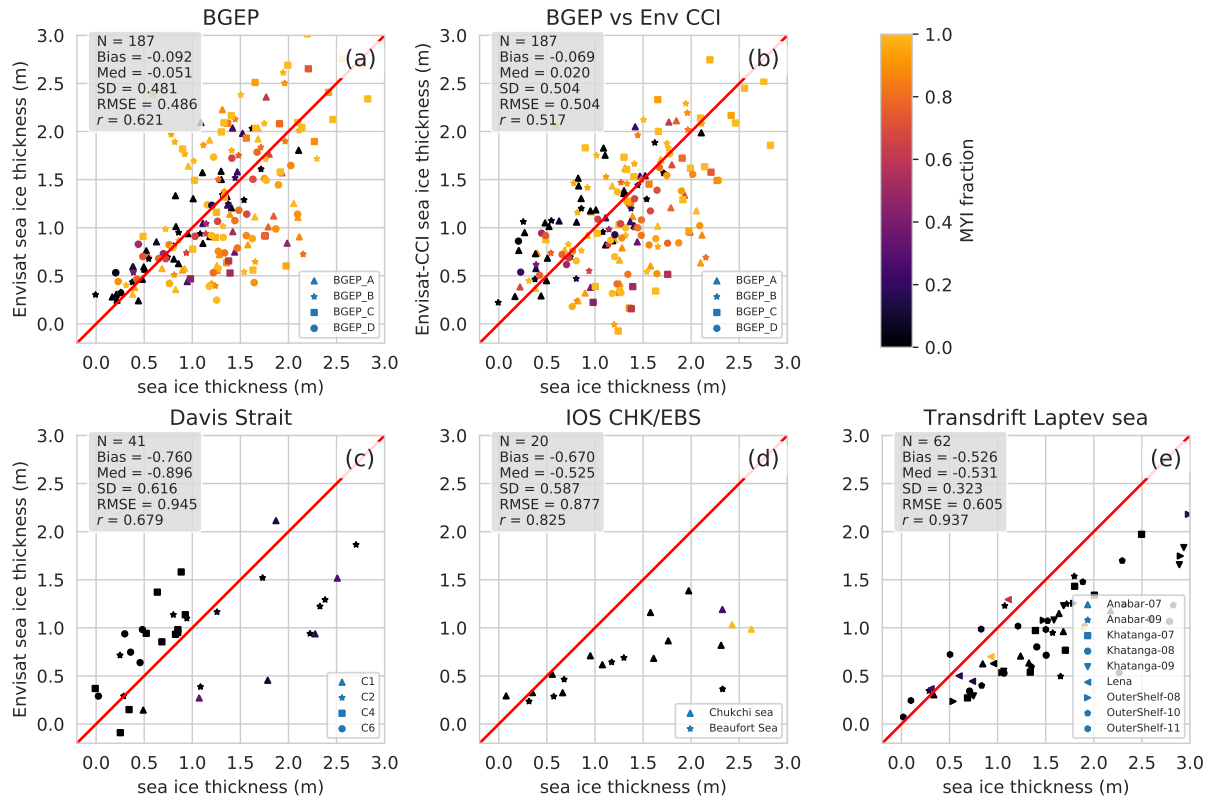
Figure 10. Comparative scatter-plots between Envisat sea ice thickness estimations and anchored moorings data sets. Each dot corresponds to a monthly averaged value. The x-axis indicates the sea ice thickness from (a) BGEP, (b) BGEP vs Env CCI, (c) Davis Strait, (ed) IOS CHK/EBS and (de) Transdrift Laptev Sea sea ice draft. The colorbar shows the MYI fraction. N is the number of the couple of values that are compared, Med refers to the Median, SD the Standard deviation, RMSE the Root Mean Square Error and r the correlation coefficient.

Figure 11. Comparative scatter-plots between ~~ERS2~~ ERS-2 sea ice thickness estimations and 3 in-situ data sets. The x-axis indicates the sea ice thickness from (a) AirEM total thickness, (b) UK/US Submarines draft and (c) Can Coast sea ice thickness. Colorbar indicates the normalized density. N is the number of the couple of values that are compared, Med refers to the Median, SD the Standard deviation, RMSE the Root Mean Square Error and r the correlation coefficient.

