Dear Editor,

We truly appreciate your and four anonymous reviewers' valuable comments and suggestions for the paper 'A comparison between Envisat and ICESat sea ice thickness in the Southern Ocean' submitted to *The Cryosphere*. We have already made a considerable revision according to these comments and suggestions, and reply to them one by one below.

Qinghua Yang and Qian Shi
On behalf of all the authors

Responses to referee #1

Dear Reviewer:

We would like to express our gratitude to you for the helpful comments to improve this manuscript. For GC1 and GC2, we modify the descriptions carefully to avoid misunderstandings following your specific comments. For GC3, we add this "total thickness" information in the data description and discussion of ICESat SIT uncertainty. For GC4, we conduct further analyses towards the FDD results combining sea ice drift information. For GC5, we switch the descriptions to active mode as you suggest.
The specific responses and revisions are shown below. They are in blue font for clarity.

Qinghua Yang, Qian Shi, Robert Ricker, Stefan Hendricks
On behalf of all the authors

Specific Comments: (I abbreviate Line with L)
Title: Since sea ice is an integral part of the Southern Ocean I suggest to use "Southern Ocean" instead of "Antarctic" ... perhaps even throughout the entire paper.
Thanks for your comments. We replaced some of the "Antarctic" with "Southern Ocean", especially in the cases like "in the Antarctic".

L50: At this point I suggest to provide a summary sentence which states that all these various data sets - despite covering limited regions and/or time periods - are extremely useful for the evaluation of models and satellite retrieval methods. I suggest to also differentiate between data sets that provide sea-ice thickness information at one fixed location (ULS) and hence allow to check the consistency over time, and data sets which have a short duration but with high resolution cover comparably large regions (e.g. Operation ice bridge or AEM) and hence allow to check the spatial variability of the sea-ice thickness retrieved from satellite data.

(1) We added a summary sentence: "Despite covering limited regions and/or time periods, all these various observational data sets are extremely useful for the evaluation of models and satellite retrieval methods." *(please see P2 line 52-53 in the revised manuscript)*
(2) We changed the expression to differentiate the data sets: "**One type of observation data** is in situ

measurements providing sea ice thickness information at one fixed location and some allow to check the consistency over time. For example, drilling data (e.g., Meiners et al., 2012) are accurate but extremely limited in temporal and spatial coverage, and hence they cannot be used to obtain an understanding of large-scale Antarctic sea ice thickness processes. Upward-looking sonars (ULS), located at 13 different sites in the Weddell Sea, provide valuable temporal evolution of sea ice draft (Harms et al., 2001; Behrendt et al., 2013a; Behrendt et al., 2013b), but a basin-wide spatial distribution cannot be derived. **The other type of data sets** has a short duration but with high resolution cover comparably large regions and hence allow to check the spatial variability of the sea-ice thickness retrieved from satellite data. ......" *(please see P2 line 36-43 in the revised manuscript)*

L50-56: I suggest to reorganize this information a bit. First of all Kurtz and Markus 2012 and Li et al. 2018 utilize laster altimetry and hence fall into what you describe in the last sentence of the lines referred to here; this should somehow be merged. Secondly, Bernstein et al. is a paper about trying to estimate sea-ice thickness in the Ross/Amundsen Sea only from a very limited set of sea-ice charts. This data does not have the same value as the data sets of the other two papers cited in the same sentence.

Thanks for your comments. We removed the references for the first sentence: "More recently, satellite remote sensing has been widely applied to investigate the spatial coverage and long-term trend of sea ice thickness in the whole Southern Ocean." Meanwhile, we cited "Kurtz and Markus, 2012" after the sentence "Satellite altimetry, including radar and laser altimetry, have been used in the Southern Ocean to retrieve sea ice thickness". *(please see P2 line 57 in the revised manuscript)*

L63/64: While I am totally fine with the sentence that snow affects radar altimetry SIT retrievals in two ways, you should first tell the reader the two ways before you come up with details of the shortcoming. First i) snow depth is required to a) correct the radar wave speed in snow and hence to appropriately convert the radar freeboard into the seaice freeboard and to b) convert sea-ice freeboard into sea-ice thickness. In both cases, but mostly in b) also the snow density plays a role. Secondly ii) the presence of snow simply modifies how the radar signal is reflected in / by the ice-snow system; the assumption of Beaven et al. is for DRY snow only. Hence, in addition to the more physical/mathematical influence of snow depth, there is this potential violation of the full-penetration assumption made by Beaven et al as is demonstrated by Willatt et al. These issues need to be specified first before you can come up with the details in Lines 65+

Thanks for your comments. We modified the description as you suggested: "Firstly, snow depth is required to correct the radar wave speed in snow and hence to appropriately convert the radar freeboard into the sea ice freeboard, as well as to convert sea ice freeboard into sea ice thickness. Secondly, the presence of snow modifies how the radar signal is reflected by the ice-snow system. Specifically, over Antarctic sea ice, the complex snow stratigraphy and frequent snow flooding associated with the formation of snow ice and superimposed ice affect radar altimetry measurements (Willatt et al., 2010), i.e. the assumption of Beaven et al. (1995) is for DRY snow only. Besides, the snow depth climatology used in the retrieval of Envisat and CryoSat-2 SIT can cause additional uncertainties due to neglecting inter-annual variability in snow depth (Bunzel et al., 2018)." *(please see P3 line 65-72 in the revised manuscript)*

L81/82: Here you please need to check recent literature because Kwok and Kacimi or Kacimi and Kwok came up with more VERY useful work based on ICESat-2 data. You should include these references here as well - and ideally already point to the fact that the coverage with ICESat-2 is much better than with ICESat.

We modified this sentence: "However, ICESat-2, which has been in orbit since 2018, provides a new source of year-round observations of total freeboard and thus better coverage than ICESat (Kwok et al., 2019; Kacimi and Kwok, 2020)." *(please see P3 line 83-85 in the revised manuscript)*

L106/107: If I am not mistaken, then the Paul et al reference point to some data analysis and algorithm development but is not specifically the reference to cite the sensor properties of Envisat RA-2. Please find a more appropriate reference which also details the footprint issue. I doubt that also Connor et al. 2009 is the adequate reference here. I am sure that are papers from the early 2000s when the altimeter was just up or about to be launched in which the system specifications are laid out well.

We removed the citation of "Paul et al., 2017" and added the citation of "Peacock and Laxon, 2004". *(please see P4 line 108 in the revised manuscript)*

L115: It might make sense to add that Laxon et al. applied this method to ERS altimeter data, i.e. the predecessor of the Envisat RA-2 instrument.

We added the information: "Sea ice thickness is retrieved from ice freeboard based on the hydrostatic equilibrium approach as first used by Laxon et al. (2003), which applying this method to ERS altimeter data, i.e., the predecessor of the Envisat RA-2 instrument." *(please see P4 line 113-115 in the revised manuscript)*

L120/121: "revised version ... Cavalieri et al (2014)" I recommend to not refer to a data set description here but refer to the main core paper of the apporach used which is the one by Markus and Cavalieri, 1998, and then it is the Comiso et al (2003) reference which points to the AMSRE sea ice processing. I suggest to make clear what the "revision" is (different tie point retrieval plus addition of retrieval errors). It would also be good if you could tell the reader on data of which years the snow depth climatology is based - because it extends well into the AMSR2 period. Finally, you may please change the URL into https://www.cen.uni-hamburg.de/icdc .

We checked the snow depth climatology data and found it only covered AMSR-E period. Therefore, we removed the AMSR2 information and modified this part as you suggested: "A snow depth climatology **(based on 2002-2011)** is employed to retrieve sea ice thickness from sea ice freeboard here (**Markus and Cavalieri, 1998; Comiso et al., 2003**). This snow-depth climatology is derived from the passive microwave sensor Advanced Microwave Scanning Radiometer-EOS (AMSR-E) for the Antarctic and is based on a revised version of the approach **with different tie point retrieval plus addition of retrieval errors** and provided by the Integrated Climate Data Center (ICDC)." *(please see P4-5 line 118-122 in the revised manuscript)*

L122/123: "the actual SIT (... mean thickness ... of the grid cell area)" --> this does not go together well. The actual SIT would be the thickness of the ice floes as they float around in the grid cell. The mean SIT takes into account that the grid cell might not be fully covered by sea ice. Hence the actual SIT is always larger or equal than the mean SIT and it is important that you write this down in a clear way.

We corrected the explanation of actual SIT: "ice thickness of the ice-covered fraction of the grid cell area". *(please see P5 line 123-124 in the revised manuscript)*

L138-140 / Eq. 3: I guess it is important to check this equation and the wording. If I am not mistaken, then the authors of these data claim on the respective web page that it is actually not the sea-ice thickness that is retrieved with this equation but it is the total (sea ice plus snow) thickness. Hence it is in a way the same type of thickness as is observed by that famous airborne EM sensor (see your introduction). In order to obtain the sea ice thickness from I retrieved using (3) one should possibly substract the snow depth and/or reformulate equation (3) such that this effect is somehow included.
We checked the equation and corrected the mistake: "And total thickness (sea ice thickness plus snow depth) can be determined from it." *(please see P5 line 139 in the revised manuscript)*

L141: Please check whether this product contains the mean gridded sea-ice freeboard or whether this is perhaps in fact the total (sea ice + snow) freeboard.
We corrected the mistake by changing "sea ice freeboard" to "total freeboard".

L147: "at more than 900 m underwater" --> I don't think that this is a relevant information because the actual sensor is mounted further up anyways - otherwise the comparably small footprint would not be possible to achieve and the footprint would possibly also change between ULS sensor locations.
Thanks for your comments. We agreed and removed this information.

L166: When I look at Fig. 8 I have difficulties to fully understand what you did. First of all, the annotation in the Figure is opposite to what you write here. Secondly, what are the start and end days for the FDD computation using, e.g. the period from FM to MJ? The same question for MJ to ON. I find it strange and not easy to understand that you kept the FDD in degrees C and did not attempt to translate this into a net ice thickness growth. With that it remains a very qualitative comparison.

(1) We added the explanation for the annotation: "We compare the dynamic FDD with the SIT variations from February/March to May/June (FMMJ) and from May/June-October/November (MJON) represented by Envisat and ICESat SIT. Specifically, FMMJ represents the differences that mean FDD/Envisat SIT/ICESat SIT in MJ minus that in FM consistent with ICESat operating periods and so does MJON." *(please see P6 line 170-172 in the revised manuscript)* For example, during the period from FM to MJ in 2004, the FDD computation is conducted by subtracting mean summer SIT (from Feb 17 to Mar 20) from mean autumn SIT (from May 18 to Jun 20).
(2) We don't translate the FDD into thickness growth because we think FDD is a robust measure for potential thermodynamic ice growth (or melt). Converting FDD into thickness would require a model and additional assumptions with uncertainties that we cannot constrain sufficiently. Therefore, we think it is reasonable to only consider FDD. We add this information in the manuscript: "We use FDD rather than converted SIT with an empirical equation because they represent the same mechanism and we cannot constrain the uncertainties sufficiently caused by additional assumptions." *(please see P10 line 276-278 in the revised manuscript)*

L167: "neglects ice growth from snowfall, freezing rain or ridging" --> I suggest to be more specific with your formulation. "snowfall" per se does not lead to ice growth. It requires the process of flooding. "freezing rain" does not trigger ice growth - at least not to my knowledge. While melting of ice crystals requires energy, formation of ice from undercooled water releases energy; hence freezing rain, although contributing millimeters of ice - mostly on top of snow - warms the snow / ice. Finally ridging is no form of ice growth. It causes dynamic thickening of the ice using ice which is already there.

