Assessment of ICESat-2 sea ice surface classification with Sentinel-2 imagery: implications for freeboard and new estimates of lead and floe geometry

Alek Petty[1], Marco Bagnardi[1], Nathan Kurtz[1], Rachel Tilling[1], Steven Fons[1], Thomas Armitage[2], Christopher Horvat[3], and Ron Kwok[4]

[1]NASA Goddard Space Flight Center
[2]Jet Propulsion Laboratory
[3]Brown University
[4]Applied Physics Laboratory, Polar Science Center, University of Washington

November 23, 2022

Abstract

NASA's Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) mission launched in September 2018 and is now providing high-resolution surface elevation profiling across the entire globe, including the sea ice cover of the Arctic and Southern Oceans. For sea ice applications, successfully discriminating returns between sea ice and open water is key for accurately determining freeboard, the extension of sea ice above local sea level, and new information regarding the geometry of sea ice floes and leads. We take advantage of near-coincident optical imagery obtained from the European Space Agency (ESA) Sentinel-2 (S-2) satellite over the Western Weddell Sea of the Southern Ocean in March 2019 and the Lincoln Sea of the Arctic Ocean in May 2019 to evaluate the surface classification scheme in the ICESat-2 ATL07 and ATL10 sea ice products. We find a high level of agreement between the ATL07 (specular) lead classification and visible leads in the S-2 imagery in these two scenes across all six ICESat-2 beams, increasing our confidence in the freeboard products and deriving new estimates of the sea ice state. The S-2 overlays provide additional evidence of the misclassification of dark leads, which are no longer used to derive sea surface in the third release (r003) ICESat-2 sea ice products. We show estimates of lead fraction and more preliminary estimates of chord length (a proxy for floe size) using two metrics for classifying sea surface (lead) segments across both the Arctic and Southern Ocean for the first winter season of data collection.

Assessment of ICESat-2 sea ice surface classification with Sentinel-2 imagery: implications for freeboard and new estimates of lead and floe geometry

A. A. Petty[1, 2], M. Bagnardi[2, 3], N. Kurtz[1], R. Tilling[1, 2], S. Fons[2, 4], T. Armitage, C. Horvat[5], R. Kwok[6]

[1] Earth System Interdisciplinary Center, University of Maryland, College Park, MD, USA

[2] NASA Goddard Space Flight Center, Greenbelt, MD, USA

[3] ADNET Systems, Inc., Bethesda, MD, USA

[4] Department of Atmospheric and Oceanic Sciences, University of Maryland, College Park, MD, USA

[5] Brown University, Providence, RI, USA.

[6] University of Washington, Seattle, WA, USA.

Corresponding author: Alek A. Petty (alek.a.petty@nasa.gov)

Key points

- Specular leads in ICESat-2 show strong agreement with leads observed in coincident Sentinel-2 imagery.
- The scenes provide further evidence of the misclassification of dark leads, which are no longer used to derive freeboard in ICESat-2 data.
- We produce preliminary estimates of lead fraction and chord length across the Arctic and Southern Ocean.

Abstract

NASA's Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) mission launched in September 2018 and is now providing high-resolution surface elevation profiling across the entire globe, including the sea ice cover of the Arctic and Southern Oceans. For sea ice applications, successfully discriminating returns between sea ice and open water is key for accurately determining freeboard, the extension of sea ice above local sea level, and new information regarding the geometry of sea ice floes and leads. We take advantage of near-coincident optical imagery obtained from the European Space Agency (ESA) Sentinel-2 (S-2) satellite over the

Western Weddell Sea of the Southern Ocean in March 2019 and the Lincoln Sea of the Arctic Ocean in May 2019 to evaluate the surface classification scheme in the ICESat-2 ATL07 and ATL10 sea ice products. We find a high level of agreement between the ATL07 (specular) lead classification and visible leads in the S-2 imagery in these two scenes across all six ICESat-2 beams, increasing our confidence in the freeboard products and deriving new estimates of the sea ice state. The S-2 overlays provide additional evidence of the misclassification of dark leads, which are no longer used to derive sea surface in the third release (r003) ICESat-2 sea ice products. We show estimates of lead fraction and more preliminary estimates of chord length (a proxy for floe size) using two metrics for classifying sea surface (lead) segments across both the Arctic and Southern Ocean for the first winter season of data collection.

1. Introduction

NASA's Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) is a new satellite mission providing unprecedented (high resolution and accuracy) profiling of the Earth's surface, especially over the high-latitude polar regions. ICESat-2 was launched successfully from Vandenberg Air Force Base on September 15th, 2018 and started acquiring science quality data on October 14th, 2018. A description of the ICESat-2 mission design and objectives can be found in Markus et al., (2017) and Neumann et al., (2019).

ICESat-2 consists of a bespoke photon-counting Advanced Topographic Laser Altimeter System (ATLAS) - a low pulse-energy laser split into a novel six-beam configuration of three beam pairs (a strong and a weak beam). At orbital velocities, ATLAS generates individual laser footprints of ~14 m (in diameter) on the Earth's surface. The high pulse repetition rate of 10 kHz means laser pulses are separated by only 70 cm along-track; a much higher resolution and sampling than NASA's previous laser altimetry mission ICESat (Zwally et al., 2002). ICESat-2 samples year-round and has only experienced minimal downtime since launch.

Various data products are now being generated routinely from IS2 which are publicly disseminated through the National Snow and Ice Data Center (NSIDC, https://nsidc.org/data/icesat-2). The primary science data product from ICESat-2 is the Level 2A ATL03 geolocated photon heights (Neumann et al., 2019) from which most of the higher-level science products are derived. For sea ice users, the primary products of interest are the L3A

along-track sea ice height and type (ATL07, https://nsidc.org/data/ATL07) and freeboard (ATL10, https://nsidc.org/data/ATL10) products. The first winter-season of Arctic sea ice heights and freeboards (October 2018 to March 2019) were presented in (Kwok, et al., 2019). Sea ice thickness has been estimated from these freeboards using external snow loading estimates from the NASA Eulerian Snow on Sea Ice Model (NESOSIM) and modified versions of the Warren climatology (Petty et al., 2020) and through the combination of ICESat-2 laser (total) freeboard with CryoSat-2 radar (ice) freeboard to derive snow depth and freeboard concurrently (Kwok et al., 2020). The surface heights and freeboards were assessed against coincident airborne elevation data collected by NASA's Operation IceBridge (Kwok et al., 2019b). The along-track heights generated from the Airborne Topographic Mapper (ATM, Studinger, 2014) flown on OIB showed strong agreement with ATL07-derived heights, while the derived freeboard from ATM showed only small (0-4 cm) differences with ATL10 freeboards, depending on the method used to derive and compare freeboard. Coincident imagery obtained during the OIB surveys by the Continuous Airborne Mapping By Optical Translator (CAMBOT) alluded to erroneous classification of dark leads in the presence of clouds (Kwok et al., 2020b). Dark leads are still included in the latest (r003) ATL07 sea ice height and surface classification product release, but have now been excluded from the derivation of freeboard in the latest (r003) ATL10 product release (the ATL07 and ATL10 products are described more in the following section). The Kwok et al., (2020b) study highlighted the significant utility of coincident high-resolution imagery for better understanding the performance of the ICESat-2 sea ice surface classification algorithm. The assessment, however, was hindered by the high consolidation of the ice pack in the region profiled during the spring 2019 OIB campaign, limiting the number of leads/open water segments and the quality of the freeboard estimates due to the lack of reliable sea surface tie-points.

The main objective of this study is to further assess the performance of the surface classification scheme in ATL07 for discriminating between sea ice and open water. We take advantage of near-coincident optical imagery from the Sentinel-2 satellite mission to assess this specific aspect of the ATL07 algorithm performance. Successfully discriminating returns between sea ice and open water has benefits beyond deriving freeboard. Additional metrics for understanding the sea ice state include concentration, lead fraction (or density) and floe size. In the marginal seas, where waves can break up and fracture the ice (Horvat et al., 2020 and

references therin), ice floes are thought to be smaller and less concentrated/consolidated (a higher lead fraction) than within the pack ice. The geometry of floes and leads is an important control on the strength of sea ice and its thermodynamic response to forcing (Feltham, 2005; Horvat et al., 2016), particularly in marginal, coastal, and seasonally-ice-covered seas. The floe size distribution (Rothrock & Thorndike, 1984) is increasingly being introduced into sea ice components of climate models (Bateson et al., 2020; Roach et al., 2018) , but only limited basin-scale observational constraints exist to-date, e.g. estimates derived from ESA's CryoSat-2 radar altimeter for the Arctic Ocean only (Horvat et al., 2019). ICESat-2 offers the exciting potential to provide new observational estimates of the floe size distribution, benefiting from the small footprint, high precision and along-track sampling rate across the multiple beams. Satellite tracks (e.g. those from ICESat-2) will profile floes at random angles and cross sections of the floe. The along-track length of a floe measured by satellite is commonly referred to as a chord length. A collection of floe chord lengths can provide statistics of floe geometry, e.g. moments of the floe size distribution and the open water fraction, under certain assumptions about the underlying floe geometry (Horvat et al., 2019).

The coincident S-2 scenes provide the ideal means for assessing new lead fraction and chord length estimates from ICESat-2 data, complementing the freeboard/thickness estimates already being generated. We choose to discuss lead fraction instead of sea ice concentration as concentration/extent was discussed in Horvat et al., (2020) and an assessment/comparison with the commonly used passive microwave-derived concentrations (taking into account the significant sampling differences) is beyond the scope of this study.

2. Data

2.1 ATL07 sea ice heights and surface classification

We primarily use data from the ICESat-2 ATL07 sea ice height and type product. A full description of the ATL07 product can be found in the Algorithm Theoretical Basis Document (ATBD, Kwok et al., 2019c) and the recent changes to this algorithm for r003 are discussed more in Kwok et al., (2020b), so here we provide only a basic overview of the methodology relevant to the results/discussion presented in this study.

ATL07 is generated by aggregating 150 photons from the ATL03 geolocated photon product (Neumann et al., 2019) independently along each of the six beams. The beams are arranged in 'strong' and 'weak' beam pairs with each beam pair separated by ~3.3 kilometers in the across-track direction and the strong/weak beams separated by ~90 m across-track and ~2.5 km along-track. A surface finding routine (ATL07/10 ATBD r003 section 4.2) first windows the photon heights around an expected sea surface then extracts a best-guess Gaussian height distribution (convolved with the expected system response) to the photon height histogram to determine 1) a single segment height, 2) an associated quality flag based on the goodness of fit and 3) associated metrics including photon rate. The heights are expressed relative to the WGS84 ellipsoid with the mean sea surface (MSS) and various time-variable geophysical corrections removed: ocean tides, solid earth tides, ocean loading, solid earth pole tides, inverted barometer. Surface-classified height segments are produced for each of the six beams (three strong and weak pairs) independently. The photon rates of the strong beam are roughly 4 times higher than those of the weak beam, which results in mean segment lengths of ~15 m for the strong beam and ~60 m for the weak beam (Kwok et al., 2019). Adding the individual laser footprint size of ~14 m to the segment length provides an estimate of the spatial resolution of the segments (i.e. a mean of ~30 m x 15 m for the strong beam and 75 m x 15 m for the weak beam).

An empirically based decision-tree algorithm is used to discriminate the returns between sea ice and open water (Kwok et al., 2016). The empirical thresholds were determined principally by 1) data collected in campaigns prior to the launch of ICESat-2 by the Multiple Altimeter Beam Experimental Lidar (MABEL) - a test-bed instrument for ICESat-2 (McGill et al., 2013); and 2) post-launch evaluation of the ATLAS performance. The three inputs to this decision tree (and the physical justifications) are as follows:

1. Photon rate (photon returns per laser pulse, the apparent reflectivity of the surface).
2. Width of the Gaussian fit to the photon height distribution (the small-scale surface roughness).
3. Background rate (deviations compared to the photon rate indicate shadows, specular reflections and/or atmospheric contamination).

The result of the decision tree determines the radiometric surface type (*height_segment_type* in the product) which includes the following surfaces: clouds, ice (gray, rough, snow covered),

specular lead and dark leads (smooth and rough). This is considered the winter-time decision tree. As discussed in the introduction, in the latest ATL07/10 data release (r003), dark leads were dropped from the derivation of sea surface due to issues with cloud attenuation, so now only specular leads are used for deriving sea surface and freeboard (Kwok et al., 2020b).

Due to the expected uncertainty in this radiometric surface classification approach (Kwok et al., 2016), a further filtering is applied based on the local height distribution:

4. Local height filter, based on the distribution of local (10 km) smooth height segments (h_{sm}) – segments with a Gaussian width < 0.13 m. Specifically the lead segment height must be between the minimum of the smooth heights (h_{sm_min}) and the maximum of either the 2nd percentile of the smooth heights or h_{sm_min} + 2 σ where σ = 2-3 cm, the expected uncertainty in surface height over smooth surfaces.

The result of the radiometric decision tree and local height filter sets the sea surface height flag (*ssh_flag, 1 = sea ice, 0 = open water*). Sea surface segments are considered candidate leads for deriving freeboard in ATL10. Note that the summertime (non-winter) decision tree simply extends the classification of ice, specular leads and dark leads as potential melt ponds (not a feature of this analysis).

Freeboards are calculated in ATL10 in 10 km along-track sections based on a reference sea surface derived from the available lead/sea surface segment heights. Consecutive lead segments are grouped together to reduce noise in the individual sea surface estimates before a single reference sea surface estimate is produced for each 10 km section. A further filtering is applied to ensure that the 10 km sections are at least 25 km away from the coast and have a concentration (from passive microwave data) > 50% - i.e. within the pack ice and away from regions thought to be more affected by waves. The 10 km reference sea surface heights must also lie within a set height window relative to the MSS (+/- 0.5 m for the Arctic and +/- 1 m for the Southern Ocean) and differences between consecutive sea surface heights must be relatively small (see the ATL07/10 ATBD for more details). After the filtering, *candidate leads* are set as leads for use in deriving freeboard (*ssh_flag* = 2) and freeboard segments are derived by differencing the heights of the sea ice surface from the local reference sea surface height. More detail about the methodology is given in Kwok et al., (2019b) and Kwok et al., (2020b).