Figure 12. Comparative scatter-plots between ~~ERS2~~ ERS-2 sea ice thickness estimations and 2 anchored moorings data sets. The x-axis shows ~~the~~ sea ice thickness ~~measurements~~ estimations from (a) IOS Beaufort Sea and (b) AWI moorings sea ice draft. The color bar indicates the respective MYI fraction. N is the number of the couple of values that are compared, Med refers to the Median, SD the Standard deviation, RMSE the Root Mean Square Error and r the correlation coefficient.
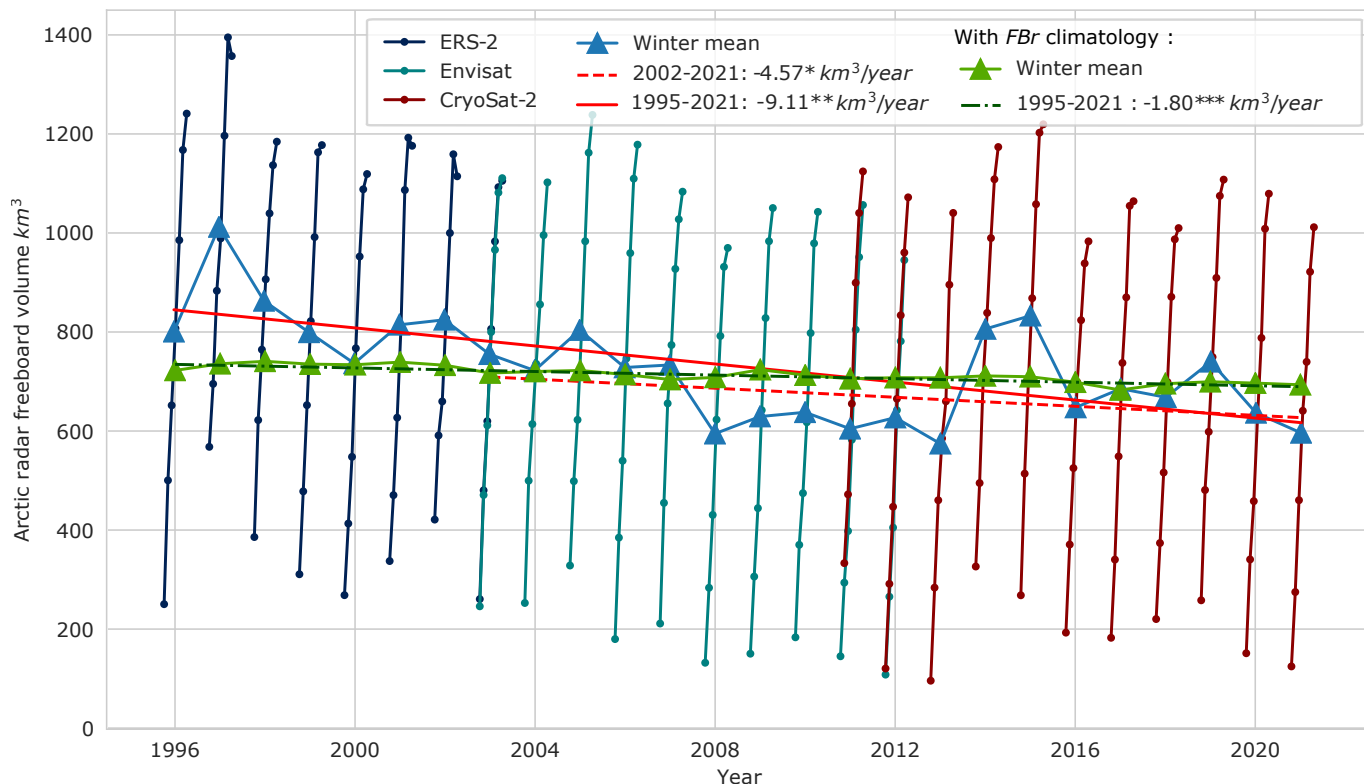
Figure 13. Time series representing radar freeboard volume up to 81.5°N for each ~~winter~~ winter month for ERS-2 in ~~orange~~dark blue, Envisat in teal and CS-2 in dark red. Blue triangles are winter mean ~~volume~~radar freeboard volumes. Red lines are linear regressions of winter mean ~~volume until~~ volumes from 2002/2003 