Landy, J. C., Petty, A., Tsamados, M., & Stroeve, J. C. (2020). Sea ice roughness overlooked as a key source of uncertainty in CryoSat-2 ice freeboard retrievals. *Journal of Geophysical Research: Oceans*, *125*(5), [e2019JC015820]. https://doi.org/10.1029/2019JC015820

University of Bristol - Explore Bristol Research
General rights

Sea Ice Roughness Overlooked as a Key Source of Uncertainty in CryoSat-2 Ice Freeboard Retrievals

Jack C. Landy[1] [ID], **Alek A. Petty[2,3]** [ID], **Michel Tsamados[4]** [ID], and **Julienne C. Stroeve[4,5]** [ID]

[1]Bristol Glaciology Centre, School of Geographical Sciences, University of Bristol, Bristol, UK, [2]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA, [3]Cryospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA, [4]Centre for Polar Observation and Modelling, Earth Sciences, University College London, London, UK, [5]Centre for Earth Observation Science, University of Manitoba, Winnipeg, Manitoba, Canada

Abstract ESA's CryoSat-2 has transformed the way we monitor Arctic sea ice, providing routine measurements of the ice thickness with near basin-wide coverage. Past studies have shown that uncertainties in the sea ice thickness retrievals can be introduced at several steps of the processing chain, for instance, in the estimation of snow depth, and snow and sea ice densities. Here, we apply a new physical model to CryoSat-2, which further reveals sea ice surface roughness as a key overlooked feature of the conventional retrieval process. High-resolution airborne observations demonstrate that snow and sea ice surface topography can be better characterized by a lognormal distribution, which varies based on the ice age and surface roughness within a CryoSat-2 footprint, than a Gaussian distribution. Based on these observations, we perform a set of simulations for the CryoSat-2 echo waveform over "virtual" sea ice surfaces with a range of roughness and radar backscattering configurations. By accounting for the variable roughness, our new lognormal retracker produces sea ice freeboards that compare well with those derived from NASA's Operation IceBridge airborne data and extends the capability of CryoSat-2 to profile the thinnest/smoothest sea ice and thickest/roughest ice. Our results indicate that the variable ice surface roughness contributes a systematic uncertainty in sea ice thickness of up to 20% over first-year ice and 30% over multiyear ice, representing one of the principal sources of pan-Arctic sea ice thickness uncertainty.

Plain Language Summary We have developed a new way of measuring sea ice thickness in the Arctic, by comparing real and simulated data from the European Space Agency satellite: CryoSat-2. Our simulations are guided by aircraft observations that demonstrate sea ice has distinct patterns of surface roughness. Traditional methods ignore or misrepresent the surface roughness, which reduces our confidence in the measured ice thickness by around one third. If we account for the roughness, however, we can extend the capability of Cryosat-2 for measuring both the thinnest smoothest sea ice and thickest roughest ice. These improvements will boost our confidence in derived estimates of the Arctic Ocean's sea ice volume budget and freshwater fluxes, while enhancing the accuracy of sea ice forecasts primed with satellite data.

1. Introduction

Basin-scale estimates of sea ice thickness have been produced since the early 1990s by satellite radar altimeters including the European Space Agency's ERS-1 and ERS-2, Envisat, and, more recently, CryoSat-2 (Laxon et al., 2003; Laxon et al., 2013; Paul et al., 2018). Since its launch in 2010, CryoSat-2 has routinely monitored the polar oceans with almost complete basin-scale coverage. The novel multilook synthetic aperture radar (SAR) capability has enabled CryoSat-2 to measure sea ice freeboard at a greater resolution and precision than preceding pulse-limited altimeters (Wingham et al., 2006). The Ku-band frequency radar of CryoSat-2 is assumed to penetrate through overlying snow and backscatter from the snow-ice interface. Sea ice freeboard, which is the portion of the floating ice that is above sea level, is then obtained from the elevation difference between radar echoes backscattered from sea ice and ocean tie-points between ice floes. These measurements of freeboard are converted to sea ice thickness by assuming hydrostatic equilibrium, taking estimates of the densities of sea ice, ocean water, and snow from the literature, and estimating the

depth of snow accumulated on the ice surface (Laxon et al., 2013; Ricker et al., 2014; Tilling et al., 2018). As the ice freeboard (not including snow cover) is typically only 5–20% of its thickness (Alexandrov et al., 2010), errors in measurements of freeboard magnify roughly 10-fold in the conversion to thickness.

Recent work has focused predominantly on the sea ice thickness error introduced by uncertain snow properties. A snow depth and density climatology (Warren et al., 1999) is still used to derive the current state-of-the-art ice thickness products (Kurtz et al., 2014; Ricker et al., 2014; Tilling et al., 2018). Snow depth has consequently been identified as a primary source of ice thickness error (Giles et al., 2008; Ricker et al., 2014). Moreover, spatiotemporal variations in snow properties, including density (Willatt et al., 2011), grain size (Kwok, 2014), and basal salinity (Nandan et al., 2017), contribute additional error to derived sea ice thickness. Ricker et al. (2014) suggested that sea ice surface roughness may also represent a significant source of ice thickness uncertainty, because the effect of roughness on CryoSat-2 radar waveforms remains unaccounted for by conventional empirical retracking algorithms. The raw range measurement between the altimeter and target is only approximate, so the radar waveform must be "retracked" to accurately identify the mean sea ice or ocean scattering surface, represented by a point on the waveform's leading edge (Quartly et al., 2019). Commonly used threshold first-maximum retracker algorithms (TFMRA) simply apply a percentage threshold of the waveform's first-maximum power return to identify this retracking point (Laxon et al., 2013; Ricker et al., 2014; Tilling et al., 2018). However, modeling studies indicate this power threshold should realistically change depending on sea ice properties, principally surface roughness at the scale of the radar footprint (Kurtz et al., 2014; Landy et al., 2019).

Three groups producing routine and publicly available sea ice thickness data from the CryoSat-2 Synthetic Aperture Interferometric Radar Altimeter (SIRAL) are the Centre for Polar Observation and Modelling (CPOM), Alfred Wegener Institute (AWI), and Goddard Space Flight Centre (GSFC). Other sea ice thickness products are available from the NASA Jet Propulsion Laboratory (JPL) (Kwok & Cunningham, 2015), Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (LEGOS) (Guerreiro et al., 2017) and the ESA Climate Change Initiative (CCI) Program (Paul et al., 2018). The first group to produce ice thickness estimates from CryoSat-2, CPOM use TFMRA with a 70% threshold for retracking sea ice floes and a curve-fitting routine for retracking leads that identifies the waveform peak (Laxon et al., 2013; Tilling et al., 2018). AWI use TFMRA with a 50% threshold for retracking both sea ice and leads (Ricker et al., 2014). GSFC optimize the fit of a physical model for the SAR altimeter echo to CryoSat-2 waveforms, with the threshold at the tracking point varying from approximately 85–95% depending on sea ice roughness at multiple scales (Kurtz et al., 2014). The empirical retracking algorithms obtain the scattering surface elevation directly from the detected waveform, rather than an ideal echo fit to the waveform, and are several orders of magnitude faster than physical retrackers (Kurtz et al., 2014). While this makes them attractive for operational sea ice thickness products, they do not account for physically realistic variations in the tracking point threshold owing to variable sea ice properties. Consequently, several studies have demonstrated that CryoSat-2 elevations from the physically based GSFC algorithm agree most accurately with ice floe and lead (ocean surface) elevations (Yi et al., 2018) and sea ice freeboards (Xia & Xie, 2018) obtained from coincident airborne lidar observations.

The GSFC model assumes a sea ice surface with an undulating Gaussian topography. However, airborne lidar data have shown that sea ice topography cannot be accurately characterized by a Gaussian model (Rivas et al., 2006). The tail of the surface height distribution can be well represented by a negative exponential function (Petty et al., 2016; Wadhams et al., 1992), with the full height distribution well represented by a lognormal model (Castellani et al., 2014; Davis & Wadhams, 1995; Landy et al., 2019; Liu et al., 2014; Tian-Kunze et al., 2014). Using a new model for the SAR altimeter echo, Landy et al. (2019) showed that —in theory—ice freeboard could be underestimated by at least 5 cm if the topography is assumed to be Gaussian rather than lognormal. The ice freeboard bias increases as the sea ice surface gets rougher, up to 20 cm for a surface roughness (e.g., height standard deviation, σ) of 0.5 m (Landy et al., 2019), translating into a "worst case" 1.5 to 2 m sea ice thickness bias.

Here, we characterize the statistical properties of sea ice surface roughness, at both the air-snow and snow-ice interfaces, using airborne observations acquired as part of NASA's Operation IceBridge (OIB) campaign. Building on these results, we apply a set of simulations from the model of Landy et al. (2019) to develop a new algorithm (hereafter the Lognormal Altimeter Retracker Model [LARM]) for retracking

CryoSat-2 radar echoes from sea ice floes and leads. We compare CryoSat-2 sea ice freeboards derived from a range of existing physical-model and empirical retracking methods to each other, and against coincident airborne freeboard observations, to ascertain the systematic uncertainty associated with the retracking step. Finally, we contrast this uncertainty with other systematic uncertainties of the sea ice thickness processing chain, introduced by snow depth or density, or ice density, for instance.

2. Characterizing the Sea Ice Surface Roughness Height Distribution

Airborne lidar and radar observations collected by OIB are used to evaluate the suitability of employing Gaussian or lognormal assumptions for the ice surface height probability density function (PDF) when modeling SAR altimeter echoes from sea ice. These two PDFs can be characterized by the single free parameter: σ, assuming zero mean, so offer a tractable solution for the waveform model optimizations (Section 3). Airborne lidar observations, such as those from the OIB Airborne Topographic Mapper (ATM), have been employed in several prior studies to examine statistical properties of the snow or sea ice surface topography (e.g., King et al., 2015; Petty et al., 2016; Rivas et al., 2006) and represent the state of the art in terms of resolution and coverage. Since the lidar observes the snow surface elevation, however, these measurements do not necessarily represent the statistical characteristics of the underlying snow-ice interface, which is ostensibly observed by the Ku-band radar of CryoSat-2. Therefore, we additionally use data from the OIB airborne Ku-band radar altimeter to directly analyze the roughness of the snow-ice interface.

2.1. OIB Lidar Data

We analyze the *ATM L1B Elevation and Return Strength V2* dataset (Studinger, 2018) (last access on 18 April 2019) collected in spring 2013 and 2014, covering 7,300 km of first-year ice (FYI) and 19,000 km of multiyear ice (MYI) across the western Arctic Ocean. The wide-scan ATM is a conically scanning laser altimeter operating at 532 nm, which provides an ~250-m swath of surface height measurements at the nominal flight altitude of ~500 m. Each shot has a footprint of around 1 m and a vertical accuracy of around 0.1 m, with higher shot density at the edge of the swath due to the conical scan pattern (Studinger, 2018). We follow the ice topography analysis approach of Petty et al. (2016, 2017) but adapt this to collate all shots within individual 1,700 m along-track sections to broadly represent a given CryoSat-2 swath ($300 \times 1,700$ m). The onboard Applanix POS/AV precision orientation system is used to delineate the data and find the bounds of each swath-like segment. The data are expressed as the relative height versus the lowest elevation in each section (data originally provided as an elevation with respect to the WGS-84 ellipsoid).