Thanks for your comments. We changed "ice growth" to "SIT changes" to avoid the misunderstanding. *(please see P6 line 166 in the revised manuscript)*

L174-177: While this is possibly a good approach it leads to the observed partly considerably larger coverage with Envisat SIT data in Figs. 4 to 6, particularly Fig. 5, which at first glance is a bit puzzling. It is of course not relevant for the comparison as long as you only consider grid cells where both, Envisat and ICESat provide values. But as shown it implies that Envisat, e.g., has much more ice in summer 2005 (Ross Sea) or 2007 (several regions) but this is just because your Envisat SIT map shows data of the entire month, e.g. April, into which an ICESat period overlaps. You could include a comment about this into your text or, alternatively, only show Envisat SIT values where both satellites provide a SIT estimate.
We considered it necessary to show more Envisat SIT data and we chose to add a discussion sentence to point out the problem: "It is noted that this approach can lead to considerably larger coverage with Envisat SIT data than ICESat, e.g., summer 2005 (Ross Sea) or 2007 (several regions) in Fig. 5." *(please see P7 line 181-182 in the revised manuscript)*

L186-189: What is the motivation to use these sea-ice concentration data which I assume are based on the ASI algorithm? If you keep this product please make sure that you refer to the algorithm name and to also provide information about the native spatial resolution of this product (which is much finer than 100 km). It might also make sense to provide the URL to the data set web page at ICDC if there is any.
Since this product is contained in the ICESat SIT data, we use it to convert the mean gridded SIT to actual SIT. We modified the introduction as you suggested: "The sea ice concentration data are derived from Special Sensor Microwave/Imager (SSM/I) and Special Sensor Microwave/Imager Sounder (SSM/IS) **based on ASI algorithm** provided by ICDC (Kaleschke et al., 2001; https://www.cen.uni-hamburg.de/en/icdc/data/cryosphere/seaiceconcentration-asi-ssmi.html) **with 12.5 km spatial resolution**, interpolated to 100 km grid NSIDC polar-stereographic grid and averaged over respective ICESat measurement periods." *(please see P7 line 190-194 in the revised manuscript)*

L207: The statement about the SIT uncertainties in the Worby 1-layer SIT data set is potentially not correct. I checked the data set and found uncertainties for both freeboard and thickness. Reading the paper Kern et al. 2016 it seems relatively clear that their computation of the SIT uncertainty included in the product is similar to their SICCI-2 SIT product from ICESat and hence based on uncertainties in densities and freeboard; only and here you are correct - snow depth uncertainty is not included. You might want to rephrase you text accordingly. Also, if I am not mistaken, then the uncertainty estimates provided with the Envisat SIT data set are possibly too large because the data set producers those days did not adequately take potential correlations between the error contribution into account. I am quite sure that, for instance, for the currently available (from AWI) CS-2 sea-ice thickness data the uncertainty is considerably smaller than for the SICCI-2 project data set and I am sure the same applies to the Envisat RA-2 data set. But you have the producers among your co-authors. So you simply need to ask!

(1) We amended the sentence about the SIT uncertainties: "ICESat SIT uncertainties are also calculated based on the uncertainties of densities and freeboard (Kern et al., 2016)." *(please see P8 line 203-204 in the revised manuscript)*
(2) We rewrote the potential reasons: "The differences in the error bars between Envisat and ICESat mainly result from the inclusion of snow depth uncertainty and lack of adequate regard for potential correlations between the error contribution." *(please see P8 line 210-212 in the revised manuscript)*

L215/216: I suggest to differentiate a bit better here between ICESat and Envisat - because Envisat provides a larger data set and hence your comparison is based on more data pairs. While not possible for ICESat it would be possible for Envisat SIT to come up with a statement about the agreemen of the seasonal cycle. Do ULS and satellite data sets provide the same seasonal cycle qualitatively?

We modified the summary sentence of the comparisons with ULS: “However, the numbers of valid data are too small to derive a reliable conclusion on the accuracy of ICESat. The comparison is based on more data pairs for Envisat, but the agreement of the seasonal cycle between ULS and satellite data sets is bad qualitatively (Fig. 3).” *(please see P8 line 216-218 in the revised manuscript)*

L221/222: "one satellite SIT grid cell is scanned only one of twice through a month" --> Please make sure to be more specific here. Not all these grid cells are covered only one / twice a month. Also this is valid for ICESat but possibly not for Envisat.

We modified this sentence: “one ICESat SIT grid cell is scanned once or twice on average through a measurement period.” *(please see P8 line 223-224 in the revised manuscript)*

L225-227: "However ... fixed ULS positions." --> While I agree that thanks to the ice motion and the integration period used the ULS point measurement kind of gains a larger representativity, it might be worthwhile to check i) how large the ice drift actually was and what their average direction was. You could use the NSIDC V4.1 sea-ice motion data set to figure this out.

As we know, ULS indeed measures the continuous ice draft in a fixed location with diameter of several meters. Considering the ice motion, ULS actually acquired dozens to hundreds of kilometers records along the trajectory of sea ice motion on monthly basis, which have enough spatial representativeness compared with ICESat-1/Envisat. Here, we track the source of sea ice that flowing over the ULS on specified month by backward tracking method based on NSIDC Pathfinder data sets. We find the ice draft records included in ULS monthly mean calculation come from a wide range area (Fig. 1). Therefore, we think this is enough to prove the heterogeneity of sea ice measured by each ULS and the validity of ULS data usage in comparison with satellite products. *(please see P8 line 229-231 in the revised manuscript)*

Besides, ULS data is one of the main reference dater sources used for assessing ice thickness from remote sensing in the past decades. ULS are used for comparison with the ice thickness derived from AVHRR (Yu and Rothrock, 1996; Drucker et al., 2003). It was also used to compare with ICESat-1 ice thickness in the Fram Strait (Spreen et al., 2009). In addition, the ULS data sets have been also used for comparison with reanalyses data in the polar region (Mu et al., 2018; Shi et al., 2021). In addition, the comparison with ULS data sets is also a convention for assessing the quality of ice thickness derived from altimeters in the European Space Agency (Kern et al., 2018).

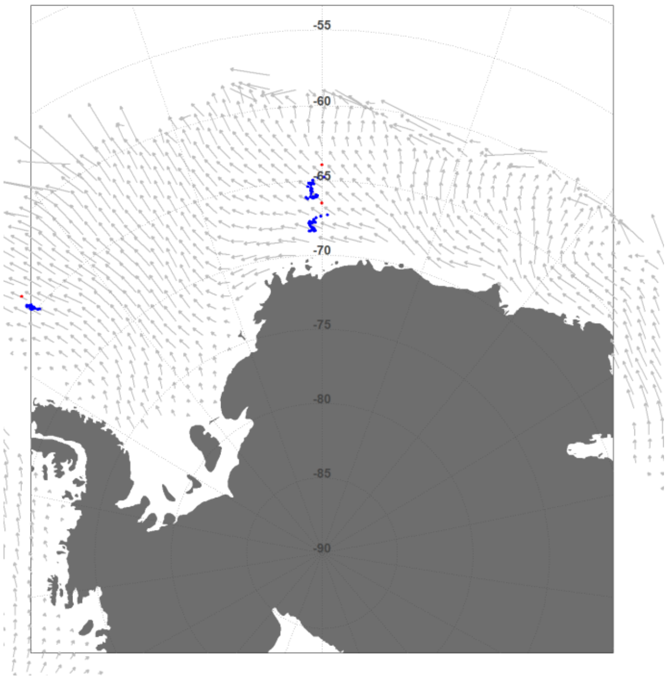

Fig. 1 The origins (30-days ago) of the sea ice (blue dots) that passing through the three ULS sites (red dots) in July 2011 by using backward tracking method based on the NSIDC v4 sea ice motion data. The grey vectors represent the monthly mean sea ice drift derived from NSIDC v4.

L237: Not clear what you mean by "The same feature is found ..." --> Are you referring to the existance of a polynya? Or are you referring to the fact that for both polynya regions, Ross Sea and Weddell Sea Envisat SIT is much higher than ICESat SIT? Please be more specific.

We clarified this sentence: "Similarly, the Ronne Ice Shelf polynya appeared only on ICESat map in 2007 but not on Envisat map." *(please see P9 line 243-244 in the revised manuscript)*

L239: "possibly fails ..." --> This is not a specific enough wording. There are two things involved with that. A) using a 100 km grid naturally results in a land mask at the same grid resolution. Hence it is very likely that the land mask used in the ICESat product extends further into the open ocean than the landmask which is used in the Envisat product. B) As stated in Kern and Spreen, it is not overly bad to not take ICESat freeboard estimates close to the coast not into acount because there the freeboard often is less accurate here compared to the open ocean due to various issues, mostly because of a lack of enough open leads detected by ICESat and hence a less accurate approximation of the local sea surface height and with that less accurate total freeboard.

We added the discussion here as you suggested: "ICESat map has a fringe with no data along most of the East Antarctic coast, which indicates that the 100 km ICESat product fails to see the sea ice close to the coast. This can be attributed to a different land mask used in the ICESat product and consideration of less accurate total freeboard there." *(please see P9 line 244-246 in the revised manuscript)*

L247/248: This apparent discrepancy could be mitigated by showing Envisat SIT only for those grid cells where ICESat has SIT values - as I mentioned earlier already. Otherwise it might be difficult to understand why the small difference between the sea-ice concentration thresholds used (60% vs. 70% ?) has such a large impact on the spatial coverage with SIT data.

Thanks for your advice. However, we think that showing larger coverage of Envisat can help achieving improved understanding of Envisat product itself. Besides, the sea ice concentration threshold for Envisat is 70 % while 60 % for ICESat. Therefore, we don't think the different thresholds are the cause of the different spatial coverage in Ross Ice Shelf Polynya.

L253-254: "probably ... resolve thick ice" --> while the statement made is correct for along-track data you need - in my eyes - to consider two issues here. The first one is that the ICESat product is gridded on a 100 km grid. Given the sparseness of ICESat overpasses with valid data such a 100 km grid SIT estimate in that region might be biased by the presence of thick landfast ice. The second one is that thanks to its finer along-track resolution ICESat can expected to be more sensitive to ocean swell. Ocean swell can result in anomalously high freeeboard values which then convert into too high sea-ice thickness values. While this is a local phenomenon again the sparseness of ICESat overpasses with valid data can results in a similar effect as described above for landfast ice.
Thank you for your comments. We considered the two issues you suggested here carefully and decided to remove this statement.

Fig. 8: I am wondering whether you could perhaps change the color table used for the FDD. It is not intuitive. A high number of FDD denotes cold conditions while a low number comparably warm conditions. I suggest you use a color table which goes from white (0 FDD) to blue (3000degC FDD). Please check whether it is common to express FDD this way. I find it strange to read about temperatures of 1500 and 3000 deg C. Also switching to the unit Kelvin does not solve the problem; ideally, as mentioned earlier, you would translate this to a net growth of sea ice (in meters). Did you check that the FDD shown for MJ-ON is in fact for that period and not for the full FM to ON period? Please note that the notation MJ-FM and ON-MJ is opposite to what you write in the text. Since you aim is to express that the maps in the right two columns show a SIT difference of, e.g. ON minus MJ you might need to invest more annotation elements to not confuse the reader.
We replotted the FDD figures by changing the color table, modifying the expression of the notations and removing the unit of FDD based on the explanation of FDD. We kept analyzing FDD instead of the converted SIT because we think FDD is a robust measure for potential thermodynamic ice growth (or melt). Besides, we conduct the forward tracking on daily FDD with the NSIDC sea ice motion data to add the dynamic effects on the purely thermodynamic growth pattern.