for dashed line and 1995/1996 for solid line, estimated trends are respectively $-4.57 \pm 8.73 \mathrm{km}^3/\mathrm{year}$ and $-9.11 \pm 5.16 \mathrm{km}^3/\mathrm{year}$. Green triangles represent winter mean radar freeboard volumes computed with a climatology of radar freeboard between 1995 and 2021, dash-dot green line is the regression for FBr volume with FBr climatology, the estimated trend is $-1.80 \pm 0.42 \mathrm{km}^3/\mathrm{year}$. *(1-p)<0.5, **(1-p)>0.99, ***(1-p)>0.999999, the probability value of the Mann-Kendall test.
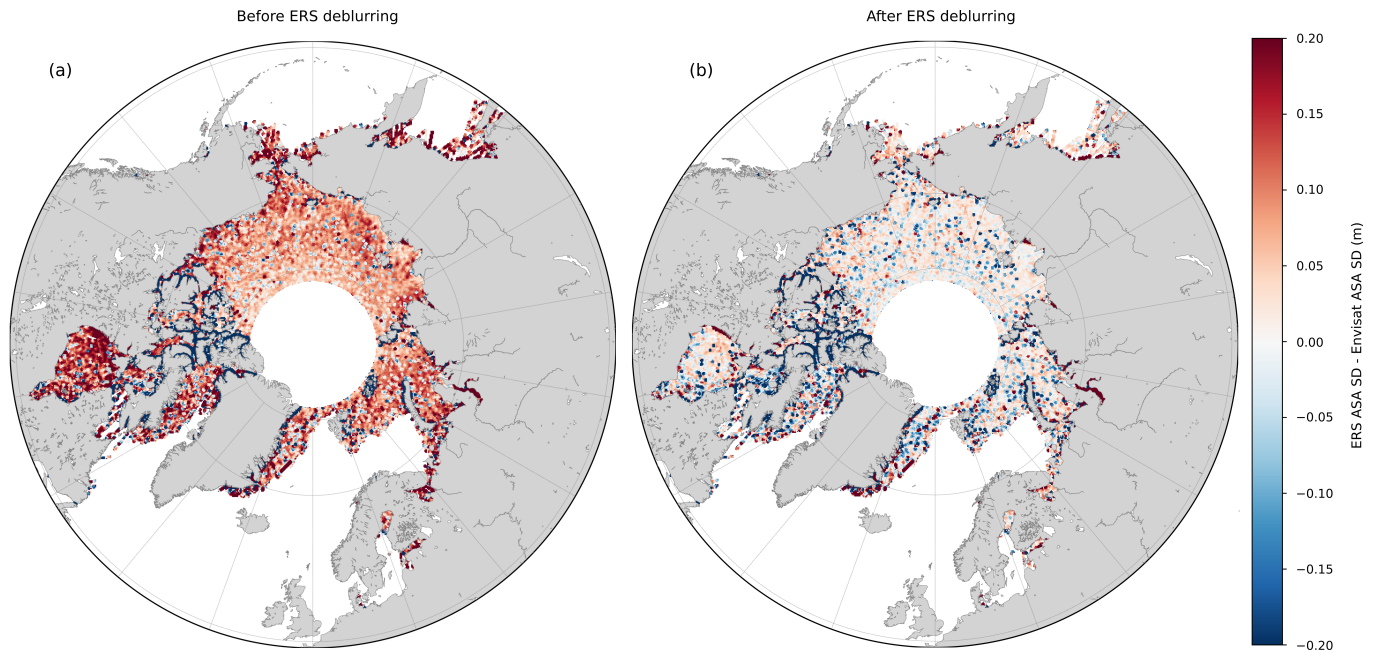
Figure A1. Comparison of the standard deviation of ASA between Envisat (cycle 12) and ERS (cycle 80) before (a) and after deblurring (b) within each grid cell of a 12.5km resolution grid. Median of the standard deviation difference for (a) is 7.5cm and 0.80 cm for (b).

