This manuscript is a methodical work that builds upon Scanlan et al. [2023] to pursue the layout of the foundations for using RSR-derived surface parameters at the GrIS. Eventually, it will become an entirely new type of measurement to assess the surface properties of the GrIS and the AIS. It is also the first study (all planetary bodies combined) that carefully confront RSR to altimetry roughness to better understand the baseline at which RSR roughness is derived. The manuscript is well written with a clear language and layout that both help its understanding by non-specialists in this intricated topic. The manuscript should definitely be considered for publication after the points below are addressed. The main improvements could come from extended discussions to better understand the possible causes and implications of some of the observations made. Of importance, roughness derived from laser altimetry measurements should be presented from a more critical standpoint to better understand the comparison with the RSR measurements.

Author Response: We'd like to begin by sincerely thanking Cyril for volunteering his time to review and comment on our manuscript. We appreciate that you find it within the scope of publication in TC and are deeply grateful for your efforts. In the following, we have gone through and responded to each of the comments raised and outlined how we have revised our manuscript in response. We believe that going through and addressing each of the comments has helped us immensely in improving the manuscript.

Our sincerest thanks once again to Cyril for taking the time to review our work and help us improve on it.

Critic of Laser Altimetry Measurements

The authors adopt a critical stance toward the RSR-derived roughness, which is appropriate given the relative novelty of this type of measurement, and it is in line with the goal of the manuscript to strengthen the understanding of the RSR-derived parameters for a confident use in polar science. To that aim, airborne and space-born laser altimetry are presented as reference measurements that have to be matched by the RSR for it to be validated. Laser altimetry is not necessarily deprived of possible bias, however, and this should be further discussed.

Laser- and RSR-derived roughness proceed from measurements implying different physical interactions with the surface. Laser altimetry is a discrete measurement of evenly-spaced signal delay translated into surface heights; while RSR-derived roughness fundamentally proceeds from the coherent summation of all the electric fields scattered back be each finite elements forming the surface within the radar footprint. Here are some thoughts that the authors could partly use if desired for further discussion in comparing the two measurements and explain their possible mismatch:

- What is the footprint size of the laser altimeters compared to the baseline measured? If the laser footprint is greater than the baseline, would that create a bias in the roughness derivation such as the observed on Fig.2?
- I suppose the laser altimetry heights are arranged along-track (i.e., a 1-D spatial direction) making any derived roughness highly sensitive to surface anisotropy. Conversely, I expect the RSR to be less sensitive to anisotropy as it "senses" the surface over a 2-D spatial footprint.

Author Response: Thank you for the comment. Cyril is correct in that we have taken the perspective of the laser altimetry results being the "truth" against which the RSR results are judged. And we agree that we have not been as critical of that assumption as is perhaps warranted. We will also raise the point that the comparison between the RSR results and the ALS versus ICESat-2 datasets are also somewhat different in that ALS is a scanning LiDAR versus the dominantly nadir pointed nature of ICESat-2. We will endeavour to address the reviewer concerns in a new sub-section to be added in Section 6.

Actions Taken: Please see the revisions made in Section 6. A new Section 6.3 *Revisiting Laser Altimetry as an Objective Dataset* has been added to the manuscript to address Dr. Grima's comments.

6.3. Revisiting Laser Altimetry as an Objective Dataset

Throughout this study, the ALS and ICESat-2 laser altimetry data and derived surface roughness results are considered the standard against which the radar altimetry RSR-based results are assessed. However, it is also worthwhile to revisit this assumption, as the nuances in the underlying datasets may affect their relative sensitivity to surface roughness conditions.

First consider the scale of the footprints relative to the RMS deviation baselines. For the airborne CryoVEx data, the footprint of individual altimetry measurements are on the order of 0.7 m in diameter (ALS), 3 m along-track by 10 m across-track (ASIRAS), and 5 m along-track by 12 m across-track (KAREN) (Skourup et al., 2019, 2021). The satellite data on the other hand, as expected have larger footprint diameters ranging between 17.5 m for ICESat-2, 1.65 km for CryoSat-2 (pulse-limited footprint; 720 km altitude, 320 MHz bandwidth), and 1.4 km for SARAL (pulse-limited footprint; 800 km altitude, 500 MHz bandwidth) (Markus et al., 2017; Steunou et al., 2015; Wingham et al., 2006). As the ALS and ICESat-2 footprints are generally smaller than posting interval for their individual datasets (i.e., 1 m by 1 m for ALS, 20 m along-track for ICESat-2 ATL06), their RMS deviation profiles are assumed to be unbiased over the baselines considered (i.e., smallest roughness baseline is greater than an individual footprint). The sole exception to this though is at crossovers where the laser (ALS or ICESat-2) has sampled the same location multiple times. Examples for the ALS are T12 and T30 in 2017 and T21 in 2019. Laser data are analysed spatially as opposed to by individual CryoVEx segment or ICESat-2 orbit number, so overlaying multiple measurements on top of one another can lead to surface elevation measurements spaced less than one footprint apart. Note though that this will only affect the smallest RMS deviations

such as those for a 0.5 m baseline reported in Figure 2 and <20 m ICESat-2 baselines in Figure 4. At larger baselines, the individual laser surface heights will continue to be more than one footprint apart and the subsequent RMS deviation profiles should not be biased by any unaccounted for large spatial sensitivities.