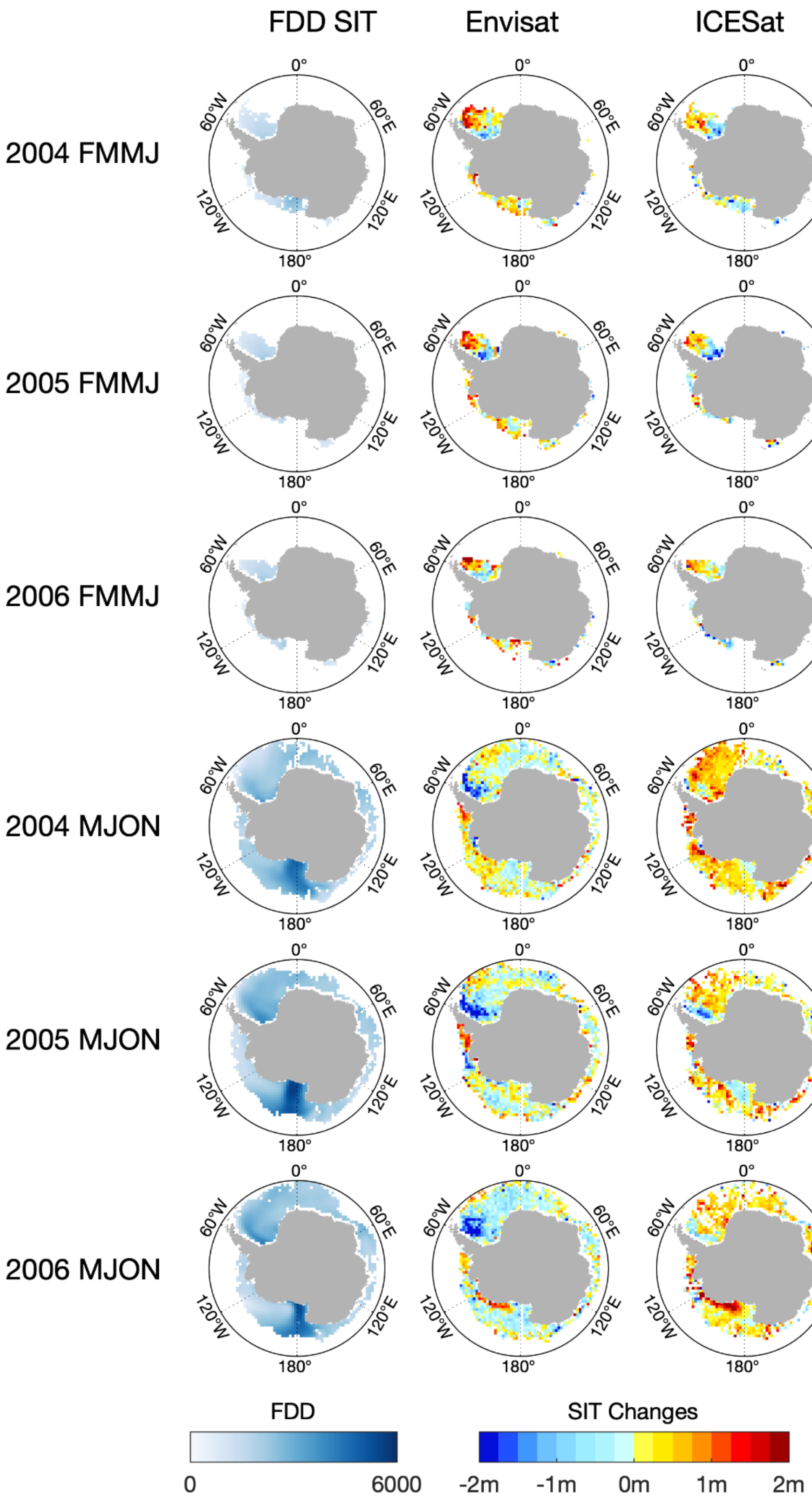

Fig. 2 The FDD differences and sea ice thickness differences from summer to autumn (FMMJ) and from autumn to spring (MJON) derived from Envisat and ICESat in 2004, 2005 and 2006. The FDD patterns are derived by forward tracking daily FDD with sea ice motion data.

L273/274: "This pattern ..." --> I suggest to add the fact that the thick ice found in the southwestern

Weddell Sea at the end of summer is advected northward. If you look at the SIT distributions it is both the tail at large SIT which is decreasing and the tail at small SIT which is increasing. In the particular case you mention here, the thick old ice is replaced by the thin younger ice formed in the polynya (plus other comparably thin ice that is recirculated from the Eastern Weddell Sea in winter.

Thanks for your comments. However, based on the results of the dynamic FDD, we find that the dynamic FDD cannot reproduce these patterns, so more mechanisms like snowfall/flooding, ridging/crack and some ice-ocean feedbacks should be examined for further study. *(please see P10 line 280-281 in the revised manuscript)*

L274-276: "The adverse ... reveal them" --> I would have wished for a more detailed discussion here because one can interpret a lot from these maps - provided one takes into account knowledge about typical snow fall patterns and ice motion. Here you could substantially add some more interesting information and interpretation to your paper.

We conduct the forward tracking on daily FDD with the NSIDC sea ice motion data to add the dynamic effects on the purely thermodynamic growth pattern (Fig. 2). We find that with the aid of sea ice motion, thick ice in the Weddell Sea and Ross Sea can be moved northward. However, the Envisat SIT decrease during MJON still cannot be explained considering the dynamic processes. Therefore, we assume the main reason of the SIT decrease is the overestimation of Envisat SIT in autumn. As for the snow fall patterns, we think it difficult to quantify the impact of snow, i.e., in which way the snow fall would lead to the sea ice thickness growth. *(please see P10 line 280-283 in the revised manuscript)*

Fig. 9: Please add to the caption what the black line and the dashed colored lines stand for. You might also give the information whether you took data from all seasons available or whether we only look at data of years 2004, 2005 and 2006 as only from these years data from all three seasons are available from ICESat.

We added the information in the caption you mentioned: "The data are taken from all seasons available." "The black line is the 1-to-1 fit line and the dashed colored lines stand for linear regression lines."

L322-326: Please note that the "nominal adjustement" suggested by Nandan et al. is derived for cases in the Arctic which might be special and not necessarily transferrable to the Southern Ocean. You could mitigate focussing too much on this exact value of 7 cm by providing a table into which you put sea-ice thickness changes in response to freeboard biases between 2 amd 10 cm in steps of 2 cm. You choose typical first-year sea-ice density. Did you expereiment with other density values to see how dominant the freeboard change is compared to a density change? You could use densities between 880 and 940 kg/m3 in steps of 20 kg/m3 to illustrate this. Why can the differences found here not also account for the differences between Envisat and ICESat in spring? And why do you consider the end of summer a season when this difference might apply?

We experimented with different density values following your suggestions and changed the way to present the sensitivity of the SIT changes to freeboard biases, snow depth biases and sea ice density. From Fig. 3, we can see that the SIT changes are more sensitive to sea ice freeboard biases than to snow depth biases. Besides, with the increase of sea ice density, the SIT changes rise. *(please see P11 line 319-330 in the revised manuscript)*

Considering the average freeboard biases, the corresponding SIT changes can be up to 0.5 m, which matches with the consistent positive differences in summer and autumn but not in spring.

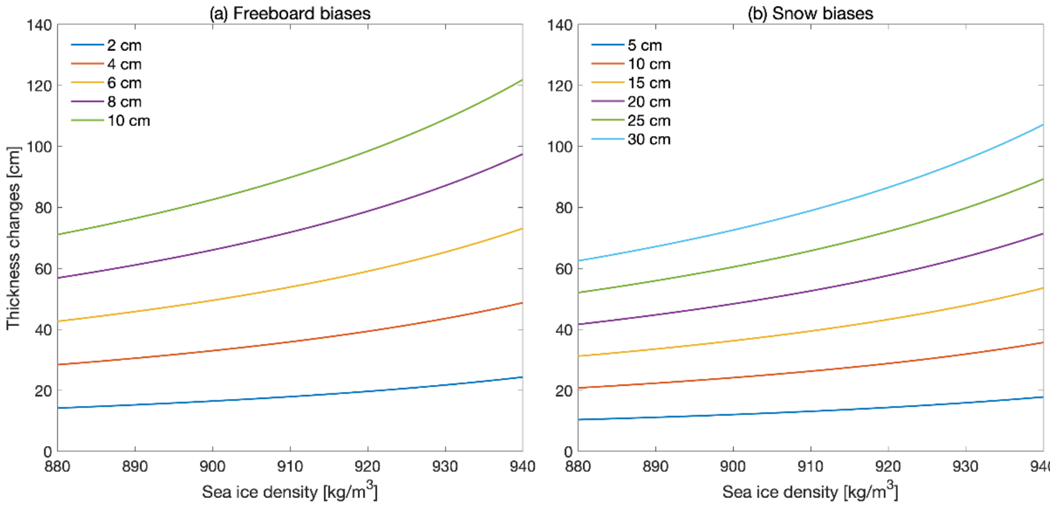

Fig. 3 Sensitivity of sea ice thickness changes to sea ice freeboard biases and snow depth biases as function of sea ice density. (a) SIT changes computed with Eq. (1) for different sea ice freeboard biases (2 cm to 10 cm). (b) Similar to (a) but computed for different snow depth biases (5 cm to 30 cm).

L356-359: Please state that you took the same values for water and sea-ice density as in Eq. 5. While your computation is of course correct, I am wondering whether the 2 cm bias assumed isn't a strong under-estimation. Yes, the analysis is based on monthly data, I agree. But the recommendation of Nandan et al you used in Eq. 5 is not tied to monthly data, is it? The monthly mean retrieval uncertainty you used should be considered the precision and not the potential bias which can be much larger - as you learned from Worby et al., Ozsoy-Cicek et al and as you could also see in the Kern and Ozsoy-Cicek paper in Remote Sensing from 2016; there we easily talk about 20 cm bias. Also te work of Kwok and Maksym from 2014 supports the notion that biases can be much higher over large regions. Hence, considering that also on a monthly scale the bias is an order of magnitude larger does not hurt and I invite you to, as suggested for Eq 5 provide a table into which you put sea-ice thickness changes in response to snow biases between 5 and 30 cm in steps of 5 cm; that would provide a much more realistic view of the potential bias due to using a snow depth data set that does not reflect the actual conditions.

We modified the analyses following your comments and the revision is shown above. *(please see P12-13 line 356-368 in the revised manuscript)*

L365-367: "While a snow ..." --> I agree to this and suggest to also stress one more time that sea-ice thickness differences you observe in your paper between different summer seasons (e.g. between Feb/Mar 2004, 2005, ... 2008) might, to a large extent, also simply be the result that the climatology does not match the actual conditions.