2.2. OIB Ku-Band Radar Data

We further analyze the *OIB Ku-Band Radar L1B Geolocated Echo Strength Profiles V2* dataset (Paden et al., 2017) (last access on 8 September 2019) collected during two airborne campaigns on 21 March 2013 and 14 March 2014, each spanning the same track across the Western Arctic Ocean from the Lincoln Sea to Point Barrow, Alaska, and covering 1,240 km of FYI and 3,085 km of MYI. The Centre for Remote Sensing of the Ice Sheets (CReSIS) Ku-band radar is an ultra-wideband radar altimeter with a central frequency of 15 GHz (Rodriguez-Morales et al., 2013). At the nominal bandwidth of 6 GHz, the flat-surface range resolution in snow (dielectric constant, $\varepsilon_s = 1.53$ for dry snow with density of 300 kg m^{-3}) is approximately 4.9 cm (Paden et al., 2017), compared to 47 cm for CryoSat-2 (Wingham et al., 2006). With such a fine range resolution, the CReSIS Ku-band radar has previously been applied to separate air-snow from snow-ice interfaces over Arctic sea ice (Kwok, 2014). Through unfocused SAR processing, raw samples are coherently stacked along-track to obtain a finer spatial resolution, before a boxcar filter is applied for a final along-track sample spacing of approximately 5 m (at nominal flight speed of 140 m s^{-1}). For a sea ice surface with zero slope, the across-track footprint is limited by the pulse duration to approximately 10–15 m depending on the surface roughness and flight altitude (Rodriguez-Morales et al., 2013).

We collate all radar samples above a threshold signal-to-noise ratio (SNR) of 10 dB along 5-km OIB tracks, providing ~1,000 samples in the best case, and vertically compensate the radar echograms for variations in the flight altitude. A 50% threshold first-maximum retracking algorithm (Landy et al., 2017) is used to estimate the elevation of the principal scattering horizon (generally assumed to be the mean level of the snow-ice interface), before a 2-km-span median filter is applied to remove residual long-wavelength undulations of the topography not corrected during aircraft compensation. Specular returns from leads within the

Table 1
Summary Statistics for Gaussian and Lognormal Model Fits to High-Resolution 1,700-m Along-Track Snow Surface Height Distributions Obtained From the OIB ATM Lidar in 2013 and 2014

2013	Num sections	Sigma (normal, m)	KS test (normal)	KS test (lognormal)	RMSE (normal)	RMSE (lognormal)	% lognormal superior fit
FYI	2,716	0.16	0.13	0.07	0.67	0.41	92
MYI	3,689	0.31	0.13	0.06	0.37	0.21	96
All	6,405	0.25	0.13	0.07	0.50	0.29	94
2014	Num sections	Sigma (normal, m)	KS test (normal)	KS test (lognormal)	RMSE (normal)	RMSE (lognormal)	% lognormal superior fit
FYI	1,313	0.20	0.14	0.08	0.57	0.38	91
MYI	7,235	0.32	0.13	0.07	0.38	0.23	94
All	9,033	0.30	0.13	0.07	0.41	0.25	94

pack ice are classified based on high waveform pulse-peakiness (Laxon et al., 2013) and backscattering coefficient (e.g., Paul et al., 2018). An estimate for the sea surface elevation fit to lead returns (Figure 2a) is finally removed, leaving a profile of snow-ice interface heights relative to their mean elevation. All sections with high aircraft altitude variations, pitch or roll, low SNR, or too few leads for a valid sea level estimate are discarded.

2.3. Gaussian and Lognormal Model Fits to Sea Ice Height Distributions

We mask all data within 25 km to the nearest coastline and use the EUMETSAT Ocean and Sea Ice Satellite Application Facilities (OSI-SAF) sea ice type product (OSI-403-c) (Breivik et al., 2012) to classify each section as either FYI or MYI. All valid samples within a lidar or radar section are used to produce normalized PDFs of the surface height, with fixed 5-cm-width bins. A least-squares method is used to fit Gaussian and lognormal models to the observed PDFs. The Kolmogorov-Smirnov (KS) test and root mean squared error (RMSE) of the model fit to the binned PDFs are used to assess goodness of fit, with a lower value indicating superior fit in both cases.

Summary statistics for more than 15,000 analyzed lidar segments and 850 radar segments are provided in Tables 1 and 2, along with examples for the snow and ice surface height PDF fits for selected segments in Figures 1 and 2, respectively. For both FYI and MYI, the lognormal model offers a better match to the underlying snow and sea ice surface elevation distributions than the Gaussian model. The KS score using either model is similar between FYI and MYI while the normalized RMSEs are clearly lower for the MYI function fits, indicating that multiyear ice roughness—in particular—is better characterized with a lognormal model. The KS test score defines the distance between an empirical distribution and reference (model) cumulative distribution. For our normalized PDFs, the absolute difference between Gaussian and lognormal test scores are consistent in Tables 1 and 2, which suggests the lognormal is superior to the Gaussian model for a similar fraction of the snow and ice surface topography segments. The lognormal model provides a superior fit than the Gaussian model to snow surface topography in 94% of cases (92% for FYI, 95% for MYI) (Table 1) and to snow-ice interface topography in 90% of cases (96% for FYI, 88% for MYI) (Table 2). These results provide

Table 2
Summary Statistics for Gaussian and Lognormal Model Fits to Medium-Resolution 5-km Along-Track Snow-Ice Interface Height Distributions Obtained From the OIB Ku-Band Radar in 2013 and 2014

2013	Num sections	Sigma (normal, m)	KS test[a] (normal)	KS test[a] (lognormal)	RMSE (normal)	RMSE (lognormal)	% lognormal superior fit
FYI	167	0.10	0.73	0.69	0.24	0.16	93
MYI	239	0.24	0.71	0.65	0.17	0.11	88
All	406	0.18	0.72	0.67	0.20	0.13	90
2014	Num sections	Sigma (normal, m)	KS test[a] (normal)	KS test[a] (lognormal)	RMSE (normal)	RMSE (lognormal)	% lognormal superior fit
FYI	81	0.14	0.77	0.67	0.23	0.16	99
MYI	378	0.22	0.71	0.64	0.18	0.13	87
All	459	0.20	0.72	0.65	0.19	0.13	89

[a]KS test statistics are higher for the radar results compared to the lidar results due to the much lower number of observations per analyzed section.
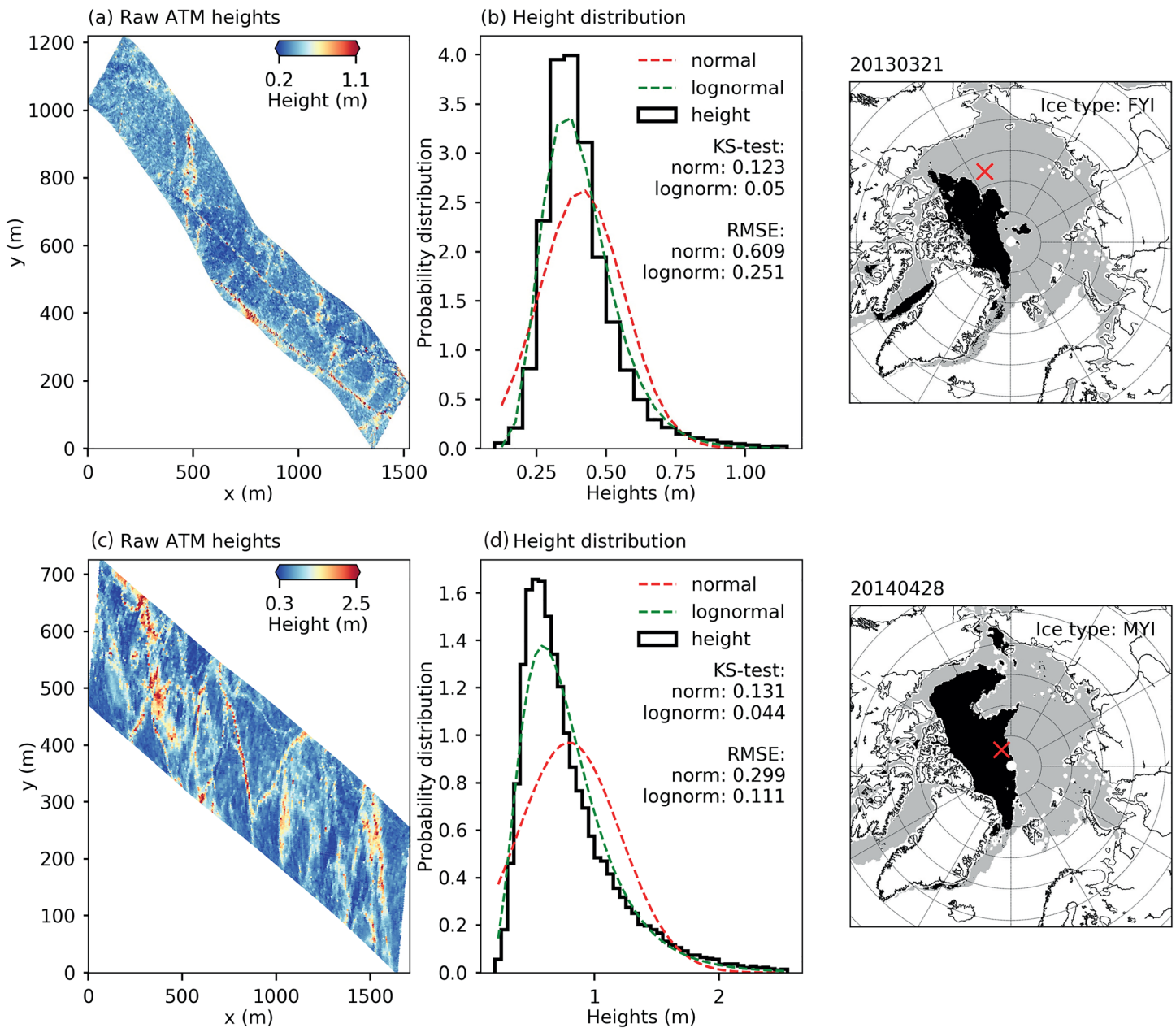
Figure 1. (a,c) Snow surface elevations from characteristic FYI and MYI 1,700-m along-track segments of the OIB ATM lidar, respectively, and (b,d) their height PDFs with corresponding best-fit Gaussian and lognormal models, and their locations inset.

strong evidence that a lognormal height distribution must be used to accurately model the radar echo backscattered from sea ice.