Table A1. ~~Radar Altimeter characteristics~~ Pulse peakiness thresholds for lead/floe classification

Mission (RA mode)	PP lead threshold	PP floe threshold
CryoSat-2 (SAR)	0.3*	0.1*
Envisat (LRM)	0.3*	0.1*
ERS-2 (LRM)	0.2839	0.1328

* Guerreiro et al. (2017)

Table A2. Radar Altimeter characteristics with 'Nb of echoes' the approximate number of individual echoes that have been summed up to deliver the 20hz or the 18hz waveforms, σ_{1b} is the estimated error on range from speckle noise

Mission (RA mode)	PRF	Data frequency	Nb of echoes	σ_{1b}
CryoSat-2 (LRM)	1.97 kHz	~~21~~ 20 Hz	94	**7** cm*
Envisat (LRM)	1.80 kHz	18 Hz	100	6.8 cm
ERS-2 (LRM)	1.02 kHz	20 Hz	50	9.6 cm

* Wingham et al. (2006)

Table A3. Monthly Statistics (average and standard deviation) on radar freeboard for each winter month from the overlap period between Envisat and CryoSat-2 and averaged for both winters (2010-2012). SD corresponds to Standard Deviation.

	Month	FBr mean (m)		FBr median (m)		FBr SD (m)		~~correlation~~ Correlation	FBr unc mean (m)		FBr unc median (m)	
		Env	CS2	Env	CS2	Env	CS2	Env/CS2	Env	CS2	Env	CS2
Winter 2010/2011	Nov	0.050	0.052	0.036	0.037	0.050	0.096	0.521	0.125	0.025	0.057	0.017
	Dec	0.052	0.057	0.041	0.048	0.050	0.102	0.447	0.101	0.025	0.051	0.016
	Jan	0.065	0.068	0.059	0.063	0.047	0.100	0.451	0.127	0.027	0.057	0.019
	Feb	0.084	0.091	0.073	0.079	0.062	0.121	0.522	0.150	0.031	0.068	0.022
	Mar	0.098	0.103	0.086	0.089	0.076	0.129	0.581	0.160	0.074	0.033	0.024
	Apr	0.112	0.116	0.096	0.097	0.088	0.135	0.643	0.158	0.035	0.080	0.025
Winter 2011/2012	Oct	0.043	0.043	0.036	0.033	0.042	0.115	0.365	0.185	0.028	0.069	0.019
	Nov	0.045	0.048	0.038	0.036	0.040	0.096	0.400	0.142	0.024	0.057	0.018
	Dec	0.051	0.052	0.042	0.044	0.054	0.111	0.510	0.114	0.025	0.050	0.017
	Jan	0.066	0.067	0.061	0.058	0.068	0.114	0.592	0.156	0.028	0.054	0.019
	Feb	0.080	0.081	0.071	0.071	0.076	0.125	0.596	0.155	0.030	0.068	0.021
	Mar	0.091	0.087	0.077	0.073	0.090	0.135	0.657	0.160	0.031	0.071	0.021
	2010-2012	0.074	0.077	0.063	0.062	0.071	0.120	0.583	0.144	0.029	0.063	0.020

Table A4. Statistics (average and standard deviation) on radar freeboard for each winter month from the overlap period between ERS-2 and Envisat and averaged for the whole winter (2002-2003). SD corresponds to Standard Deviation.

	Month	FBr mean (m)		FBr median (m)		FBr SD (m)		~~correlation~~ Correlation	FBr unc mean (m)		FBr unc median (m)	
		ERS-2	Env	ERS-2	Env	ERS-2	Env	ERS-2/Env	ERS-2	Env	ERS-2	Env
Winter 2002/2003	Oct	0.061	0.058	0.049	0.043	0.044	0.048	0.881	0.176	0.152	0.087	0.066
	Nov	0.074	0.073	0.048	0.044	0.064	0.070	0.937	0.145	0.134	0.075	0.062
	Dec	0.076	0.074	0.050	0.048	0.067	0.074	0.900	0.146	0.115	0.075	0.064
	Jan	0.083	0.082	0.065	0.065	0.056	0.068	0.862	0.164	0.148	0.078	0.071
	Feb	0.095	0.093	0.075	0.074	0.068	0.074	0.895	0.172	0.156	0.082	0.076
	Mar	0.105	0.103	0.085	0.084	0.074	0.092	0.834	0.167	0.168	0.087	0.081
	Apr	0.112	0.112	0.094	0.096	0.076	0.088	0.845	0.168	0.157	0.093	0.087
	2002-2003	0.090	0.088	0.073	0.071	0.068	0.079	0.876	0.164	0.148	0.082	0.073

Table A5. Statistics of sea ice thickness difference between Envisat and each validation data set.

	Campaign	Bias (m)	Median (m)	Standard Deviation (m)	RMSE (m)	Correlation
FBr	Envisat CCI	-0.007	-0.007	0.052	0.053	0.712
Sea Ice Thickness	ICESat-1	~~-0.211~~ -0.197	~~-0.125~~ -0.116	~~0.901~~ 0.906	~~0.926~~ 0.928	~~0.495~~ 0.497
	OIB	~~-0.321~~ -0.352	~~-0.245~~ -0.249	~~0.828~~ 1.018	~~0.888~~ 1.077	~~0.422~~ 0.332
	Air EM	~~-0.367~~ -0.288	~~-0.567~~ -0.517	~~0.955~~ 0.991	~~1.022~~ 1.032	~~0.071~~ 0.047
	Can Coast	~~0.088~~ 0.089	~~0.044~~ -0.001	~~0.615~~ 0.632	~~0.621~~ 0.637	~~0.396~~ 0.406
	UK and US submarines	~~-0.373~~ -0.391	~~-0.265~~ -0.280	~~0.507~~ 0.532	~~0.629~~ 0.659	~~0.565~~ 0.548
	~~IMB -0.248 -0.252 1.201 1.226 0.055~~ BGEP	~~0.035~~ -0.092	~~0.035~~ -0.051	~~0.456~~ 0.481	~~0.453~~ 0.486	~~0.633~~ 0.621
	Davis Strait	~~-0.634~~ -0.760	~~-0.736~~ -0.896	~~0.519~~ 0.616	~~0.791~~ 0.945	~~0.739~~ 0.679
	IOS Chukchi and Beaufort sea	~~-0.516~~ -0.670	~~-0.372~~ -0.525	~~0.493~~ 0.587	~~0.702~~ 0.877	~~0.849~~ 0.825
	Transdrift Laptev sea	~~-0.372~~ -0.526	~~-0.381~~ -0.531	~~0.260~~ 0.323	~~0.444~~ 0.605	~~0.946~~ 0.937

Table A6. Statistics of sea ice thickness difference between ERS-2 and each validation data set.

