Comment on tc-2021-227
Anonymous Referee #3

Referee comment on "A comparison between Envisat and ICESat sea ice thickness in the Southern Ocean" by Jinfei Wang et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2021-227-RC3, 2021

*General comments*

This manuscript presents an intercomparison of sea ice thickness derived from Envisat and ICESat satellite altimetry for the Southern Ocean, compared with sea ice thickness derived from draft observations from upward looking sonars (ULS) in the Weddell Sea. The comparison suggests a better agreement between ICESat and ULS sea ice thickness and differences between ICESat and Envisat, including differences in resolving the Ross Ice Shelf polynya, thicker ice in summer from Envisat than from ICESat, and differences in SIT variations from autumn to spring. The results are compared and discussed with the help of accumulative freezing-degree-days (FDD).

The intercomparison of the ESA SICCI sea ice thickness products with independent sea ice thickness observations is very useful and contributes to our understading of the uncertainties and limitations of the datasets. Recently more focus has been on intercomparing the more recent CryoSat-2 part of this dataset, so this intercomparison of the Envisat period is particularly welcome.

I would judge this manuscript good on the criteria for originality and significance. However, improvements can be made in the categories for scientific and presentation quality. I have some doubts on the methods used for retrieving sea ice thickness from freeboard observations. Why is a different threshold sea ice concentration used for both altimeters, and why is the updated method by Li et al. (2018) for ICESat not used? I also have doubts about comparing and discussing the differences due to the sensors and the snow depth by just looking at the sensitivity of the hydrostatic balance equation. It would be good to see an analysis of what the actual effect of using different snow depth products would be.

The manuscript would also improve from a thorough read-through and rewrite of sentences that are now confusing or even causing misunderstanding. I have laid out my main comments and suggested technical corrections below.

*Specific comments*
GC1: Why not use same snow product and methods for ICESat and Envisat? Now the discussion of the differences due to sensors and snow depth only include a sensitivity of the hydrostatic equilibrium to snow depth/freeboard, but not the actual effect. It would be an option to calculate sea ice thickness from ICESat with the AMSR-E snow depths as well so you can compare what part of the difference is a direct effect from the difference in sensors and what is caused by the difference in snow depth. I understand that this involves quite some more work, but I think a the statement that is now made in the summary (L406-408) is a bit strong for the amount of proof you have for this, as you've not made the actual comparison.

GC2: Why is a different sea ice concentration threshold used for ICESat (60%, L143) than for Envisat (70%, L123)?

GC3: In lines 381-386 you introduce an improvement of the method you have used to obtain ICESat sea ice thickness. What is the reason for not using this improved method?

GC4: L229: 'an overall comparison between Envisat and ICESat effective SIT'. In the methods it said the Envisat SIT product 'represents the actual SIT (i.e., mean thickness of the ice-covered fraction of the grid cell area)' (L122-123) and that the ICESat SIT product is the 'effective sea ice thickness (i.e., mean thickness per grid cell including open water areas)' (L141-142). Are these two products compared here? This would not be a fair comparison, as the effective sea ice thickness is by definition going to be thinner than the actual sea ice thickness. If you are comparing actual sea ice thickness products please clarify here and in the methods section.

GC5: What values or products have been used for water, snow, and ice density in the calculations of sea ice thickness with Eq 1, Eq 2, and Eq 3? Are they the same for Envisat and ICESat? If not, discuss the effect on the results.

GC6: ULS and satellite altimetry SIT distributions would be interesting to see as well, if possible.

GC7: There are some significant issues with interpreting sentences throughout the manuscript. I have added some key examples below, but the clarity of the manuscript could improve from a thorough read-through.

L77-78: 'Several freeboard- ... compared (Kern et al., 2016).’ This sentence feels unrelated to the rest of the paragraph and is therefore confusing. If you want to go into this you need to explain the different retrieval algorithms. But I think it’s better to leave this sentence out of the introduction and leave this to the methods (as you’ve explained this more clearly in section 2.2).

L148-149: 'The signals ... travel time.' This sentence makes it sound like travel time is used to differentiate observations of sea ice bottom vs. sea surface, but I think you are
trying to say that the distance is determined from the travel time measurement. It also
sounds like only two measurements are made, one from the sea ice bottom and one from
the sea surface. Please rewrite this.
L254: 'ICESat is more sensitive to thick ice than Envisat', but the Envisat SIT product is
thicker than ICESat? You describe this bias well in section 4.1, but here it's a bit
confusing, as you seem to say that ICESat should show thicker ice.
L262: 'Envisat has a positive difference with respect to ICESat'. I do not understand what
this means.
Have a look at the suggestions for technical corrections too.

*Technical corrections*

L24: 'while the uncertainties of *the* snow depth product are' or 'while the uncertainties
of snow depth product*s* are'

L30: 'it is still unclear if ... sea ice thickness'. Change 'also associated with' to
'accompanied by', these changes do not have to be related (or associated) but can be
separate.

L39-40: 'from the ASPeCt can provide' change to 'by the ASPeCt expert group can provide'

L42: 'airborne electromagnetic data which measure total freeboard', data don't measure
things, maybe rephrase.

L52: Remove 'basically', this sounds very unscientific.

L54-55: Consider more recent studies that have retrieved Antarctic sea ice thickness, e.g.

L83-84: 'also how the different ... distribution.' Very vague, what are 'the different
retrieval methods', ICESat and Envisat?