Expanding beyond individual footprints to consider coverage, only for the CryoVEx case do the 100 m-wide ALS swaths cover the ASIRAS and KAREN radar footprints completely. In this case, it is then possible to be relatively certain that both datasets are responding to the same surface conditions. The same cannot be said for the satellite datasets. In general, the ICESat-2 surface elevations are predominantly sensitive to along-track conditions. While there are ideally surface elevations from all six of the across-track beams (subject to ICESat-2 data quality control, see Section 2.2), the spatial sampling of surface roughness will always be denser in the along-track direction. The RSR results on the other hand, represent the collected response of all scatterers within the broader illuminated radar footprint (Grima et al., 2012, 2014a, 2022). Even though attempts are made to ensure that the satellite laser and radar data being compared come from the same region (e.g., Figure 5), the sampling of the surface within those regions is not necessarily equivalent. ICESat-2 RMS deviations will be more strongly affected by any anisotropic surface conditions, whereas the CryoSat-2 and SARAL RSR results are based on a more complete two-dimensional view of the surface. Considering data on a monthly time interval further negates possible impacts of the different orbital designs and repeat cycles (91 days, 369 days, and 35 days for ICESat-2, CryoSat-2 and SARAL respectively).

In summary, while biases stemming from non-negligible laser footprints are considered minimal and care is taken to ensure overlapping measurements for each comparison, the different spatial footprints and sensitivities to anisotropy may still influence surface roughness derived from laser datasets and its comparison to the radar results.

Other Comments

l.129. I understand the pulses are sumed along-track. Could you comment on how this could affect the Pc/Pn ratio and the related roughness derivation? Could that affect the mismatch between the empirical and analytical RSR roughness?

Author Response: Thank you for the comment. Cyril is correct in that the CryoSat-2 LRM and SARAL data have undergone along-track summing. Unfortunately, minimally processed, un-stacked waveforms are not available for these low-rate mode data products (but are for CryoSat-2 when it is operating in its SARIn mode). We agree that the implications of this on the sensitivity to surface roughness and especially the correlation length were not expressed in the initial submission. We will revise the manuscript to address this point specifically.

Actions Taken: Please see the revisions have been made in Section 3.2.

Once coherent (Pc) and incoherent (Pn) are determined, deriving surface roughness requires adopting a representative backscattering model. The simplest implementation of the RSR technique (Grima et al., 2012, 2014b, a; Scanlan et al., 2023) assumes incoherent *backscattering from the surface follows the Small Perturbation Model (SPM) (Ulaby et al., 1982). In this case, surface RMS height can be derived following*

$$\sigma_h = \frac{\lambda e^{P_n/2P_c}}{4\pi\sqrt{P_c/P_n}},$$

where λ is the signal wavelength [m]. The validity bounds of the SPM are kσ_h<0.3 and kl<3 where k is the radar wavenumber [m-1] and l is the surface roughness correlation length [m]. The analytical relationship for deriving RMS height from the RSR results in Equation 3 is based on the assumption that the surface roughness correlation length (i.e., the length scale over which the roughness occurs) is large relative to the radar footprint and can be neglected (Grima et al., 2012, 2014a). ***A subtle feature of this approach for both the CryoSat-2 LRM and SARAL results is that the along-track stacking inherent in the L1B and SGDR datasets (see Section 2.2) makes the RSR results more sensitive to the surface roughness correlation length in the along-track direction (Grima et al., 2014a). Stacking preferentially enhances reflections from tilted roughness elements fore and aft of the stacking midpoint, increasing their contribution to the total received power and accentuating the along-track correlation length. However, any impact due to CrySat-2 LRM and SARAL stacking is assumed to be minimal.***

In~~Note that in~~ contrast to when deriving surface dielectric permittivities from the RSR results, the absolute calibration of the RSR coherent powers is not directly required to produce a roughness estimate (Grima et al., 2012, 2014a, b; Scanlan et al., 2023). However, as an additional check on the overall applicability of the RSR approach, CryoVEx RSR results have been calibrated using the contemporaneous in-situ measurements.

l.185. Could you recall to the reader why the echo is corrected from the nadir surface slope? For planetary radar sounders, this is usualy not done because the radiation pattern is not directive wrt. to the surface slope. If the transmitted signal is directive, the slope correction might work up to a certain angle depending of the beam shape (unless you also correct for the beam gain).

Author Response: Thank you for the comment. The large-scale slope correction comes from a preceding study (Scanlan et al., 2023) because in the very initial implementation, the RSR results were heavily influenced by the surface slope. A strong association was observed between the echo powers extracted from the satellite altimetry waveforms and the large-scale tilt of the GrIS surface. For example, a ~15 dB drop in SARAL echo power was observed for surface slopes between 0.1° and 1° (Supplementary Figure S1 in Scanlan et al., 2023). This led to a clear pattern in coherent powers which inversely matched the slope conditions (higher Pc in the interior and lower Pc around the margins). And once calibrated, the RSR inversion results were unrealistic (e.g., dielectric permittivities markedly exceeding 3.15 in the GrIS interior). Because the initial results seemed to be so dominated by larger-scale topography, the decision was made to account for it at the echo power level as opposed to at the RSR inversion stage. That being said, the reviewer's point is well taken and revisiting this decision and how it flows through to the final results would be a worthwhile exercise.

