General Comments

This manuscript presents an interesting study into using the leading edge width (LeW) of CryoSat-2 measurements between 2011 and 2021 to investigate the long-term characteristics of Greenland firn conditions. The authors begin their manuscript with an introduction to the importance of understanding melt events, their effects across the Greenland Ice Sheet (GrIS) and how they intend to approach the problem using remote sensing data supported by in situ measurements and numerical climate model outputs. In Section 2, the authors present the various datasets considered in their study. Section 3 then outlines how the data are used, combined and the types of analyses the authors perform; the results of which are presented in Section 4. Finally, in Section 5, the authors reflect on the implications of the results and how future studies could expand on them, before the main conclusions are outlined in Section 6.

Overall, the authors have analyzed and presented a substantial amount of data. They take a very thorough approach to assessing the long-term spatial patterns in CryoSat-2 LeW by incorporating another satellite altimeter (i.e., ICESat-2), other derived satellite and airborne datasets (i.e., roughness, radar-laser offsets, topography), in-situ measurements (i.e., densities), and model results (i.e., densities, meltwater content, and firn air content). The scope of bringing all the data together is impressive and I think very exciting direction of work. Relatedly however, presenting so much data requires a very clear narrative and defined structure to help a general and non-specialized reader avoid being lost in all the details. My main comment on the manuscript is that I found this aspect of it to be underdeveloped, which could hamper the impact of the results.

While reading through study and going through all the results, I found it hard at times to get a sense of how all the different pieces fit together. I think the manuscript would benefit from a clearer statement of the central hypothesis and the physical reasoning behind it that would then frame the overall study. In the Introduction, the authors state the importance of melt and refreeze events and postulate that LeW could be used to study long-term patterns. What I think is missing though is how LeW is affected by melting/refreezing events. What is changing in the firn and how does that affect the radar signals? Concepts of surface and volume scattering as well as refrozen layers appear repeatedly later on in the manuscript, but I think explicit descriptions of what the authors mean by these, their linkage to the physical state of the firn and what that means for the CryoSat-2 would substantially help fortify the overall narrative. This bridging between radar theory and more classical glaciological concepts will also strengthen the impact the manuscript will have by really outlining how all these pieces fit together and what the results mean. Some of the specific comments below will also be in this direction.

I hope the authors find the following comments constructive as they work towards revising their manuscript.

Specific Comments

1) As the manuscript primarily centers on CryoSat-2 LeW results, I would suggest the authors consider revising the title to “Assessing spatio-temporal variability of firn scattering over Greenland with CryoSat-2”. I understand that the inclusion of ICESat-2 data makes a case for multiple altimeters, but my impression is that the ICESat-2 data are more complementary to the

main CryoSat-2 dataset. In a similar way to the MAR and IMAU climate model data, ICESat-2 data appear to be used more to help explain trends in the CryoSAt-2 data, not necessarily as the primary data source themselves.

2) Line 57 "The LeW is adopted as it is sensitive to volume scattering …" Line 67 "… we have to understand both volume scattering and surface scattering …" These are two instances where a more explicit statement of what the authors mean by volume and surface scattering could help improve the overall framing of the study. How/why LeW is sensitive to these two concepts and what on the surface and in the firn contributes to them? I think it would broaden the reach of the manuscript by removing the hurdle of needing to be familiar with nuanced radar theory concepts and motivate exactly why the specific model outputs are chosen for comparison with the LeW results in the latter stages of the manuscript.

3) The authors dedicate Lines 35-51 motivating CryoSat-2 and LeW as a metric for studying firn. I recommend the authors consider expanding more clearly on the motivations for using the other datasets (e.g., ICESat-2, in-situ densities, dz, roughness, topography, and model results) to help explain the LeW results. What aspect of the LeW signal are these datasets being used to interpret? I found Lines 52-69 to be confusing as it was not always clear how these different datasets all supported the LeW analysis.

4) I recommend the authors consider reducing the number of adverbs (e.g., furthermore, finally, in addition, therefore, additionally, etc.) used to start sentences to make them more direct and impactful.