176 **2.2 Sentinel-2 imagery**

177 Sentinel-2 is a constellation of two twin satellites, Sentinel-2A (S-2A) and Sentinel-2B
178 (S-2B), operated by the European Space Agency (ESA) and launched in June 2015 and March
179 2017, respectively. The satellites host the MultiSpectral Instrument (MSI), which provides 13
180 reflective-wavelength bands in the wavelength region between 443 nm and 2202 nm (visible,
181 near-infrared, short-wave infrared). Depending on the band, spatial resolution varies between 10,
182 20, and 60 m, and scenes are ~ 110 km x 110 km. For a given area, the shift by 180 degrees
183 between the two sun-synchronous polar orbits and a 290 km wide swath guarantee a revisit time
184 of 5 days at the equator, which improves to one image per day at higher latitudes due to
185 overlapping swaths with different viewing angles. Systematic global coverage of land surfaces
186 and coastal waters by S-2 imagery is available between 84 ^{o}N and 56 ^{o}S. However, additional
187 imagery is available near and over coastal regions of Antarctica, up to 73 ^{o}S.

188 We searched for coincident S-2 and ICESat-2 data over sea-ice covered portions of the
189 Arctic and Southern oceans. Our search algorithm matched the footprint of all S-2 images with a
190 nominal cloud coverage of less than 10% acquired during fall, winter, and spring months (fall
191 2018 to spring 2019) within the search regions (Arctic: Sep-May / Antarctic: Mar-Nov) with the
192 ICESat-2 reference ground track (RGT) times and locations (data from https://icesat-
193 2.gsfc.nasa.gov/science/specs). We searched for overlapping data pairs with an acquisition time
194 difference of less than 2 hours to mitigate for ice drift. The selected images were checked for
195 quality, presence/absence of sea ice, actual cloud cover, and for the extent of the overlap with
196 ICESat-2 data, to obtain a catalog of data pairs suitable for our analyses. No images for winter
197 months (e.g. December/January in the Northern Hemisphere) are available due to the lack of sun
198 illumination. The optimal scenes found in this search were a ~50 km section in the Lincoln Sea
199 of the Arctic Ocean in May 2019 (i.e. the end of winter, time difference of 94 minutes) and the
200 Western Weddell Sea of the Southern Ocean in March 2019 (i.e. the start of the austral winter,
201 time difference of only 7 minutes), as shown in Figure 1. This Arctic Ocean scene occurs late in
202 the winter/spring season, after typical dates of melt onset in the Arctic. However, the location is
203 in the coldest/thickest ice of the Arctic and no surface melt is visible in the imagery. Visual
204 inspection of the scenes showed good coincidence and no obvious issues of drift misalignment.

205 ## 3. Methods

206 ### 3.1 Classification evaluation

207 To compare the IS-2 ATL07/10 data with the S-2 imagery we use a simple nearest
208 neighbor interpolation scheme (the nearest geocoded pixel of the imagery to the given ATL07/10
209 beam segment) to produce a coincident profile of S-2 surface reflectance. In particular, we use
210 the red band, which is available at the highest 10 m spatial resolution and convert the Digital
211 Number (DN) from Level-2A surface reflectance data products to spectral radiance (DN/10,000).
212 We compare these profiles qualitatively as the lack of perfect time coincidence with the imagery,
213 together with the possible impact of ice drift and the contrasting resolutions of the data, make it
214 challenging to carry out a more robust quantitative assessment.

215 ### 3.2 Lead fraction

216 To estimate lead fraction, we take the segment length weighted ratio of sea surface
217 segments to the total length of valid segments in 10 km along-track sections across the combined
218 three strong beams (we combine data from the 3 beams before sectioning the data). Note that the
219 IS-2 measurements include three beam pairs spread over ~6.6 km across-track, so for this to be
220 treated as a two-dimensional lead fraction estimate one needs to assume that the underlying ice is
221 isotropic and homogenous over this 10 km x 6.6 km window, something we plan to test more in
222 future work. The along-track lead fraction is calculated as

223 $$L_f = \frac{\sum_{N_{sl}} l_s^i}{\sum_{N_{all}} l_s^i}, \quad (1)$$

224 where l_s is the segment length, N_{all} is the total number of ice and lead segments in the given
225 section (10 km along-track in this study), N_{sl} is the total number of lead segments (specular only
226 in r003) and *i* is the height segment index.

227 The ATL07/10 sea surface classification approach is designed to be strict – i.e. possible
228 leads are potentially thrown away by the height percentile filter (step 4 in Section 2.1) to reduce
229 the likelihood of ice segments being erroneously used to derive a reference sea surface. This
230 approach makes sense when one considers the high sensitivity of the freeboard estimate to errors
231 in the reference sea surface height calculation. However, for ice floe/lead geometry analyses,

such a strict filter could result in an underestimate of lead fraction (and an overestimate of chord length, discussed in the following section). We therefore adopt two approaches for deriving lead fraction: i) $L_f{}^{v1}$ where we use the ATL07 sea surface flag for candidate leads (*ssh_flag = 1*) and ii) $L_f{}^{v2}$ where we use the ATL07 specular lead classification ($2 \leq$ *height_segment_type* ≥ 5) together with a less strict 10 km local height percentile filter (20% instead of 2%, see step 4 in Section 2.1) to determine the sea surface segments. In summary $L_f{}^{v1}$ is derived using the *ssh_flag* from the product, whereas $L_f{}^{v2}$ is derived using a higher 20% local height filter. This is a relatively crude way of exploring the sensitivity of these lead estimates to the underlying classification algorithm which we will aim to expand on in future work.

3.3 Chord length

To calculate chord length, we devised a simple algorithm that splits the along-track segment data into floe (or chord) groups based on the sea surface classification flag. The sea surface segments are discarded. Similar to lead fraction, we use two approaches here; $C_l{}^{v1}$ utilizes the sea surface flag (*ssh_flag*) and $C_l{}^{v2}$ uses the sea surface classification derived in this study using the 20% local height filter. In both cases a 'floe group' needs to include at least 3 height segments and have a maximum spacing between consecutive segments less than 300 m. 300 m is the upper end of the tail of the distribution of segment lengths for the strong beams (Kwok et al., 2019). Data gaps between segments are caused primarily by atmospheric scattering (e.g. by clouds), which could result in erroneously high chord lengths, especially as clouds can form preferentially over leads (although this is still being debated, e.g. Li et al., 2020). We also discard any floe groups which are smaller than 60 m, which we take to be the minimum resolved chord length due to considerations of footprint size and segment length, and groups which are longer than 50 km. We add 15 m, the approximate photon footprint resolution, to all groups to finally derive estimates of chord length. We run this analysis for each of the three strong beams independently.

Note that we do not provide estimates of lead length in this study as estimates were provided in Kwok et al., (2019) and this is thought to depend more on the resolution of the ATL07 data, so investigations focused on the underlying ATL07 algorithm may be needed to produce more reliable lead length estimates in future work.

For both the lead fraction and chord length estimates, we bin the along-track data to a 25 km x 25 km polar stereographic grid (EPSG:3411 for the Arctic and EPSG:3412 for the Southern Ocean) using data collected from the first Arctic (November 1, 2018 to April 30, 2019) and austral (May 1, 2019 to September 30, 2019) winters.

4. Results

4.1 Comparisons with Sentinel-2 imagery

The S-2/ATL07 comparison for the ~50 km profile within the western Weddell Sea on March 17th, 2019 is shown in Figure 2. This cropped S-2 image depicts a scene of mixed ice surfaces - large consolidated floes, small broken up floes, and the occasional lead opening. It also features very flat/thin ice for the first ~20 km and thicker/rougher ice (relative heights ~2/2.5 m higher) for the remaining 30 km along-track. The final ~1 km of the scene features clouds, as can be seen more clearly in the zoomed-out Figure 1. The ATL07 radiometric surface classification scheme detected 218 specular lead segments and 8 dark lead segments for this scene, with 196 candidate (and indeed utilized) sea surface segments. The ATL07/10 lead classification shows strong agreement with the S-2 imagery. As most of the lead segments are specular, the photon rate shows corresponding spikes in these same locations, along with drops in the background rate (as expected from background photons scattering away from the detector over specular surfaces). The scene highlights the ability of ICESat-2 to detect both narrow (10s of meters) and wide (100s of meters) openings in the ice cover. Dark lead classified segments are produced at ~7 km and ~46 km along-track, with both appearing erroneous. The latter coincides clearly with the presence of clouds in the S-2 imagery, further confirmed by the attenuation in photon rate around this part of the scene, similar to the Operation IceBridge comparisons given in Kwok et al., (2020b). The ~7 km along-track dark lead is harder to diagnose, as this is associated with a sharp drop in photon rate, as one would expect from a lead, but no obvious lead or cloud is visible in the imagery. The derived lead fractions and chord lengths in this scene are also shown in Figure

2 (panel 6). The lead fractions are estimated as 1.47% (v_1) and 1.64% (v_2), while the mean chord lengths are estimated as 2.76 km (v_1) and 2.32 km (v_2).

The S-2/ATL07 comparison for the ~50 km profile within the Lincoln Sea, Arctic Ocean on May 25th, 2019 is shown in Figure 3. The S-2 image depicts a scene of large consolidated floes, small broken up floes/leads, but also a much larger (> 1 km) lead opening. No clouds are visible in the imagery (or implied by attenuations in the photon rate). The heights extend from just over 0 m to ~2-2.5 m. The ATL07 radiometric surface classification scheme detected 166 specular leads and 0 dark leads for this scene, with 73 candidate (and again utilized) sea surface segments. The ATL07/10 lead classification again shows strong agreement with the S-2 imagery. Zooming in on the large lead in the S-2 image (not shown) highlights the presence of young ice formation on the right (further along-track) side of the lead, which explains some of the drops in photon rate and the absence of lead classifications. It is encouraging to note that the refrozen lead of low height/freeboard at ~2 km along-track is not classified as a specular or dark lead. The lead fractions for this scene are estimated as 2.12% (v_1) and 6.45% (v_2), while the chord lengths are estimated as 1.83 km (v_1) and 2.13 km (v_2). The difference between the v_1 and v_2 estimates are higher, and we can see this is mainly driven by the inclusion (or absence) of the segments within the large lead opening. Including the extra sea surface segments by using a higher (20%) height threshold in v_2 increases lead fractions, as expected, however also results in the more counter intuitive result of increasing the mean chord length as the ice floe groupings remain too small to be classified as a floe and are simply discarded. At ~10 km along-track, the S-2 scene shows some small but highly consolidated ice floes, with no leads detected, highlighting the challenge of lead/floe detection in the more consolidated ice regimes.

Further examples for these two scenes of both the strong and weak beams are given in the Supplementary Information (Figures S1 to S10). These additional beam comparisons show again a remarkably high level of agreement with the S-2 imagery in terms of the lead classification and further examples of lead fraction and chord length estimates. No obvious degradation of the classification performance is visible in the weak beam comparisons, although we generally see fewer leads and lower lead fractions than with the strong beams.

4.2 Basin scale assessments

To better understand the sea ice classification and to provide context for the basin-scale lead fraction/chord length estimates shown next, we first show basin-scale maps of key lead classification metrics from ATL07 across the Arctic and Southern Ocean for the winter study periods (Arctic: November 1, 2018 to April 30, 2019; Southern Ocean: May 1 2019 to September 30, 2019). Figure 4 shows the radiometric lead fraction compared to all (segment length weighted) segments, the ratio of radiometric specular lead segments to dark lead segments, the fraction of specular lead segments that become sea surface segments (*ssh_flag* > 0) and the fraction of sea surface (lead) segments compared to all segments – i.e. the lead fraction. The radiometric lead and sea surface lead segments generally follow the spatial pattern expected from our past knowledge of the sea ice state, increases in leads (declines in concentration) towards the ice edge. In both hemispheres dark leads make up a significant (~50%) fraction of the total number of radiometric leads. In both hemispheres a significant (>70%) fraction of the specular leads pass the height filter and are assigned as sea surface segments. There is a clear increase in the fraction of discarded specular leads along the ice edge in both hemispheres, but also within the Canadian Arctic Archipelago. As ATL10 applies a stricter 50% concentration filter from passive microwave data to avoid wave contamination (ATL07 uses 15%), and a 25 km coastal mask, many of these regions adjacent to the ice edge (in both hemispheres) and near coastlines will not be processed into freeboard. Sea ice concentrations from the monthly (final and near-real time) NSIDC Climate Data Record (CDR, Meier et al., 2017; Meier et al., 2017) averaged over this same time period across both hemispheres are given in Figure 5 for context.

Figure 6 shows the 25 km x 25 km gridded lead fractions calculated using the *ssh_flag* (${L_f}^{v1}$, as in Figure 4d), and also using the 20% height filter (${L_f}^{v2}$). ${L_f}^{v2}$ consistently results in higher lead fractions as expected, especially around the ice edge. We also note the increased lead fractions in areas of known polynya formation - the North Water Polynya to the northwest of Greenland and, to a lesser degree, Terra Nova Bay Polynya in the northern Ross Sea. Neither are as visible in the passive microwave-derived concentrations (Figure 5). Mean Arctic Ocean lead fraction are given as 1.06 % (${L_f}^{v1}$) and 1.80 % (${L_f}^{v2}$). Mean Southern Ocean lead fractions are given as 0.66% (${L_f}^{v1}$) and 1.24% (${L_f}^{v2}$). The Southern Ocean lead fractions, while lower on average, show increased lead fractions within the pack ice (compare with sea ice concentration in

Figure 5), especially in the eastern Southern Ocean. This is expected considering the open ocean boundaries and the general divergent (away from the coast) Antarctic ice drifts. The results also highlight the sensitivity of these estimates to this one aspect of the ATL07 lead finding algorithm. Lead fractions within the pack ice of the Central Arctic are very low (< 1%) which we discus more in the discussion/summary section.