Tables 1 and 2 also reveal the similarity of kilometer-scale roughness statistics for air-snow and snow-ice interfaces of the Arctic sea ice cover. For instance, the lognormal model provides a superior fit than the Gaussian model to the snow and ice topography in a consistent ~90% of cases. The average standard deviation of the height distribution (σ) is significantly higher for the snow surface (0.28 m) than the ice surface (0.19 m), although we did not analyze identical sections between the two instruments. Since the snow and sea ice surfaces exhibit similar roughness properties, our results support the argument that snow surface roughness is primarily controlled by the roughness of the underlying sea ice (e.g., Doble et al., 2011; Iacozza & Barber, 1999; Kurtz et al., 2013), for instance, by the spacing and height of deformed ice and pressure ridges.
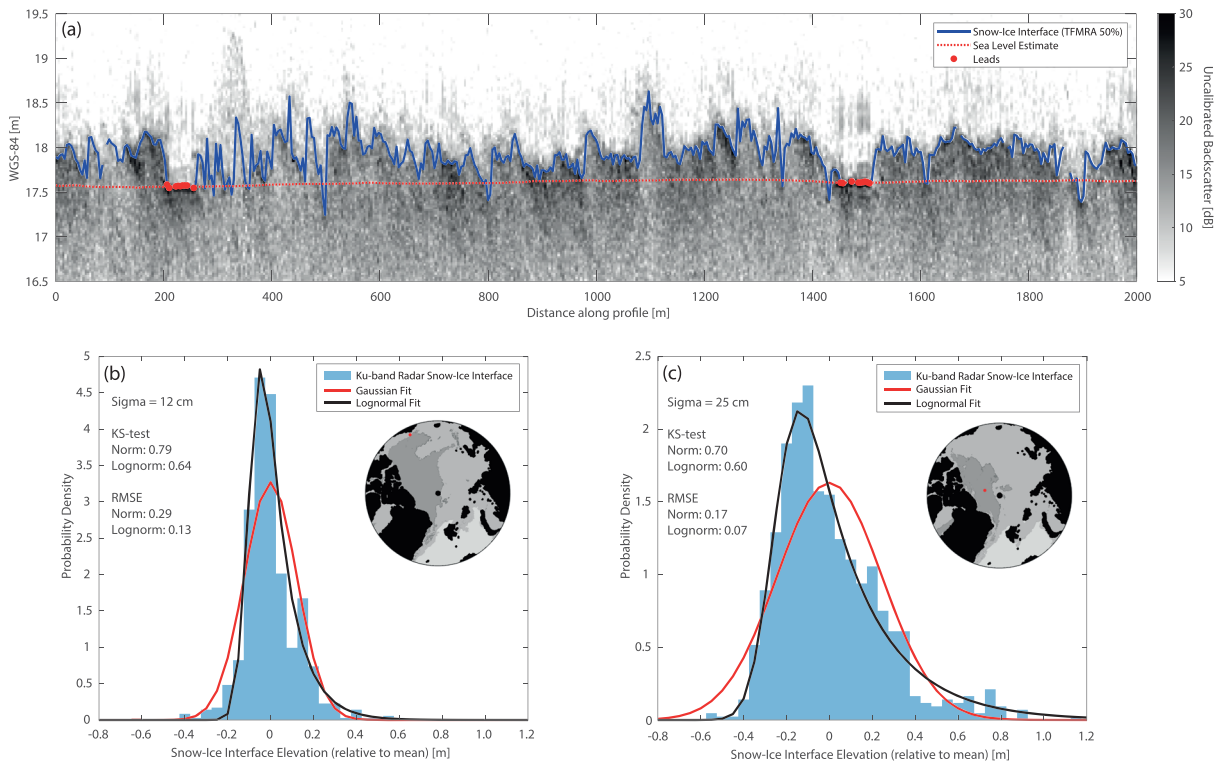
Figure 2. (a) Echogram from the OIB Ku-band radar over MYI in the Lincoln Sea March 2013, compensated for aircraft altitude changes and relative to the WGS-84 ellipsoid, including the retracked elevation of the snow-sea ice interface, samples classified as leads, and an estimate for the sea surface height. Minor (5–10 dB) backscattering from the snow surface and volume is apparent above the retracked sea ice surface and evidence for radar sidelobes is apparent above and below the leads. Height PDFs from characteristic (b) FYI on 14 March 2014 and (c) MYI on 21 March 2013, 5-km along-track segments of the radar with corresponding best-fit Gaussian and lognormal models, and their locations inset.

3. Application of a Numerical SAR Altimeter Echo Model for Retracking CryoSat-2 Waveforms

We use the facet-based numerical model of the delay-Doppler SAR altimeter echo from snow-covered sea ice presented in Landy et al. (2019) to generate ideal waveforms for comparison against CryoSat-2. The model can simulate echoes from a tetrahedral mesh of virtual sea ice surfaces generated from statistical descriptions of the ice surface roughness. The final echo integrates the total backscattered power from all facets of the sea ice surface height distribution illuminated by the CryoSat-2 antenna footprint. As previous studies, we make the necessary assumption that the principal scattering horizon of the CryoSat-2 radar originates from the snow-ice interface (Laxon et al., 2013), so we assume zero surface or volume scattering from the snowpack in our simulations. Only CryoSat-2 observations from October to April each year are processed.

3.1. SAR Echo Model Simulations

Statistical rough surfaces with Gaussian or lognormal height PDF and prescribed roughness parameters: rms height σ and correlation length l are generated through the techniques presented in Landy et al. (2019), and each is converted to a triangular irregular network (TIN). The surface area is constant in all simulations: 400 m along the track of the satellite and 8,000 m across-track, based on the illuminated area of the Doppler beam-limited footprint along-track and several times the pulse-limited footprint across-track (Wingham et al., 2006). The resolution of the statistical surfaces was a constant 5 m to ensure the surface roughness (at multiple scales) was accurately characterized (Landy et al., 2015; Landy et al., 2019). To generate a lookup table of ideal echoes from which to fit the CryoSat-2 waveforms, we modified two parameters of the surface: σ, at 1 cm increments from 0 to 100 cm, and the subgrid "radar-scale" sea ice roughness (i.e., roughness at a scale below the length of a 5-m facet), s_{rms}, at 0.25 mm increments from 0 to 6 mm. The latter range has been determined from terrestrial lidar measurements of natural and artificially grown sea ice
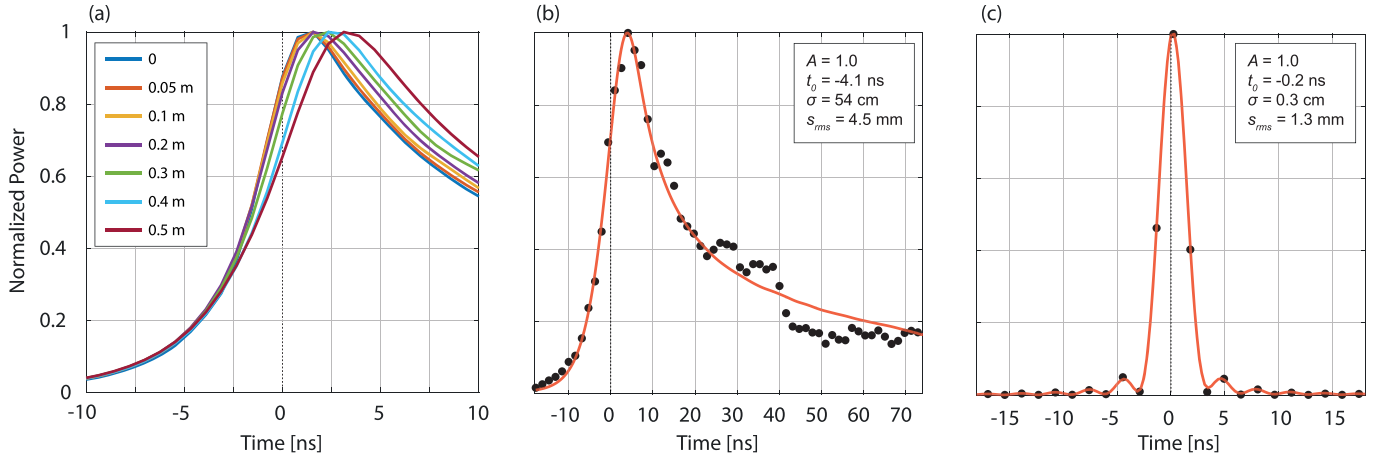
Figure 3. (a) CryoSat-2 SAR-mode waveforms simulated for sea ice surfaces with lognormal height distribution, varying height standard deviation σ, and fixed radar-scale roughness s_{rms} of 3 mm. Example model fits to CryoSat-2 SAR-mode waveforms for (b) sea ice floe and (c) lead, with the four optimized parameters. The tracking points ("epochs") are represented by the vertical lines at $t = 0$ in each plot.

(Landy et al., 2015; Landy, Isleifson, et al., 2015). Where $\sigma < 0.02$ m and $s_{rms} < 1$ mm, the waveform is representative of a specular lead-type echo (Kurtz et al., 2014). All other surface parameters were kept constant, including the correlation length l at 5 m, following Rivas et al. (2006). This matrix of roughness parameters produced a lookup table of 2,222 simulated echoes (e.g., Figure 3a), repeated for both Gaussian and lognormal surfaces.

Echo range bins were oversampled by five times in each simulation to accurately characterize the waveform leading-edge and peak and ensure that fits to CryoSat-2 waveforms were not degraded by poorly resolved model echoes. All other model satellite and antenna parameters were kept as in Table II in Landy et al. (2019). Because the rough surface generator produces a single random version of a surface, for each combination of σ and s_{rms}, we ran the echo model for 100 iterations to get a representative statistical sample of echoes and calculated the final echo as the mean of these iterations.

One significant modification was made to the model equations outlined in Landy et al. (2019) to produce more realistic simulations of the sea ice surface backscatter coefficient σ^0. In Landy et al. (2019), we made the simple assumption that the snow-ice interface was a purely incoherent (diffusely) scattering surface at the Ku-band frequency of CryoSat-2. Snow-ice interface $\sigma^0(\theta)$ for the incidence angles illuminated by the CryoSat-2 footprint, θ, was simulated from the integral equation model (Fung, 1994). However, we now take the more realistic approach to model $\sigma^0(\theta)$ from the snow-ice interface as the combination of incoherently scattered and coherently reflected power, accounting for expected near-nadir reflection of the altimeter beam over relatively smooth sea ice at Ku-band (Fetterer et al., 1992). As described in Landy et al. (2019) for modeling coherent reflection from leads, the fraction of the backscattered power reflected coherently from the snow-ice interface is

$$\omega_{coh} = exp\left(-4k_0^2 s_{rms}^2 cos^2\theta\right),\tag{1}$$

where k_0 is the radar carrier wavenumber (284.3 m^{-1}). For s_{rms} toward the smoother end of observed roughness heights over sea ice (at a length scale <5 m), ω_{coh} is up to 0.72 at near-nadir θ, verifying the importance of including coherent scattering contributions when modeling Ku-band altimeter backscattering from sea ice facets. As in Landy et al. (2019), the coherent backscattering coefficient for a flat, relatively smooth snow-ice interface, inclined at facet incidence angle θ, can be written in the form (Fung & Eom, 1983)

$$\sigma_{coh}^0 = \left(\frac{R_0}{\beta_c}\right)^2 exp\left(-4k_0^2 s_{rms}^2\right) exp\left(-\frac{\theta^2}{\beta_c^2}\right),\tag{2}$$

where R_0 is the Fresnel reflection coefficient at nadir incidence and β_c is the effective width of the

Table 3
Waveform Simulation Parameter Bounds

Parameter	Initial point	Lower bound	Upper bound
A	Waveform peak	Waveform peak	Waveform peak * 1.5
t_0	Time at waveform peak	Time at first bin	Time at final bin
σ	σ at waveform PP value	0	100 cm
s_{rms}	s_{rms} at waveform PP value	0	6 mm

angular extent of the coherent backscatter component (kept identical to the angular interval between successive bursts of the CryoSat-2 Doppler beams). Thus, the total snow-ice interface $\sigma_{si}^0 = \sigma_{coh}^0 + \sigma_{inc}^0$, where the incoherent scattering component is obtained from the integral equation model. The effect of changing small-scale snow-ice interface roughness s_{rms} for a sea ice surface with fixed, very smooth topography ($\sigma = 1$ cm) is shown in Figure S1 in the supporting information. The echo becomes more sharply peaked as s_{rms} decreases, with the transition in peakiness occurring rapidly as the fraction ω_{coh} switches from near-zero to >0.5 at a radar-scale roughness depending on frequency. When σ increases above ~10 cm, modifying the small-scale roughness has little effect on the echo shape because local facet incidence angles increase well above zero and σ_{coh}^0 becomes negligible outside the point perpendicular to the antenna boresight.

