Reviewer 1, Anonymous Referee

Review of “Sensitivity of CryoSat-2 Arctic sea-ice volume trends on radar waveform interpretation” by Ricker et al.

This study concerns the use of new CryoSat-2 radar altimetry data to calculate trends in Arctic sea ice thickness and volume, and their associated uncertainties. The authors assess a range of the leading-edge retracker thresholds as applied to CryoSat-2 radar waveforms and investigate the resulting range in both sea ice thickness, and volume, that arise with respect to a particular threshold setting. The goal of the study is to “isolate and quantify the effect of SAR waveform interpretation from other uncertainties in the freeboard to thickness conversion”, and is the first study to directly address this issue. The authors find that while the absolute magnitude of the CryoSat-2 thickness estimates is impacted by the choice of threshold setting, the overall trends in sea ice volume over the CryoSat-2 measurement period remain consistent, regardless of the retracker threshold setting. The authors conclude that the seasonal evolution of snow on sea ice should be considered when choosing an appropriate retracker threshold.

I recommend this paper for publication after the authors address the following comments, as well as the more detailed minor revisions listed below.
Some of the manuscript’s authors have provided a public release of CryoSat-2 sea ice thickness products through the Alfred Wegener Institute “meereisportal”. However the online dataset is not explicitly mentioned in this manuscript. Can the authors clarify if any/one of the algorithms assessed here is used in the generation of the online sea ice thickness product? Many other investigators are currently using the publicly available datasets such that it would be extremely helpful to be able to point to this (or another?) published paper to reference the online data, and to understand its accuracy and limitations. If methodologies differ significantly, that might also be addressed here.

We thank the reviewer for his comments. We added a citation of the “meereisportal” where we directly link to the products that are provided there. We basically used the same processing chain, but recently implemented some improvements. We will therefore update the data that they coincide with the results of this paper.

Can the authors describe in more detail the choice of waveform fitting routine (grey line in Fig. 3) utilized for the lead waveforms? The authors state that it is a “result of linear interpolation and smoothing”. Please provide more sufficient details on this approach. The waveform fit (or ‘model’, e.g. grey line, Fig 3b) appears to be at least a full range-bin wider (a half range bin on the leading- and trailing-edges) than the waveform itself. This misfitting will cause a bias in the estimated lead/sea surface height elevation. I suspect if the lead waveform fitting routine were to work more efficiently, the difference between the elevation tracking pointing associated with each of the TFMRA thresholds would be very close, and differences would be negligible.

There has been a bug in the plotting routine for Fig. 3 regarding the normalization. This is fixed now and shows that the interpolated and smoothed waveform matches the original waveform at the leading edge. We also added some more explanation to the text.

P1839 L4-5: Can the authors explain how the open water/lead elevations are interpolated – i.e. what interpolation scheme is used, and what is the result of the low-pass 25 km filter? Does the routine result in one SSA estimate per 25 km grid cell or results sampled at a higher resolution along-track?

We retrieve the SSA track-wise from linear interpolation between the detected leads. When we observe a dense cluster of leads, we obtain significant noise on the SSA doing the linear interpolation (due to the speckle noise that is affecting each measurement). Therefore we apply a running mean with 25 km width to reduce the noise on the SSA. We also added some more explanation to the text.

The thickness results over multi-year ice are confounding, particularly those results presented in Figure 8 and Table 3. Why is MYI sea ice thickness thicker for the TFMRA50 threshold than for the TFMRA40 threshold, particularly in March 2013? This is counterintuitive based on the waveform examples provided in Fig. 3 and the threshold re-tracking points on the leading edge.Moreover, the title rows in Table 3 are confusing, e.g. the second row of "FI(m)" seems redundant and misaligned with the results. Please revise the layout of the table so that the results are more clearly presented.

Only for the TFMRA40 retrieval we applied a penetration correction term to the freeboard, since from the comparison with the laser altimetry data we observed that for the 40 % threshold, the radar is not penetrating the snow completely. In contrast we assumed a full penetration of the snow for the TFMRA50 and TFMRA80 retracker in consistence with Laxon et al. (2013) and Kurtz et al. (2014). Assuming only partial penetration leads to a lower ice thickness compared to the assumption that the freeboard represents the ice freeboard. We acknowledge that these assumptions were confusing for the reader. Therefore we now consider the freeboard without any corrections and also compare the freeboard retrievals of the 50 and 80 % threshold with the laser altimetry data.

P1842 L8-9: What is the "ICESat/GLAS" surface-type mask? Why is it applied to the CryoSat-2 data in this study?

We used this mask to exclude the Canadian Archipelago but additionally excluded the Baffin bay and now also the area between Svalbard and Sewernaja Semlja, since there are only few or no measurements that are considered for the W99 fit (see Warren et al. (1999)). Only snow depth data from regions where the W99 fit bases on measurements should be used.