Figure 7 shows the 25 km gridded chord lengths calculated using the *ssh_flag* (C_l^{v1}) and the 20% height filter (C_l^{v2}). Mean Arctic Ocean chord lengths are given as 5.56 km (C_l^{v1}) and 5.23 km (C_l^{v1}). Mean Southern Ocean chord lengths are given as 5.43 km (C_l^{v1}) and 4.55 km (C_l^{v2}). As in the lead fractions, there is high spatial variability not captured by these simple hemispheric means. Lower chord lengths are produced within the eastern Arctic Ocean (compared to the western Arctic Ocean) and also in the eastern Southern Ocean (compared to the western Southern Ocean). The spatial variability is generally higher than the lead fraction as only small changes in lead fraction can have a large impact on the derived chord lengths. The Arctic shows a sharp increasing gradient within the more central Arctic region. Chord lengths are high (consistently greater than 10 km) in the central Arctic, which we discuss more in the following section.

Figure 8 shows the probability distribution of the individual chord lengths for the Arctic and Southern Ocean using the C_l^{v1} and C_l^{v2} chord length estimates. The data are plotted on a log-log scale to highlight the power law nature of the underlying distributions. The consistency of the distributions across hemispheres and lead finding algorithm (*v1* and *v2*) is encouraging. The distributions become more variable at the tail of the distribution (> 10 km) due to the lower sampling rate of these higher chord lengths. While log-log distributions can help visually highlight power-law like behavior in empirical data, more robust statistical tests are needed to truly test for whether data is well characterized by a power-law, especially due to the issues with the less well-observed tails of the distribution (Clauset et al., 2009; Stern et al., 2018, Horvat et al., 2019). We plan to explore this in future work as we compare against the chord length estimates from other observed data, as discussed below.

5. Discussion and summary

The coincident Sentinel-2 (S-2)/ICESat-2 scenes provide crucial validation of the ICESat-2 sea ice classification procedure, with the caveat that these represent only a small

fraction of the available sea ice data produced from ICESat-2 to-date. The specular lead classification in the sea ice products shows strong agreement with the imagery across all beams, while the small quantity of dark leads found in these scenes generally appear erroneous, providing further evidence to that presented in Kwok et al., (2020b) of their more uncertain reliability. More coincident S-2 scenes across different sea ice regimes (time and space) would help provide further insight into the potential utility of the dark leads, especially as they make up a significant fraction (~50-60%) of the total number of radiometric leads detected by the current ATL07 algorithm. As discussed in Kwok et al., (2020b), the exclusion of dark leads from the freeboard calculation has the downside of reducing coverage, especially in some of the more consolidated sea ice regimes, providing motivation for continued study to ensure that this exclusion is not overly restrictive.

It is interesting to note that the ATL07 segments over the relatively wide (> 1 km) lead in our Arctic Ocean S-2 scene (Figure 3) were consistently classified as specular. Intuitively we might expect this to be considered a 'dark lead' as the larger opening increases the fetch, but in-fact the returns are consistently specular, until the photon rate drops drastically (to less than that over the snow covered ice) as the ice starts to refreeze, with these segments classified as ice. Similarly, the very thin cracks shown in the imagery (on the order of meters to tens of meters, i.e. the segment resolution of the ATL07 data) were consistently classified as ice. It is unclear from the S-2 imagery (10 m resolution) the extent to which these leads/cracks have refrozen. Higher resolution imagery (e.g. from Digital Globe's Worldview imagery) could provide useful further insight here and will be explored in future work.

The local height filter included in the lead (sea surface) finding algorithm further limits the number of leads identified. This appears reasonable when the goal is accurately determining a reference sea level for freeboard but less ideal for deriving sea ice state information, e.g. lead fraction and chord length. This study presented a crude approach to relax this height filter (increasing the percentile threshold from 2% to 20%), which is worthy of further investigation as we seek to increase the utility of the ICESat-2 sea ice data.

Validation, or at least comparisons, of our basin-scale ICESat-2 estimates of lead fraction and chord length with existing observational estimates are still needed. Basin-scale lead fraction estimates have been produced from various satellite sensors, e.g. NASA's Moderate Resolution

Imaging Spectroradiometer (MODIS) (Hoffman et al., 2019; Willmes & Heinemann, 2016) and Advanced Microwave Scanning Radiometer (AMSR-E) (Röhrs & Kaleschke, 2012), and also ESA'S CryoSat-2 and Envisat radar altimeters (Röhrs & Kaleschke, 2012; Tilling et al., 2019). Chord length distributions have also been produced from a compilation of more sporadic satellite/airborne imagery estimates (Stern et al., 2018). However, comparisons with these data are hindered by the considerable differences in spatial/temporal sampling, and also the different interpretations and definitions of leads across sensors and algorithms. As discussed earlier, the lead classification algorithm in ATL07 appears to require clear sections of open water to trigger a specular lead classification. Therefore, newly refrozen leads, which may be defined as leads in other (generally coarser resolution) products, may be simply classified as ice. The radar altimeter-derived estimates of lead fraction and chord length from CryoSat-2 (Tilling et al., 2019, Horvat et al., 2019) provide arguably the most similar dataset for comparison/assessment. The CryoSat-2 lead fractions and chord lengths shown in Tilling et al., (2019) appear consistently higher and lower, respectively, than the ICESat-2 results presented here. For example, Tilling et al., (2019, Figure 1) shows the percentage of CryoSat-2 waveforms classified as lead (a quantity related to, but not the same as lead fraction) as ~10 % in the Central Arctic but up to ~60 % in the more peripheral seas, while the chord lengths (Tilling et al., 2019, Figure 2) are ~3 km in the Central Arctic and ~1 km in the peripheral seas. In contrast, the chord lengths presented in this study are consistently > 10 km within the Central Arctic. There is better agreement on the general spatial pattern of increasing lead fractions/decreasing chord lengths in the peripheral seas, however. Radar altimeters such as CryoSat-2 are highly sensitive to the presence of specular leads within the radar swath and have the added benefit of being largely unaffected by clouds. However, the CryoSat-2 footprint is much larger (~400 m along track and up to ~15 km across track) than ICESat-2 (mean of ~30 m x 15 m for a given ATL07 strong beam segment). Future work will aim to better assess, and hopefully reconcile, the lead fraction and chord length estimates from these two missions, taking these significant sampling differences into account. We also hope to explore further statistical testing of the power-law hypothesis for the floe length distribution, which was recently challenged in Horvat et al., (2019) using CryoSat-2 chord length estimates. This reconciliation is further motivated by the need to provide reliable observational constraints on the floe size distributions being incorporated into sea ice components of Global

Climate Models (Horvat, et al., 2019). Airborne data from NASA's Operation IceBridge campaigns, which combine high-resolution imagery and laser altimetry, should prove invaluable.

Our hope is that ICESat-2 sea ice data can be used to provide routine, and reliable, basin-scale measurements of lead fraction, ice concentration and chord length estimates, in addition to its primary mission requirement of delivering accurate estimates of sea ice height and freeboard. Additional ICESat-2 sea ice algorithm testing and development is needed to further improve the classification accuracy, which can be guided by additional comparisons with imagery (satellite and airborne) and other airborne and satellite datasets. The ICESat-2 lead classification algorithm utilizes fixed empirical thresholds (discussed in Section 2.1) which can easily be tuned/calibrated as needed, as can other elements of the algorithm including the height percentile threshold (which we simply explored in this study) and the photon aggregations. These thresholds also need to be better explored in terms of appropriate strong/weak beam differences. Efforts are also on-going to provide further insight into the dark leads and their potential reintroduction by utilizing a new cloud filter in the sea ice algorithm (Kwok et al., 2020b). However, this assumes that the dark lead classifications unaffected by clouds are more reliable, which is still unclear. The results and discussion presented here also raise the issue of how best to classify newly refrozen leads or gray ice in the sea ice products, or even where exactly to draw the line between sea ice and open water – another key consideration as we seek to increase the utility of the ICESat-2 sea ice products.

Acknowledgements

We would like to thank the entire ICESat-2 project for their continued efforts in delivering and maintaining the high-quality sea ice data analyzed in this study.

Data availability

Sentinel-2 imagery was derived from Copernicus data obtained at https://schihub.copernicus.eu. The ICESat-2 ATL07 (sea ice height and type) and ATL10 (freeboard) data products used here are available through the NSIDC (https://nsidc.org/data/atl07 & https://nsidc.org/data/atl10). Our lead fraction and chord length data will be made available through the NASA Goddard Space Flight Center's Cryospheric Sciences Lab website on completion of peer review.

References

Bateson, A. W., Feltham, D. L., Schröder, D., Hosekova, L., Ridley, J. K., & Aksenov, Y. (2020). Impact of sea ice floe size distribution on seasonal fragmentation and melt of Arctic sea ice. *The Cryosphere*, *14*(2), 403–428. https://doi.org/10.5194/tc-14-403-2020

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703. https://doi.org/10.1137/070710111

Feltham, D. L. (2005). Granular flow in the marginal ice zone. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *363*(1832), 1677–1700. https://doi.org/10.1098/rsta.2005.1601

Hoffman, J. P., Ackerman, S. A., Liu, Y., & Key, J. R. (2019). The Detection and Characterization of Arctic Sea Ice Leads with Satellite Imagers. *Remote Sensing*, *11*(5), 521. https://doi.org/10.3390/rs11050521

Horvat, C, Roach, L. A., Tilling, R., Bitz, C. M., Fox-Kemper, B., Guider, C., et al. (2019). Estimating the sea ice floe size distribution using satellite altimetry: theory, climatology, and model comparison. *The Cryosphere*, *13*(11), 2869–2885. https://doi.org/10.5194/tc-13-2869-2019

Horvat, C., Blanchard-Wrigglesworth, E., & Petty, A. (2020). Observing Waves in Sea Ice With ICESat-2. *Geophysical Research Letters*, *47*(10), e2020GL087629. https://doi.org/10.1029/2020GL087629

Horvat, C., Tziperman, E., & Campin, J.-M. (2016). Interaction of sea ice floe size, ocean eddies, and sea ice melting. *Geophysical Research Letters*, *43*(15), 8083–8090. https://doi.org/10.1002/2016GL069742

Kwok, R., Cunningham, G. F., Hoffmann, J., & Markus, T. (2016). Testing the ice-water discrimination and freeboard retrieval algorithms for the ICESat-2 mission. *Remote Sensing of Environment*, *183*, 13–25. https://doi.org/10.1016/j.rse.2016.05.011

Kwok, R., Cunningham, G., Hancock, D., Ivanoff, A., & Wimert, J. (2019). Ice, Cloud, and Land Elevation Satellite-2 Project: Algorithm Theoretical Basis Document (ATBD) for Sea Ice Products,. https://icesat-2.gsfc.nasa.gov/science/data_products.

Kwok, R., Kacimi, S., Markus, T., Kurtz, N. T., Studinger, M., Sonntag, J. G., et al. (2019). ICESat-2 surface height and sea-ice freeboard assessed with ATM lidar acquisitions from Operation IceBridge. *Geophysical Research Letters*, *44*, 11228– 11236. https://doi.org/10.1029/2019GL084976

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*(10), 6942–6959. https://doi.org/10.1029/2019JC015486

Kwok, R., Kacimi, S., Webster, M. A., Kurtz, N. T., & Petty, A. A. (2020). Arctic Snow Depth and Sea Ice Thickness From ICESat-2 and CryoSat-2 Freeboards: A First Examination. *Journal of Geophysical Research: Oceans*, *125*(3), e2019JC016008. https://doi.org/10.1029/2019JC016008

Kwok, R., Petty, A. A., Bagnardi, M., Kurtz, N. T., Cunningham, G. F., & Ivanoff, A. (2020). Refining the sea surface identification approach for determining freeboards in the ICESat-2 sea ice products. *The Cryosphere Discussions*, 1–18. https://doi.org/10.5194/tc-2020-174

Li, X., Krueger, S. K., Strong, C., Mace, G. G., & Benson, S. (2020). Midwinter Arctic leads form and dissipate low clouds. *Nature Communications*, *11*(1), 206. https://doi.org/10.1038/s41467-019-14074-5

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., et al. (2017). The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. *Remote Sensing of Environment, 190*, 260–273. https://doi.org/10.1016/j.rse.2016.12.029

McGill, M., Markus, T., Scott, V. S., & Neumann, T. (2013). The Multiple Altimeter Beam Experimental Lidar (MABEL): An Airborne Simulator for the ICESat-2 Mission. *Journal of Atmospheric and Oceanic Technology*, *30*(2), 345–352. https://doi.org/10.1175/JTECH-D-12-00076.1

Meier, W. N., Fetterer, F., & Windnagel, A. K. (2017). Near-Real-Time NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 1. https://doi.org/10.7265/N5FF3QJ6

Meier, W. N., Fetterer, F., Savoie, M., Mallory, S., Duerr, R., & Stroeve, J. (2017). NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 3. https://doi.org/10.7265/N59P2ZTG

Neumann, T. A., Martino, A. J., Markus, T., Bae, S., Bock, M. R., Brenner, A. C., et al. (2019). The Ice, Cloud, and Land Elevation Satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System. *Remote Sensing of Environment, 233*, 111325. https://doi.org/10.1016/j.rse.2019.111325

528 Petty, A. A., Kurtz, N. T., Kwok, R., Markus, T., & Neumann, T. A. (2020). Winter Arctic Sea
529 Ice Thickness From ICESat-2 Freeboards. *Journal of Geophysical Research: Oceans*,
530 *125*(5), e2019JC015764. https://doi.org/10.1029/2019JC015764

531 Roach, L. A., Horvat, C., Dean, S. M., & Bitz, C. M. (2018). An Emergent Sea Ice Floe Size
532 Distribution in a Global Coupled Ocean-Sea Ice Model. *Journal of Geophysical*
533 *Research: Oceans*, *123*(6), 4322–4337. https://doi.org/10.1029/2017JC013692

534 Röhrs, J., & Kaleschke, L. (2012). An algorithm to detect sea ice leads by using AMSR-E
535 passive microwave imagery. *The Cryosphere*, *6*(2), 343–352. https://doi.org/10.5194/tc-
536 6-343-2012

537 Rothrock, D. A., & Thorndike, A. S. (1984). Measuring the sea ice floe size distribution. *Journal*
538 *of Geophysical Research: Oceans*, *89*(C4), 6477–6486.
539 https://doi.org/10.1029/JC089iC04p06477

540 Stern, H. L., Schweiger, A. J., Zhang, J., & Steele, M. (2018). On reconciling disparate studies of
541 the sea-ice floe size distribution. *Elem Sci Anth*, *6*(1), 49.
542 https://doi.org/10.1525/elementa.304

543 Studinger, M. (2014). IceBridge Narrow Swath ATM L1B Elevation and Return Strength,
544 Version 2. *Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center*.
545 https://doi.org/10.5067/CXEQS8KVIXEI.