3.2. Optimization Algorithm for Fitting CryoSat-2 Waveforms

Observed CryoSat-2 waveforms are truncated to remove range bins below the noise floor (Ricker et al., 2014) and significant secondary peaks on the waveform trailing edge are filtered out. These secondary peaks are caused by strong reflection from specular targets, such as leads or smooth regions of ice, well outside the nadir point of the radar footprint. As the modeled echoes represent ideal targets with statistically defined roughness and uniform scattering properties, we cannot fit and extract useful information from these secondary peaks, instead focusing on optimizing the fit to the waveform leading edge. Secondary peaks are identified from the first derivative of the waveform power and a simple, conservative filter is used to remove the peaks while retaining the full waveform. We found the process of "cleaning" the waveform trailing edge to vastly improve the speed and quality of model fits.

We use a least-squares fitting procedure to optimize the functional form of the modeled sea ice echo to observed CryoSat-2 waveforms, following a similar method to Kurtz et al. (2014). Linear interpolation coefficients are precomputed for the Gaussian and lognormal echo lookup tables, so that echoes can be rapidly generated for any queried combination of roughness parameters. A least-squares fitting routine based on the bounded trust region reflective algorithm (implemented through the *MATLAB* function *lsqnonlin*) is used to minimize the difference between the model fit and each CryoSat-2 power waveform. To prevent the algorithm converging on a local minimum, the finite difference step size is kept relatively small and maximum number of iterations large. If the fit does not converge within the iteration limit, the waveform is discarded.

Following methods adopted for retracking ocean echoes (Dinardo et al., 2018; Ray et al., 2015), the optimization minimizes four free parameters: A, the scaled waveform amplitude, t_0, the tracking point of the mean radar scattering surface (or "epoch"), σ, the surface topography rms height, and s_{rms}, the small-scale roughness (Figures 3b and 3c). The pulse-peakiness (PP) parameter is calculated for each CryoSat-2 waveform, following the method of Laxon et al. (2013), and compared to a table of PP values for model waveforms in the lookup table. Initial guesses and bounds for the free parameters are given in Table 3.

The quality of an echo fit, if the optimization algorithm converges to a solution, is determined from the coefficient of determination r^2 and the RMSE between model fit and CryoSat-2 waveform. A waveform fit is accepted if the $r^2 > 0.95$, or if the $r^2 > 0.90$ and the RMSE <10%. For instance, in March 2015, 26% of SAR and SARIn waveforms north of 50° and in the sea ice zone (SIC >70%) were discarded by applying these criteria and the filters described above.

A waveform optimization typically converges to a solution within 0.01–0.2 s on a high-specification desktop computer, so the physical model-fitting routine is ~2 orders of magnitude slower than threshold first-maximum retracker algorithms (TFMRA). We have retracked the entire 2010–2018 CryoSat-2 SAR and SARIn archive, above 50° North and greater than 10 km from land, simultaneously applying algorithms for the physical echo model with Gaussian and lognormal assumptions for the topography, TFMRA with 50% and 70% thresholds, and a peak-identifying algorithm for lead returns (Tilling et al., 2018) to enable a direct comparison between methods.

4. Comparison of Physical-Model and Threshold-Based Retracking Algorithms

Empirical retrackers, such as the TFMRA, apply a constant threshold of the CryoSat-2 waveform's leading-edge power to estimate the sea ice surface mean elevation. However, our physical model for the CryoSat-2 echo waveform suggests that the mean elevation (tracking point) actually exists at a different location on the leading edge depending on the sea ice surface roughness. Across a typical range of sea ice surface roughness heights from 5 to 55 cm (the approximate range of σ values determined from the OIB data analyzed in section 2.3), the tracking point in our lookup table of Gaussian echoes varies from 95 to 85% of the leading-edge power, verifying the results of Kurtz et al. (2014). The tracking point shifts closer to the waveform peak (i.e., 100% normalized power) as σ approaches zero. Indeed, returns from leads—with almost zero roughness—have a modeled retracking point >95% (Figure 3c). Across the same range of σ from 5 to 55 cm, the retracking point in our lookup table of Lognormal echoes varies from 95% to 60% (Figure 3a), indicating that the more realistic Lognormal model predicts the true sea ice mean elevation to vary by considerably more than the Gaussian model suggests. The model results indicate that heavily deformed sea ice, with surface roughness >50 cm, should have a retracking point around 60% (e.g., Figure 3b), whereas for newly formed sea ice with little roughness it would be around 95%.

Figure 4 illustrates the pan-Arctic patterns of the offsets between gridded sea ice and sea surface elevations measured by each retracker, averaged for March over the CryoSat-2 record. Over smooth FYI in the Eastern Arctic and marginal seas the Lognormal and Gaussian retrackers produce elevations within 2 cm, whereas over rougher MYI in the Central Arctic the Gaussian retracker produces elevations more than 6 cm lower than LARM (Figure 4e). The regional pattern of this offset is consistent with regional variability in sea ice surface roughness characteristics (Figure 1). The Lognormal roughness assumption enables mean footprint sea ice elevation to accurately increase over deformed and MYI, in a way that is limited for a Gaussian retracker. Sea level estimates from the Lognormal and Gaussian retrackers are almost identical, with a mean difference of only −0.2 cm (Figure 4l).

Both threshold-based algorithms produce estimates for the mean sea ice elevation that are higher than LARM (Figures 4f and 4g). This is because the lognormal roughness model suggests the true retracking threshold should only reduce to 70% of the waveform's leading-edge power when the roughness height σ increases above 40 cm. The offset between the 70% threshold retracker and LARM is only 1–2 cm over MYI in the Central Arctic, because in this region the mean sea ice roughness is generally between 30 and 50 cm, so that a retracking threshold of 70% is quite realistic (Figure 4f). However, physical models indicate that a 70% retracking threshold is unrealistic over smooth FYI (Kurtz et al., 2014; Landy et al., 2019), with the threshold likely >85% (Ricker et al., 2015), so the retracker overestimates mean elevation within the FYI zone by 4–6 cm with respect to LARM.

The offsets between sea ice elevations measured by LARM and the 50% threshold retracker are considerably larger, with the threshold algorithm overestimating LARM by up to 30 cm over rough MYI (Figure 4g). Such large offsets in sea ice elevation will not translate directly into an ice freeboard bias, as freeboard is estimated from the difference between ice and sea levels (see section 5), with the method for retracking the sea level varying between groups too. Although there are significant offsets between the relative sea surface heights measured through these different methods (Figures 4m and 4n), the offsets are predictably near-constant across the Arctic, having no regional pattern. In contrast, the offsets in retracked sea ice elevation do vary regionally (Figures 4e–4g), reflecting the pan-Arctic pattern of sea ice surface roughness (Landy et al., 2015). This has a direct impact on the CryoSat-2 radar freeboard derived from the relative difference between sea ice and sea surface elevations.

5. Uncertainties in Sea Ice Freeboard Introduced by Ice Surface Roughness

Each of the developers of available CryoSat-2 sea ice thickness datasets use a different combination of algorithms for classifying waveforms, retracking sea ice and lead elevations, for applying geophysical corrections to ice and sea levels, filtering poor quality observations, and gridding final radar freeboard estimates. Rather than comparing gridded ice freeboard products, our approach here is to reproduce only the retracking steps of other group's processing chains to enable direct evaluation of the retracker algorithms on derived sea ice
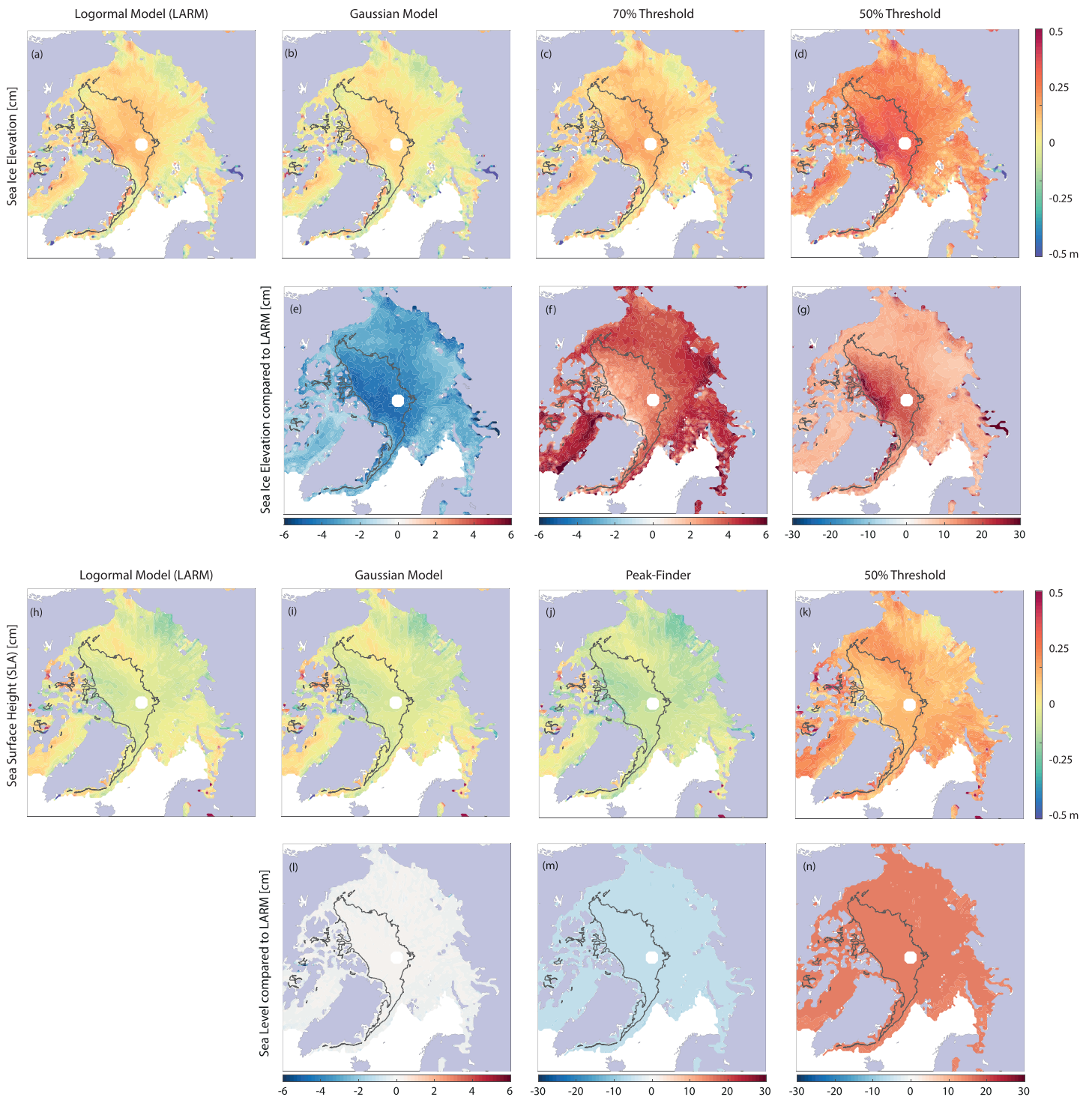
Figure 4. March mean radar sea ice surface elevations obtained from (a) the lognormal model retracker (LARM), (b) the Gaussian model retracker, (c) the 70% threshold retracker, and (d) the 50% threshold retracker, with respect to the DTU15 mean sea surface height for the full 2011–2018 CryoSat-2 record. (e–g) show the sea ice elevation offsets between each retracker minus LARM. March interpolated sea level anomalies (SLAs) from (h) LARM, (i) the Gaussian model, (j) the peak-finder retracker, and (k) the 50% threshold retracker, for 2011–2018, with respect to the DTU15 mean sea surface. (l–n) show the sea level offsets between each retracker minus LARM. Note the varying ranges between offset plots. The average MYI edge is highlighted in gray. Plots (e–g) and (l–n) are reproduced for November in Figure S2.

freeboard. Although we have reproduced the retracking algorithms as closely as possible, it is therefore important to note that our freeboard estimates are not intended to and will not exactly emulate those from other groups, due to the various nuances in the respective processing chains.