5) To be more specific on the types of GrIS changes of interest in this study, I recommend the authors re-phrase Line 70 from "… assess long-term changes over the …" to "… assess long-term surface changes over the …"

6) In Section 2.1, I recommend the authors include more detail on the nature of the CryoSat-2 LRM data and what differentiate them from other CryoSat-2 data products (e.g., what is unique/different in their acquisition/data processing?).

7) Line 89, please include the range resolution of CryoSat-2.

8) In Figure 1, I'd ask the authors to consider including the $b_{0.99}$ and $b_{0.01}$ values for each waveform as well as map (perhaps as an insert) of where these two locations are in Greenland are.

9) Line 94, what high-resolution DEM model is used?

10) In Line 97, the authors state that they used July measurements as indicative of post-melt conditions but there is no way for the reader to assess if melting has ceased at these locations by the time the data were acquired; especially knowing how extreme the melt extents observed in the summer of 2012 were. I would recommend the authors provide further support for this statement or consider using data from later in the year.

11) I am not sure I fully understand the context for why two different grid resolutions (50x50 and 25x25) are used. I suggest the authors clarify this point.

12) Line 115. Do all CryoSat-2 measurements in a given month have a corresponding ICESat-2 measurement within 50 m or are their spatial gaps? I'd also recommend the authors provide their reasoning for choosing 50m when the footprint of CryoSat-2 LRM data is much larger.

13) With how Sections 2, 3 and 4 are structured, the CryoSat-2/ICESat-2 results from Figure 2 and Lines 122-128 seem to be more suited to Section 4 than Section 2. I understand they are used again in Section 2.4, but could Section 2.4 be treated more abstractly by referring to a subsurface depth extent to be determined later? The current placement seems to interrupt the flow of describing all the individual datasets considered.

14) Lines 134-136. I recommend the authors elaborate a bit more on how "computational efficiency" necessitates using both the 100m and 1km ArcticDEMs in these two instances. What about these specific applications makes the use of two different DEMs more efficient?

15) Figures 2, 3, 4, 6, 7, and 9. I recommend the authors elaborate why they used a DEM from Helm et al. (2014a, b) as their basemap instead of one of the ArcticDEMs they use in their analysis. Also, I'd recommend including a colorbar for the elevations the first time it is used.

16) Line 146. I recommend the authors clarify how the weights are determined in their weighted average densities.

17) I recommend the authors consider better motivating the inclusion of the IMAU FAC. FAC is a column-integrated measurement (Line 168) whereas LeW derived from CryoSat-2 is seemingly only sensitive to the upper few meters (Figure 2). Why would these two datasets derived over different depth ranges be considered comparable?

18) Line 168. The "but" in "… 1.5 m but the FAC …" can be removed.

19) I recommend the authors expand on why these particular in-situ firn density measurements are used instead of the more comprehensive SUMup dataset (i.e., Vandecrux et al. 2023)? Furthermore, why is it necessary for firn density profiles to contain the 2012 melt year (Line 185)?

Vandecrux, B. et al. The SUMup collaborative database: Surface mass balance, subsurface temperature and density measurements from the Greenland and Antarctic ice sheets (1912- 2023). Arctic Data Center https://doi.org/10.18739/A2M61BR5M (2023).

20) Line 212. These 10 DEM elevation groups have not been mentioned yet, so I do not follow how they can be "aforementioned". I recommend the authors clarify this statement.

21) Line 230-231. These seem to be the elevation bands mentioned in Line 212. I recommend the authors clarify why they include elevation bins down to 100 m elevation. It is my understanding that the study only considers CryoSat-2 LRM data which cover the high-elevation interior portion of the GrIS.

22) Line 243. I recommend the authors clarify the "Following …" used to start this section. The previous two analyses described in Section 3.1 and 3.2 use to 25x25 km grid. The adoption of the 50x50 km grid here seems to be a marked departure from what has occurred previously as opposed to following/continuing.

23) Section 3.3. I recommend the authors clarify which months are included in their analysis of long-term variations. As it reads, it seems as though June-December LeW data are not represented (average is derived between January and May, Lines 243-244). What motivates this choice and why are Fall/early winter data not considered? If the goal is to avoid melt being present in the snow, would focusing on the full non-melt season (e.g., Oct.-Apr.) be more appropriate as opposed to following calendar years?