Actions Taken: Because this points to future work beyond the consideration of the immediate results, we have included a pointer to this specific problem when discussing future development directions at the end of Section 6.1.

Even though the RSR surface roughness results do not appear to be relevant as direct inputs for current SMB modelling, that does not mean they are without future intrinsic value by themselves. ***First,*** ~~*For example,*~~ *there may be a role for directly using the radar-derived surface roughness estimates to refine the retracking of the radar waveforms and improving surface height determinations.* ***Second*** ~~*Furthermore*~~*, there are clear spatial heterogeneities in the RSR results (Figures 8 and 9) that warrant further investigation and may shed light on the nature of GrIS surface conditions.* ***Third, they represent a baseline for interpreting the RSR results in the context of other radar backscattering models (Fung and Chen, 2004; Ulaby et al., 1982) and revisiting some of the underlying decisions (e.g., correcting echo powers for nadir slopes prior to RSR processing).*** *Lastly, the ever-increasing confidence in our ability to reliably observe GrIS surface properties from CryoSat-2 and SARAL surface echo powers provides a foundation to continue applying and adapting these techniques to earlier satellite remote sensing datasets (e.g., ERS-1, ERS-2, ENVISAT); thereby extending our observational timeseries.*

l.207. Maybe recall that this condition on the correlation coefficient is arbitrary, bit it enables to discard some of those terrains with more than one roughness regime in order to comply with the assumption behind the HK statistics.

Author Response: Thank you for the comment. We agree that the assumptions implicit in the definition of "quality" when it comes to assessing the RSR results are not being communicated as clearly as they should be. We will revise the manuscript to make this more explicit.

Actions Taken: Please see the additions made in Section 3.2.

Two metrics are used to quality control the RSR results: first, the distance to the furthest surface echo power measurement considered, and second, the correlation coefficient between the observed surface echo power histogram and the statistical fit. The former is irrelevant for the CryoVEx data as we mandate using the closest 1000 data points surrounding the in-situ locations regardless of how far they may be. For the satellite results, search radii of 50 km, 25 km, and 40 km are used for the CryoSat-2 LRM, CryoSat-2 SARIn, and SARAL results respectively (Table 1). The CryoSat-2 LRM and SARAL maximum search radii are the same as those used in Scanlan et al., (2023), while the CryoSat-2 SARIn radius is extended as more datapoints are being considered. A minimum correlation coefficient of 0.96 is required for all RSR results (Grima et al., 2012, 2014b, a; Scanlan et al., 2023). ***This threshold is used to select only locations where the envelope of observed surface echo powers can be well-described by the homodyned K-distribution, which includes implicit assumptions for how radar energy is scattered from the surface (Grima et al., 2014a). As such, locations that do not meet this correlation coefficient threshold are not necessarily of poorer quality but require a different interpretation of the observed echo powers. This study though focuses solely on locations well-described by the homodyned K-distribution statistics.***

In addition to revisions made in Section 3.2, we have also modified the language in the Figure 9 caption to better reflect the nuances in the quality control assessment.

Figure 9: Locations of >2σ outliers between the wavelength-scale RMS deviations projected from ICESat-2 and a) CryoSat-2 and b) SARAL surface roughness estimates. The locations are plotted on top of maps showing the number of months with valid (i.e., quality-controlled) RSR observations for the period 2013-2018 (72 months). While some outlying roughness mismatches occur closer to the boundaries of the various datasets, there is a cluster in SE Greenland at ~3000 m elevation in the vicinity of the ice divide that corresponds with a zone of~~anomalously low quality~~ RSR results ***that do not meet the quality control criteria****. The impact of CryoSat-2 and SARAL orbital designs can be seen in the spatial patterns (CryoSat-2 SARIn latitudinal stripping and SARAL hatching) in the southern portions of the ice sheet.*

Fig.2.c-d and the related discussion. I'm intrigued about this scale-invariant roughness for the laser. Would it be possible for it to be a bias due to the technique? This mode happend for very low baselines (< 10m). How does this compare with the laser footprint?

Author Response: Thank you for the comment. This is related to one of the main points Cyril raised previously, so please see our response above. However, we do not think there is a footprint-introduced bias in the ALS data as the stated footprint is less than one meter at the nominal flying altitude of 300 m. So small baselines down to one meter should be reliably resolved by the laser scanner. The seemingly scale invariant roughness at the T30 and T41 seems to be a highly localized feature unique to these locations.

Actions Taken: Please see the revisions made in response to the reviewer's previous 'Critic of Laser Altimetry Measurements' comment.

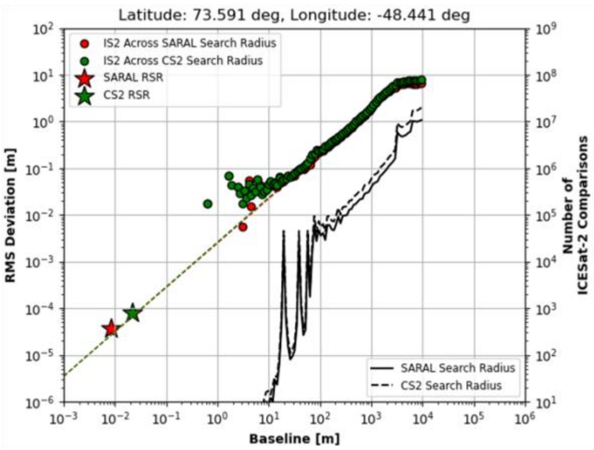
l.251-53. Could you discuss this observation? I wonder if there is a physical sense that could explain that this 200-700m baseline is better for radar roughness downscaling. Is this baseline range related to the radar footprint somehow? This is the tricky part for the radar. Any individual scattering is usually associated to a wavelength-scale baseline (although this is a just common theoretical assumption), but the coherent wave front recorded at the antenna is the summation of all those scattereres from within a much wider footprint. In the end, the scale of the footprint migth also have a role in the derived roughness.