We refer to the reproduced GSFC method as the Gaussian model. CPOM use TFMRA with a 70% threshold for retracking sea ice floes and a curve-fitting routine for retracking leads that identifies the waveform peak (Tilling et al., 2018), which we refer to as Threshold70+Peaks. AWI use TFMRA with a 50% threshold for retracking both sea ice and leads (Ricker et al., 2014), which we refer to as Threshold50. A simple multi-parameter classification technique, based on waveform pulse-peakiness (Kurtz et al., 2014; Laxon et al., 2013), stack standard deviation (Laxon et al., 2013) and backscatter coefficient (Paul et al., 2018), is used to separate sea ice and lead returns (Quartly et al., 2019). We apply identical classifications, waveform filtering, geophysical corrections and sea level tie-point interpolation schemes between all methods (Landy et al., 2017). In reality, these steps vary between different group's processing chains. Radar freeboard is calculated from the elevation difference between sea ice floes and sea level. To provide context for our findings, the patterns of radar freeboard from each of the official group's products alongside the 'emulated' versions using our own processor are presented for March 2015 in Figure S3. Our emulated freeboards reproduce the general patterns of each group's true product and, importantly, the relative offsets between freeboards are consistent across the four products.

Winter sea ice radar freeboard from the LARM retracker exhibits the expected pattern of higher radar freeboards (15–35 cm) in the MYI zone north of Greenland and the Canadian Archipelago, with lower radar freeboards (5–20 cm) in the seasonal ice zone and marginal seas (Figure 5a). The Gaussian model produces a comparable regional ice freeboard distribution (Figure 5b); however, FYI freeboards are 0–3 cm lower and, most importantly, freeboards in the MYI zone are 2–6 cm lower than LARM (Figure 5e). This translates into a reasonably consistent 10–20% underestimation across the Arctic (Figure 5h). Both physical retrackers estimate similar elevations for the sea level (Figure 4l), so the ice freeboard distribution closely reflects the observed offsets in measured sea ice elevations (Figure 4e). Over the entire March 2011–2018 record, the mean CryoSat-2 radar freeboard difference when using a Gaussian-based physical retracker compared to LARM is −1.3 cm (−12%) over FYI and −2.7 cm (−15%) over MYI.

The Threshold70+Peaks method uses different thresholds for retracking sea ice and lead waveforms, with a constant 16.26 cm bias deducted to align sea ice elevations to sea level (Tilling et al., 2018), producing generally thinner radar freeboard estimates than the LARM algorithm (Figure 5c). The freeboard offsets between LARM and this method exhibit a similar spatial pattern across the Arctic (Figure 5f), with a March mean freeboard difference of −1.3 cm (−11%) over FYI and −1.3 cm (−7%) over MYI. By deducting a constant bias to align sea ice floe elevations to sea level, the Threshold70+Peaks method coincidentally accounts for some of the expected roughness-induced variation in the retracking threshold between FYI and MYI zones. This is because the constant correction has a relatively larger impact on reducing the freeboard of thinner FYI compared to thicker MYI. However, around the ice pack margins the 16 cm correction appears to reduce ice elevation retrievals excessively, with the estimated freeboard underestimating LARM by >20% (Figure 5i).

The Threshold50 algorithm uses a fixed threshold at 50% of the waveform leading-edge power for retracking both sea ice and lead returns, which produces relatively low radar freeboards over the FYI zone (Figure 5d) —within 2 cm of the LARM and Gaussian Model results (Figure 5g). However, applying the 50% threshold over rough sea ice leads to a systematically higher estimate for the ice elevation (Figure 4g), producing 2–12 cm (5–25%) higher freeboards over the MYI zone than the physical retrackers (Figures 5g and 5j). The March mean radar freeboard difference for the Threshold50 algorithm compared to LARM is −0.3 cm (−3%) over FYI and +2.4 cm (+12%) over MYI.

The key take-away from Figure 5 is that between the four algorithms, with all steps of the processing chain equal except for the retracking method, the March radar freeboards have an uncertainty range of, on average, 15% over FYI and 29% over MYI. These values constitute systematic uncertainties which cannot be reduced by sample averaging and will translate directly into systematic sea ice thickness uncertainties, once the radar freeboard is converted to thickness.
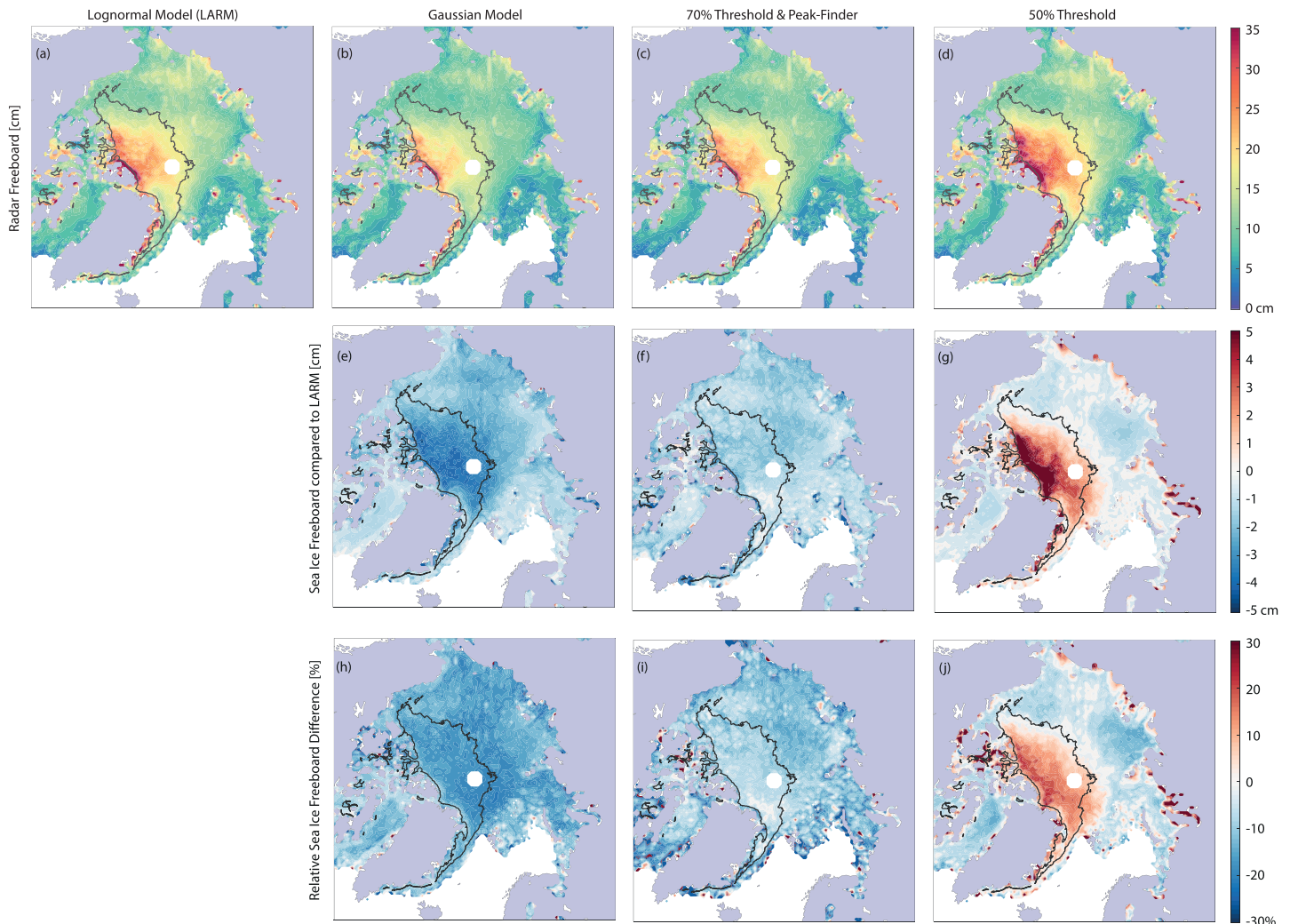
Figure 5. March mean CryoSat-2 radar freeboard [cm] obtained from (a) the Lognormal Altimeter Retracker Algorithm (LARM), (b) the Gaussian model, (c) TFMRA70+Peaks, with a 16.26 cm bias correction, and (d) TFMRA50. Absolute and relative freeboard differences between each retracker algorithm and LARM are provided in (e–g) and (h–j), respectively. The MYI edge is highlighted in gray. The same plots for mean November 2010–2018 freeboard are shown in Figure S4.

6. Independent Validation of Sea Ice Freeboard Retrievals

Although the LARM and GSFC algorithms are based on theoretically more-realistic representations of variable sea ice backscattering properties than empirical threshold-based algorithms, few studies have directly intercompared algorithm performance with respect to independent sea ice observations (Xia & Xie, 2018; Yi et al., 2018). Yi et al. (2018) compared lead and sea ice elevations and derived ice freeboards over several CryoSat-2 tracks, with an identical processing chain save for varied waveform retrackers. Here we reprocess sea ice freeboards from each of the four algorithms introduced above and compare them to broadly coincident OIB observations from 27 airborne campaigns in 2011, 2012, and 2013. For this analysis we use the *L4 Sea Ice Freeboard, Snow Depth, and Thickness V1* dataset (Kurtz et al., 2015), which covers the period 2009–2013. Only CryoSat-2 observations acquired ±3 days either side of an OIB campaign are used in this analysis. Following Xia and Xie (2018), we use snow depths from OIB to correct ATM elevations to the snow-ice interface and to correct CryoSat-2 range for the delayed Ku-band wave propagation within the snowpack (Kwok & Cunningham, 2015). (We use a snow density-dependent correction based on the equation for wave propagation in snow in Ulaby et al., 1982, as described in Mallett et al., 2020.) The processing chains for each CryoSat-2 dataset are identical except for the four different retracking algorithms. Both the OIB and
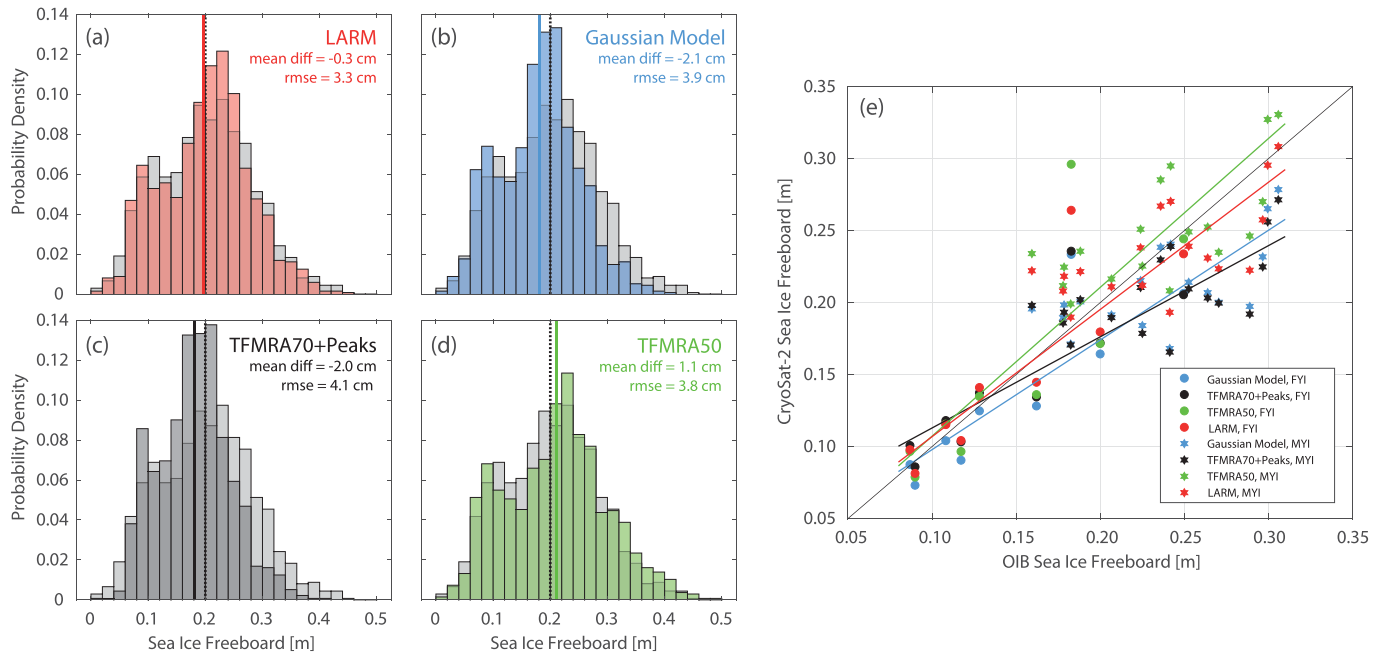
Figure 6. Probability density functions of sea ice freeboard obtained from CryoSat-2, via (a) LARM, (b) the Gaussian model, (c) TFMRA70+Peaks, and (d) TFMRA50, overlaid on ice freeboard PDFs from the L4 OIB sea ice data product (gray background) (Kurtz et al., 2015), within 25 km EASE-grid cells. Insets are the mean differences and root-mean square errors of paired gridded CryoSat-2 and OIB samples. (e) Intercomparison of sea ice freeboard from each CryoSat-2 retracking technique versus mean ice freeboard measured during each of the 27 OIB airborne campaigns. For this analysis, the CryoSat-2 data have been processed using coincident OIB observations of snow depth.