	Campaign	Bias (m)	Median (m)	Standard Deviation (m)	RMSE (m)	Correlation
Sea Ice Thickness	Air EM	~~0.702~~ 0.816	~~0.737~~ 0.869	~~0.814~~ 0.846	~~1.107~~ 1.175	~~0.253~~ 0.260
	UK and US submarines	~~-0.279~~ -0.217	~~0.584~~ -0.163	~~0.647~~ 0.567	~~0.720~~ 0.607	0.740
	Can Coast	~~0.118~~ 0.096	~~0.063~~ 0.047	~~0.626~~ 0.616	~~0.636~~ 0.623	~~0.596~~ 0.602
	IOS Beaufort sea	~~-0.740~~ -0.782	~~-0.527~~ -0.583	~~0.778~~ 0.781	~~1.072~~ 1.103	~~0.671~~ 0.677
	AWI moorings	~~-0.598~~ -0.662	~~-0.593~~ -0.735	~~0.343~~ 0.499	~~0.677~~ 0.807	~~0.759~~ 0.662

Appendix B: List of Acronyms

Table B1. List of Acronyms

ACRONYMS	Description
AGC	Automatic Gain Control
ASD	Altimetric Snow Depth
CHK	Chukchi
CS-2	CryoSat-2
EBS	Eastern Beaufort Sea
Env	Envisat
ERS	European Remote-Sensing Satellite
FB	Sea ice Freeboard
FBr	Radar Freeboard
FBt	Total Freeboard
FYI	First-year Ice
HTL	Height Tracking Loop
ILA	Ice Level Anomaly
LRM	Low Resolution Mode
MSS	Mean Sea Surface
MYI	Multiyear Ice
MLP	Multi-layer Perceptron
NN	Neural Network
PRF	Pulse Repetition Frequency
RA	Radar Altimeter
RMSE	Root Mean Squared Error
SAR	Synthetic Aperture Radar
SARM	Synthetic Aperture Radar Mode
SARIn	Synthetic Aperture Radar Interferometric
SD	Standard Deviation
SIT	Sea Ice Thickness
SLA	Sea Level Anomaly
STL	Slope Tracking Loop
TFMRA	Threshold First-Maximum Retracker Algorithm
ULS	Upward Looking Sonar

590

References

Alexandrov, V., Sandven, S., Wahlin, J., and Johannessen, O. M.: The relation between sea ice thickness and freeboard in the Arctic, The Cryosphere, 4, 373–380, https://doi.org/10.5194/tc-4-373-2010, 2010.

Andersen, O., Stenseng, L., Piccioni, G., and Knudsen, P.: The DTU15 MSS (Mean Sea Surface) and DTU15LAT (Lowest Astronomical Tide) reference surface, http://lps16.esa.int/, eSA Living Planet Symposium 2016 ; Conference date: 09-05-2016 Through 13-05-2016, 2016.

Belter, H. J., Janout, M. A., Hölemann, J. A., and Krumpen, T.: Daily mean sea ice draft from moored upward-looking Acoustic Doppler Current Profilers (ADCPs) in the Laptev Sea from 2003 to 2016, https://doi.org/10.1594/PANGAEA.912927, type: data set, 2020.

Brockley, D. J., Baker, S., Femenias, P., Martinez, B., Massmann, F.-H., Otten, M., Paul, F., Picard, B., Prandi, P., Roca, M., Rudenko, S., Scharroo, R., and Visser, P.: REAPER: Reprocessing 12 Years of ERS-1 and ERS-2 Altimeters and Microwave Radiometer Data, IEEE Transactions on Geoscience and Remote Sensing, 55, 5506–5514, https://doi.org/10.1109/TGRS.2017.2709343, 2017.

Brodzik, M. J., Billingsley, B., Haran, T., Raup, B., and Savoie, M. H.: EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets, ISPRS International Journal of Geo-Information, 1, 32–45, https://doi.org/10.3390/ijgi1010032, 2012.

Calafat, F., Cipollini, P., Bouffard, J., Snaith, H., and Féménias, P.: Evaluation of new CryoSat-2 products over the ocean, Remote Sensing of Environment, 191, 131–144, https://doi.org/https://doi.org/10.1016/j.rse.2017.01.009, 2017.

Carrere, L., Lyard, F., Cancet, M., and Guillot, A.: FES 2014, a new tidal model on the global ocean with enhanced accuracy in shallow seas and in the Arctic region, p. 1, 2015.

Cavalieri, D., Parkinson, C., Gloersen, P., and Zwally, H. J.: Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 1, https://doi.org/10.5067/8GQ8LZQVL0VL, type: dataset, 1996.

Crameri, F.: Scientific colour maps, https://doi.org/10.5281/ZENODO.1243862, language: en, 2021.

Crameri, F., Shephard, G. E., and Heron, P. J.: The misuse of colour in science communication, Nature Communications, 11, 5444, https://doi.org/10.1038/s41467-020-19160-7, 2020.