24) Lines 261-263. I recommend the authors be more specific on where on the GrIS they are referring to. Are the number they state representative of the ice sheet as a whole or only a portion of it?

25) Figure 4. I suggest the authors be specific with the LeW time periods behind the data presented here. Do they match the time periods shown at the top of the plot or are they those outlined in Section 3.3?

26) Line 266. I have a hard time following the logic behind this statement because there isn't a really clear statement of how/why LeW is sensitive to volume scattering. Is increased volume scattering expected to increase or decrease LeW? Figure 1 would imply a positive correlation but, to me, here it seems to imply the opposite (reduced scattering (implying reduced LeW) due to subsurface high-density layers).

27) Figure 6. I recommend the authors consider including select representative 2D histograms directly comparing dh and LeW in addition to the correlation coefficient maps. I think this would give a sense on if the data are clustered or the range over which they co-vary against one another.

28) Line 291. Could the authors expand on this point and elaborate on how surface scattering effects the LeW/dh correlation? Is it because the OCOG retracker becomes less sensitive in rough areas?

29) Line 293. I'd recommend the authors be very careful with the statement that penetration depth increases because LeW increases. LeW is an interpretation of an observed signal. If there was no volume scattering in the subsurface, the signal would penetrate as just deeply but no reflected power would exist at that point in the waveform, so LeW would only be a function of surface roughness. The depth to which it is possible for a radar to say something about the subsurface is a function of both how radar is designed and operated (e.g., transmit power, noise levels, data processing) as well as the structure and makeup of the surface and subsurface. All of these would affect at what point the SNR of a reflection from the subsurface would reach 0 dB. In light of this, would it be a more appropriate/accurate option to use "radar-laser height offsets" as opposed to "penetration depths"?

30) Line 314. I am confused by the statement here of a notable recovery in firn conditions and what is on Line 266 where the authors state firn recovery is not reflected. I recommend the authors clarify the distinction/difference between these two seemingly conflicting results.

31) A general comment on the Figures, but I'd ask the authors to consider using different colormaps for different variables. The same red-to-blue colormap is used in Figures 4, 6, 7, 8, 9,

and 10 even though the variable being plotted changes; sometimes an absolute value is shown and sometimes a difference. I would also recommend that when presenting data on a map, the authors label their colorbars to make it explicit what variable is being shown.

32) Lines 338-339. I recommend the authors provide more explanation regarding why regular, annual melt-refreeze cycles are less impactful on volume scattering compared to more intermittent events.

33) In the Discussion section, the authors devote the first paragraph to contrasting their results against those of Rutishauser et al. (2024). The authors compare the results in terms of their spatial patterns, but I would also suggest the authors consider the nature of the underlying radar measurements as well. The OIB MCoRDS radar operates in a much different frequency range compared to the CryoSat-2 SIRAL altimeter. What affect will that have on the resulting data and interpretations that could be assumed to be responding to more or less the same near-surface stratigraphy?

34) Also in the Discussion section, I would also suggest the authors be more specific with what they expect can be gained from integrating radar measurements at other frequencies (Lines 413-415)? MCoRDS data are substantially different from CryoSat-2 but, as outlined in the previous comment, frequency-dependent impacts are not discussed. How can improved results in complex surface and volume scattering areas be improved by adopting more frequencies? At the same time, I'd ask the authors to consider what this means for future dual-frequency radar altimeters such as ESA CRISTAL which will operate Ku- and Ka-band altimeters simultaneously.

35) Figure 10. I recommend the authors elaborate more on the specific elevation intervals presented in 10b, 10d, and 10e. Why are these specific intervals chosen and what additional information do they add? Thes subplots and the specific elevation intervals, not only deviate from every plot that has been shown so far but they are not even mentioned in the main text. What overall purpose do they serve?

36) Line 433. Here "sub-surface" appears with a hyphen, while through the rest of the manuscript it is written as "subsurface".

36) Line 441. "stratigraphy" is misspelt in the *Code and data availability* section.