L88: 'the former inter-comparison study', which study is this?
L147: change 'underwater' to 'below sea level'

L150-151: 'once several minutes', do you mean 'every several minutes'? Please rewrite and maybe be more specific (what is several minutes)?

L153: seasons -> season

L166: 'FM-MJ and MJ-ON'. I guess you are referring to February/March-May/June and May/June-October/November. Please specify the first time you mention these abbreviations.

L182: 'Before ... first.' Repetitive, just use 'before' or 'first'.

L193: Remove 'during the comparison'.

L197: 'We provide ... SIT products.' Rewrite this sentence. I would suggest something like 'The error bars in the figure show the uncertainty estimates of/from the SIT products'.

L197-200: 'The Envisat SIT ... Li et al., 2018).' Move these sentences to the methods? Also: I think adding an estimate of the ULS uncertainty to Figure 3 as well would improve the interpretation of this figure. You mentioned an estimate of the ULS uncertainty in L152-154. You now mention when the error bars of the altimetry sensors do not overlap with the ULS points, but it would be interesting to see if they do overlap with the ULS error bars.

L207: Why are the uncertainties of freeboard and snow depth not considered for the ICESat SIT uncertainties?

L208-209: 'ICESat does not capture ... on thicker ice.' I'm not sure where I can see this in Figure 3?

L210: 'error bars can cover' -> remove 'can'

L210: 'However, since many contributions are not well characterized and quantified'. What
contributions is this about and how are they not well characterized and quantified?

L225-226: 'considering the typical sea ice motion'. Briefly characterize this typical sea ice motion (fast, direction?), so the reader can see why the monthly average ULS SIT can be referred to as a spatial average value.

L235: What are 'the ship-based observations'? This is not introduced in the paper before.

L237: change 'feature' to 'dissimilarity' or another more descriptive word.

L249: 'but with thickness estimates of up to 1.5 m'. Make sure it is clear to the reader that this is thinner than elsewhere.

L264: 'the two datasets coincide with each other', this sounds a bit like they temporally coincide instead of the distributions being similar (which is I think what you want to say here). Please rewrite.

L269-270: 'We calculate the period-average SIT from the model'. This might be my lack of experience with freezing-degree-days: the FDD in Figure 8 and Table 6 show the total negative temperatures between these months right? I do not understand how it shows SIT. I understand that FDD and SIT are related but I don't see how the model actually calculates average SIT? Please make this more clear in the methods. If 'the model' is not FDD, maybe specify what model you mean?

L271: 'Envisat SIT has opposite developments from ICESat and FDD during MJ-ON'. Envisat and ICESat do not really show the opposite? They both show the strongest thinning in the western Weddell Sea and both show thickening near the coast in the Amundsen Sea. Please rewrite this to describe the difference, I think something like that Envisat shows more thinning all around the Southern Ocean and ICESat generally more thickening?

L271, Figure 8, and Table 6: Please be consistent in how you refer to these periods (MJ-ON or ON-MJ and FM-MJ or MJ-FM). I would suggest for summer to autumn you use FM-MJ (instead of the subtraction MJ-FM you used in Figure 8) as this order is more intuitive.

L272: 'both products', which two products? Envisat and ICESat, or satellite altimetry and FDD?
L274: 'The adverse patterns', adverse (preventing success or development; harmful; unfavourable) might not be the right word here?

L279-280: 'the regression lines have large positive intercepts in all three seasons, indicating that Envisat SIT tends to exceed ICESat SIT for thin ice'. I can see in Figure 9 that this is true, but the latter does not necessarily follow from the former. A large positive intercept could also be caused by Envisat SIT being lower than ICESat SIT for thick ice. Again, in the figure I can see this is not the case here, but maybe just rephrase the explanation to just say 'For all five locations, Envisat SIT tends to exceed ICESat SIT for thin ice', without referring to the intercept.

L281: change 'splashes' to 'cloud' which is more often used to describe a collection of points in scatterplots. 'Exceed' in what way? Envisat or ICESat or both?

L284-285: 'The numbers in the last ... values per season'. This might be something to replace to the caption of the table. Also, in the table it does not look like this is in the last column, but in the first row?

L294-295: 'it is known that ... homogenous stratigraphy'. This statement could use a citation.

L296: 'considering the large ... of about 70 m'. Maybe specify that the pulse-limited footprint is Envisat and the laser beams ICESat.

L341: maybe just say 'may come from the AMSR-E snow depth' here as you haven't yet discussed why it might be biased.

L347-348: 'the differences that AMSR-E snow depths minus the ASPeCt observations are positive ...', rephrase this sentence to something like 'AMSR-E snow depth minus the ASPeCt observations is positive'

L349: 'the satellite passive microwave snow depth'. Maybe introduce AMSR-E as a passive microwave sensor in the methods, so readers that don't know the AMSR-E snow depth product know what you are referring to here.

L351: '... lead to underestimations' and '... lead to overestimations', under- and overestimations of what? SIT?
L357: The retrieval uncertainty of AMSR-E?

L357-358: 'suggesting that sea ice thickness change is insensitive to the snow depth', I would suggest change to 'the sensitivity is low', as SIT does change with snow depth, just not by a lot.

L363-364: 'The usage of snow climatology allows reducing the relative uncertainties', it's a bit unclear what these 'relative uncertainties' are and how they are reduced.

L389: Remove 'firstly'

L392: change 'not comparable to' to 'overestimating' or something else more descriptive of the difference between the two.