We conduct a comparison between the usage of snow depth climatology and the real snow depth, and the results are presented as: “To further quantify the differences between snow depth climatology and actual snow depth contributions, we conduct the retrieval of Envisat SIT by replacing the snow depth climatology with SICCI AMSR-E snow depth on Envisat level-3 sea ice freeboard data and keeping remainder the same values. The new Envisat SIT is compared with ICESat SIT and the variations of their differences are shown in Fig. 4. This figure reveals that the impacts of snow depth climatology are larger in the Amundsen-Bellingshausen Sea and the Western Weddell Sea compared to other sectors. Among the three seasons, the variations are larger in summer, partly accounting for the differences between Envisat and ICESat SIT." *(please see P13 line 369-380 in the revised manuscript)*

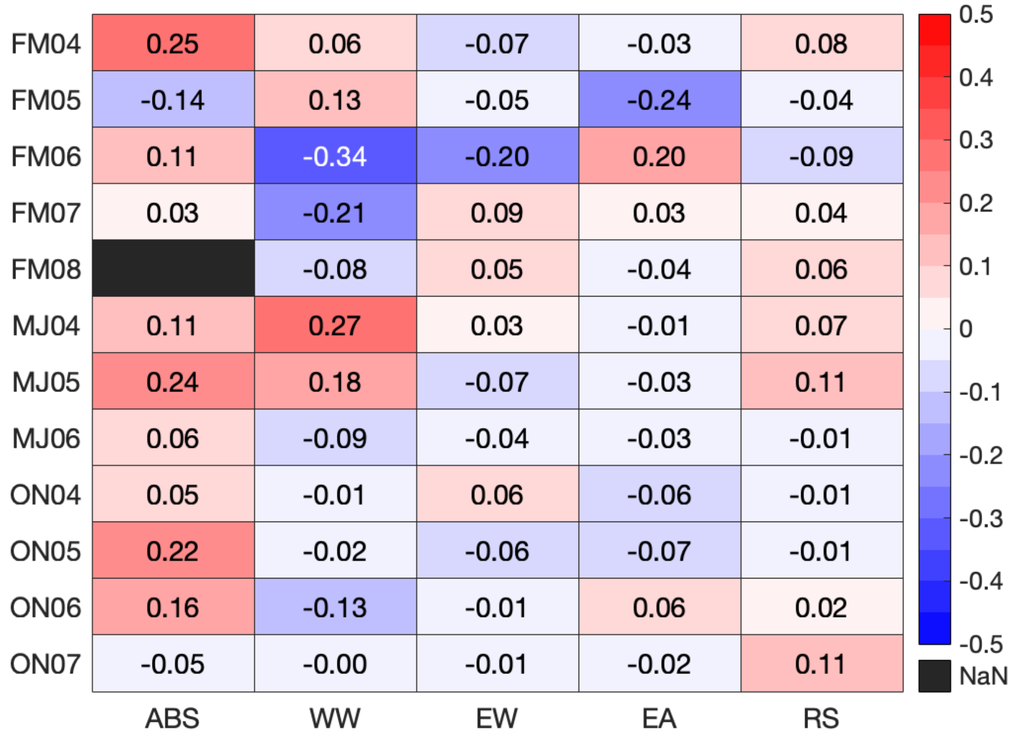

Fig. 4 Changes in the differences between Envisat and ICESat SIT for each comparison period and each region under the experiment of the snow depth climatology impacts.

L374-386: You might want to mention here that possibly the approach by Kern et al. (2016) is providing the total (sea ice plus snow depth) thickness. Taking this into account, the actual 1-layer sea-ice thickness values shown in this paper would possibly even be a bit smaller - with the respective consequences for your results. See also my comments in the context of Eq. 3.

We added this discussion: "Besides, this method is actually providing the total (sea ice plus snow depth) thickness. Taking this into account, the actual ICESat SIT shown in this paper would possibly even be a bit smaller." *(please see P14 line 390-392 in the revised manuscript)*

L402-404: "Compared to the FDD ..." --> In order to make this quite general statement you should investigate these maps in more detail and write more text in the respective section. See also my comments about your usage of FDD.

We modified the statement based on the new results we achieve: "Compared to the FDD results in 2004, 2005 and 2006, we find that the Envisat SIT decrease during MJON cannot be explained considering the dynamic processes and we assume the main reason of the SIT decrease is the overestimation of Envisat SIT in autumn." *(please see P14 line 417-419 in the revised manuscript)*

Editoral remarks / Typos:

L31: Actually, to obtain the sea-ice volume you need to combine the sea-ice thickness with the sea-ice area. I strongly recommend to change the working accordingly.

We changed “sea ice extent” to “sea ice area” since the slight increase and the abrupt decrease occur on sea ice area (Maksym, 2019).

L57++: Please check the paper for the typo: CyroSat-2. It needs to read "CryoSat-2"

We corrected these mistakes.

L104: "aboard on" --> either "aboard" or "on".

We deleted “on” here.

L112/113: "The delay correction ... " I suggest to delete this sentence here and instead add it in the discussion section when you discuss error sources / the uncertainties of the Envisat data.

We moved this sentence to the discussion section as you suggested.

L129: As ICESat is not operating anymore it is grammatically possibly more correct to write "lasted" instead of "lasts".

We corrected these mistakes.

L153/154: "The uncertainty ... height calibration" --> I suggest to rewrite this: "The uncertainty in summer is smaller than in other seasons because open water occurs more frequently in the ULS footprint and with that the estimate of the sea surface height is more accurate.

We modified this sentence as you suggested.

L241/242: "However, ... near zero." -->perhaps better: "However, these differences have to be seen in the light of the standard deviations of ~0.6 m."

We modified this sentence as you suggested.

L258/259: "According to Table 5 ..." --> you could point out better that DESPITE the large difference and RMSD the correlation is actually the highest of the three seasons investigated.

We modified this sentence: “According to Table 5, despite the largest mean difference in autumn of 0.57 m and large RMSD of 0.47 m, the correlation is actually the highest of the three seasons investigated of 0.71.”

L281: What are "splashes"?

We replaced “splashes” with “cloud” here.

L294/295: "though it is known ..." --> please support this knowledge with respective references.

We added the reference here: Willatt et al., 2010.

L296: "footprint of" --> "footprint of the radar altimeter of"

We modified this expression.

L372: The perfect place for the Kwok and Maksym paper from 2014 (JGR-Oceans I think) and possibly for one of his more recent papers where he looked into ICESat-2.

We added the references here: Kwok and Maksym, 2014; Kacimi and Kwok, 2020.

L389/390: --> This sentence reads a bit strange in the context of what follows. My suggestion: "In this study, we compare estimates of the sea-ice thickness obtained from satellite altimeter observations by Envisat RA-2 (radar) and ICESat GLAS (laser) in the Southern Ocean."

We modified this sentence as you suggested.

L391: "Envisat-ULS" --> please make sure the reader understands the "-" as a minus so that it is clear that ULS sea-ice thickness values are smaller than Envisat (and ICESat) values. Currently, this is not clear from the text.

We spelled out the "-" here to make it clearer.

L392: "The results ..." --> I don't understand this sentence in the context of the previous one. Consider to remove.

We removed this sentence.

L394/395: "According ..." --> three time usage of difference / different. Consider to rephrase.

We rephrased this sentence as: “According to the results, the differences between Envisat and ICESat sea ice thickness vary in each season, year and region”.

L395/396: "difference of ... between Envisat SIT minus ICESat SIT" reads strange. Please consider re-phrasing. I note: In contrast to L391 here you spell out the "-".

We modified this sentence as: “More specifically, the smallest monthly average difference (SD in brackets) for Envisat SIT minus ICESat SIT exists in spring of 0.00 m (0.39 m), while larger differences (SD) exist in summer and autumn by 0.52 m (0.68 m) and 0.57 m (0.45 m), respectively.”

L406-408: You might want to re-phrase this sentence after you have considered by comments in the context of Eq. 5 and 6.

We modified the sentence as: “Through the sensitivity experiments, we find that Envisat SIT changes are more sensitive to sea ice freeboard biases than to snow depth biases. Besides, with the increase of sea ice density, the SIT changes rise.”

Figure 3: I suggest that you avoid to write "ENV-ULS" and the like because it is easily misinterpreted as a difference Envisat SIT minus ULS SIT which I doubt is the quantity you are showing here.

Thanks for the comments but here “ENV-ULS” does represent the difference between Envisat SIT and ULS SIT.

Responses to referee #2

Dear Reviewer:

We would like to express our gratitude to you for the comments to improve this manuscript. However, we need to clarify that the purpose of this paper is to give a comprehensive and statistical comparison between Envisat and ICESat sea ice thickness data. Only when the significant differences are admitted, the importance of dealing with the uncertainties of these products is revealed. Besides, we have already discussed the probable causes of the differences in section 4 and we supply more experiments following your suggestions.

The specific responses and revisions are shown below. They are in blue font for clarity.

Qinghua Yang, Qian Shi, Robert Ricker, Stefan Hendricks
On behalf of all the authors

Major concerns:

L90-92. I think the comparison of the two SIT products with ULS is not appropriate since the single measurement point (6-8 m) cannot represent a grid with 50 km or even 100 km. Moreover, only the uncertainty of sea ice draft derived with ULS 5-12 cm is presented (L152-153), the uncertainty of SIT derived with Eq. 4 is missing and Fig.3 also lacks error bars for ULS, thus making the comparison unreliable. "Both Envisat and ICESat SIT have been interpolated onto each ULS location in the nearest neighbour way" (L183-184) further introduces huge uncertainties. Based on these considerations, it is not recommended to use ULS as a comparison data source. ULS can be used if the Envisat or ICESat footprints spatio-temporally coincide with it, and the uncertainty of SIT derived with ULS is clear.

Thanks for your comments. However, we think that you are biased in denying the feasibility of using ULS data as comparison data with ICESat-1/Envisat due to their relatively narrow footprint. As we know, ULS indeed measures the continuous ice draft in a fixed location with a diameter of several meters. Considering the ice motion, ULS acquired dozens to hundreds of kilometers records along the trajectory of sea ice motion on a monthly basis, which have enough spatial representativeness compared with ICESat-1/Envisat. Here, we track the source of sea ice that flows over the ULS in a specified month by backward tracking method based on NSIDC Pathfinder data sets. We find the ice draft records included in ULS monthly mean calculation come from a wide range area (Fig. 1). Therefore, we think this is enough to prove that the spatial representativeness of the monthly average ULS data can be compared with that of ICESat-1/Envisat. *(please see P8 line 229-231 in the revised manuscript)*

Besides, ULS data was generally used for ice thickness comparison in the previous studies. ULS is used for comparison with the ice thickness derived from AVHRR (Yu and Rothrock, 1996; Drucker et al., 2003). It was also used to compare with ICESat-1 ice thickness in the Fram Strait (Spreen et al., 2009). In addition, the ULS data sets have also been used for comparison with reanalyses data in the polar region (Mu et al., 2018; Shi et al., 2021). In addition, the comparison with ULS data sets is also a convention for assessing the quality of ice thickness derived from altimeters in the European Space Agency (Kern et al., 2018).

In summary, we think that the reason for rejecting us due to the spatial representativeness of ULS ice thickness is untenable. Previous studies (referred to above) have shown that using ULS for validation of satellite-derived sea-ice thickness data sets can be considered as state of the art.

