Response to Anonymous Referee #2

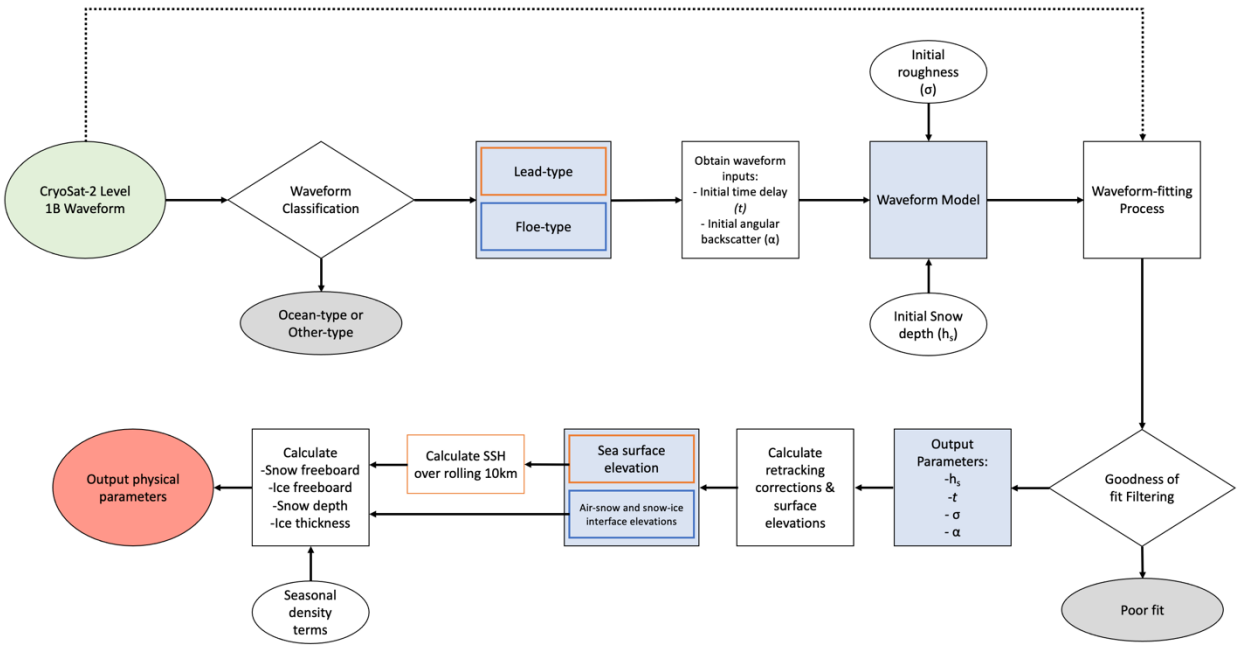
This paper presents an important work to not only fill in the gaps of the ICESat/ICESat-2 observations on the Antarctica sea ice thickness, but also provide a continuous ice thickness time series showing obvious seasonal cycle characteristics with CryoSat-2 from 2010 to 2021. The authors utilize a physical model and a waveform fitting method that they developed in their previous work to get snow depth and total freeboard, then the sea ice thickness and volume. This work provides a sea ice thickness dataset that could be merged with that derived with ICESat/ICESat-2 to produce a longer-term observations of circum-Antarctica sea ice, which would greatly promote global climate change studies. However, there are some concerns that need to be clarified by the authors before publication.

Dear Reviewer,

Thank you so much for your comments on this manuscript. We appreciate the time you put in as well as your contributions to improving this study.
Below, you will find our responses (blue) to your comments (black). We agree with many of your concerns, and have outlined ways in which we will improve the revised manuscript.

Thanks again,
Steven Fons, Nathan Kurtz, and Marco Bagnardi

Major:
It is not clear at which step what parameters are estimated. Generally, there are several parameters involved here: total freeboard (air-snow interface elevation-derived), ice freeboard (snow-ice freeboard elevationderived), snow depth, and ice thickness. L181 states the ice freeboard and snow depth are output parameters produced directly by CS4WFA? However, L292-293 shows the snow depth are derived from subtracting snow-ice freeboard elevation from air-snow interface elevation. So, my question is what is the input and output between each step in this study? At which step comparing with what reference dataset? I suggest supplementing a flow chart to make it clearer.
Thanks for the suggestion, and apologies for the confusion. Some of these processes can be described in different ways. For example, the snow depth **is** an output parameter from the model, but is identical to subtracting the interface elevations and therefore can be “calculated” either way. Additionally, the snow-ice interface tracking point is output from the model, which is used to find the snow-ice interface elevation (via equation 5) and then ice freeboard. The updated manuscript will better describe some of these processes. In addition, we will add in a flowchart similar to the following:



Where ovals show input/output data (green is start, red is end, white are inputs, grey is discarded), diamonds show decision/filtering processes, white squares show calculations/other processes, and blue squares show 'milestones' (inset orange squares for lead-type milestones, blue for floe-type milestones).