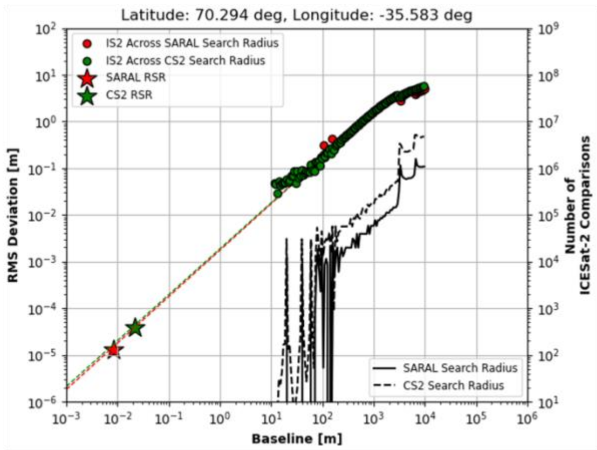
Author Response: Thank you for the comment. The radar footprint radii range from meter-scale for the airborne CryoVEx case to ~1.5 km for the satellites (i.e., CryoSat-2 and SARAL). All the while the alignment with the 200-700 m ICESat-2 roughness baseline seems to be consistent for both the airborne and satellite cases.

Why the RSR results seem to be keying off that 200-700 m interval is something we have thought long about but have yet to come to a firm conclusion. But it is important to consider that focusing too much on 200-700 m interval specifically may be a bit of a misnomer. That specific baseline range is taken because it is where we believe we have both reliable ALS and

ICESat-2 RMS deviation (Allan) profiles (e.g., Figure 4). This is because the along-track ICESat-2 ATL06 data cannot provide insight at baselines shorter than the 20 m posting interval and we only consider ALS data within 1 km of the in-situ locations (capping maximum investigable baseline).

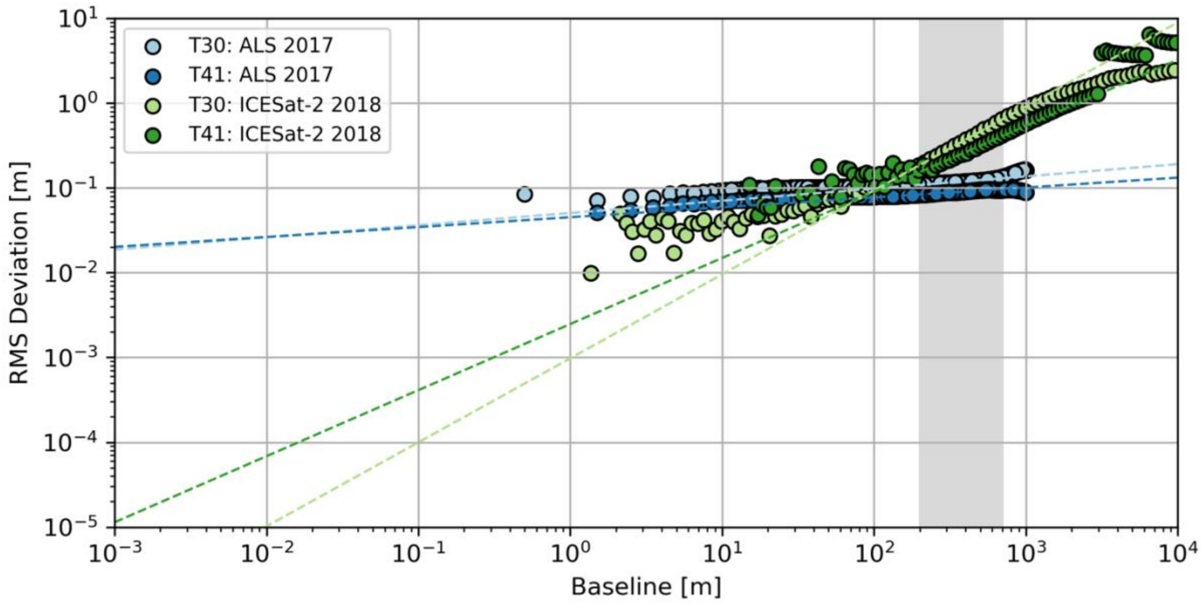
If we take a closer look into some of the individual ICESat-2 Allan profiles (two more examples shown below), we can see that the linear behaviour observed between 200 and 700 m can extend well beyond this range, to baselines of ~3 km (at which there seems to be another breakpoint). Here the baselines are on the order of the CryoSat-2 and SARAL footprint diameters but are substantially greater than the airborne CryoVEx footprints. Note also, if we push the ICESat-2 data to as short of baselines as possible using data near orbit intersections (hence there are only a small number of comparative data points; black solid and dashed lines in the figures below), we do see that the RMS deviations start to level off a bit similar to what is observed in the ALS data.