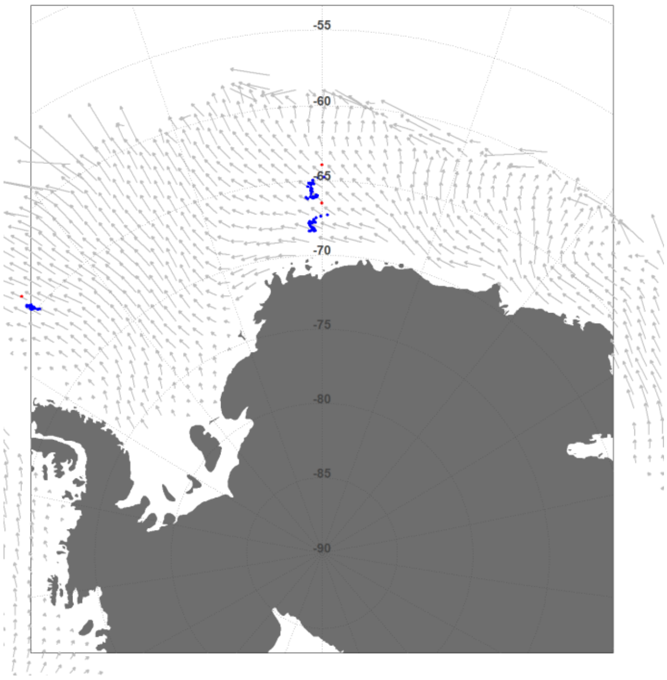

Fig. 1 The origins (30-days ago) of the sea ice (blue dots) that passing through the three ULS sites (red dots) in July 2011 by using backward tracking method based on the NSIDC v4 sea ice motion data. The grey vectors represent the monthly mean sea ice drift derived from NSIDC v4.

The difference between the Envisat-based actual SIT, i.e., the mean thickness of the icecovered fraction of the grid cell area (without open water areas) (L122-123), and the ICESat effective sea ice thickness, i.e., mean thickness per grid cell including open water areas (L141-142), is not tackled nor discussed for the two datasets.

We point out the different thickness representations for Envisat and ICESat. And we choose to compare the effective sea ice thickness during the intercomparison process. We have clarified this in the paper: "The effective Envisat SIT is calculated by multiplying the SIC contained in the data for each grid from OSI-SAF Global Sea Ice Concentration (OSI-409) and the OSI-SAF Global Sea Ice Concentration continuous reprocessing offline product (OSI-430) (http://osisaf.met.no)." (*please see P8-9 line 234-236 in the revised manuscript*).

Considering the huge differences between Envisat and ICESat SIT products (as can be seen in Fig. 9 and Table 7), the main object of this work should not stay at just comparing those products, but concentrating on the qualitative and quantitative analysis of the causes leading to the differences. Currently, these issures are only simply discussed in Section 4. Following works may be considered by the authors:

L253-254 About the sentence "Probably inferring that …" Is it really the key reason for SIT overestimation of Envisat than ICESat in autumn? The similar doubt also appears in summer (L262-263).

L21 and L256-257. Why on earth the mean Envisat SIT decreases while the mean ICESat SIT increases from autumn to spring? This should be supported with supplement experiments.

L360-361. "The largest effect might not come from the impact of ice deformation on the snow-depth retrieval but might be due to the difference between actual snow depth from that represented by the climatology.” Can the influence of climatology quantified?

I didn’t see solid evidences supporting the statement “The potential overestimation of sea ice freeboard caused by range biases accounts for much of the differences between Envisat and ICESat SIT in summer and autumn, while the biases of snow depth are not the dominant cause of the differences.”

(1) We realized that the statement is correct only for along-track data. Firstly, given the sparseness of ICESat overpasses with valid data such a 100 km grid SIT estimate in that region might be biased by the presence of thick landfast ice. Besides, ocean swell can result in anomalously high freeboard values which then convert into too high sea-ice thickness values. While this is a local phenomenon, the sparseness of ICESat overpasses with valid data can results in a similar effect as for landfast ice. Therefore, we considered the two issues here carefully and decided to remove this statement.

(2) We conduct the forward tracking on daily FDD with the NSIDC sea ice motion data to add the dynamic effects on the purely thermodynamic growth pattern (Fig. 2). We find that with the aid of sea ice motion, thick ice in the Weddell Sea and Ross Sea can be moved northward. However, the Envisat SIT decrease during MJON still cannot be explained considering the dynamic processes. Since the dynamic FDD cannot reproduce these patterns, more mechanisms like snowfall/flooding, ridging/crack and some ice-ocean feedbacks should be examined for further study. However, one thing we can give a speculation based on the analyses in autumn and the regular rule during freezing seasons is that the main reason for Envisat SIT overall decrease during MJON is the overestimation of Envisat SIT in autumn. As for the snow fall patterns, we think it difficult to quantify the impact of snow, i.e., in which way the snow fall would lead to the sea ice thickness growth. *(please see P10 line 275-281 in the revised manuscript)*

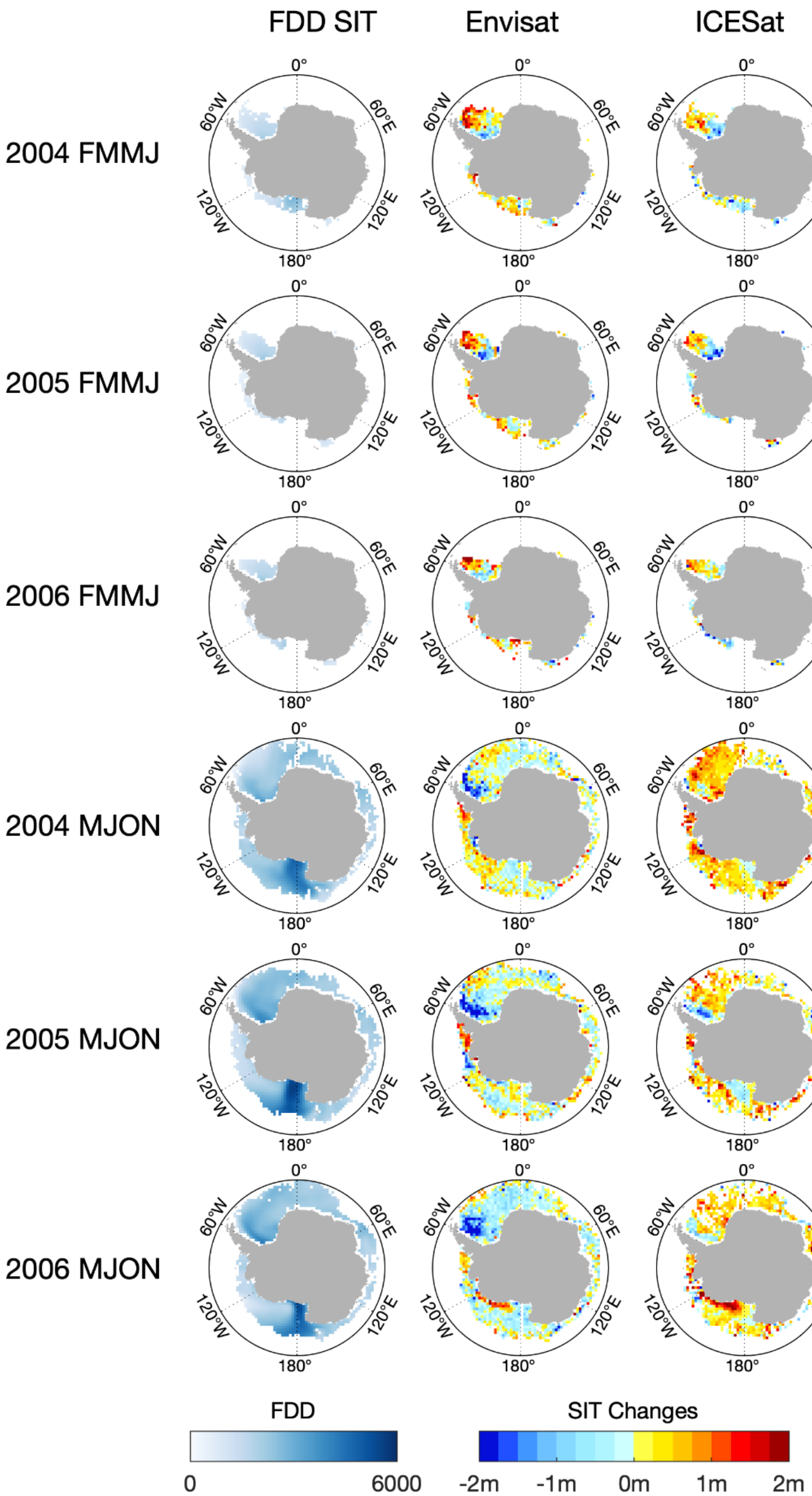

Fig. 2 The FDD differences and sea ice thickness differences from summer to autumn (FMMJ) and from autumn to spring (MJON) derived from Envisat and ICESat in 2004, 2005 and 2006. The FDD patterns are derived by forward tracking daily FDD with sea ice motion data.

(3) We quantify the contribution of the usage of snow depth climatology instead of actual snow depth

during the Envisat SIT retrieval. We redo the retrieval of Envisat SIT by replacing the snow depth climatology with SICCI AMSR-E snow depth on level-3 sea ice freeboard data. The new Envisat SIT is compared with ICESat SIT and the variations of their differences are shown in Fig. 3. This figure reveals that the impacts of snow depth climatology are larger in the Amundsen-Bellingshausen Sea and the Western Weddell Sea compared to other sectors. Among the three seasons, the variations are larger in summer, partly accounting for the differences between Envisat and ICESat SIT. *(please see P13 line 375-380 in the revised manuscript)*

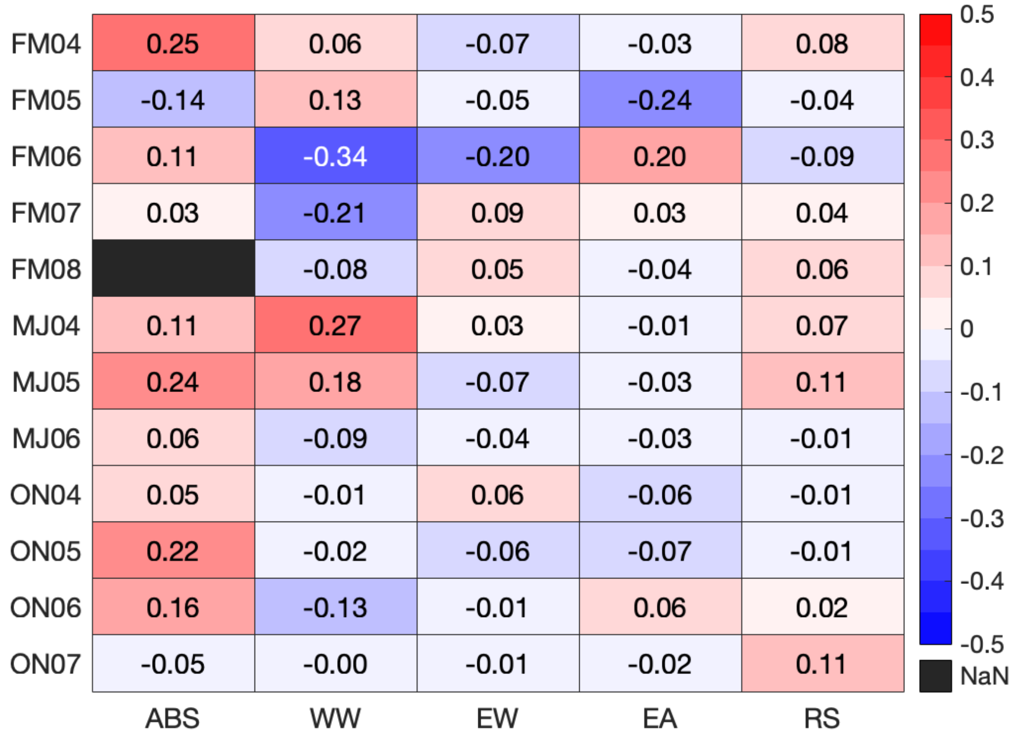

Fig. 3 Changes in the differences between Envisat and ICESat SIT for each comparison period and each region under the experiment of the snow depth climatology impacts.