CryoSat-2 observations of sea ice freeboard are sampled onto the same 25-km grid and compared. Ice freeboards from all coincident grid cells in all campaigns are compared in Figures 6a–6d, whereas for Figure 6e, we compare the mean freeboard per campaign as calculated by OIB and CryoSat-2.
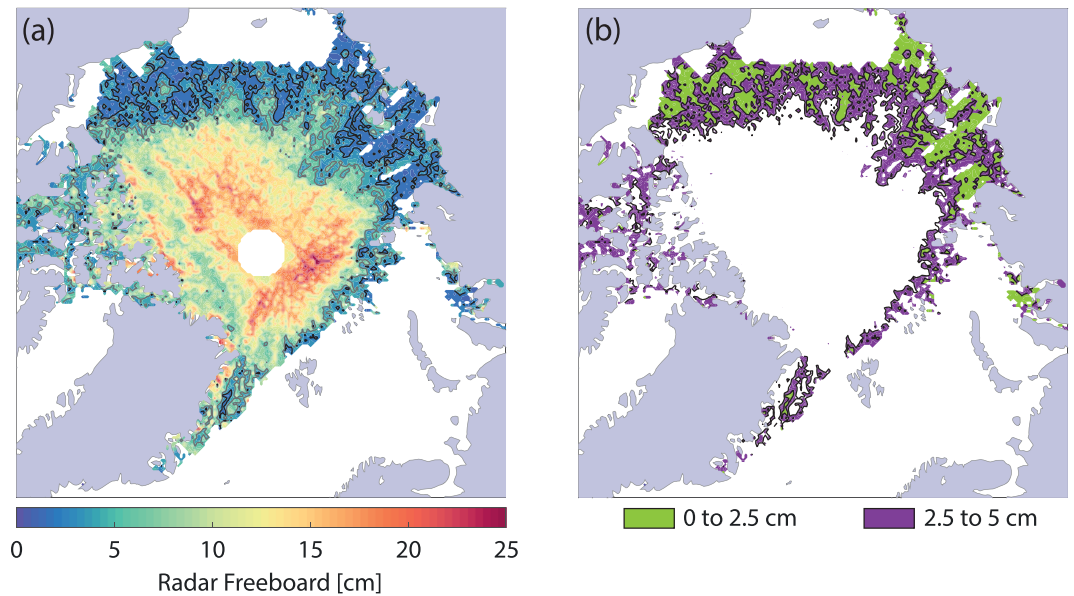

Figure 7. The lower limit on detectable sea ice radar freeboards when applying the LARM algorithm to CryoSat-2 in October 2015. (a) Mean 25-km gridded radar freeboards below 5 cm are enclosed by the gray contour, and radar freeboards below 2.5 cm are enclosed by the black contour. (b) Radar freeboards between 0 and 2.5 cm are shown in green, and freeboards between 2.5 and 5 cm in purple.
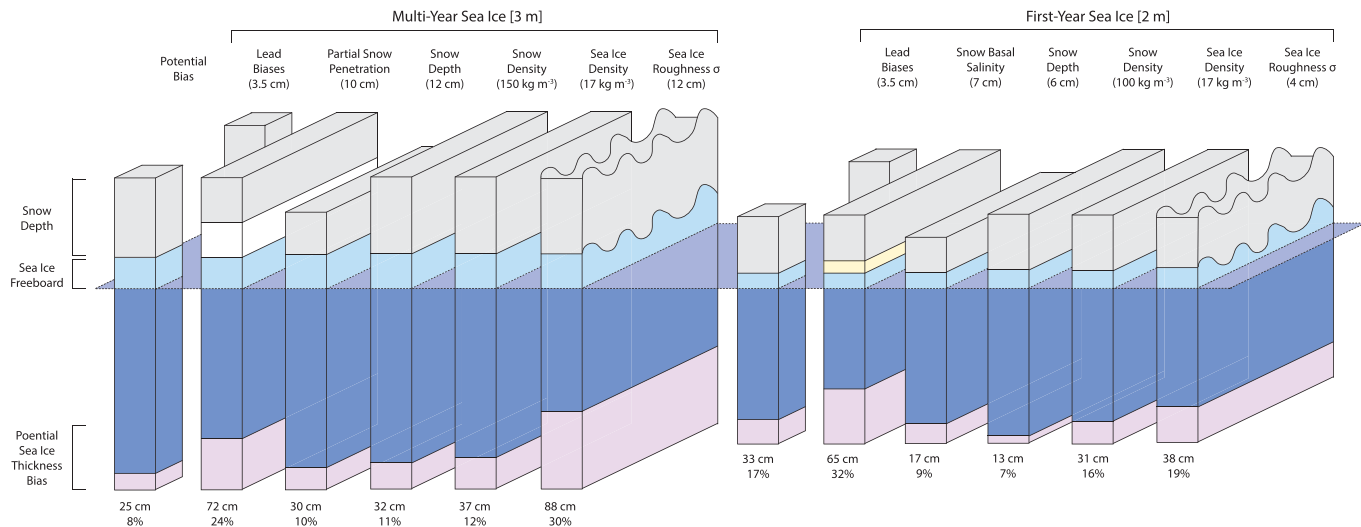
Figure 8. Schematic diagram of the potential systematic bias in CryoSat-2-derived sea ice thickness introduced from each principal source of uncertainty: Partial penetration into the snowpack on multiyear ice, snow depth, snow density, partial penetration into snow on first-year ice owing to snow basal salinity, sea ice density, and sea ice surface roughness. Potential ice thickness biases are illustrated in pink.

PDFs of the 25-km gridded samples highlight the differences between CryoSat-2 waveform retracking approaches. The RMSE of paired samples from OIB and CryoSat-2—when processed with LARM—is lowest of the four algorithms at 3.3 cm (Figure 6a), between 13% and 20% smaller than the RMSE of the alternative algorithms. In our comparison, the Gaussian model and Threshold70+Peaks methods do not appear to capture the thickest sea ice in the upper tail of the OIB ice freeboard distribution (Figures 6b and 6c), underestimating mean OIB freeboards by 2.1 and 2.0 cm, respectively. The TFMRA50 method produces ice freeboards capturing the thick ice tail of the OIB distribution (Figure 6d) but appears to overestimate mean OIB freeboard by 1.1 cm.

It is important to note in this analysis, however, that the uncertainties of individual OIB snow depth or ice freeboard samples are each around 5–6 cm (Kurtz et al., 2015). The OIB snow depths have also exhibited low biases over rough sea ice (King et al., 2015). Although we use coinciding OIB snow depths to convert the CryoSat-2 radar freeboards to sea ice freeboards, so can reasonably discount systematic biases in the snow depth data, it is unlikely that the OIB freeboard uncertainties are uncorrelated along the aircraft track enabling errors to be reduced by sample averaging. Consequently, the differences between the four CryoSat-2 retracking algorithms shown in Figure 6 are likely to be below the uncertainty or "noise floor" of the independent OIB observations.

In Figure 6e, the average sea ice freeboard for all FYI or MYI sections within a single OIB campaign are compared to coinciding CryoSat-2 observations, obtained from each of the four retracking approaches. Variability between the ice freeboards derived from each method is more than three times larger over MYI (standard deviation [SD] = 2.0 cm) than over FYI (SD = 0.6 cm), indicating that the choice of retracking algorithm is more important for retrieving accurate ice freeboards over MYI. The least-square trend between the LARM and OIB freeboards is closest to one-to-one (Figure 6e), confirming that LARM captures the dynamic range of Arctic sea ice freeboards from thin and smooth FYI up to thick and rough MYI (Figure 6a), with a mean difference to the OIB freeboard of only −0.3 cm.

7. Discussion and Conclusions

Another example of the advantages of applying a physically based approach for retracking CryoSat-2 observations is evident over thin sea ice. It has been conventionally estimated that CryoSat-2 has a minimum ice thickness retrieval limit of ~0.5 m, owing to the challenge measuring radar freeboards thinner than ~5 cm through radar altimetry (Ricker et al., 2017). However, our model simulations indicate that smooth newly-formed sea ice, with a roughness height σ of only 1–3 cm, should produce radar echoes with a

retracking point >95% (Landy et al., 2019). A fixed 50% or 70% retracking threshold for sea ice floes therefore imposes a strict but artificial lower limit on the detectable ice freeboard. A roughness-based correction would need to be removed from the retracked sea ice floe elevations to accurately resolve the freeboards of thinner ice floes (Laforge et al., 2020). Based on the range difference between modeled echoes from smooth thin sea ice with a >95% retracking threshold and from leads with a ~98% threshold (Figure 3c), this places a theoretical lower detection bound on the ice freeboard retrieval of approximately 2.5 cm. As evidence for this improved detection limit, we observe mean radar freeboards as low as 2–3 cm covering regions of several thousand square kilometers at the pack ice margins during October, calculated from the LARM algorithm (e.g., Figure 7). In the absence of snow cover, this yields a minimum detectable ice thickness of around 0.25 m.