All this being said, we agree that some of the nuances around the 200-700 m interval need to be addressed more explicitly in the manuscript. We have incorrectly implied it to be a 'fixed' interval, while it is better thought of as a consequence related to the nature of the datasets we are looking at. While we cannot tie the apparent sensitive roughness scale concretely to the radar footprint (the CryoVEx footprints are much smaller), we can state that based on the observations, the minimum consistent baseline the RSR results appear sensitive to is on the order of 100's of meters. We will make numerous revisions in the manuscript to try and make this clearer.

Actions Taken: We have revised the ICESat-2 RMS deviation profiles in Figure 4 to illustrate some of the more nuanced aspects of these data, especially when pushing the data as far as possible at small baselines.



The Figure 4 caption has been modified as follows.

Figure 4: The comparison of March/April 2017 ALS (blues) and November 2018 ICESat-2 (greens) RMS deviation profiles centred on locations T30 (light) and T41 (dark) along the EGIG line. There is a substantial discrepancy between the two sets of RMS deviation profiles in the overlapping baseline range (100 m to 1 km) further confirming that the local regions surrounding T30 and T41 are markedly smoother than those further afield. The anomalous behaviour in the ICESat-2 RMS deviation profiles at ***both small (i.e., <80 m) and long*** *~~large~~ (i.e., >4 km) baselines is related to the quick drop-off in the number of comparable surface elevations.*

In-line with the revision of Figure 4, the following revisions have been made in Section 4.1.

Amongst the ALS results, the 2017 T30 and T41 RMS deviation profiles (Figure 2c) are unique, in that they do not exhibit the increased scale-dependent roughness behaviour at longer baselines. Instead, their RMS deviation profiles are flat and monotonic (i.e., not piecewise linear). The reason for this change in surface roughness behaviour is due to the ALS data surrounding T30 and T41 preferentially covering extremely smooth local areas of the GrIS. Figure 3 presents the local surface elevations (Figures 3a and 3b) along with the height deviations (elevations minus the constant location-specific background plane; Figures 3c and 3d) surrounding the T30 and T41 ALS datasets. Elevations are taken from the 10 m ArcticDEM mosaic (Porter et al., 2018) and the background plane is defined using all 10 m ArcticDEM data within 20 km of the CryoVEx in-situ measurement location. It is clear that the topographic variability across each of these sites is very small at T30 and essentially non-existent at T41. It is then not surprising that the corresponding RMS deviation profiles (Figure 2c) do not exhibit the increase in RMS deviation at large horizontal baselines that is observed at the other CryoVEx locations. To further emphasise the smoothness of the GrIS near T30 and T41, Figure 4 compares the ALS RMS deviation profiles with those derived from all ICESat-2 surface elevations within 25 km and 35 km of T30 and T41, respectively. We must use ICESat-2 data that are further away from the T30 and T41 sites because these locations are between ICESat-

2 orbital ground tracks. When considering surface topography over a broader regional area (i.e., ICESat-2), the stronger scale dependency in surface roughness is once again observed. Less scale dependency exists in the ALS RMS deviation profiles between 100 m and 1 km because T30 and T41 are sited in a locally very smooth portion of the GrIS (Figure 3). ~~*It is also worth noting that the breakpoint in the RMS deviation profiles at baselines between 100 and 200 m is not being captured by the 20 m posting ICESat-2 ATL06 data. Therefore, without the small baseline insight gained from the airborne CryoVEx ALS data, the RSR surface roughness results could have been misinterpreted as a direct continuation of the monotonic ICESat-2 RMS deviation profiles.*~~

In addition to the return of scale-dependent roughness when considering the broader regions surrounding T30 and T41, the ICESat-2 results in Figure 4 also demonstrate two other notable points. The first is that when the ICESat-2 baseline is pushed to its shortest limit (i.e., considering closely spaced surface elevations at orbit crossovers), the RMS deviation profile appears to flatten. While there are only a small handful of surface elevation measurements at these short baselines (e.g., 10's of points separated by 10 m compared to 10,000's of points separated by 20 m), the ICESat-2 results do seem to exhibit the same less scale-dependent roughness pattern as has been observed in the ALS results (Figures 2c and 2d). And again, as with the ALS data, the transition in the RMS deviation profile occurs in the 100 m range. The second notable point is that the piecewise linear, scale-dependent behaviour in the RMS deviation observed in the ALS data between 200 and 700 m baselines (Figures 2c and 2d) appears to continue well beyond that range, to upwards of roughly 3 km in the ICESat-2 results. Recall that only ALS data within 1 km of the in-situ measurement location underlie the RMS deviation profiles in Figure 2. The 200 to 700 m interval used in the projection of the RMS deviation profiles to the radar wavelengths is used solely because it is an interval common to both the ALS and ICESat-2 results. The scale-dependent behaviour in the surface roughness appears to extend to much longer baselines. Lastly, while the ICESat-2 RMS deviation profiles to being to suffer data availability issues, there does appear to be another marked transition to less scale-dependent roughness at the longer (i.e., >4 km) baselines.

The following revision has been made in Section 4.2.