In Section 4, authors compare snow depth and thickness with other datasets. why not snow freeboard included in comparison?

This is a fair question. As mentioned in the text, a comprehensive snow freeboard comparison was done using ICESat-2 data in Fons et al. (2021; *Earth and Space Science*). We point readers there for this comparison, but I understand that it still feels lacking in this work. The other reviewer had similar thoughts, so we plan to add in a snow freeboard comparison to Operation IceBridge data from a flight in the Weddell Sea in October 2010 to the revised version.

L179. About the statement 'Fit parameters are discarded if the result is a "poor fit"', how much data have been discarded and what kind of data are discarded?

Good question, this is info that should have been included. The number varies from orbit-to-orbit, but on average 86% of waveforms are kept (meaning 14% are filtered out). That number is higher for floe-type waveforms (21% filtered out). Only <1% of lead-type waveforms are filtered out. These numbers tell us that waveforms that are filtered out are typically "messy", meaning that they have multiple peaks (due to scattering from off-nadir leads) and/or don't fit the typical profile of a return floe-type waveform. Lead waveforms are typically specular and without many off-nadir peaks, and therefore are less likely to be filtered.

We will add these values to the revised manuscript.

The similar question for L213, for what types of waveform, there is no good fit? This is important because it may better inspect the proposed method.
Good question, with a similar answer to above. Poor-fitting waveforms typically have multiple peaks, usually brought on by off-nadir leads (Kurtz et al. 2013, Tilling et al. 2018). This is a common problem with many CryoSat-2 retracking methods, as the wide footprint is susceptible to returns from far off-nadir. We do understand that filtering these waveforms out could potentially bias the retrievals, especially for sea ice close to leads. However, quantifying this bias would be very difficult. Overall, we will add in some info towards this point in the revised version.

Minor:
L1." a physical waveform model and a waveform-fitting method to estimate the snow depth and snow freeboard" is misleading. What is the relationship between the physical waveform model and the waveform-fitting method and are they used for snow depth and snow freeboard respectively or as a whole? Are they the comprehensive method called CS4WFA? The sentence is ambiguous.
You are right that they are all part of a comprehensive method that we're calling CS2WFA for simplicity. We construct a model and use the waveform-fitting approach to estimate output parameters, which includes snow depth and snow freeboard. We can modify this in the revised version to something along the lines of: "We estimate the snow depth and snow freeboard of Antarctic sea ice using a comprehensive retrieval method (hereafter referred to as CS2WFA) that consists of a physical waveform model and a waveform-fitting process to fit the modeled waveform to CryoSat-2 data. "

L3. Is the thickness in "snow depth and thickness" snow or sea ice thickness?
It is referencing the sea ice thickness, and we will add this in for more clarification.

L9-10. Some findings are expected after the sentence. For example, "…, showing the interannual differences between the two kinds of satellites". Or add "Results show that" before "Reconciling…"
Fair point. We plan to revise the abstract in the updated manuscript based on both reviewers' comments, however, if this sentence remains it will change it to something like: "We place these thickness estimates in the context of a longer-term, snow-freeboard-derived, laser-radar sea ice thickness time series that began with ICESat and continues with ICESat-2, and show interannual differences between these satellites."

L80-82. There may be misstatements on those literatures:
The approach "Worby" in Kern 2016 is a static ratio between sea ice thickness and snow depth, which are seasonal empirical values from ASPeCt, which are ship-based observations.
Li 2018 is a dynamic ratio between snow depth and sea ice thickness, which are initial guess from empirical equations between the total freeboard and snow/ice thickness by Ozsoy-Cicek Burcu, et al; Sea ice thickness retrieval algorithms based on in situ surface elevation and thickness values for application to altimetry;Journal of Geophysical Research: Oceans;2013, 118 (8):3807-3822. Xu 2022 is an improved version of Li 2018 with a similar strategy.

Wang 2022 uses the same 'Worby' method, i.e., the static ratio, in Kern 2016 for ICESat, while uses snow depth from AMSR-E/AMSR-2 for Envisat-based sea ice thickness retrieval.
So, I suggest the sentences to be rewritten. For example:
"With some empirical parameters, sea ice thickness can be estimated with ICESat/ICESat-2 alone. Regarding sea ice as a single ice/snow layer, the Worby method (Kern et al., 2016) uses a static ice-snow ratio which are seasonal empirical values from the ASPeCt program (Worby et al., 2008). The one-layer method (OLM, Li et al., 2018) and its improved model (OLMi, Xu et al., 2021) are proposed with a dynamic ice-snow ratio for each footprint measurement based on an initial guess from the empirical relationship between snow depth/ice thickness and snow freeboard (Burcu, et al., 2013)."
Thanks for this suggestion. The idea was to categorize all of these based on their use of a snow depth/ice thickness ratio, but we do see how the way it is currently written simplifies the description of these studies. We also appreciate the suggestion, and will change this in the revised manuscript to something very similar:
"Other studies have used a ratio between the snow depth and sea ice thickness to estimate the ice thickness from laser altimetry. The Worby method (from Kern et al. 2016) used a static snow-ice ratio derived from seasonal empirical values from the ASPeCt program (Worby et al. 2008), while the one-layer method (OLM, Li et al., 2018) and its improved model (OLMi, Xu et al., 2021) use a dynamic snow-ice ratio for each footprint measurement based on an empirical relationship between snow depth/ice thickness and snow freeboard (Burcu, et al., 2013)."