(4) We conclude the sensitivity of the SIT changes to freeboard biases, snow depth biases and sea ice density in Fig. 4 by analyzing Eq. (1):

$$I = \frac{F\rho_{water} + S\rho_{snow}}{\rho_{water} - \rho_{ice}} \quad (1)$$

The sensitivities to freeboard biases and to snow depth biases are calculated by:

$$\frac{dI}{dF} = \frac{\rho_{water}}{\rho_{water} - \rho_{ice}} \quad (2)$$

$$\frac{dI}{dS} = \frac{\rho_{snow}}{\rho_{water} - \rho_{ice}} \quad (3)$$

From Fig. 4, we can see that though the magnitudes of the resulting thickness changes are quite similar, the SIT changes are more sensitive to sea ice freeboard biases than to snow depth biases. Besides, with the increase of sea ice density, the SIT changes rise. For typical sea ice freeboard biases (7 cm for the Arctic nominal adjustment suggested by Nandan et al. (2017, 2020)), the sea ice density variations induce the thickness changes ranging from ~0.5 m to ~0.8 m. For typical snow depth biases (20 cm for the monthly mean retrieval uncertainty in Kern and Ozsoy-Cicek (2016)), the thickness changes from ~0.4 m to ~0.7 m. **Although this sensitivity analysis is not solid enough for the explanation for the SIT differences in three seasons, it can provide a reasonable conjecture that freeboard biases are the main cause of the positive differences in summer and autumn.** *(please see line 319-330 and 356-368 in the revised manuscript)*

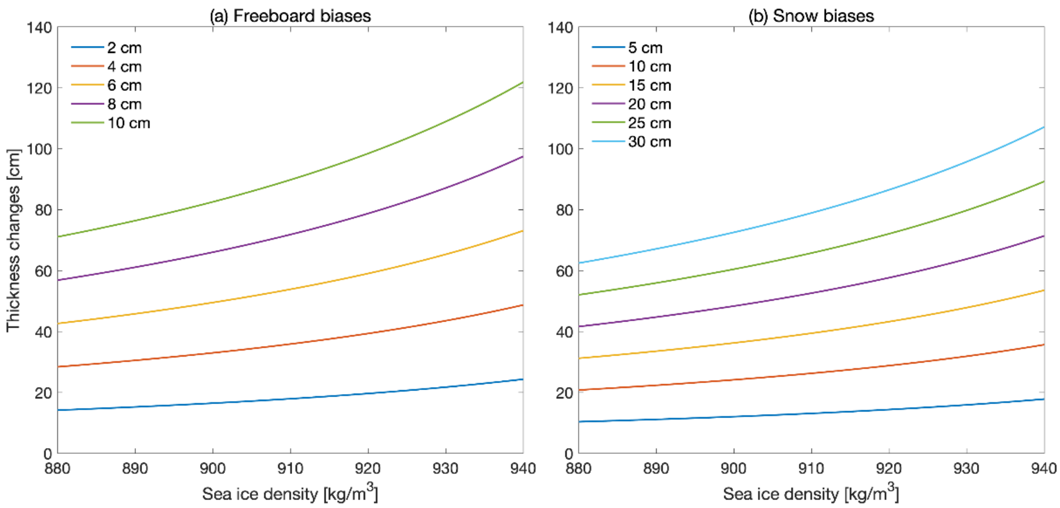

Fig. 4 Sensitivity of sea ice thickness changes to sea ice freeboard biases and snow depth biases as function of sea ice density. (a) SIT changes computed with Eq. (1) for different sea ice freeboard biases (2 cm to 10 cm). (b) Similar to (a) but computed for different snow depth biases (5 cm to 30 cm).

L124 The sea ice thickness derived with the modified ice density approach, i.e., Eq.3 can be considered to be updated to the new OLMi method (Xu, et al. (2021). "Deriving Antarctic Sea-Ice Thickness from Satellite Altimetry and Estimating Consistency for NASA's ICESat/ICESat-2 Missions." Geophysical Research Letters. http://dx.doi.org/10.1029/2021GL093425), which showed the modified ice density approach in Kern et al. (2016) would overestimate SIT.

Thanks for your information. We conduct the comparison between the Envisat SIT and the new ICESat SIT derived by Xu et al. (2021). Figure 5 shows consistent positive variations, with larger ones in summer, especially in the Amundsen and the Bellingshausen Sea and the Western Weddell Sea. However, we do not aim to choose the best ICESat SIT product with the most real SIT, but investigate the causes of the differences between Envisat and ICESat SIT, and how different sensors and retrieval methods are represented in the SIT fields. In addition, the empirical approaches used by Xu et al. (2021) were developed from a suite of historic in-situ observations of freeboard, snow depth and sea-ice thickness which in a way have the character of climatology as well. Therefore, we keep using the ICESat product from Kern et al. (2016) and discuss its uncertainties in section 4.3.

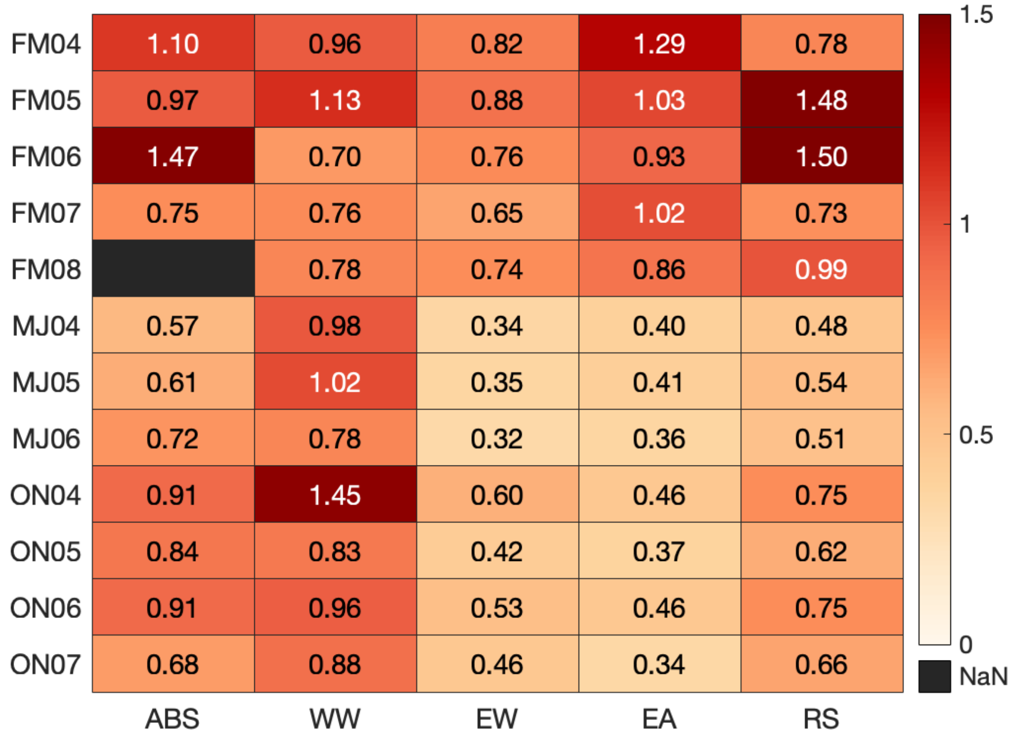

Fig. 5 Changes in the differences between Envisat and ICESat SIT for each comparison period and each region under the experiment of the new OLMi ICESat SIT.

Minor concerns:

L22-24 Please quantify the percentage of the uncertainties caused by the radar backscatter and snow depth products respectively.

The new sensitivity analyses are shown above. Since the exact freeboard biases and snow depth biases cannot be quantified, we can only achieve a general SIT change coming from either of them.

L64 “the radar altimetry SIT retrievals” to “SIT retrieval by the radar altimetry”

We modified it as you suggested.

Are the densities used in Eq. 1 and Eq.2/3 the same? If not, how does it influence the SIT retrieved by the two sensors?

Yes, the densities are similar. The ρ_water in Eq.1 is 1024 kg m^{-3} while in Eq.3 is 1023.9 kg m^{-3}. The ρ_ice in Eq.1 is 916.7 kg m^{-3} while in Eq. 2 is 915.1 kg m^{-3}. The ρ_snow in Eq.1 and Eq.2 are the same of 300 kg m^{-3}.

L166 and L271 MF-MJ or MJ-MF? MJ-ON or ON-MJ? Please unify them throughout the paper, such as those 'MJ-ON' (in the text) or 'ON-MJ' (Fig. 8).

We unified the descriptions to “FMMJ” and “MJON” throughout the paper.

Why is it called snow depth climatology (L66, L118), snow-depth climatology (L119), or snow climatology (L363), and what is the real difference between them and the actual snow depth? Besides, what is the meaning of “have the character of a climatology” (L386)?

We unified the descriptions to “snow depth climatology”. This snow depth climatology is an average snow depth based on 2002-2011 and it neglects the interannual variations of the snow depth. “Have the character if a climatology" means that the sea ice thickness derived from Li et al. (2018) is still affected by the usage of climatology data.

L270 "from the model" is unclear.

We corrected this sentence to: "We calculate the period-average FDD corresponding to ICESat operating periods for the same spatial coverage."

L274-276 I don't think it is an adverse pattern comparing MJ-ON with FM-MJ. Please also make "different abilities" clear.

We replaced "adverse" with "opposite". "Different abilities" represent the ability to detect small-scale deformation processes.

L284-285 the weighted average is in the first row instead of in the last column?

We meant to explain the numbers in the last column and we moved this sentence to the caption in Table 7: "N is the numbers of comparison pairs, taking into account the actual number of values per season."

L379-381. The sentence "Therefore, …the ice-snow column" is hard to understand. For example, "underestimations of sea ice and snow observations" is not clear, is it sea ice thickness and snow depth underestimation? What is the "apparent ice density"?

Since visual ship-based observations of sea ice thickness and snow depth are used in the ICESat SIT retrieval, the underestimations of these data can have effects on the modified ice-snow density. We amended the sentence to: "Therefore, the largest uncertainty of ICESat comes from the potential underestimations of the ship-based sea ice thickness and snow depth observations for the computation of the bulk density of the ice-snow column (Kern et al., 2016)."

Fig.8 Suggest to use the same Antarctica background (in grey) as that in the other figures such as Fig. 4/5/6 since we can notice the big blank area along the Ross Sea coast in this figure.

We modified this figure as shown in Fig. 2.

Table 4 what's N? It should be introduced in the title. Same happens in Table 5/6/7.

N is the number of comparison pairs. We added this introduction in the captions of the tables.

Table 5. I suggest to also compute the difference between ENV and ICE at grid scale instead of just subtract with the computed statistical values (the "Difference" column). I mean, the mean of the third column of Figure 4/5/6 should be computed. Based on the figures, I think the two values would be different.

The averages of the difference patterns are calculated only for the grid cells with both available Envisat and ICESat SIT. Therefore, the two ways would lead to the same results.

Table 6. "sea ice thickness differences" should be followed by "with standard deviation in brackets".

We added this information in the caption.

References:

Yu, Y., and Rothrock, D. A. (1996), Thin ice thickness from satellite thermal imagery, J. Geophys. Res., 101(C11), 25753– 25766, doi:10.1029/96JC02242.

Drucker, R., Martin, S., and Moritz, R. (2003), Observations of ice thickness and frazil ice in the St. Lawrence Island polynya from satellite imagery, upward looking sonar, and salinity/temperature moorings, J. Geophys. Res., 108, 3149, doi:10.1029/2001JC001213, C5.