546 Tilling, R., Ridout, A., & Shepherd, A. (2019). Assessing the Impact of Lead and Floe Sampling
547 on Arctic Sea Ice Thickness Estimates from Envisat and CryoSat-2. *Journal of*
548 *Geophysical Research: Oceans*, *124*(11), 7473–7485.
549 https://doi.org/10.1029/2019JC015232

Willmes, S., & Heinemann, G. (2016). Sea-Ice Wintertime Lead Frequencies and Regional Characteristics in the Arctic, 2003–2015. *Remote Sensing*, *8*(1), 4. https://doi.org/10.3390/rs8010004

Zwally, H. J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., et al. (2002). ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *Journal of Geodynamics*, *34*(3–4), 405–445. https://doi.org/10.1016/S0264-3707(02)00042-X

Figures

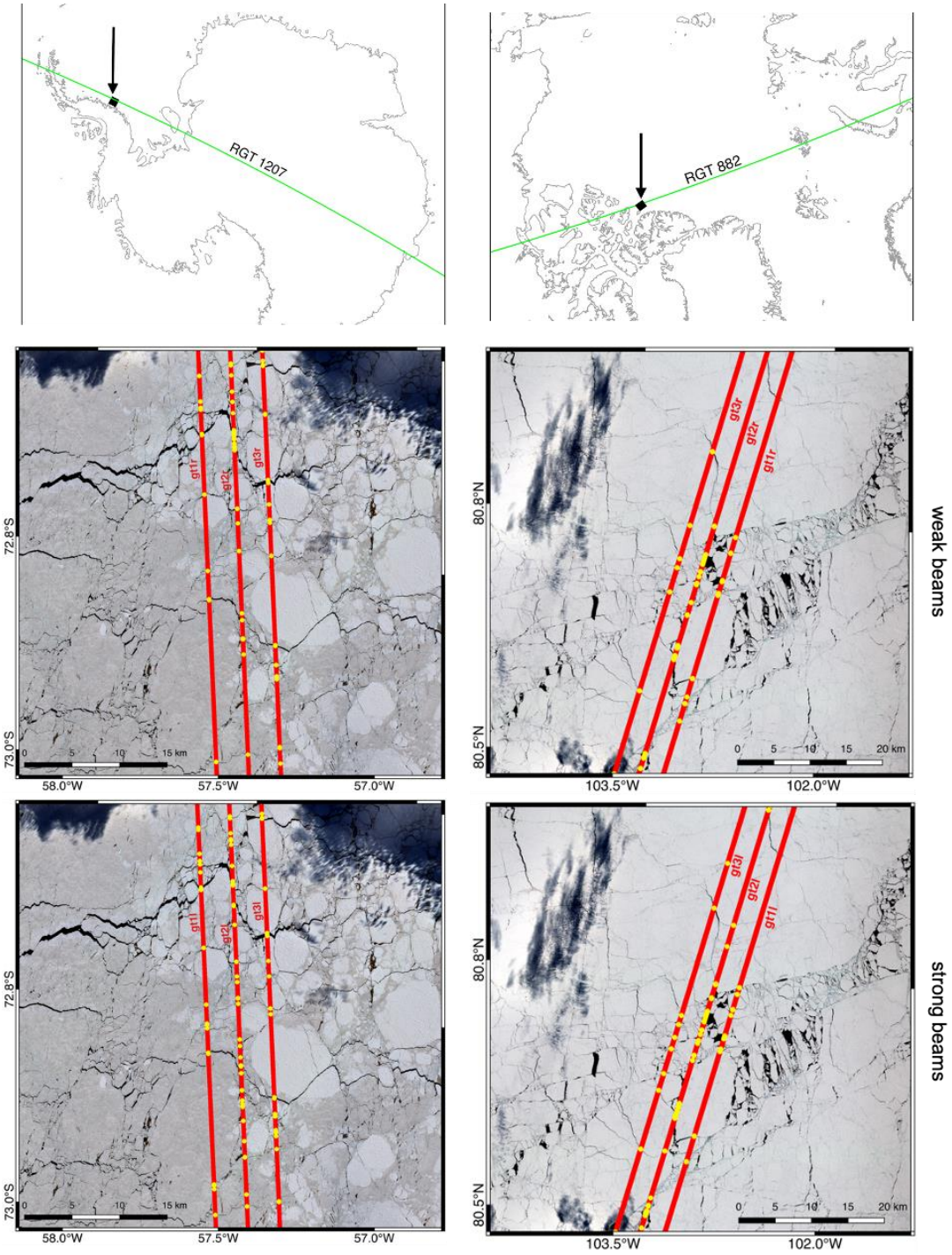


Figure 1: (bottom left) ICESat-2 ATL07 data overlaid on Sentinel-2 RGB image for near-coincident overpass in the (left column) Western Weddell Sea, Southern Ocean, on March 17th, 2019 and (right column) Lincoln Sea of the Arctic Ocean on May 25th, 2019. The time difference between the ICESat-2 and Sentinel-2 overpass for the Weddell Sea scene is ~7

minutes (ICESat-2 acquisition at 12:29 UTC, Sentinel-2 acquisition at 12:22 UTC), time difference for the Lincoln Sea scene is ~94 minutes (ICESat-2 acquisition at 00:36 UTC on 26 May 2019, Sentinel-2 acquisition at 23:01 UTC on 25 May 2019). Red segments are classified as sea ice (*ssh_flag* = 0), yellow segments are classified as leads (*ssh_flag* ≥ 1). The ICESat-2 footprint is not to scale.

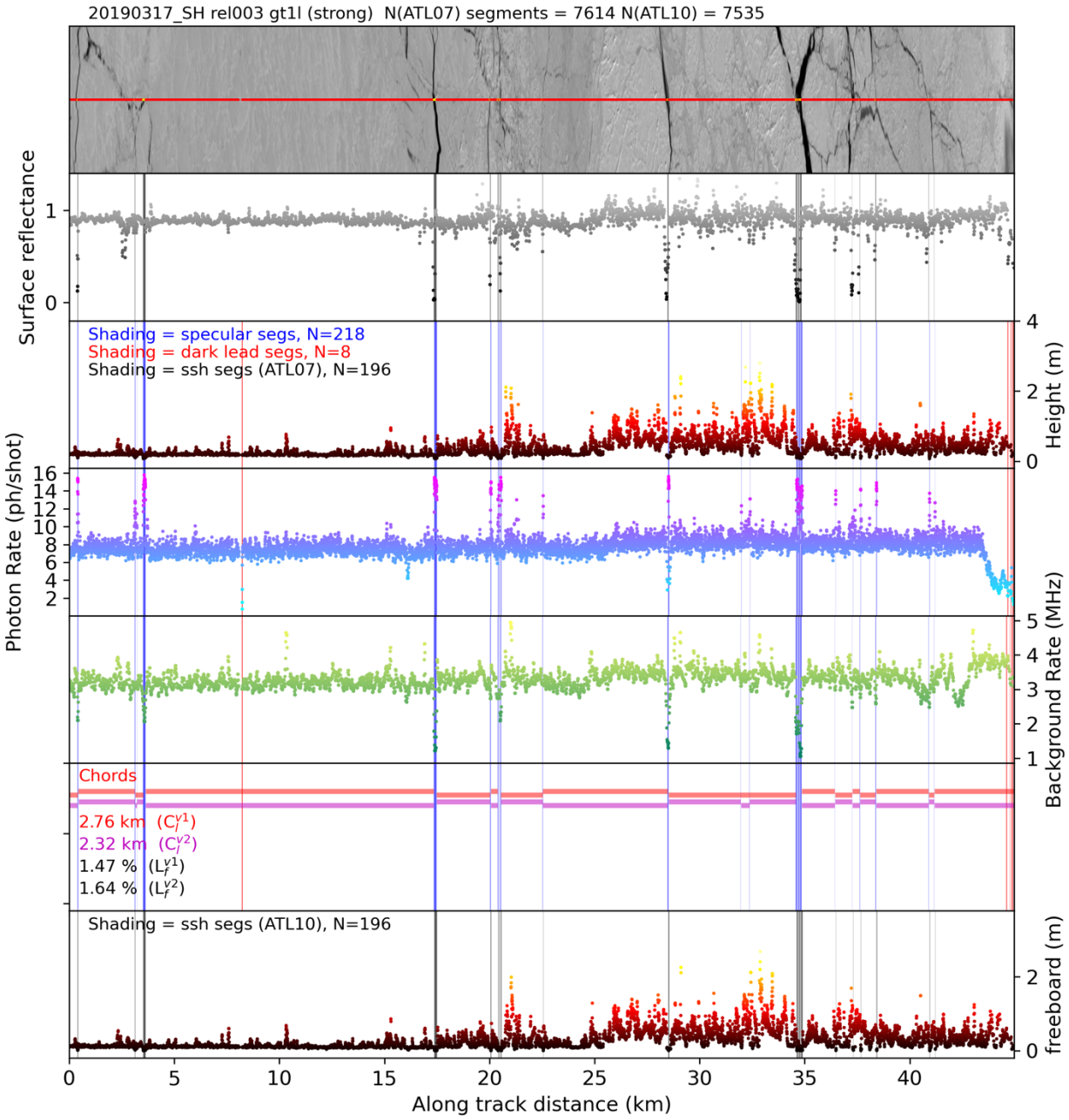
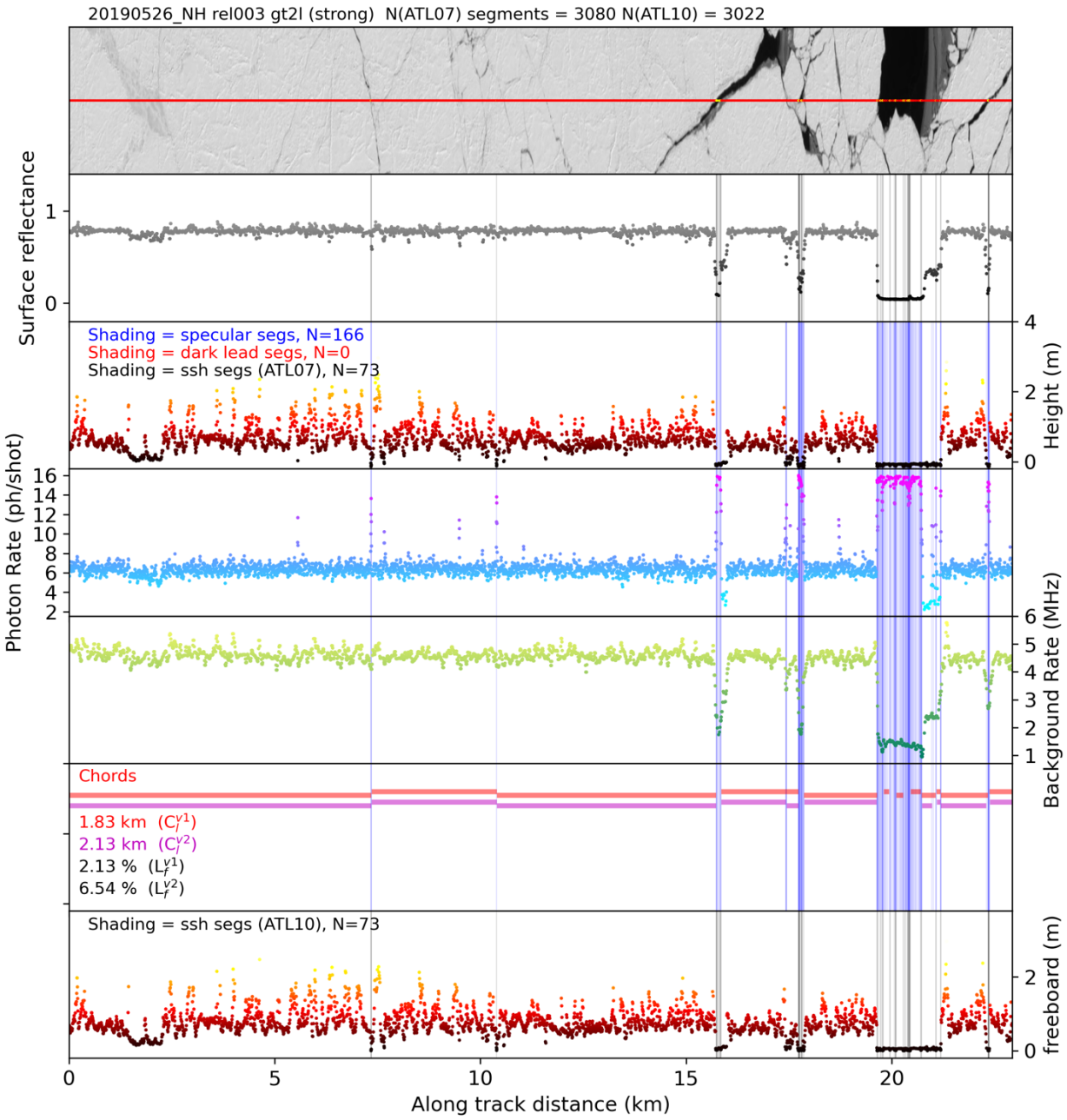


Figure 2: ICESat-2 ATL07 data from the strong beam *gt1l* overlaid on Sentinel-2 RGB imagery for a ~ 50 km profile in the Western Weddell Sea on March 17th, 2019 (see Figure 1 for location). (top panel) Sentinel image with ICESat-2 profile overlaid showing the sea surface flag (*ssh_flag*, red = ice, yellow = sea surface); (second panel) surface reflectance calculated from the red band of the Sentinel-2 image from nearest neighbor pixels to the ICESat-2 profile; (panels 3-5) segment height, photon rate and background rate respectively from ATL07. Grey shading in panels 2-3 indicate candidate leads (*ssh_flag* = 1) while in panels 4-5, blue shading indicates specular lead classifications (*height_segment_type* = 2 to 5) and red shading indicates dark lead classifications (*height_segment_type* = 6 to 8). Panel 7 shows the derived chord lengths as horizontal bars (red = v_1, magenta = v_2) with each chord grouping shifted vertically to indicate the groups, and statistics of the mean chord lengths and lead fractions using these two approaches. Panel 8 shows the freeboard in ATL10, with the shading indicating actual leads used to derive reference sea surface (*ssh_flag* = 2). N indicates the number of segments in ATL07 and ATL10 as specified.