In lieu of presenting hundreds of individual ICESat-2 RMS deviation profiles together with the CryoSat-2 and SARAL RMS heights from Scanlan et al. (2023) (i.e., akin to Figures 2c and 2d) and following on from what has been learned from the CryoVEx results, Figure 6 presents the ***comparison***~~*comparisons*~~ *of the RSR RMS heights, and the wavelength-scale RMS deviations projected from a linear fit to the RMS deviation profiles for baselines between 200 and 700 m for 328 locations across the GrIS.* ***As presented in relation to Figure 4, the 200-700 m interval is selected due to its general representativeness of the piecewise linear, scale-dependent roughness behaviour observed in both the ALS and ICESat-2 data. It should not be considered a uniquely fixed interval.*** *These 328 locations have been pseudo-randomly selected based solely on considerations for the computational load when performing the point-to-point surface deviation comparison as part of the RMS deviation profile calculation (i.e., ≤55,000 ICESat-2 surface elevations). To ease the comparison, all surface roughness estimates* *(i.e., CryoSat-2 and SARAL RSR RMS heights or ICESat-2 RMS deviations projected to the CryoSat-2 and SARAL wavelength scale) have been normalized by the radar signal wavelength. While there may be the suggestion of possible linear relationship between the radar- and laser-derived surface roughness estimates, the is clearly no 1:1 agreement. This appears to be in part due to a floor in the RSR results as they consistently fail to recover the smallest ICESat-2 RMS deviations. The mean absolute error between the two sets of surface roughness estimates is 0.0308 λ for CryoSat-2 and 0.0346 λ for SARAL.*

The following changes have been made in Section 6.1.

Instead, for the Ku- and Ka-band airborne and satellite radar altimetry data, RSR surface roughness is best interpreted ***not as the true wavelength-scale RMS deviation, but the projection of the scale-dependent behaviour observed at baselines between hundreds of metres and a few kilometres to the wavelength scale*** *~~as the wavelength-scale projection of the true RMS deviation profile behaviour observed between 200 and 700 m baselines~~. The implication is then that the SARAL and CryoSat-2 RSR surface roughness results only have physical meaning far beyond the individual roughness feature scales currently considered critical in heat flux and SMB* ***studies****~~modelling~~. As such, they have no direct role to play in the improvement of current GrIS SMB modelling. The relevance only at long baselines is likely also the reason why the surface roughness timeseries presented in Figure 10 do not exhibit the strong seasonal variability that has been reported in derivations of aerodynamic roughness lengths (Smeets and van den Broeke, 2008; van Tiggelen et al., 2021, 2023).*

The following revisions have been made in the Section 7.

Against this backdrop, surface roughness derived from CryoVEx radar altimetry surface echo powers do not align with ***a continuation****~~the extrapolation~~ of the LiDAR RMS deviation profiles to the wavelength scale. In fact, they appear to align much better with the extrapolation of* ***the consistent*** *piecewise linear portion of the RMS deviation profiles* ***observed at****~~for~~ baselines between* ***hundreds of metres and a few kilometres****~~200 and 700 meters~~. Building on the CryoVEx results, the direct comparison between extrapolated ICESat-2 surface roughness RMS deviations (from the piece-wise linear portion between 200 and 700 meters) and previously published CryoSat-2 and SARAL radar surface echo powers derived using an analytical backscattering model reveals that the radar-based results tend to overestimate surface roughness.*

The observed sensitivity of the spaceborne RSR results to surface roughness ***that is at its smallest, hundreds of meters in scale****~~at hundreds of meters baselines~~ suggests they are not well-suited to being incorporated in current surface mass balance (SMB) modelling as these models rely on the roughness of individual meter-scale features such as hummocks or sastrugi.*

l.310. Throughout the paper, the RMS deviation (aka. Allan deviation) and RMS heigth appears to be used interchangeably. I understand the authors do know they are different, but it is not clearly highligthed in the text and the reader could be misled, especially when those two parameters are compared on the same 1:1 plot (Fig.6b). A brief discussion on the difference between RMS deviation and RMS heigth should be added to understand if one should expect any bias in that figure or not.

Author Response: Thank you for the comment. Confusion is understandable and is a result of imprecision on our part. The RMS heights presented here are those coming directly from the previous publication (i.e., Scanlan et al., 2023) where the SPM was used in the inversion of the RSR results. We will check the manuscript to make sure the context is clear whenever specific terminologies are used. We will also include a brief description of the RMS height and its main salient difference compared to the RMS deviation (i.e., no consideration of scale).

Actions Taken: The following revision has been made in Section 3.2.