It is clear from Figure 5 that, with the waveform classifications and geophysical corrections equal between methods, different retracking algorithms still have a significant impact on the measured sea ice freeboard. Despite the pan-Arctic mean radar freeboard only varying by 4–6 cm between methods (section 5) the freeboard pattern, and asymmetry between FYI and MYI zones in particular, can vary considerably. We follow the AWI processing chain (Ricker et al., 2014) to estimate the variations in ice thickness and volume introduced by these uncertainties in retracking approach. This includes using adapted snow depth and density estimates from (Warren et al., 1999) and correcting for the delayed Ku-band wave propagation speed in snow. For the Gaussian Model, pan-Arctic mean sea ice thickness for March 2011–2018 is 15 and 19 cm lower than the LARM algorithm, for FYI and MYI, respectively. For the Threshold70+Peaks method, mean ice thickness is 14 and 9 cm lower than LARM. And finally, the Threshold50 method is 1 cm lower than LARM for FYI and 17 cm larger than LARM for MYI. Total sea ice volume (above a latitude of 70°) compared to LARM is -605 km^3 (-8%) for the Gaussian Model, -430 km^3 (-5%) for Threshold70+Peaks, and $+130$ km^3 ($+3\%$) for Threshold50. These variations in ice thickness and volume will be different between months and years, but the range of the percentage uncertainty is similar to the estimate in Kwok and Cunningham (2015). Systematic uncertainties in ice thickness/volume of this magnitude could have severe implications for calculating the seasonal and interannual sea ice thickness/volume budget, including ice growth and melt rates (e.g., Kwok, 2018; Petty et al., 2018; Stroeve et al., 2018), fluxes of sea ice exported from the Arctic via Fram Strait (Ricker et al., 2018), for initializing dynamical sea ice forecasting systems (Day et al., 2014; Schröder et al., 2019), and for constraining freshwater fluxes into the Arctic Ocean (Bacon et al., 2015).

Sallila et al. (2019) discovered a similar pattern of pan-Arctic offsets between the CPOM and AWI sea ice thickness products to our emulated freeboards, with the MYI consistently thicker and FYI consistently thinner in the AWI product. They also found the GSFC ice thickness product to be universally thicker than the other products, which contradicts our finding that minimum radar freeboard was derived from the Gaussian model retracker. However, results from Yi et al. (2018) and Xia and Xie (2018) demonstrate that GSFC sea ice freeboards are significantly lower than coincident freeboards from AWI and CPOM algorithms, corroborating our findings (Figure S3) and suggesting that thicker ice in the GSFC product (Kurtz et al., 2014) is likely introduced in the processing chain for converting ice freeboard to thickness.

It is still an open question as to how much retracker/roughness-based uncertainties compare to other potential sources of uncertainty in the estimation of ice thickness from satellite radar altimeter observations, especially considering the lack of validation data to constrain the relative errors. Here we attempt to estimate the systematic uncertainty budget of Arctic sea ice, evaluating our new results in the context of those from the literature. It is important to emphasize that systematic uncertainties only manifest as biases when they can be identified as such against a reference "truth" and we are not therefore suggesting that one sea ice thickness product is necessarily more accurate than others. Rather, these estimates characterize realistic upper limits of the main physical parameters *potentially* introducing systematic (i.e., signed) bias to sea ice thickness retrievals. Uncertainties are intended to be applicable at the scale of a gridded sea ice thickness product from CryoSat-2, with >10-km grid cells, rather than individual ~300-m along-track radar samples which could have far larger errors.

The following principal sources of systematic uncertainty are considered: (i) partial wave penetration into the snowpack on MYI, for instance, due to metamorphic snow features (King et al., 2018; Kurtz et al., 2014; Ricker et al., 2015; Tonboe et al., 2010; Willatt et al., 2009; Willatt et al., 2011); (ii) snow depth and (iii)

density (Alexandrov et al., 2010; Kern et al., 2015; Lawrence et al., 2018; Ricker et al., 2014; Rostosky et al., 2018; Tonboe et al., 2010; Warren et al., 1999; Zygmuntowska et al., 2014); (iv) partial penetration into the snowpack on FYI, for instance due to brine wicking-induced snow basal salinity (Nandan et al., 2017; Willatt et al., 2011); (v) sea ice density (Alexandrov et al., 2010; Tonboe et al., 2010; Zygmuntowska et al., 2014); (vi) sea level bias resulting from off-nadir ranging to leads (Armitage & Davidson, 2013) and low along-track lead densities (Kwok & Cunningham, 2015); and finally (vii) sea ice surface roughness (Hendricks et al., 2010; Ricker et al., 2014) (results presented in this study). Different sources of uncertainty may also be correlated with each other rather than entirely independent (Lawrence et al., 2018). Systematic sea ice thickness uncertainties are calculated by multiplying the component errors with the partial derivative of the respective variable (e.g., Ricker et al., 2014).

Estimates for individual component uncertainties range from 7–32% over FYI and from 8–30% over MYI (Figure 8). Systematic uncertainties introduced in the retracking step have the potential to bias sea ice thickness estimates up to 20% over FYI and up to 30% over MYI (Figure 8), with regional patterns depending on the ice surface roughness (Figures 5e–5j). For FYI with a thickness of 2 m, this translates into an uncertainty of around 40 cm, and for MYI with a thickness of 3 m, an uncertainty of around 90 cm. Over FYI, the 20% roughness uncertainty contributes more than the combined systematic errors from snow depth and density but is lower than the potential uncertainty introduced by unknown snow basal salinity. Over MYI, the 30% roughness uncertainty contributes greater systematic uncertainty than any other single source of error.

Given the error that sea ice roughness can potentially introduce into thickness estimates, our results demonstrate the importance of determining a consistent and rigorous approach for extracting information, including freeboard and the roughness itself, from SAR altimeter observations of sea ice. It is anticipated that sea ice will generally become "smoother" as the Arctic's remaining MYI is replaced by FYI (Martin et al., 2016; Rampal et al., 2011), but with regions of the ice cover under compression more easily deformed and ridged. This must be accounted for if we are to accurately estimate sea ice thickness from future altimeter missions. We hope our findings will inform scenario testing and developments in the lead up to new missions, such as the candidate ESA dual Ka-/Ku-band Copernicus Polar Ice Snow and Topography Altimeter (CRISTAL) mission. It is clear from our results that Arctic radar and sea ice freeboards still require thorough evaluation, before focus moves onto evaluating the derived sea ice thickness. We also aim to build on this knowledge of surface roughness and satellite radar retracking to better understand CryoSat-2 returns over the Antarctic sea ice pack (which can feature regions of high surface roughness) (Fons & Kurtz, 2019) and in future efforts to compare and reconcile the sea ice thickness data produced from NASA's new ICESat-2 laser altimetry mission (Kwok et al., 2019).

References

Alexandrov, V., Sandven, S., Wahlin, J., & Johannessen, O. (2010). The relation between sea ice thickness and freeboard in the Arctic. *The Cryosphere*, *4*(3), 373–380.

Armitage, T., & Davidson, M. (2013). Using the interferometric capabilities of the ESA Cryosat-2 mission to improve the accuracy of sea ice freeboard retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(1), 529–536.

Bacon, S., Aksenov, Y., Fawcett, S., & Madec, G. (2015). Arctic mass, freshwater and heat fluxes: Methods and modelled seasonal variability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*(2052).

Breivik, L., Eastwood, S., & Lavergne, T. (2012). Use of C-band scatterometer for sea ice edge identification. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(7), 1669–2677.

Castellani, G., Lüpkes, C., Hendricks, S., & Gerdes, R. (2014). Variability of Arctic sea-ice topography and its impact on the atmospheric surface drag. *Journal of Geophysical Research: Oceans*, *119*, 6743–6762. https://doi.org/10.1002/2013JC009712

Davis, N., & Wadhams, P. (1995). A statistical analysis of Arctic pressure ridge morphology. *Journal of Geophysical Research*, *100*(C6), 10,915–10,925.

Day, J., Hawkins, E., & Tietsche, S. (2014). Will Arctic sea ice thickness initialization improve seasonal forecast skill? *Geophysical Research Letters*, *41*, 7566–7575. https://doi.org/10.1002/2014GL061694

Dinardo, S., Fenoglio-Marc, L., Buchhaupt, C., Becker, M., Scharroo, R., Joana Fernandes, M., & Benveniste, J. (2018). Coastal SAR and PLRM altimetry in German Bight and West Baltic Sea. *Advances in Space Research*, *62*(6), 1371–1404. https://doi.org/10.1016/j.asr.2017.12.018

Doble, M., Skourup, H., Wadhams, P., & Geiger, C. (2011). The relation between Arctic sea ice surface elevation and draft: A case study using coincident AUV sonar and airborne scanning laser. *Journal of Geophysical Research*, *116*, C00E03. https://doi.org/10.1029/2011JC007076

Fetterer, F. M., Drinkwater, M. R., Jezek, K. C., Laxon, S. W. C., Onstott, R. G., & Ulander, L. M. H. (1992). Sea ice altimetry. In F. Carsey (Ed.), *Microwave remote sensing of sea ice. Geophysical monograph* (62 ed., pp. 111–135). Washington D.C: American Geophysical Union. https://doi.org/10.1029/GM068p0111

Fons, S., & Kurtz, N. (2019). Retrieval of snow freeboard of Antarctic sea ice using waveform fitting of CryoSat-2 returns. *The Cryosphere*, *13*(3), 861–878.

Fung, A., & Eom, H. (1983). Coherent scattering of a spherical wave from an irregular surface. *IEEE Transactions on Geoscience and Remote Sensing*, *31*(1), 68–72.

Fung, A. K., (1994). Microwave scattering and emission models and their applications. s.l.:Artech House, Inc..

Giles, K. A., Laxon, S. W., & Ridout, A. L. (2014). Circumpolar thinning of Arctic sea ice following the 2007 record ice extent minimum. *Geophysical Research Letters*, *35*, L22502. https://doi.org/10.1029/2008GL035710

Guerreiro, K., Fleury, S., Zakharova, E., Kouraev, A., Rémy, F., & Maisongrande, P. (2017). Comparison of CryoSat-2 and ENVISAT radar freeboard over Arctic sea ice: Toward an improved Envisat freeboard retrieval. *The Cryosphere*, *11*(5), 2059–2073. https://doi.org/10.5194/tc-11-2059-2017

Hendricks, S., Stenseng, L., Helm, V., & Haas, C. (2010). Effects of surface roughness on sea ice freeboard retrieval with an airborne Ku-band SAR radar altimeter. Honolulu, USA, IEEE Geosci. Rem. Sens. Symposium, pp. 3126–3129.

Iacozza, J., & Barber, D. (1999). An examination of the distribution of snow on sea-ice. *Atmosphere-Ocean*, *37*(1), 21–51.

Kern, S., Khvorostovsky, K., Skourup, H., Rinne, E., Parsakhoo, Z. S., Djepa, V., et al. (2015). The impact of snow depth, snow density and ice density on sea ice thickness retrieval from satellite radar altimetry: Results from the ESA-CCI Sea Ice ECV Project Round Robin Exercise. *The Cryosphere*, *9*(1), 37–52. https://doi.org/10.5194/tc-9-37-2015

King, J., Howell, S., Derksen, C., Rutter, N., Toose, P., Beckers, J. F., et al. (2015). Evaluation of Operation IceBridge quick-look snow depth estimates on sea ice. *Geophysical Research Letters*, *42*, 9302–9310. https://doi.org/10.1002/2015GL066389

King, J., Skourup, H., Hvidegaard, S. M., Rösel, A., Gerland, S., Spreen, G., et al. (2018). Comparison of freeboard retrieval and ice thickness calculation from ALS, ASIRAS, and CryoSat-2 in the Norwegian Arctic to field measurements made during the N-ICE2015 expedition. *Journal of Geophysical Research: Oceans*, *123*, 1123–1141. https://doi.org/10.1002/2017JC013233

Kurtz, N., Galin, N., & Studinger, M. (2014). An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting. *The Cryosphere*, *8*, 1217–1237.

Kurtz, N., Studinger, M., Harbeck, J., Onana, V., & Yi, D. (2015). IceBridge L4 Sea Ice Freeboard, Snow Depth, and Thickness, *Version 1,* Boulder, Colorado USA: S.n.

Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., et al. (2013). Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data. *The Cryosphere*, *7*(4), 1035–1056. https://doi.org/10.5194/tc-7-1035-2013

Kwok, R. (2014). 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

Kwok, R. (2018). Arctic Sea ice thickness, volume, and multiyear ice coverage: Losses and coupled variability (1958–2018). *Environmental Research Letters*, *13*(10).

Kwok, R., & Cunningham, G. (2015). Variability of Arctic sea ice thickness and volume from Cryosat-2. *Philosophical Transactions of the Royal Society A*, *373*.

Kwok, R., Markus, T., Kurtz, N. T., Petty, A. A., Neumann, T. A., Farrell, S. L., et al. (2019). Surface height and sea ice freeboard of the Arctic Ocean from ICESat-2: Characteristics and early results. *Journal of Geophysical Research: Oceans*, *124*, 6942–6959. https://doi.org/10.1029/2019JC015486

Laforge, A., Fleury, S., Dinardo, S., Garnier, F., Remy, F., Benveniste, J., et al. (2020). Toward improved sea ice freeboard observation with SAR altimetry using the physical retracker SAMOSA+. *Advances in Space Research*. https://doi.org/10.1016/j.asr.2020.02.001

Landy, J., Ehn, J., & Barber, D. (2015). Albedo feedback enhanced by smoother Arctic sea ice. *Geophysical Research Letters*, *42*, 710–714.

Landy, J., Isleifson, D., Komarov, A., & Barber, D. (2015). Parameterization of centimeter-scale sea ice surface roughness using terrestrial LiDAR. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(3), 1271–1286.

Landy, J., Komarov, A., & Barber, D. (2015). Numerical and experimental evaluation of terrestrial LiDAR for parameterizing centimeter-scale sea ice surface roughness. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(9), 4887–4898.

Landy, J., Tsamados, M., & Scharien, R. (2019). A facet-based numerical model for simulating SAR altimeter echoes from heterogeneous sea ice surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(7), 4164–4180.

Landy, J. C., Ehn, J. K., Babb, D. G., Thériault, N., & Barber, D. G. (2017). Sea ice thickness in the Eastern Canadian Arctic: Hudson Bay Complex & Baffin Bay. *Remote Sensing of Environment*, *200*, 281–294. https://doi.org/10.1016/j.rse.2017.08.019

Lawrence, I. R., Tsamados, M. C., Stroeve, J. C., Armitage, T. W. K., & Ridout, A. L. (2018). Estimating snow depth over Arctic sea ice from calibrated dual-frequency radar freeboards. *The Cryosphere*, *12*(11), 3551–3564. https://doi.org/10.5194/tc-12-3551-2018

Laxon, S., Peacock, N., & Smith, D. (2003). High interannual variability of sea ice thickness in the Arctic region. *Nature*, *425*(6961), 947–950. https://doi.org/10.1038/nature02050

Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., et al. (2013). CryoSat-2 estimates of Arctic sea ice thickness and volume. *Geophysical Research Letters*, *40*, 732–737. https://doi.org/10.1002/grl.50193

Liu, C., Chao, J., Gu, W., Li, L., & Xu, Y. (2014). On the surface roughness characteristics of the land fast sea-ice in the Bohai Sea. *Acta Oceanologica Sinica*, *33*(7), 97–106. https://doi.org/10.1007/s13131-014-0509-3

Mallett, R. D. C., Lawrence, I. R., Stroeve, J. C., Landy, J. C., & Tsamados, M. (2020). 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*(1), 251–260. https://doi.org/10.5194/tc-14-251-2020

Martin, T., Tsamados, M., Schroeder, D., & Feltham, D. (2016). The impact of variable sea ice roughness on changes in Arctic Ocean surface stress: A model study. *Journal of Geophysical Research: Oceans*, *121*, 1931–1952. https://doi.org/10.1002/2015JC011186

Nandan, V., Geldsetzer, T., Yackel, J., Mahmud, M., Scharien, R., Howell, S., et al. (2017). Effect of snow salinity on CryoSat-2 Arctic first-year Sea ice freeboard measurements. *Geophysical Research Letters*, *44*, 10,419–10,426. https://doi.org/10.1002/2017GL074506

Paden, J. et al., 2017. IceBridge Ku-Band Radar L1B Geolocated Radar Echo Strength Profiles, Version 2, Boulder, Colorado, USA: NASA National Snow and Ice Data Center Distributed Active Archive Center.

Paul, S., Hendricks, S., Ricker, R., Kern, S., & Rinne, E. (2018). 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*(7), 2437–2460. https://doi.org/10.5194/tc-12-2437-2018

Petty, A., Holland, M., Bailey, D., & Kurtz, N. (2018). Warm Arctic, increased winter sea ice growth? *Geophysical Research Letters*, *45*, 12,922–12,930. https://doi.org/10.1029/2018GL079223

Petty, A., Tsamados, M., & Kurtz, N. (2017). Atmospheric form drag coefficients over Arctic sea ice using remotely sensed ice topography data, spring 2009–2015. *Journal of Geophysical Research: Earth Surface*, *122*, 1472–1490. https://doi.org/10.1002/2017JF004209

Petty, A. A., Tsamados, M. C., Kurtz, N. T., Farrell, S. L., Newman, T., Harbeck, J. P., et al. (2016). Characterizing Arctic sea ice topography using high-resolution IceBridge data. *The Cryosphere*, *10*(3), 1161–1179. https://doi.org/10.5194/tc-10-1161-2016

Quartly, G. D., Rinne, E., Passaro, M., Andersen, O. B., Dinardo, S., Fleury, S., et al. (2019). Retrieving sea level and freeboard in the Arctic: A review of current radar altimetry methodologies and future perspectives. *Remote Sensing*, *11*(7), 881. https://doi.org/10.3390/rs11070881

Rampal, P., Weiss, J., Dubois, C., & Campin, J. (2011). IPCC climate models do not capture Arctic sea ice drift acceleration: Consequences in terms of projected sea ice thinning and decline. *Journal of Geophysical Research*, *116*, C00D07. https://doi.org/10.1029/2011JC007110

Ray, C., Martin-Puig, C., Clarizia, M. P., Ruffini, G., Dinardo, S., Gommenginger, C., & Benveniste, J. (2015). SAR altimeter backscattered waveform model. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(2), 911–919. https://doi.org/10.1109/TGRS.2014.2330423

Ricker, R., Girard-Ardhuin, F., Krumpen, T., & Lique, C. (2018). Satellite-derived sea ice export and its impact on Arctic ice mass balance. *The Cryosphere*, *12*(9), 3017–3032.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., & Davidson, M. (2014). Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation. *The Cryosphere*, *8*(4), 1607–1622. https://doi.org/10.5194/tc-8-1607-2014

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J., & Haas, C. (2017). A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data. *The Cryosphere*, *11*(4), 1607–1623. https://doi.org/10.5194/tc-11-1607-2017

Ricker, R., Hendricks, S., Perovich, D. K., Helm, V., & Gerdes, R. (2015). Impact of snow accumulation on CryoSat-2 range retrievals over Arctic sea ice: An observational approach with buoy data. *Geophysical Research Letters*, *42*, 4447–4455. https://doi.org/10.1002/2015GL064081

Rivas, M., Maslanik, J., Sonntag, J., & Axelrad, P. (2006). Sea ice roughness from airborne LIDAR profiles. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(11), 3032–3037.

Rodriguez-Morales, F., Gogineni, S., Leuschen, C. J., Paden, J. D., Li, J., Lewis, C. C., et al. (2013). Advanced multifrequency radar instrumentation for polar research. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(5), 2824–2842.

Rostosky, P., Spreen, G., Farrell, S. L., Frost, T., Heygster, G., & Melsheimer, C. (2018). Snow depth retrieval on Arctic sea ice from passive microwave radiometers—Improvements and extensions to multiyear ice using lower frequencies. *Journal of Geophysical Research: Oceans*, *123*, 7120–7138. https://doi.org/10.1029/2018JC014028

Sallila, H., Farrell, S., McCurry, J., & Rinne, E. (2019). Assessment of contemporary satellite sea ice thickness products for Arctic sea ice. *The Cryosphere*, *13*(4), 1187–1213.

Schröder, D., Feltham, D. L., Tsamados, M., Ridout, A., & Tilling, R. (2019). New insight from CryoSat-2 sea ice thickness for sea ice modelling. *The Cryosphere*, *13*(1), 125–139. https://doi.org/10.5194/tc-13-125-2019

Stroeve, J., Schroder, D., Tsamados, M., & Feltham, D. (2018). Warm winter, thin ice? *The Cryosphere*, *12*(5), 1791–1809.

Studinger, M., (2018). IceBridge ATM L1B Elevation and Return Strength, Version 2. *[2013-2014]*, Boulder, Colorado, USA: NASA National Snow and Ice Data Center Distributed Archive.

Tian-Kunze, X., Kaleschke, L., Maaß, N., Mäkynen, M., Serra, N., Drusch, M., & Krumpen, T. (2014). SMOS-derived thin sea ice thickness: Algorithm baseline, product specifications and initial verification. *The Cryosphere*, *8*(3), 997–1018. https://doi.org/10.5194/tc-8-997-2014

Tilling, R., Ridout, A., & Shepherd, A. (2018). Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data. *Advances in Space Research*, *62*(6), 1203–1225.

Tonboe, R., Pedersen, L., & Haas, C. (2010). Simulation of the CryoSat-2 satellite radar altimeter sea ice thickness retrieval uncertainty. *Canadian Journal of Remote Sensing*, *36*(1), 55–67.

Ulaby, F., Moore, R., & Fung, A. (1982). *Microwave remote sensing: Active and passive*. Boston, MA: Artech House.

Wadhams, P., Tucker, W. B. III, Krabill, W. B., Swift, R. N., Comiso, J. C., & Davis, N. R. (1992). Relationship between sea ice freeboard and draft in the Arctic Basin, and implications for ice thickness monitoring. *Journal of Geophysical Research*, *97*(C12), 20,325–20,334. https://doi.org/10.1029/92JC02014

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., & Colony, R. (1999). Snow depth on Arctic Sea ice. *Journal of Climate*, *12*(6), 1814–1829. https://doi.org/10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2

Willatt, R., Giles, K. A., Laxon, S. W., Stone-Drake, L., & Worby, A. P. (2009). Field investigations of Ku-band radar penetration into snow cover on Antarctic sea ice. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(1), 365–372.

Willatt, R., Laxon, S., Giles, K., Cullen, R., Haas, C., & Helm, V. (2011). Ku-band radar penetration into snow cover on Arctic sea ice using airborne data. *Annals of Glaciology*, *52*(57), 197–205. https://doi.org/10.3189/172756411795931589

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., et al. (2006). CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields. *Advances in Space Research*, *37*(4), 841–871. https://doi.org/10.1016/j.asr.2005.07.027

Xia, W., & Xie, H. (2018). Assessing three waveform retrackers on sea ice freeboard retrieval from Cryosat-2 using Operation IceBridge Airborne altimetry datasets. *Remote Sensing of Environment*, *204*, 456–471.

Yi, D., Kurtz, N., Harbeck, J., Kwok, R., Hendricks, S., & Ricker, R. (2018). Comparing coincident Elevation and freeboard from IceBridge and five different CryoSat-2 Retrackers. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(2), 1219–1229.

Zygmuntowska, M., Rampal, P., Ivanova, N., & Smedsrud, L. (2014). Uncertainties in Arctic Sea ice thickness and volume: New estimates and implications for trends. *The Cryosphere*, *8*(2), 705–720.