585

586 **Figure 3**: As in Figure 2 but for a profile in the Lincoln Sea of the Arctic Ocean on May 26th,
587 2019 for the strong beam *gt2l* (see map in Figure 1).

588

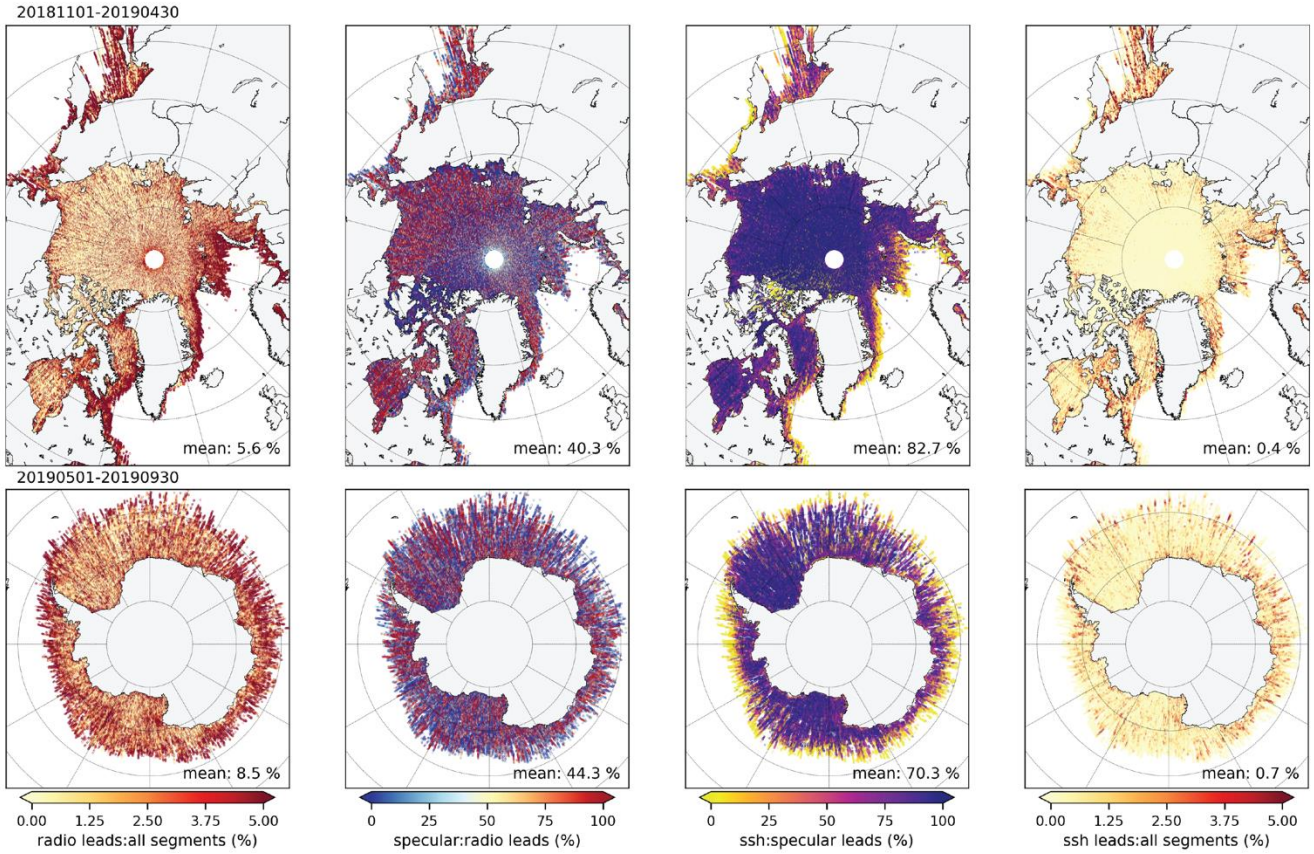


Figure 4: (first column) percent of all ATL07 segments classified as radiometric leads (*height_segment_type* = 2 to 9), (second column) percent of radiometric leads classified as specular (*height_segment_type* = 2 to 5), (third column) percent of specular leads that become sea surface segments (*ssh_flag* ≥ 1), (fourth column) percent of all height segments that become sea surface segments. (Top row) Arctic Ocean, November 1 to April 30, 2019, (bottom row) Southern Ocean, May 1 to September 30, 2019. All these percentages are calculated in 10 km along-track sections across the three strong beams and are segment length weighted.

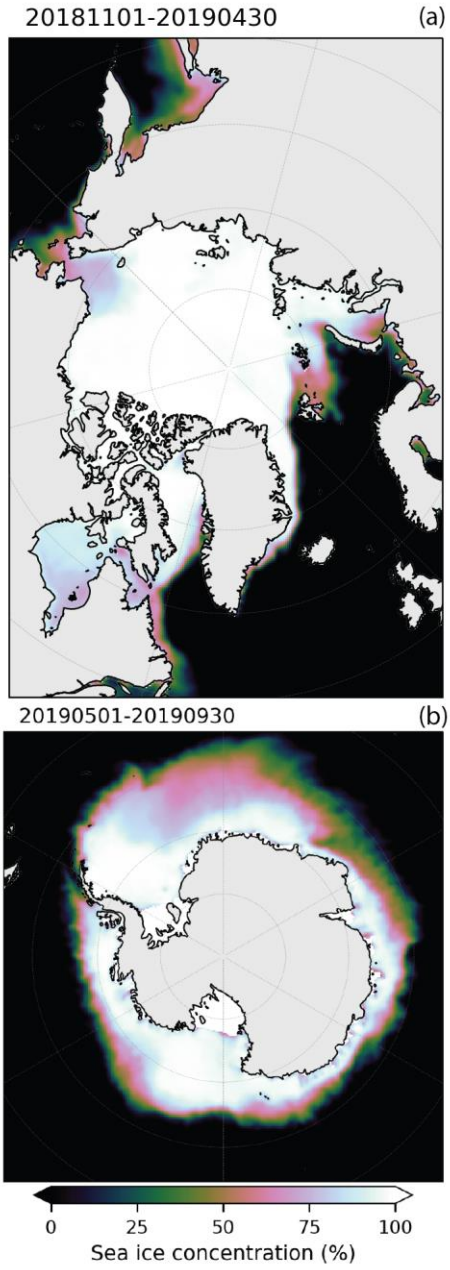


Figure 5: Mean winter sea ice concentration derived from passive microwave data (Climate Data Record, CDR) for the (top) Arctic Ocean, November 1, 2018 to April 30, 2019 and (bottom) Southern Ocean, May 1 to September 30, 2019.

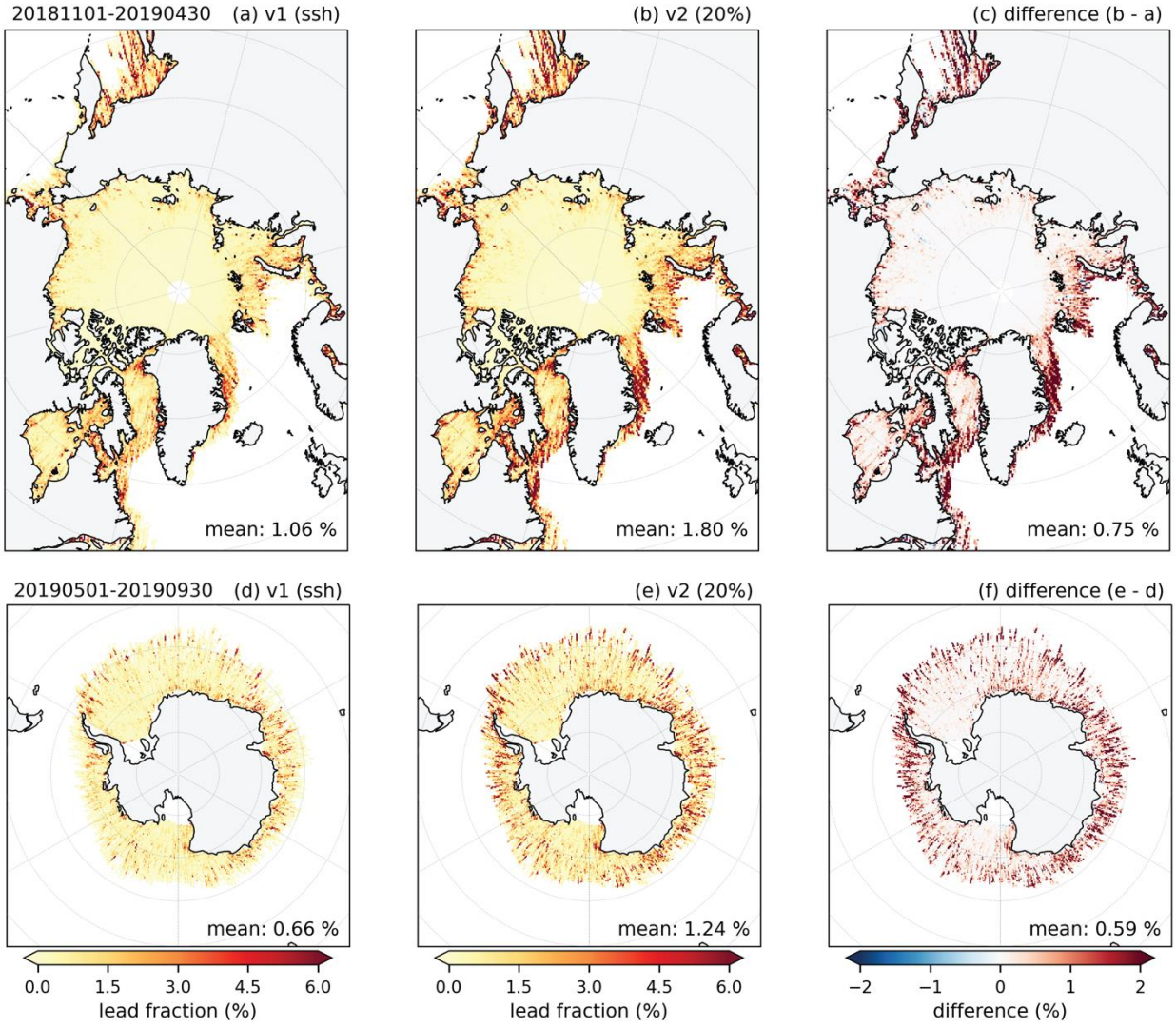


Figure 6: Lead fraction using (left) the *ssh_flag* ($L_f{}^{v1}$) and (middle) using the higher 20% height filter ($L_f{}^{v2}$) and the difference (b minus a) on the right. Data are for the (top) Arctic Ocean, November 1 to April 30, 2019, (bottom) Southern Ocean, May 1 to September 30, 2019. All the fractions are calculated in 10 km along-track sections and are segment length weighted. Data have binned to a 25 km x 25 km polar stereographic grid.

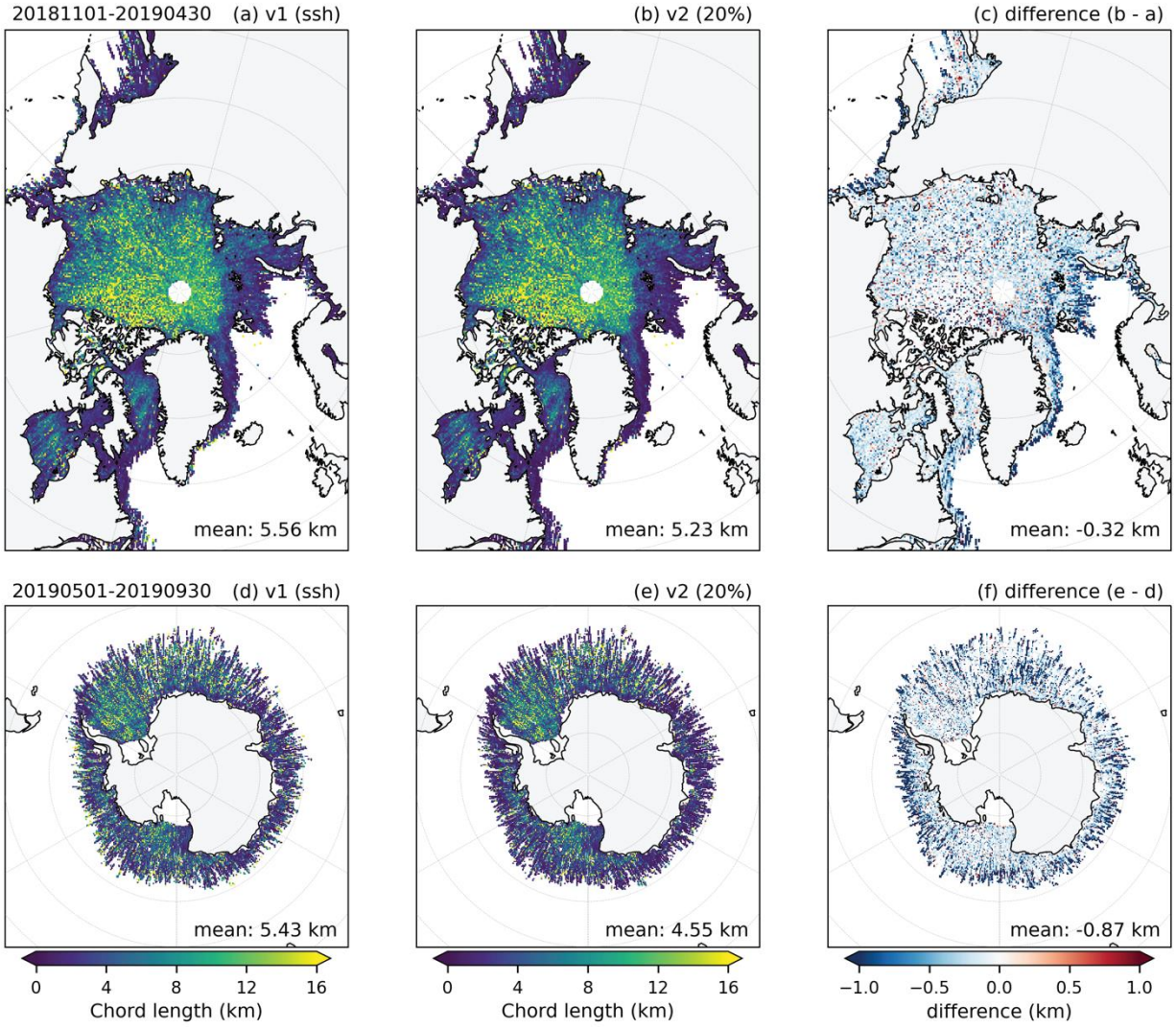


Figure 7: As in Figure 5 but for chord length (left) C_l^{v1} and (middle) C_l^{v2} and difference (right).