Drucker, R., Martin, S., and Moritz, R.: Observations of ice thickness and frazil ice in the St. Lawrence Island polynya from satellite imagery, upward looking sonar, and salinity/temperature moorings, 108, 3149, https://doi.org/10.1029/2001JC001213, 2003.

Garnier, F., Fleury, S., Garric, G., Bouffard, J., Tsamados, M., Laforge, A., Bocquet, M., Fredensborg Hansen, R. M., and Remy, F.: Advances in altimetric snow depth estimates using bi-frequency SARAL and CryoSat-2 Ka–Ku measurements, The Cryosphere, 15, 5483–5512, https://doi.org/10.5194/tc-15-5483-2021, 2021.

Giles, K. A., Laxon, S. W., and Worby, A. P.: Antarctic sea ice elevation from satellite radar altimetry, Geophysical Research Letters, 35, L03 503, https://doi.org/10.1029/2007GL031572, 2008.

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://doi.org/10.5194/tc-11-2059-2017, 2017.

Haas, C., Lobach, J., Hendricks, S., Rabenstein, L., and Pfaffling, A.: Helicopter-borne measurements of sea ice thickness, using a small and lightweight, digital EM system, 67, 234–241, https://doi.org/10.1016/j.jappgeo.2008.05.005, 2009.

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://doi.org/10.5194/tc-8-1539-2014, 2014.

Hendricks, S., Paul, S., and Rinne, E.: ESA Sea Ice Climate Change Initiative (Sea_Ice_cci): Northern hemisphere sea ice thickness from the Envisat satellite on a monthly grid (L3C), v2.0, Centre for Environmental Data Analysis (CEDA), https://doi.org/10.5285/F4C34F4F0F1D4D0DA06D771F6972F180, 2018.

Kendall, M. G.: Rank correlation methods., 1948.

635 Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, CoRR, 2014.

Krishfield, R. A., Proshutinsky, A., Tateyama, K., Williams, W. J., Carmack, E. C., McLaughlin, F. A., and Timmermans, M.-L.: Deterioration of perennial sea ice in the Beaufort Gyre from 2003 to 2012 and its impact on the oceanic freshwater cycle: SEA ICE IN THE BG FROM 2003 TO 2012, Journal of Geophysical Research: Oceans, 119, 1271–1305, https://doi.org/10.1002/2013JC008999, 2014.

Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, 640 freeboard, and snow depth products from Operation IceBridge airborne data, 7, 1035–1056, https://doi.org/10.5194/tc-7-1035-2013, 2013.

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, https://doi.org/10.5194/tc-8-1217-2014, 2014.

Kwok, R.: Simulated effects of a snow layer on retrieval of CryoSat-2 sea ice freeboard, Geophysical Research Letters, 41, 5014–5020, https://doi.org/10.1002/2014GL060993, 2014.

645 Kwok, R. and Haas, C.: Effects of radar side-lobes on snow depth retrievals from Operation IceBridge, Journal of Glaciology, 61, 576–584, https://doi.org/10.3189/2015JoG14J229, 2015.

Kwok, R., Kurtz, N. T., Brucker, L., Ivanoff, A., Newman, T., Farrell, S. L., King, J., Howell, S., Webster, M. A., Paden, J., Leuschen, C., MacGregor, J. A., Richter-Menge, J., Harbeck, J., and Tschudi, M.: Intercomparison of snow depth retrievals over Arctic sea ice from radar data acquired by Operation IceBridge, The Cryosphere, 11, 2571–2593, https://doi.org/10.5194/tc-11-2571-2017, 2017.

650 Laforge, A., Fleury, S., Dinardo, S., Garnier, F., Remy, F., Benveniste, J., Bouffard, J., and Verley, J.: Toward improved sea ice freeboard observation with SAR altimetry using the physical retracker SAMOSA+, Advances in Space Research, p. S0273117720300776, https://doi.org/10.1016/j.asr.2020.02.001, 2020.

Landy, J. C., Tsamados, M., and Scharien, R. K.: A Facet-Based Numerical Model for Simulating SAR Altimeter Echoes From Heterogeneous Sea Ice Surfaces, IEEE Transactions on Geoscience and Remote Sensing, 57, 4164–4180, 655 https://doi.org/10.1109/TGRS.2018.2889763, 2019.

Landy, J. C., Petty, A. A., Tsamados, M., and Stroeve, J. C.: Sea Ice Roughness Overlooked as a Key Source of Uncertainty in CryoSat-2 Ice Freeboard Retrievals, Journal of Geophysical Research: Oceans, 125, https://doi.org/10.1029/2019JC015820, 2020.

Landy, J. C., Dawson, G. J., Tsamados, M., Bushuk, M., Stroeve, J. C., Howell, S. E. L., Krumpen, T., Babb, D. G., Komarov, A. S., Heorton, H. D. B. S., Belter, H. J., and Aksenov, Y.: A year-round satellite sea-ice thickness record from CryoSat-2, Nature, 609, 517–660 522, https://doi.org/10.1038/s41586-022-05058-5, number: 7927 Publisher: Nature Publishing Group, 2022.