P1842 L13-14: What motivated the choice of a boxcar average to interpolate data across the 2 degree polar hole in the CryoSat-2 data? I don't understand the result of a boxcar interpolation scheme in this case. Is the approach actually a linear interpolation, since you only have one data point across either side of the polar hole for a given CryoSat orbit? Have you checked the results against other interpolation schemes – e.g. a weighted average?

We former used a linear interpolation to close the gap at the pole. We then used a boxcar average to smooth the interpolated area. This is now obsolete since we excluded the sea-ice volume considerations to focus more on the difference of the freeboard retrievals from different thresholds.

P1846 L3-5: the authors state that the ALS system provides "high-precision and high-resolution measurements", but then go on to state that "the accuracy for the range measurements is about a few cm". The specific details of the lidar system – elevation accuracy, precision, and measurement resolution, should all be included here so that the CryoSat-2 elevation measurements can be placed in the context of lidar data.

The precision of the laser measurement itself is very high indeed (in the range of mm). But the positioning of the aircraft that is needed for data processing introduces some uncertainties that increase with the flight distance. More details can be found in *Airborne lidar measurements for Cryosat validation* by Forseberg et al. (2002) and in the ARCGICE final reprot *Combination of Spaceborne, Airborne and In-Situ Gravity Measurements in Support of Arctic Sea Ice Thickness Mapping* by Forseberg et al. (2007).

Figures 6 and 7 appear to suggest that the long repeat period of the CryoSat-2 orbit, with non-uniform monthly subsampling of the Arctic basin, partially contributes to the radar freeboard uncertainty term. An assessment of radar freeboard uncertainty indicates a peculiar pattern wherein uncertainties align along lines of longitude (e.g. Fig 6b), a result that is not easy to explain by the physical nature of the sea ice pack. The fraction of detected leads also indicates a similar pattern. Although the authors describe these patterns in the conclusion section, a re-evaluation of your gridding routine/25 km grid cell size to account for the non-uniform monthly sub-sampling of the Arctic Ocean due to the CryoSat-2 orbit appears necessary. Since the goal of the study is to "isolate and quantify the effect of SAR waveform interpretation from other uncertainties in the freeboard to thickness conversion" the magnitude of uncertainty associated with orbit subsampling should be addressed with respect to the uncertainties arising from retracking thresholds, and ideally should be removed before considering the volume trends.

The orbit subsampling and the procedure of gridding is definitely an item for future investigations, but not in the scope of this paper. Furthermore this is rather a question of how to interpret the monthly average. These uncertainties do not result directly from the CS-2 measurements or the freeboard-to-thickness conversion.

Section 5: How does the resolution of the ASCAT backscatter map relate to the ice- type mask utilized in this analysis? Figure 6 would indicate that the higher backscatter associated with MYI features are not accurately represented by the "MYI mask", especially over the specific features that the authors identify. How do the authors justify the use of the coarser resolution mask rather than ASCAT backscatter to delineate ice type? Indeed, a comparison of the FYI and MYI ice thickness results using both methods would be insightful.

We use the OSI SAF ice-type product for discrimination between FYI and MYI. It also uses the ASCAT backscatter as input data. The "MYI mask" might have been confusing. We actually interpolate the OSI SAF ice type product on each CS-2 track. When we grid the data, we also grid the ice type for each CS-2 measurement and obtain a monthly mean ice type, related to the CS-2 product. The term "mask" might be confusing here since we only use the gridded ice type to visualize the ice type on the maps. However, we observe that the shape of the CryoSat-2 waveforms also indicates the ice type and we might use this information in the future for data processing.

Minor revisions:
Re-check English usage and grammar throughout P1832 L2: "last decades" – clarify which decades P1832 L13: change to "on the order of"
P1833 L8: change "multi-seasonal" to "multi-year"
P1834 L1: What is the reference for "large footprint of the order of 10km and an orbit coverage limited to 82.5 oN"? Check this statement. The literature suggests a surface footprint size approximately 2–10 km in diameter on sea ice (e.g. Connor et al., 2009 and references therein). ERS-1, -2 and Envisat have coverage to a latitudinal limit of 81.5 oN.
P1834 L25: Change to "Laxon et al. (2013) used a leading-edge threshold retracker, while in a recent study. . ."
P1837 L4: change "exemplary" to "example" or "typical" P1838 L13: Change "refused" to "discarded" or "removed"

P1839 L18: change "inside the snow layer" to "through the snow layer"
P1841 L6: include the word "modified" before "Warren snow climatology"
P1841 L21: Insert "Consistent with the approach of Laxon et al. (2013)" before "we use ice densities of . . ."
Table 2: Rather than the term "variable", the range of values considered for each parameter would be much more helpful. Please revise.

The minor revisions have been considered and were incorporated into the manuscript.