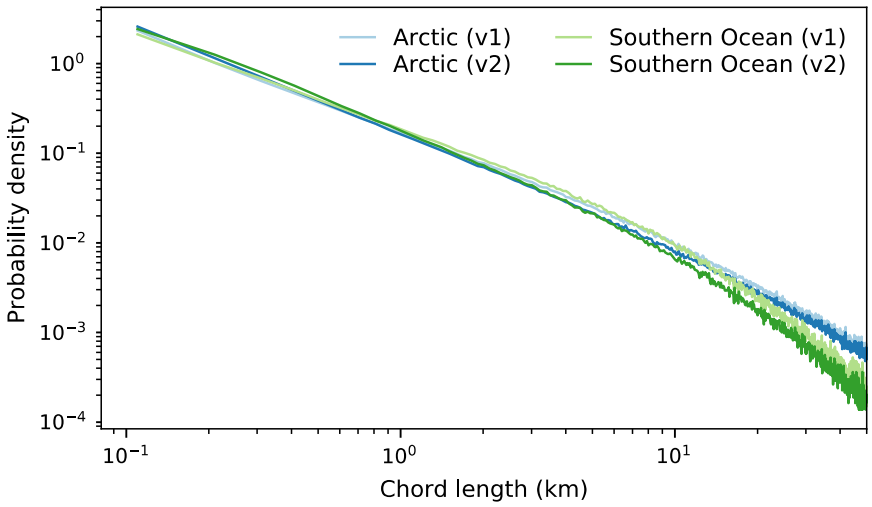


Figure 8: Probability distributions of the Arctic and Southern Ocean chord lengths (v1: C_l^{v1} and v2: C_l^{v2}) for the same winter-time periods as in Figure 7. Distributions are plotted on a log-log scale.

Supplementary information for "Assessment of ICESat-2 sea ice surface classification with Sentinel-2 imagery: implications for freeboard and new estimates of lead and floe geometry" by Petty et al.

Corresponding author: Alek A. Petty (alek.a.petty@nasa.gov)

Included in this supplementary document are further Sentinel-2/ICESat-2 overlays across the two scenes for the other 5 ICESat-2 beams not shown in the main manuscript.