Spreen, G., Kern, S., Stammer, D., and Hansen, E. (2009), Fram Strait sea ice volume export estimated between 2003 and 2008 from satellite data, Geophys. Res. Lett., 36, L19502, doi:10.1029/2009GL039591.

Mu, L., Losch, M., Yang, Q., Ricker, R., Losa, S. N., and Nerger, L. (2018). Arctic-wide sea ice thickness estimates from combining satellite remote sensing data and a dynamic ice-ocean model with data assimilation during the CryoSat-2 period. J. Geophys. Res., 123, 7763– 7780, doi: 10.1029/2018JC014316

Shi, Q., Yang, Q., Mu, L., Wang, J., Massonnet, F., and Mazloff, M. R.: Evaluation of sea-ice thickness from four reanalyses in the Antarctic Weddell Sea, The Cryosphere, 15, 31–47, https://doi.org/10.5194/tc-15-31-2021, 2021.

Kern, S., Khvorostovsky K., and Skourup, H.: D4.1 Product Validation & Intercomparison Report (PVIR-SIT), available at: http://icdc.cen.uni-hamburg.de/fileadmin/user_upload/ESA_Sea-Ice-ECV_Phase2/SICCI_P2_PVIR-SIT_D4.1_Issue_1.1.pdf, 2018.

Responses to referee #3

Dear Reviewer:

We would like to express our gratitude to you for the helpful comments to improve this manuscript. We have carefully modified the discussion and the expression following your comments. The specific responses and revisions are shown below. They are in blue font for clarity.

Qinghua Yang, Qian Shi, Robert Ricker, Stefan Hendricks
On behalf of all the authors

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.

Thanks for your comments. We retrieve a new ICESat SIT using the same snow depth climatology product and the same hydrostatic equilibrium approach as Envisat SIT in response to your advice. The new ICESat SIT is compared with Envisat SIT and the changes to the differences are shown in Fig. 1. Compared to Fig. 4-6 in the manuscript, we can find that the positive differences in the Weddell Sea in summer and autumn increase, while the rest of the differences mainly decrease.

However, we don't think this experiment is available to distinguish the impacts from the sensors and from snow depth, since the new ICESat SIT still employs a possibly biased snow depth product. It can only be used to clarify the differences between the usage of hydrostatic equilibrium retrieval method and the modified density retrieval method.

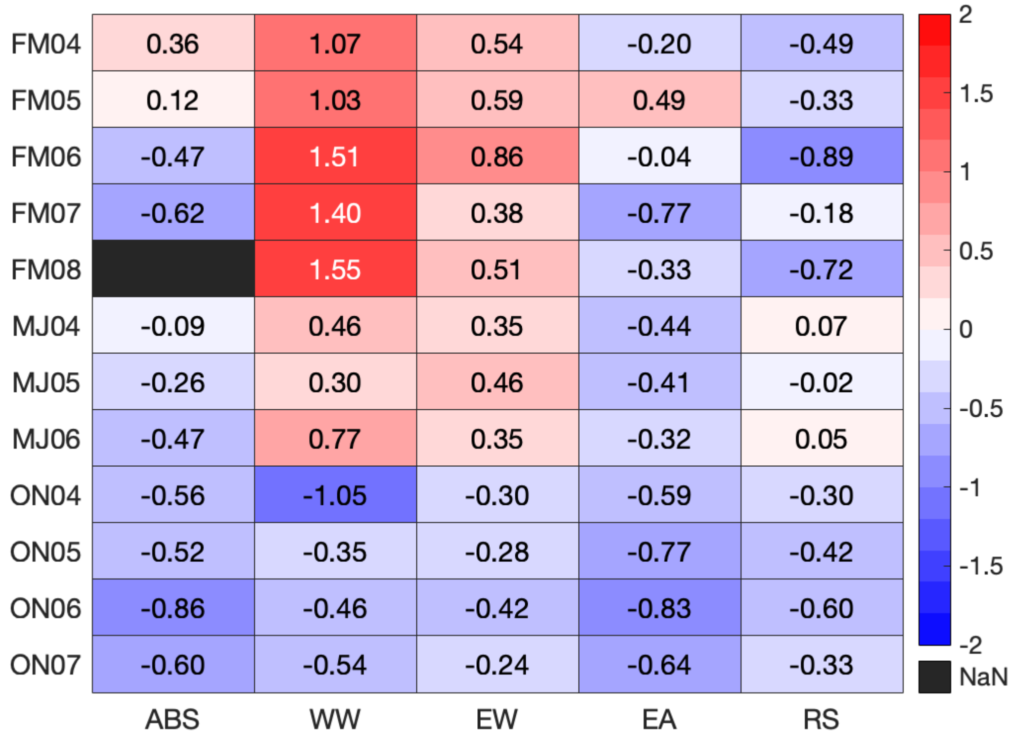

Fig. 1 Changes in the differences between Envisat and ICESat SIT for each comparison period and each region under the experiment of the new ICESat SIT.

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

The usage of different SIC thresholds is because of the different thresholds used in the retrieval of the two data sets. Envisat SIT employs a SIC threshold of 70% during the retrieval while the ICESat SIT uses 60%. Only areas with sea ice concentrations greater than the threshold are considered a valid area for detection of leads and sea ice. We also tested the difference between using 60% and 70% SIC threshold for ICESat during the comparison with Envisat SIT. According to Table 1, this different threshold does not play an important role in the results of this paper. D(60) refers to Envisat minus ICESat (ENV-ICE) applying 60% SIC threshold for ICESat, while D(70) refers to ENV-ICE when SIC threshold for ICESat is 70%. Since the ice concentration gradients are usually quite steep, there will not be a lot of area with values 60% < SIC < 70%.

Table 1. Statistical results of the comparison between Envisat SIT and ICESat SIT using 60% and 70% SIC threshold at each operating period.

	ON04	ON05	ON06	ON07	FM04	FM05	FM06	MA07	FM08	MJ04	MJ05	MJ06
D(60) (m)	0.00	0.05	-0.19	0.14	0.89	0.74	0.47	0.61	0.92	0.61	0.55	0.60
D(70) (m)	0.00	0.06	-0.21	0.15	0.79	0.66	0.43	0.61	0.89	0.60	0.55	0.61

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?

We investigated the ICESat product that Li et al. (2018) produces by comparing with ICESat from Kern et al. (2016) and the ULS SIT used in this study. From Fig. 2 we can see that the differences between two ICESat products are small in general, with some larger differences in the West Weddell Sea and Amundsen Sea. Table 2 shows that compared with ULS SIT, ICESat SIT from Kern et al. (2016) performs even better. Based on these analyses, we think that the inter-comparison results with Envisat SIT in this study are not affected by the choice of ICESat product. More importantly, we do not aim to choose the best ICESat SIT product with the most real SIT in this study, but investigate the causes of the differences between Envisat and ICESat SIT, and how different sensors and retrieval methods are represented in the SIT fields. In addition, the empirical approaches used by Li et al. (2018) were developed from a suite of historic in-situ observations of freeboard, snow depth and sea-ice thickness which in a way have the character of climatology as well. Therefore, our work is still based on the data produced by Kern et al. (2016).

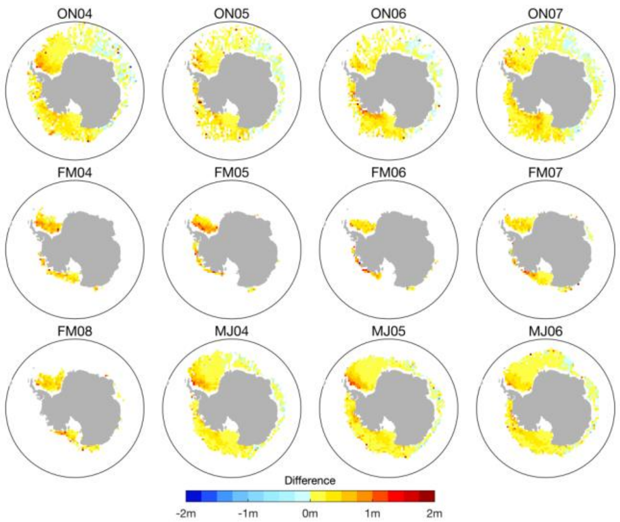

Fig. 2 Maps of differences that ICESat SIT from Kern et al. (2016) minus ICESat SIT from Li et al. (2018) at each operating period.

Table 2. The differences and RMSD between ULS SIT and the two ICESat SIT at each site. ICE(K) refers to ICESat SIT from Kern et al. (2016) and ICE(L) refers to ICESat SIT from Li et al. (2018).

	ICE(K)-ULS		ICE(L)-ULS	
	D (m)	RMSD (m)	D (m)	RMSD (m)
Site 207	0.60	0.66	0.68	0.34
Site 229	-0.04	0.07	0.17	0.15
Site 231	0.27	0.34	0.33	0.44

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 ares)' (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.

During the comparison with ULS observations, we compare Envisat and ICESat actual SIT to exclude zero thickness measured by ULS. Therefore, we divide ICESat SIT by the sea ice concentration contained in the ICESat data for each grid.

Then, we compare the effective SIT of Envisat and ICESat during the intercomparison work by multiplying Envisat SIT by the sea ice concentration contained in the Envisat data. We added the information in section 3.2: "The effective Envisat SIT is calculated by multiplying the sea ice concentration contained in the data for each grid which come from OSI-SAF Global Sea Ice Concentration (OSI-409) until April 16, 2015 and the OSI-SAF Global Sea Ice Concentration continuous reprocessing offline product (OSI-430) afterwards (http://osisaf.met.no)." *(please see P8-9 line 234-236 in the revised manuscript)*

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.

The ρ_water in Eq.1 is 1024 kg m^{-3} while in Eq.3 is 1023.9 kg m^{-3}. The ρ_ice in Eq.1 is 916.7 kg m^{-3} while in Eq. 2 is 915.1 kg m^{-3}. The ρ_snow in Eq.1 and Eq.2 are the same of 300 kg m^{-3}.

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

Thanks for your advice. However, since there are little corresponding ULS data, the large-scale SIT distributions cannot be presented.

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.

We checked through the manuscript and made our best to interpret every sentence clearly.

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 algortihms. 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).

We agreed and removed this sentence as you suggested.

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.

We rewrote the sentence: "The sensors transmit sound pulses upwards with a footprint of 6–8 m in diameter and the signals are reflected either by the sea ice bottom or the sea surface, **yielding two-way travel time which can be converted into distances.**" *(please see P6 line 148-150 in the revised manuscript)*

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.

We realized that the statement is correct only for along-track data. Firstly, given the sparseness of ICESat overpasses with valid data such a 100 km grid SIT estimate in that region might be biased by the presence of thick landfast ice. Besides, ocean swell can result in anomalously high freeboard values which then convert into too high sea-ice thickness values. While this is a local phenomenon, the sparseness of ICESat overpasses with valid data can results in a similar effect as for landfast ice. Therefore, we considered the two issues here carefully and decided to remove this statement.

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.

We wrote the sentence: "In summer, the agreement between Envisat SIT and ICESat SIT is not good, mainly due to their different performances on thick ice above 3 m." *(please see P10 line 269-270 in the revised manuscript)*

Techinical corrections

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

We modified this sentence: "while the uncertainties of the snow depth product are not the dominant cause of the differences".

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 seperate.

We modified this sentence: "However, it is still unclear if the recent increase in Antarctic sea ice area is also accompanied by a similar change in sea ice thickness."