P18. Figure 7. What caused the uncertainty difference between different sea sectors?
How is the pan-Antarctic uncertainty computed?
The difference in uncertainty in the different sectors comes from the difference in the snow freeboard, snow depth, and snow depth uncertainty between the sectors, since these are the variables in Equation 7 that vary spatially. The pan-Antarctic uncertainty is calculated simply as an average of the uncertainty of all grid cells basin-wide. This is a simplistic way to do this, but provides some estimate of where we expect the uncertainty to be. We plan to modify this figure in the revised manuscript (to answer the other reviewer's comments), but will also include a better mention of how the pan-Antarctic uncertainty is calculated.

L156. Add Ψ after "a physically-modeled waveform" for Eq.2
Good idea – this will be added.

L158. What is hsd in eq.4? If it is snow depth, should it be hs according to Eq.1 and Table 1?
That is correct, and will be changed in the revised version.

L187 Rn should be R0 here?
Yes, thanks for catching this. It should indeed be R0, and we will make that change in the revised version.

Title of Figure 2. For "the daily linear interpolation", to avoid misunderstanding that the dashed lines arederived with daily data, it may be better written as "the linearly interpolated dashed

lines between the midpoint dates are used as daily density estimates for the thickness calculation."
Thanks for the suggestion - that is a small but important distinction. We will make these changes in the caption of Figure 2.

L262, hfs should be hfs. Is there system uncertainty for snow freeboard in Eq.7?
Thanks for pointing out the formatting error – we will make that change.
Equation 7 is based on Petty et al. (2020) and Ricker et al. (2014), and is a simplistic way to provide some measure of uncertainty. It does not include an estimate of systematic uncertainty for freeboard, since it is not reliably known. While we acknowledge that an estimate of snow freeboard uncertainty would help to better constrain the total thickness uncertainty, the current equation is used until we have a better idea of the snow freeboard uncertainties. We will add in more information about the limitations of this uncertainty estimate in the revised version.

L290. Section 4.4 missed in this sentence.
Thanks – we will update this sentence to include a reference to section 4.4.

P15. Figure 4. Are the vertical dashed lines mean values? Can the mean, std, and model values in eachseason/month be put on the map? The same with Figure 5.
The position of figure 4 is better to be close to the text in Section 4.1. The position of Figure 8 can also be adjusted.
That is correct – the vertical dashed lines are the mean values, which is mentioned in the caption. We will add the requested values to the plots in both figure 4 and figure 5. I agree with the positioning, and we will do our best to keep figures near relevant text. However, this is something that will change in the copyediting process and is somewhat beyond our control.

L366. The exception is not only Ross Sea. Am-Bel is 4 cm thicker in spring than summer.
Thanks for pointing this out – we will update the text to this point. The other reviewer also had comments on this figure, so this section – including the sentence you mentioned – will be modified in the revised manuscript.

L384. The last word "that" is 'than'?
This will be changed.

L395. "thinner that" to "thinner than"
We will make this change in the revised manuscript. Thanks for catching these!

P21. Figure 9. Use different colors to differentiate decreasing and increasing regions instead of the uniform grey?
Good idea – these were meant to just signal statistically significant trends, but I like the suggestion of coloring by their region-average trend. We will add in these colorings in the revised manuscript provided that the figure remains in the revised version.

L476. About the discussion of the radar-laser difference in this section, another possible reason may be the high spatial resolution of ICESat/ICESat-2 than the grid/segment-averaged coarse resolution CryoSat-2. The lower resolution may smooth higher or lower signals values.
Yes, this is a valid point and was not adequately mentioned in the original version. We will add in some information regarding the resolution differences between the two satellites.
(Additionally, this difference was explored further in a previous work - Fons et al.2021; Earth and Space Science - so we will also add in a reference to that work as well.)

P25. The legend in Figure 10 is "Sea ice thickness derived with CryoSat-2 snow depth" instead of "CryoSat2 snow depth"?
This will be changed – thanks.

L701-703. The citation of Meredith 2019 can be improved according to their suggestions: https://www.cambridge.org/core/books/ocean-and-cryosphere-in-a-changing-climate/polarregions/8D76B8865B796C16991F7A9FB6271C2D.
Thanks for pointing this out. As they recommended, the new citation will be changed to: (Intergovernmental Panel on Climate Change (IPCC), 2022).