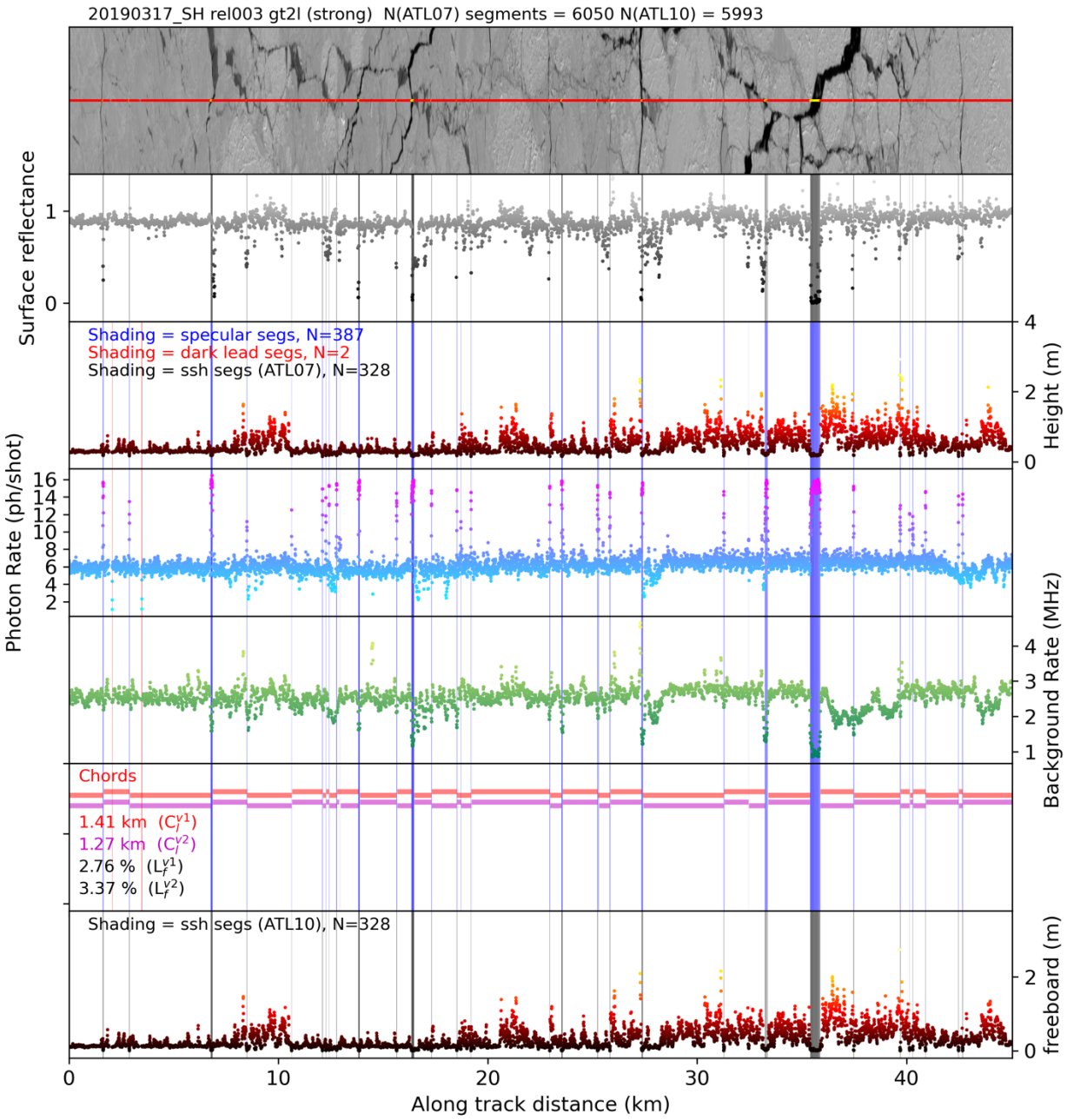


Figure S1: As in Figure 2 but the strong beam *gt2l*

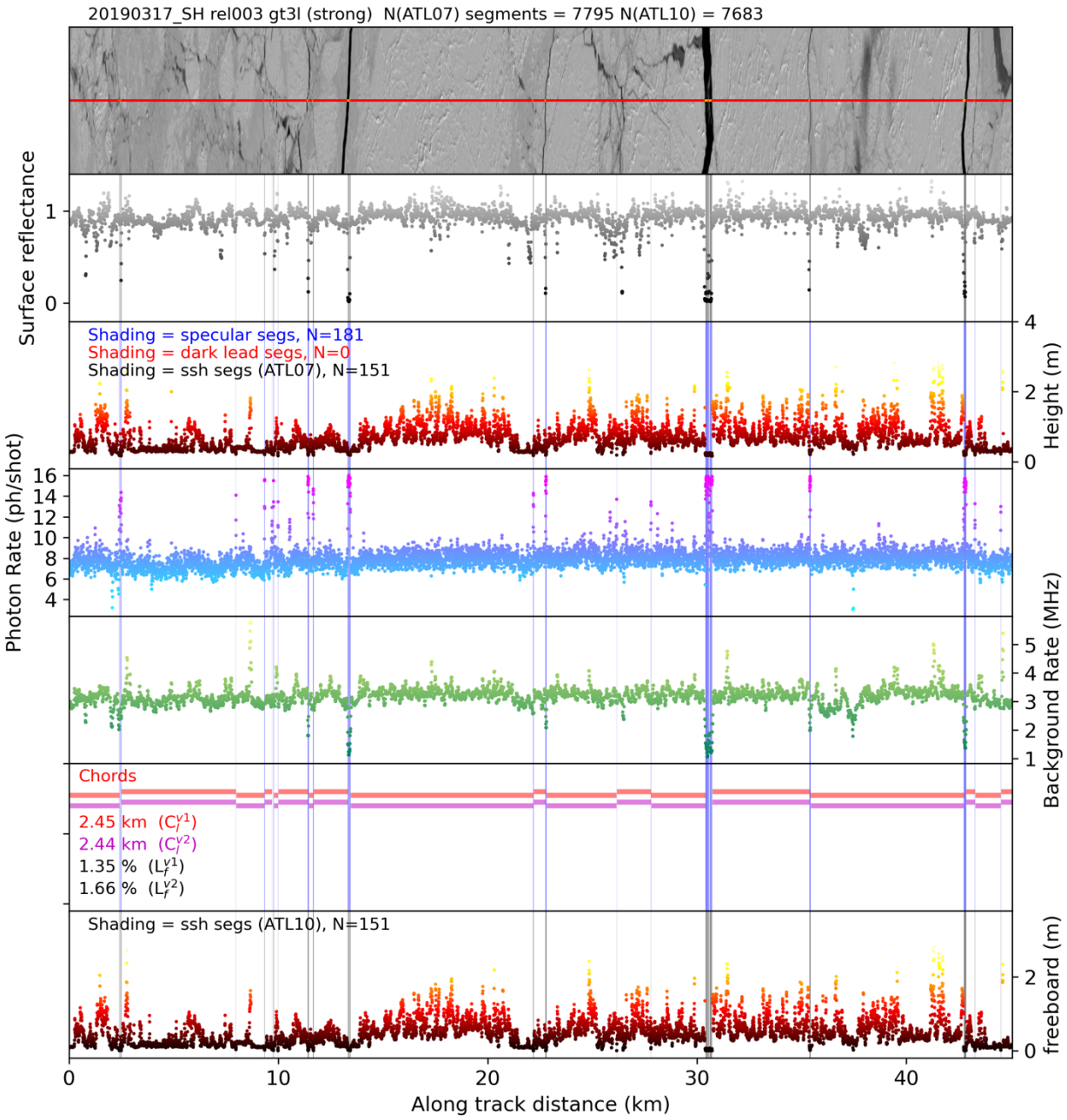


Figure S2: As in Figure 2 but the strong beam *gt3l*

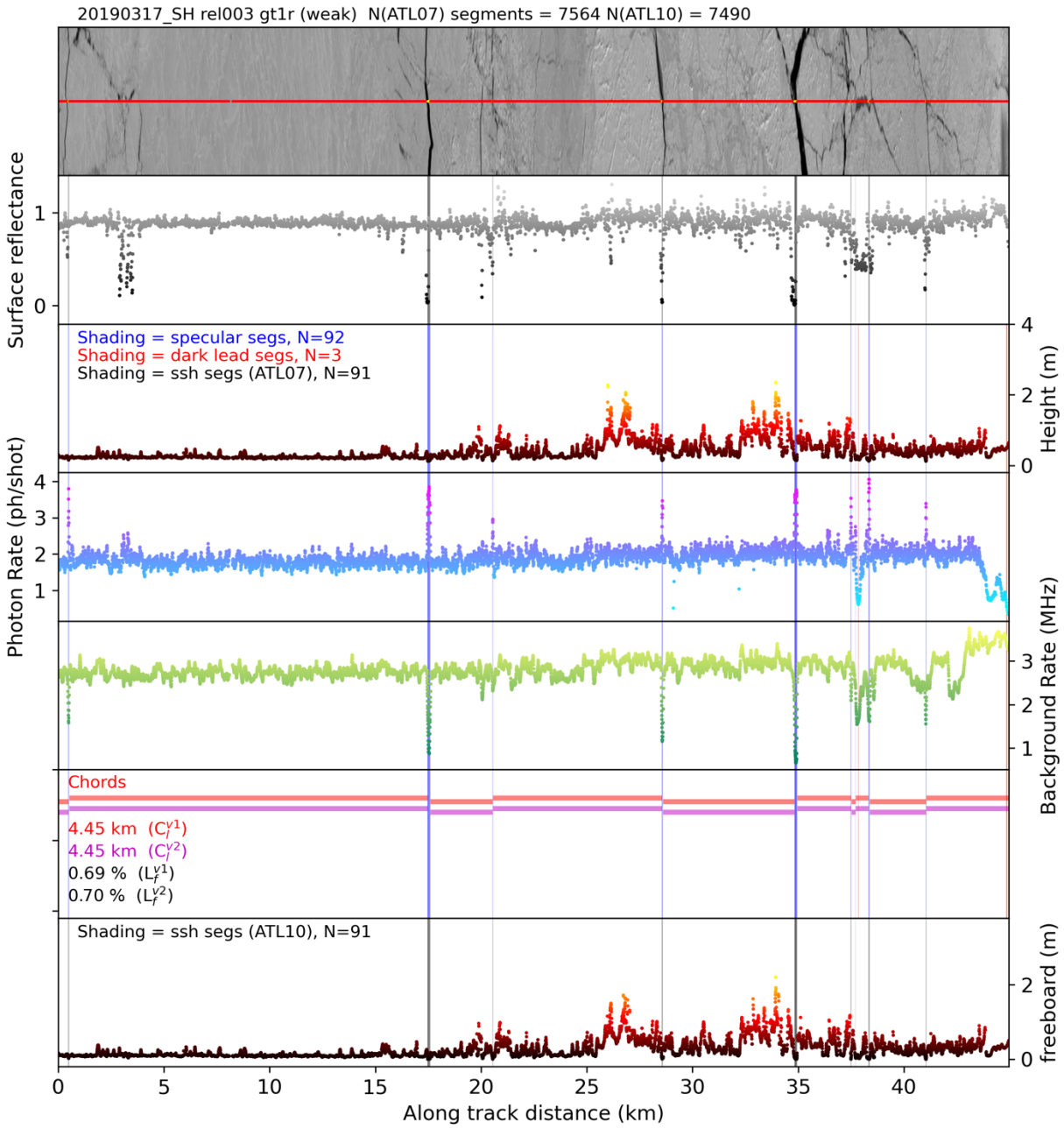


Figure S3: As in Figure 2 but the weak beam *gt1r*

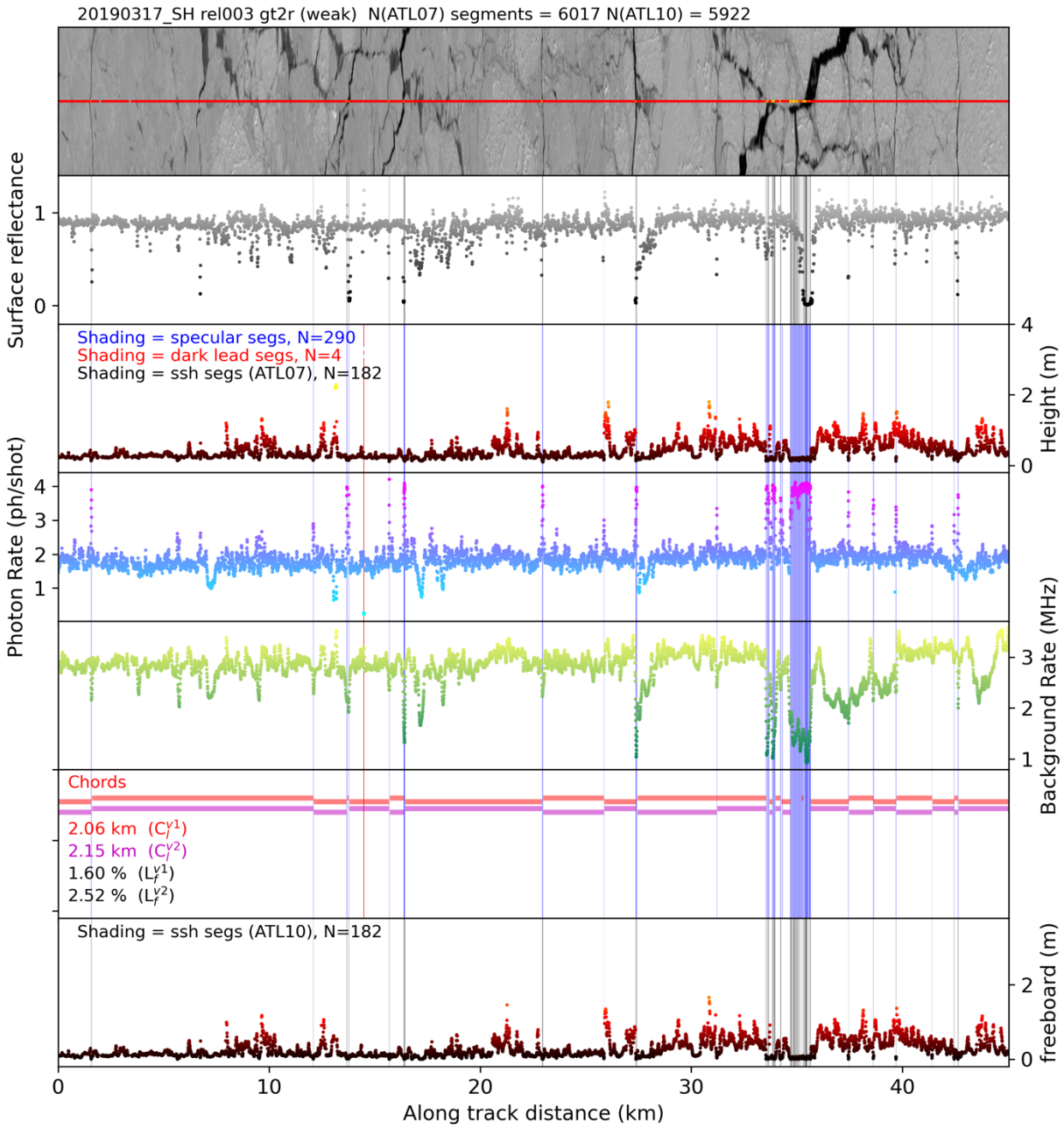


Figure S4: As in Figure 2 but the weak beam *gt2r*

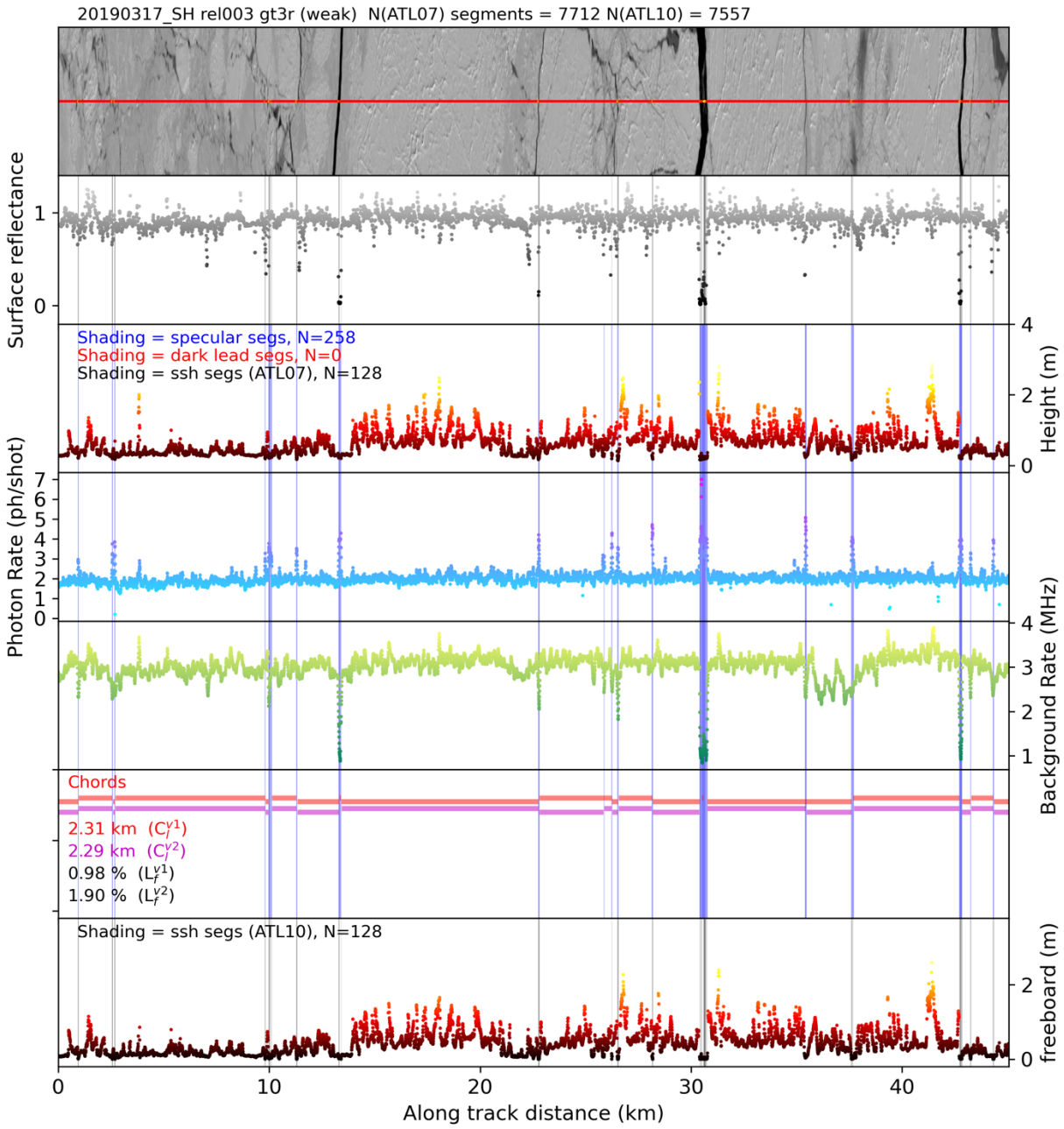


Figure S5: As in Figure 2 but the weak beam *gt3r*