In this case, surface RMS height can be derived following

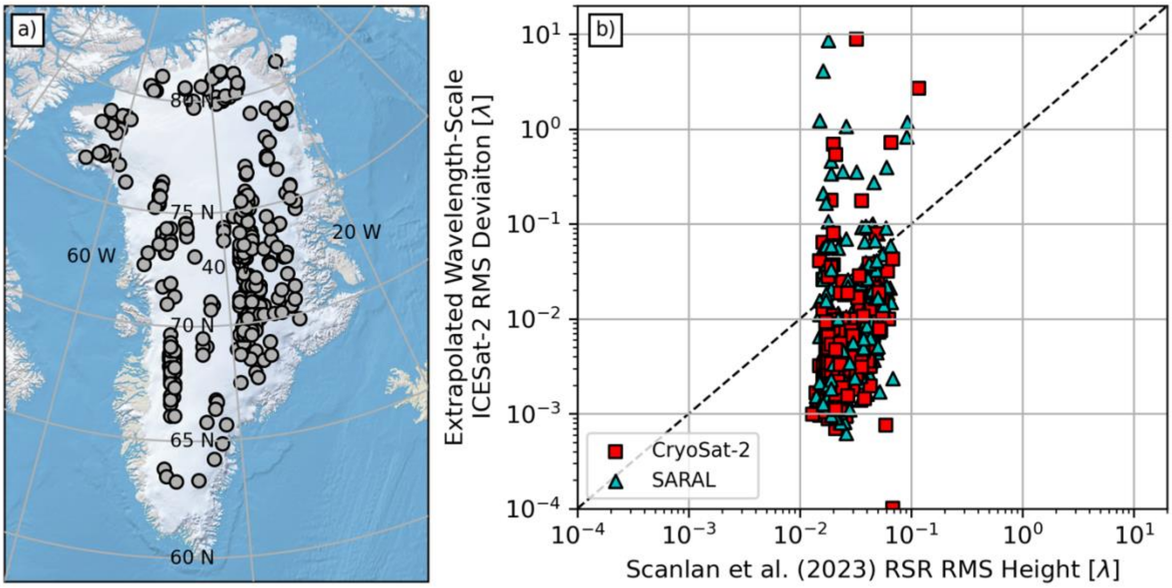
$$\sigma_h = \frac{\lambda e^{P_n/2P_c}}{4\pi\sqrt{P_c/P_n}}, \qquad (3)$$

where λ is the signal wavelength [m]. ***In contrast to RMS deviation (Equation 1), RMS height is simply the standard deviation of the surface heights (after removing the mean) with no consideration of horizontal scale (Shepard et al., 2001).*** *The validity bounds of the SPM are kσ_h<0.3 and kl<3 where k is the radar wavenumber [m-1] and l is the surface roughness correlation length [m].*

The following revision has been made in Section 4.2.

In lieu of presenting hundreds of individual ICESat-2 RMS deviation profiles together with the CryoSat-2 and SARAL RMS heights ***from Scanlan et al. (2023)*** *(i.e., akin to Figures 2c and 2d) and following on from what has been learned from the CryoVEx results, Figure 6 presents the* ***comparison****~~comparisons~~ of* ***those initial****~~the~~ RSR RMS* ***height estimates from Equation 3****~~heights,~~ and the wavelength-scale RMS deviations projected from a linear fit to the RMS deviation profiles for baselines between 200 and 700 m for 328 locations across the GrIS.*

We have revised the horizontal axis label of Figure 6 to more clearly express the provenance of these data.



The following revision has been made to the Figure 6 caption as well.

Figure 6: Panel a) presents the locations of 328 pseudo-randomly chosen locations across the GrIS where in Panel b) ***Scanlan et al. (2023)*** *RSR surface roughness results (CryoSat-2 as squares and SARAL as triangles) are compared to wavelength-baseline projected RMS deviations from ICESat-2. While there is a general positive associated between the two sets of roughness estimates, the RSR results do not reliably recover the smallest ICESat-2 roughness levels. The mean absolute error between the ICESat-2 and the CryoSat-2 and SARAL RSR-based wavelength-normalized roughness estimates are 0.0308 λ and 0.0346 λ respectively.*

Eq.4. Although the methodology to obtain such relationship could be used elsewhere, its coefficients migth not. It is then necessary to discuss the perceived validity domain of this equation to avoid the community to use it with any radar system and in any environment.

Author Response: Thank you for the comment. We agree that, this point should be made explicit in the manuscript. We will revise the manuscript to include it.

Actions Taken: The following revisions have been made in Section 5.1.

Applying this empirical relationship to deriving surface roughness estimates from the RSR outputs yields the comparison against the projected ICESat-2 RMS deviations presented in Figure 8. Comparing Figures 6b and Figure 8, surface roughness produced using the empirical mapping relation (Equation 5) clearly produces a better match than the analytical model (Equation 3). Quantitatively, the mean absolute errors between the ICESat-2 and RSR-based wavelength-normalized roughness estimates are reduced from 0.0308 λ (CryoSat-2) and 0.0346 λ (SARAL) for the analytical model to 0.0119 λ and 0.0174 λ using the empirical model; the substantial reduction indicating a much better agreement between the radar and laser roughness estimates. Furthermore, the difference between ICESat-2 and RSR-based empirical *surface roughness clusters around zero for both CryoSat-2 (Figure 8b) and SARAL (Figure 8c); whereas the analytical approach led to consistently greater RSR surface roughness.* ***A****~~It should be noted that a~~* *similar study* ***only with****~~using~~* *a smaller number of locations in December 2018, also observed an improvement in mean absolute error using the revised empirical RSR-roughness model (Equation 5).* ***Expanding more broadly, the form of Equation 5 and the procedure for developing it could be adapted to other RSR implementations on Earth as well as beyond, but care will have to be taken to ensure the coefficients are appropriate as they may vary in different contexts/applications.***

Fig.8. It sounds too me there is also another outlier group of points between 0.01. and 0.1m, just below the bottom dotted line. Does this group of points have a spatial cohesivness?

Author Response: Thank you for the comment. The locations plotted in Figure 9 are not discriminated between those above 2-sigma or below, all outliers are plotted together. We see now that should have been communicated more clearly. We expect most of the clustering in SE Greenland will be the >2-sigma outliers since there are just more of them. There does not seem to be any obvious cluster that is not on a data coverage boundary besides that in the southeast.