Laxon, S.: Sea ice altimeter processing scheme at the EODC, International Journal of Remote Sensing, 15, 915–924, 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, 425, 947–950, https://doi.org/10.1038/nature02050, 2003.

665 Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, The Cryosphere, 9, 269–283, https://doi.org/10.5194/tc-9-269-2015, 2015.

Lindsay, R. and Schweiger, A. J.: Unified Sea Ice Thickness Climate Data Record, 1947 Onward, Version 1, https://doi.org/10.7265/N5D50JXV, publisher: NSIDC, 2013.

Liston, G., Stroeve, J., and Itkin, P.: Lagrangian Snow Distributions for Sea-Ice Applications, https://doi.org/10.5067/27A0P5M6LZBI, type: dataset, 2020a.

Liston, G. E., Itkin, P., Stroeve, J., Tschudi, M., Stewart, J. S., Pedersen, S. H., Reinking, A. K., and Elder, K.: A Lagrangian Snow-Evolution System for Sea-Ice Applications (SnowModel-LG): Part I—Model Description, 125, https://doi.org/10.1029/2019JC015913, 2020b.

Mallett, R. D. C., Lawrence, I. R., Stroeve, J. C., Landy, J. C., and Tsamados, M.: Brief communication: Conventional assumptions involving the speed of radar waves in snow introduce systematic underestimates to sea ice thickness and seasonal growth rate estimates, The Cryosphere, 14, 251–260, https://doi.org/10.5194/tc-14-251-2020, 2020.

Mann, H. B.: Nonparametric tests against trend, Econometrica: Journal of the econometric society, pp. 245–259, 1945.

Meier, W. N., P. D., Farrell, S., Haas, C., Hendricks, S., Petty, A., Webster, M., Divine, D., Gerland, S., Kaleschke, L., Ricker, R., Steer, A., Tian-Kunze, X., Tschudi, M., and Wood, K.: NOAA Arctic Report Card 2021: Sea Ice, https://doi.org/10.25923/Y2WD-FN85, publisher: United States. National Oceanic and Atmospheric Administration. Office of Oceanic and Atmospheric Research. Global Ocean Monitoring and Observing (GOMO) Program, 2021.

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: ARCTIC SEA ICE: REVIEW OF RECENT CHANGES, Reviews of Geophysics, 52, 185–217, https://doi.org/10.1002/2013RG000431, 2014.

Melling, H.: Ice Draft and Ice Velocity Data in the Beaufort Sea, 1990-2003, Version 1, https://doi.org/10.7265/N58913S6, publisher: NSIDC, 2008.

Nandan, V., Geldsetzer, T., Yackel, J., Mahmud, M., Scharien, R., Howell, S., King, J., Ricker, R., and Else, B.: Effect of Snow Salinity on CryoSat-2 Arctic First-Year Sea Ice Freeboard Measurements: Sea Ice Brine-Snow Effect on CryoSat-2, Geophysical Research Letters, 44, 10,419–10,426, https://doi.org/10.1002/2017GL074506, 2017.

National Snow and Ice Data Center: Submarine Upward Looking Sonar Ice Draft Profile Data and Statistics, Version 1, https://doi.org/10.7265/N54Q7RWK, 2006.

Paul, S., Hendricks, S., Ricker, R., Kern, S., and Rinne, E.: Empirical parametrization of Envisat freeboard retrieval of Arctic and Antarctic sea ice based on CryoSat-2: progress in the ESA Climate Change Initiative, The Cryosphere, 12, 2437–2460, https://doi.org/10.5194/tc-12-2437-2018, 2018.

Peacock, N. R.: Arctic sea ice and ocean topography from satellite altimetry, Ph.d. thesis, University of London, 1998.

Peacock, N. R. and Laxon, S. W.: Sea surface height determination in the Arctic Ocean from ERS altimetry, Journal of Geophysical Research: Oceans, 109, https://doi.org/https://doi.org/10.1029/2001JC001026, 2004.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Perovich, D., Richter-Menge, J., and Polashenski, C.: Observing and understanding climate change: Monitoring the mass balance, motion, and thickness of Arctic sea ice, http://imb-crrel-dartmouth.org/archived-data/, 2022.

Poisson, J.-C., Quartly, G. D., Kurekin, A. A., Thibaut, P., Hoang, D., and Nencioli, F.: Development of an ENVISAT Altimetry Processor Providing Sea Level Continuity Between Open Ocean and Arctic Leads, IEEE Transactions on Geoscience and Remote Sensing, 56, 5299–5319, https://doi.org/10.1109/TGRS.2018.2813061, 2018.