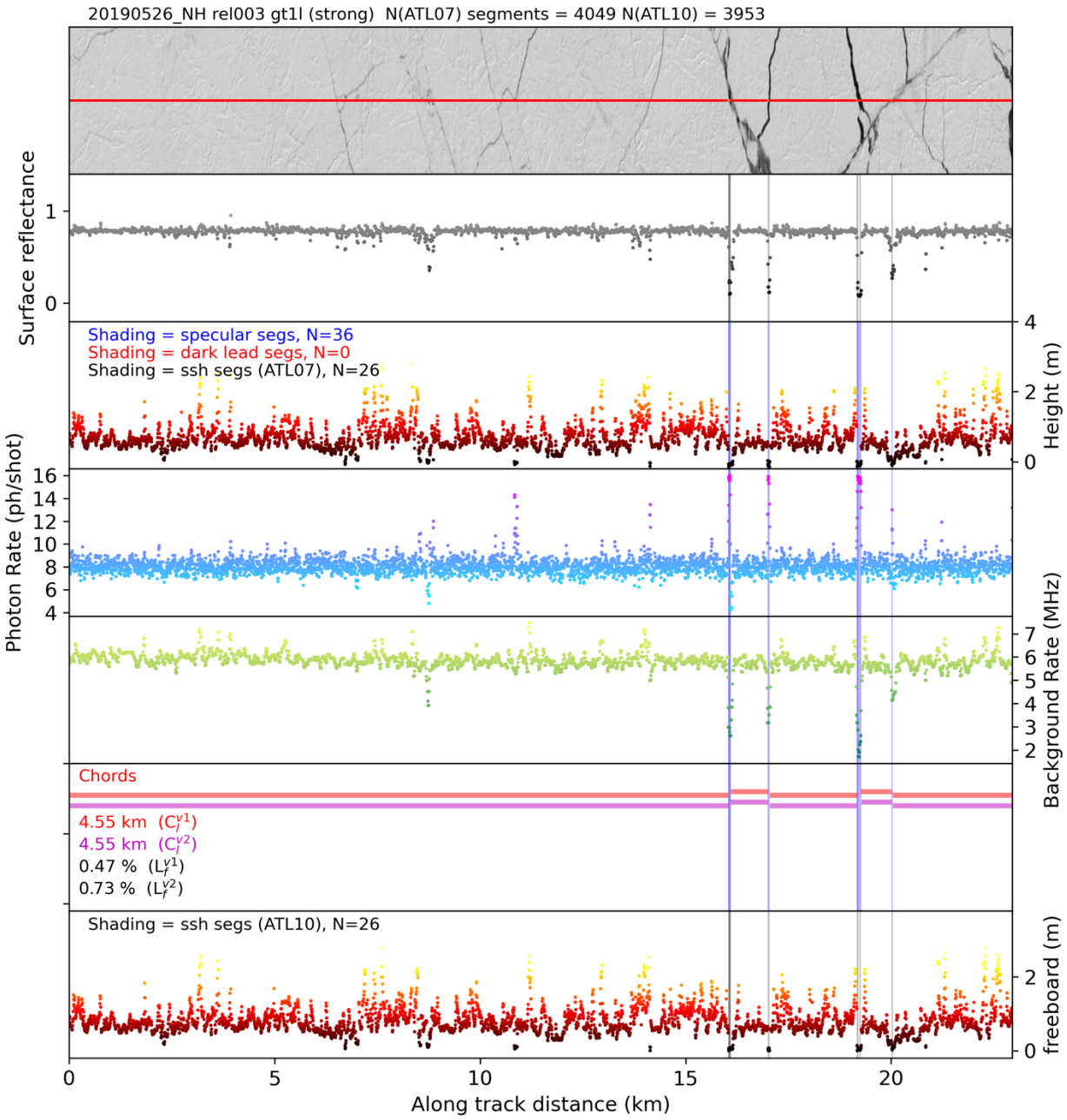


Figure S6: As in Figure 3 but the strong beam *gt1l*

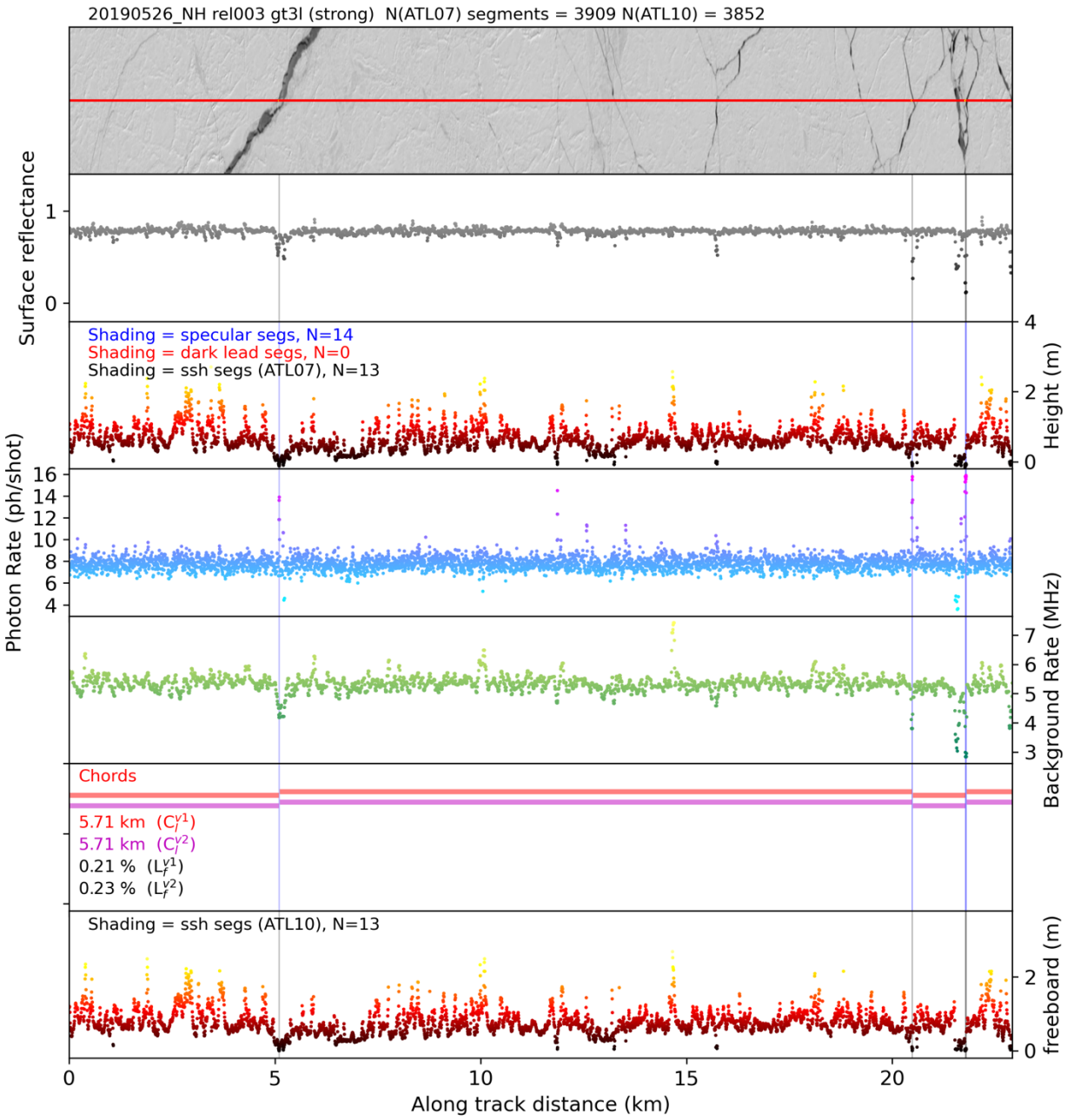


Figure S7: As in Figure 3 but the strong beam *gt3l*

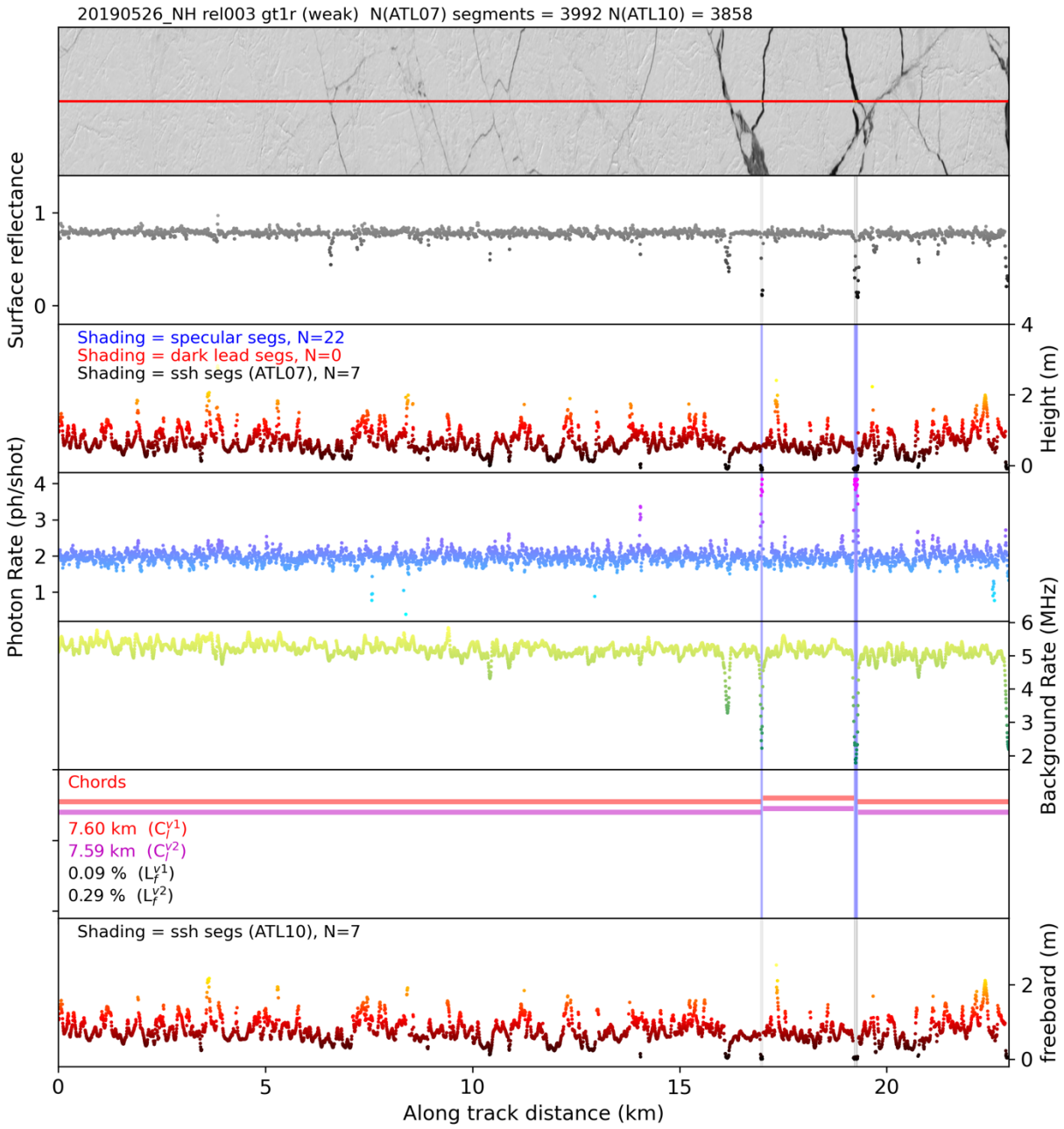


Figure S8: As in Figure 3 but the weak beam *gt1r*

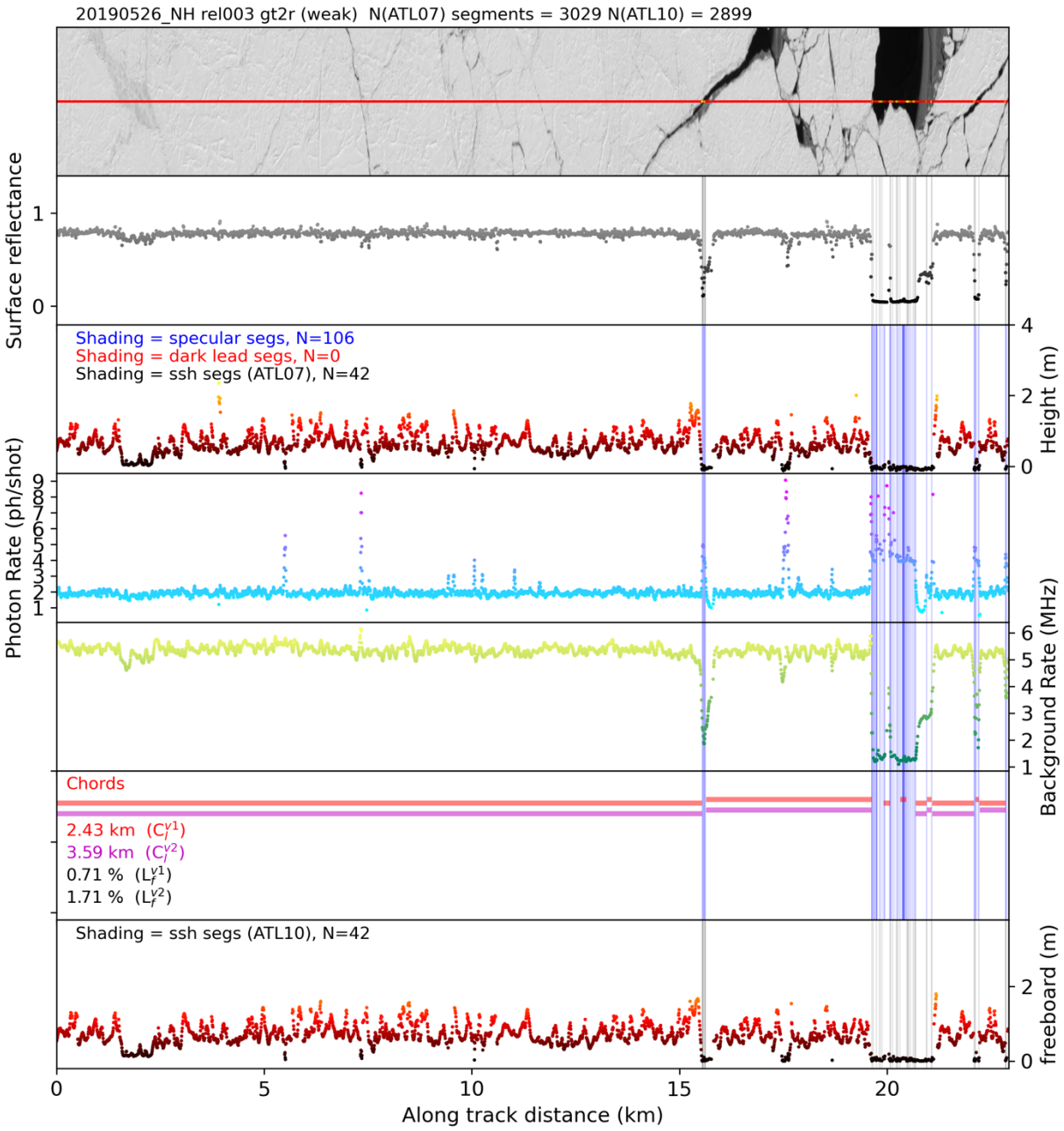


Figure S9: As in Figure 3 but the weak beam *gt2r*

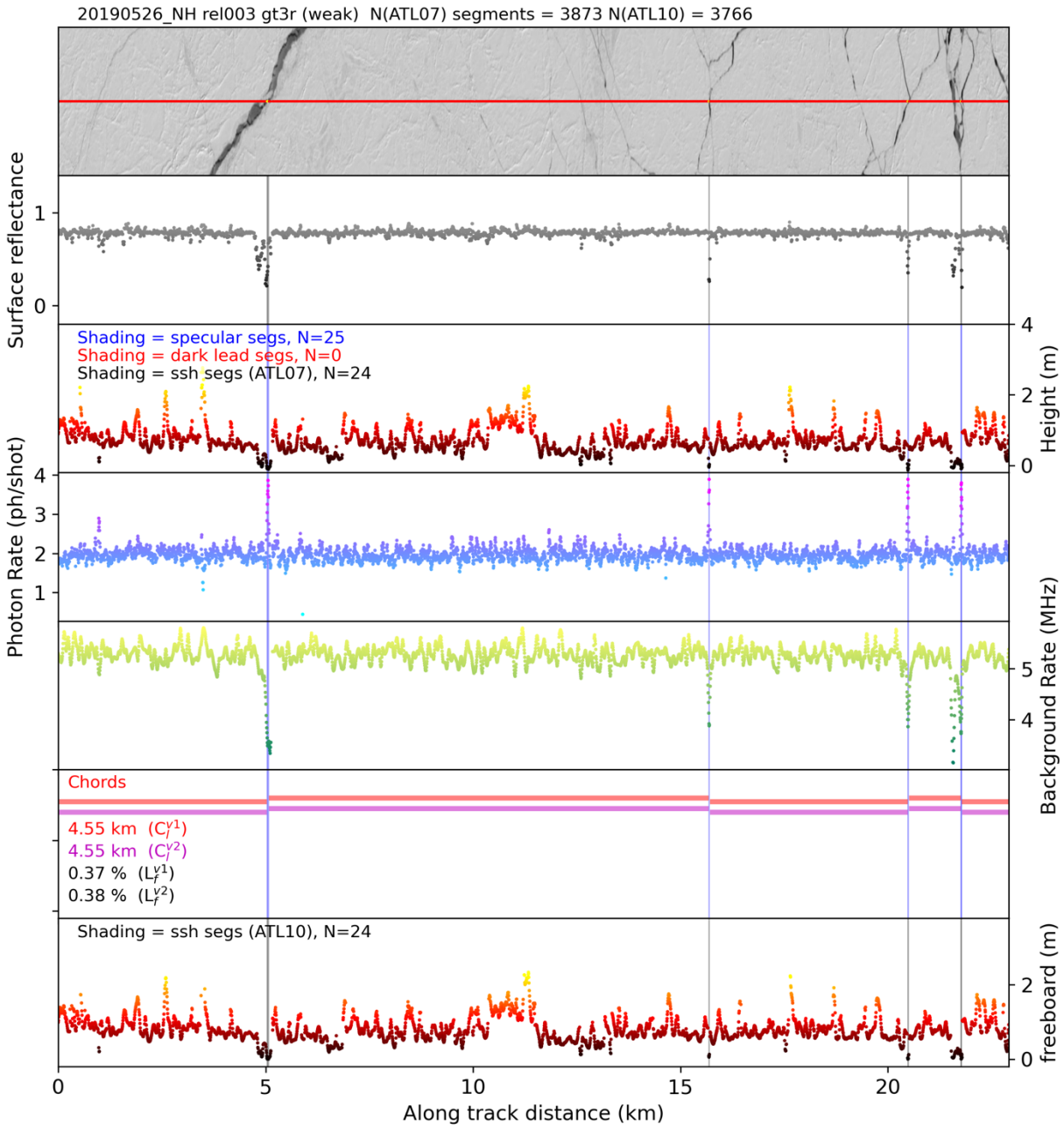


Figure S10: As in Figure 3 but the weak beam *gt3r*