Actions Taken: The following revision has been made in Section 5.1.

An interesting feature present in the radar/laser surface roughness comparison of Figure 8 is that substantial disagreements (i.e., >2σ outliers) between the radar and laser altimetry surface roughness estimates are not symmetric and mainly occur above the dashed 1:1 line (i.e., an ICESat-2 surface roughness greater than that of either CryoSat-2 or SARAL). When looking at where ***all of*** *these outliers occur spatially across the GrIS (Figures 9a and 9b), there is a clear clustering of locations in SE Greenland. That some outlying surface roughness results can be found around the GrIS periphery or at the boundary of the different CryoSat-2 acquisition modes is not unexpected, as this is where the RSR technique is known to struggle with more spatially heterogeneous surfaces and where there are fewer data enveloping a specific location (Scanlan et al., 2023).*

We have also made the following revision to the Figure 9 caption.

Figure 9: Locations of ***all*** *>2σ outliers between the wavelength-scale RMS deviations projected from ICESat-2 and a) CryoSat-2 and b) SARAL surface roughness estimates. The locations are plotted on top of maps showing the number of months with valid (i.e., quality-controlled) RSR observations for the period 2013-2018 (72 months). While some outlying roughness mismatches occur closer to the boundaries of the various datasets, there is a cluster in SE Greenland at ~3000 m elevation in the vicinity of the ice divide that corresponds with a zone of RSR results that do not meet the quality control criteria. The impact of CryoSat-2 and SARAL orbital designs can be seen in the spatial patterns (CryoSat-2 SARIn latitudinal stripping and SARAL hatching) in the southern portions of the ice sheet.*

l.380-84. The authors suggestion is completly valid. I would like to propose additional explanations that could change the Pc/Pn ratio used to derive the radar roughness. The authors are free to consider them or not:

- A differential volume scattering between two terrains will also affect the Pc/Pn ratio. Volume scattering will add incoherent energy.
- If there is a strong roughness anysotropy, the along-track measurements of laser altimetry migth differ from the radar sensitive to more spatially distributed scatterers.
- Thin layering (firn crust?) could change the coherent energy.
- When the surface is too rough, the surface mean slopes will start to scatter most of the energy off-nadir. The Pn measured at the antenna is then an underestimate of the integrated incoherent energy arising from that surface

Author Response: Thank you for the comment. First off, re-reading these sentences and in light of some of Dr. Grima's other comments, we do think the language here needs to be cleaned up. Second, we think these are great alternative suggestions that could explain the anomalous results in this SE portion of the GrIS.

Actions Taken: Please see the revisions …

That being said, the cluster in SE Greenland is surprising as it occurs across a high elevation and inland portion of the ice sheet. Interestingly, the SE Greenland cluster of roughness mismatches corresponds to a location where ~~valid (i.e., quality control passing)~~ monthly (2013-2018) RSR results ***meeting the quality control criteria (Section 3.2) are amongst*** *~~seem to be~~ the rarest. This suggests that the GrIS surface in this area may be unique in some way that continuously affects the RSR results. Based on the ICESat-2 results from Figure 8, one possible explanation could be that this area is substantially rougher than the inland GrIS as a whole and yields distributions of radar altimetry surface echo powers that cannot be* ***cleanly*** *fit by a single homodyned K-distribution probability density function~~; thereby causing the RSR technique to fail~~.* ***Other alternative explanations that could incite changes in the distribution of surface echo powers include a local variation in volume scattering affecting the amount of diffuse scattering and firn crusts/thin layering affecting the specular component.*** *However, as this study focuses on understanding the RSR roughness results, a deeper assessment of the root cause for why the RSR technique seems to experience issues in this area is left to future work.*

l.391. The RSR is not especially lower quality, it is indicative of some sort of additional surface properties that is not caught by the laser (they are hard to disentangle, though) as mentioned in l.457-58.

Author Response: Thank you for the comment. We agree that this was a poor choice of words on our part. Along with the Dr. Grima's earlier comment on what the 0.96 correlation coefficient actually implies about the results, we will revise the manuscript to try and make this clear.

Actions Taken: The following revisions have been made to the Figure 9 caption.

*Figure 9: Locations of >2σ outliers between the wavelength-scale RMS deviations projected from ICESat-2 and a) CryoSat-2 and b) SARAL surface roughness estimates. The locations are plotted on top of maps showing the number of months with valid (i.e., quality-controlled) RSR observations for the period 2013-2018 (72 months). While some outlying roughness mismatches occur closer to the boundaries of the various datasets, there is a cluster in SE Greenland at ~3000 m elevation in the vicinity of the ice divide that corresponds with a zone of~~anomalously low quality~~ RSR results **that do not meet the quality control criteria**. The impact of CryoSat-2 and SARAL orbital designs can be seen in the spatial patterns (CryoSat-2 SARIn latitudinal stripping and SARAL hatching) in the southern portions of the ice sheet.*

l.481. larger "than".

Author Response: Thank you for the comment. This was a typo that made it through our proof-reading and will be corrected

Actions Taken: Please see the revision has been made in Section 6.2.

*Based on Figures 7 and 8, that the analytical permittivities are larger **than** ~~the~~empirical results is not unsurprising.*