Raney, R.: A delay/Doppler radar altimeter for ice sheet monitoring, in: 1995 International Geoscience and Remote Sensing Symposium, IGARSS '95. Quantitative Remote Sensing for Science and Applications, vol. 2, pp. 862–864, IEEE, Firenze, Italy, https://doi.org/10.1109/IGARSS.1995.521080, 1995.

Rheinlænder, J. W., Davy, R., Ólason, E., Rampal, P., Spensberger, C., Williams, T. D., Korosov, A., and Spengler, T.: Driving Mechanisms of an Extreme Winter Sea Ice Breakup Event in the Beaufort Sea, Geophysical Research Letters, 49, https://doi.org/10.1029/2022GL099024, 2022.

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, 8, 1607–1622, https://doi.org/10.5194/tc-8-1607-2014, 2014.

Roca, M., Laxon, S., and Zelli, C.: The EnviSat RA-2 Instrument Design and Tracking Performance, IEEE Transactions on Geoscience and Remote Sensing, 47, 3489–3506, https://doi.org/10.1109/TGRS.2009.2020793, 2009.

Rothrock, D. A. and Wensnahan, M.: The Accuracy of Sea Ice Drafts Measured from U.S. Navy Submarines, 24, 1936–1949, https://doi.org/10.1175/JTECH2097.1, 2007.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, Journal of the American statistical association, 63, 1379–1389, 1968.

Stammer, D.: Satellite altimetry over oceans and land surfaces, Earth observation of global changes, 2018.

Stroeve, J. and Notz, D.: Changing state of Arctic sea ice across all seasons, Environmental Research Letters, 13, 103 001, https://doi.org/10.1088/1748-9326/aade56, publisher: IOP Publishing, 2018.

Stroeve, J., Liston, G. E., Buzzard, S., Zhou, L., Mallett, R., Barrett, A., Tschudi, M., Tsamados, M., Itkin, P., and Stewart, J. S.: A Lagrangian Snow Evolution System for Sea Ice Applications (SnowModel-LG): Part II—Analyses, Journal of Geophysical Research: Oceans, 125, https://doi.org/10.1029/2019JC015900, 2020.

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.

Theil, H.: A rank-invariant method of linear and polynomial regression analysis, Indagationes mathematicae, 12, 173, 1950.

Tilling, R., Ridout, A., and Shepherd, A.: Assessing the Impact of Lead and Floe Sampling on Arctic Sea Ice Thickness Estimates from Envisat and CryoSat-2, Journal of Geophysical Research: Oceans, 124, 7473–7485, https://doi.org/https://doi.org/10.1029/2019JC015232, 2019.

Tschudi, M., Meier, W. N., Stewart, J. S., Fowler, C., and Maslanikand, J.: EASE-Grid Sea Ice Age, https://doi.org/10.5067/UTAV7490FEPB, type: dataset, 2019.

Ulaby, F. T., Moore, R. K., and Fung, A. K.: Microwave remote sensing: Active and passive, Volume 3-From theory to applications, vol. 3, 1986.

Wadhams, P.: Arctic sea ice morphology and its measurement, Arctic Technology and Policy, edited by: Dyer, I. and Chryssostomidis, C, Hemisphere Publishing Corp., Washington, DC, 1984.

Wadhams, P. and Horne, R. J.: An Analysis Of Ice Profiles Obtained By Submarine Sonar In The Beaufort Sea, Journal of Glaciology, 25, 401–424, https://doi.org/10.3189/S0022143000015264, 1980.

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 Ice, Journal of Climate, 12, 1814–1829, https://doi.org/10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2, 1999.

Wensnahan, M.: Sea-ice draft from submarine-based sonar: Establishing a consistent record from analog and digitally recorded data, Geophysical Research Letters, 32, L11 502, https://doi.org/10.1029/2005GL022507, 2005.

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., de Chateau-Thierry, P., Laxon, S. W., Mallow, U., Mavro-

745   cordatos, C., Phalippou, L., Ratier, G., Rey, L., Rostan, F., Viau, P., and Wallis, D. W.: CryoSat: A mission to determine the fluctuations
in Earth's land and marine ice fields, 37, 841–871, https://doi.org/10.1016/j.asr.2005.07.027, 2006.

Witte, H.: AWI Moored ULS Data, Greenland Sea and Fram Strait, 1991-2002, https://doi.org/10.7265/N5G15XSR, publisher: NSIDC,
2005.

Yi, D. and Zwally, H. J.: Arctic Sea Ice Freeboard and Thickness, Version 1, https://doi.org/10.5067/SXJVJ3A2XIZT, type: dataset, 2009.

750 Zwally, H. J., Yi, D., Kwok, R., and Zhao, Y.: ICESat measurements of sea ice freeboard and estimates of sea ice thickness in the Weddell
Sea, Journal of Geophysical Research, 113, C02S15, https://doi.org/10.1029/2007JC004284, 2008.