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

We modified this sentence.

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

We changed "measure" to "provide".

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

We removed this word.

L54-55: Consider more recent studies that have retrieved Antarctic sea ice thickness, e.g. Kurtz & Markus (2012) and Kacimi & Kwok (2020).

We added Kacimi & Kwok (2020) here and removed Zwally et al. (2008).

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

Yes, the different retrieval methods that Envisat and ICESat SIT products use have impacts on their differences.

L88: 'the former inter-comparison study', which study is this?

The study is Kern et al. (2016), and we modified the sentence to make it clear: "Based on the former inter-comparison study (Kern et al., 2016), we choose the ICESat sea ice thickness data derived from the modified ice density approach for comparison."

L147: change 'underwater' to 'below sea level'

We agreed with your comment but we found it unnecessary to mention the mooring location in detail.

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

We modified the sentence: "The intervals of sea ice draft measurements are between 3 and 15 minutes from November 1990 to March 2008."

L153: seasons -> season

We corrected this issue.

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.

We modified the sentence: "We compare FDD with the SIT variations from February/March to May/June (FMMJ) and from May/June-October/November (MJON) represented by Envisat and ICESat SIT."

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

We deleted "first".

L193: Remove 'during the comparison'.

We removed this phrase.

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'.

We rewrote the sentence as: "The error bars in the figure show the uncertainty estimates 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

We added the ULS uncertainty of ± 0.05 m following Belter et al. (2020) in the figure.

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.

We find that the ULS uncertainty cannot explain these differences.

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

Snow depth uncertainty is not included because the ICESat SIT retrieval method does not require additional snow-depth information. However, we checked Kern et al. (2016) and their computation of the SIT uncertainty included in the product is based on uncertainties in densities and freeboard. Also, the uncertainty estimates provided with the Envisat SIT data set are possibly too large because the data set producers those days did not adequately take potential correlations between the error contribution into account. Therefore, we rewrote the sentence as: "The large differences in the error bars between Envisat and ICESat mainly result from the inclusion of snow depth uncertainty in Envisat SIT, and lack of adequate regards for potential correlations between the error contribution."

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

We rewrote the sentence as: "In the eastern Weddell Sea (at sites 229 and 231), ICESat has a few overestimations while Envisat has larger overestimations, but the Envisat error bars cover almost all the observations."

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

We removed "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?

These uncertainty contributions include spatial and temporal variability on snow depth as well as snow and sea ice density. Few information about these data exists in the Antarctic. Besides, the coverage of sea ice type (first and multi-year) products is incomplete for the Envisat observation period.

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.

We added the sea ice motion information in the supplement. With the sea ice motion data from NSIDC, the 30-day origins of the sea ice passing the three ULS sites in July 2011 is shown in Fig. S1 and it proves the heterogeneity of sea ice measured by each ULS and the validity of ULS data usage in comparison with satellite products.

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

The ship-based observations are the ASPeCt data from Worby et al. (2008). According to their Table 3, the average ice thickness in spring West Pacific is 0.68 m, smaller than Envisat and ICESat SIT in our study.

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

We changed "feature" to "dissimilarity".

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.

We added a sentence to clarify it: "Envisat detects sea ice in the Ross Sea all the years, but with thickness estimates of up to 1.5 m, much larger than expectant seasonal ice thickness."

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.

We rewrote the sentence: "In spring, the two data sets have similar distributions, represented by similar mean and modal thicknesses."

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?

FDD is calculated by daily degrees below freezing summed over the total number of days the temperature was below freezing. According to Lebedev (1938), a simple model is constructed to produce sea ice thickness:

$\text{Thickness (cm)} = 1.33 * \text{FDD (°C)}^{0.58}$

Note that the calculated thickness only accounts for the freezing of sea water and excludes ice variations from snowfall, freezing rain or ridging. However, we don't translate the FDD into thickness growth in this study because i) we think FDD is sufficient to stand for thickness growth and ii) using the very simple translation equation adds uncertainties into our analyses.

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?

We modified the description: "As both satellite products present decreasing SIT in the southern Weddell Sea and increasing SIT in the northwest Weddell Sea during FMMJ, Envisat SIT shows more thinning all around the Southern Ocean and ICESat SIT generally shows more thickening during MJON."

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

Thanks for your advice. We used FMMJ and MJON to replace the previous usage.

L272: 'both products', which two products? Envisat and ICESat, or satellite altimetry and FDD?

Envisat and ICESat. We changed it to "both satellite products" to clarify.

L274: 'The adverse patterns', adverse (preventing success or development; harmful; unfavourable) might not be the right word here?

We changed "adverse" to "opposite".

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.

We rewrote the sentence as you suggested: "For all five locations, Envisat SIT tends to exceed ICESat SIT for thin ice."

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?

We changed "splashes" to "cloud". This kind of "exceed" reveals that mean ICESat SIT are nearly constant through all three seasons in the western Weddell Sea, while mean Envisat SIT are noticeably larger in summer and autumn

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?

We moved this sentence to the caption. We meant to explain the numbers in the last column and we modified the description: "The numbers in the last column of the table are the sample sizes of the comparisons, taking into account the actual number of values per season."

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

We added a citation: Willatt et al. (2010).

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

We modified the sentence: "considering the Envisat large pulse-limited footprint of about 2–10 km and smaller footprint of ICESat laser beams of about 70 m".

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

We removed "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 oservations is positive'

We rephrased the sentence as you suggested.

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.

Thanks for your advice. We have introduced AMSR-E snow depth in section 2.1 and 4.2. We just added the passive microwave information in section 2.1: "This snow-depth climatology is derived from the passive microwave sensor Advanced Microwave Scanning Radiometer-EOS (AMSR-E) and AMSR-2 for the Antarctic".

L351: '... lead to underestimations' and '... lead to overestimations', under- and overestimations of what? SIT?

We meant to say under- and overestimations of snow depth and we added "of snow depth" in the sentence.

L357: The retrieval uncertainty of AMSR-E?

Yes. According to Kern et al. (2015) the average monthly retrieval uncertainty of AMSR-E is 2 cm. However, we realized this value is the precision but not the potential bias which can be much larger. Therefore, we checked sea-ice thickness changes in response to snow biases between 5 and 30 cm in steps of 5 cm.

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.

We modified the sentence: "However, compared with Fig. 10a, SIT changes are more sensitive to freeboard biases than to snow biases."

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.

Generally, using a snow climatology for converting ice freeboard to thickness neglects any interannual snow variability. It can reduce the actual snow depth biases to some extent.

L389: Remove 'firstly'

We removed it.

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

Thanks for your advice, but we delete this sentence since the previous sentence has stated the result of the comparison with ULS.

Responses to referee #4

Dear Reviewer:

We would like to express our gratitude to you for the comments to improve this manuscript. According to your and other reviewers' comments, we have conducted further research on the issues that you suggest. Please find the specific responses and revisions shown below. They are in blue font for clarity.

Qinghua Yang, Qian Shi, Robert Ricker, Stefan Hendricks
On behalf of all the authors

Major comments:
-One of my main concerns has to do with the actual validity and usefulness of the comparison between the satellite estimates and the ULS data. As clearly stated by the authors, there are significant differences in temporal and spatial sampling. The authors even point out that the results are not consistent. I believe it would be more beneficial to the paper to focus solely on the intercomparison between Envisat and ICESat data.

Thanks for your comments. However, we think that you are biased in denying the feasibility of using ULS data as comparison data with ICESat-1/Envisat due to their relatively narrow footprint. As we know, ULS indeed measures the continuous ice draft in a fixed location with a diameter of several meters. Considering the ice motion, ULS acquired dozens to hundreds of kilometers records along the trajectory of sea ice motion on a monthly basis, which have enough spatial representativeness compared with ICESat-1/Envisat. Here, we track the source of sea ice that flows over the ULS in a specified month by backward tracking method based on NSIDC Pathfinder data sets. We find the ice draft records included in ULS monthly mean calculation come from a wide range area (Fig. 1). Therefore, we think this is enough to prove that the spatial representativeness of the monthly average ULS data can be compared with that of ICESat-1/Envisat. *(please see P8 line 229-231 in the revised manuscript)*

Besides, ULS data was generally used for ice thickness comparison in the previous studies. ULS is used for comparison with the ice thickness derived from AVHRR (Yu and Rothrock, 1996; Drucker et al., 2003). It was also used to compare with ICESat-1 ice thickness in the Fram Strait (Spreen et al., 2009). In addition, the ULS data sets have also been used for comparison with reanalyses data in the polar region (Mu et al., 2018; Shi et al., 2021). In addition, the comparison with ULS data sets is also a convention for assessing the quality of ice thickness derived from altimeters in the European Space Agency (Kern et al., 2018).

In summary, we think that the reason for rejecting us due to the spatial representativeness of ULS ice thickness is untenable. Previous studies (referred to above) have shown that using ULS for validation of satellite-derived sea-ice thickness data sets can be considered as state of the art.

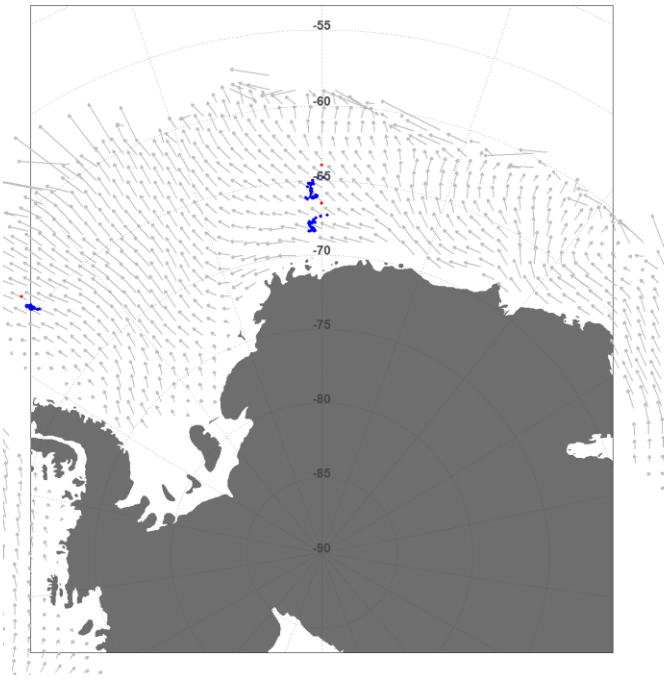

Fig. 1 The origins (30-days ago) of the sea ice (blue dots) that passing through the three ULS sites (red dots) in July 2011 by using backward tracking method based on the NSIDC v4 sea ice motion data. The grey vectors represent the monthly mean sea ice drift derived from NSIDC v4.

- Another major concern is the way that the comparison between the Envisat and ICESat-2 SIT is carried out. I think the paper would be more robust if a comparison of the actual freeboards and snow depths (total freeboards for ICESat) was introduced. The assumption made on snow depth can have a huge impact on the mean and variability of the derived sea ice thickness.

Thanks for your comments. Indeed, the involvement of snow depth can have a huge impact on the retrieved sea ice thickness. However, the purpose of this paper is to give a comprehensive and statistical comparison between Envisat and ICESat sea ice thickness data, and to highlight the importance of dealing with the possible biases of these products. Additionally, comparing the total freeboard still needs an additional snow depth product, since the radar altimeter aboard Envisat detects sea ice freeboard. Actually, we have discussed the possible impacts of the snow depth usage on the SIT retrieval in the paper.

- While the authors explored the possible causes of the observed differences between the two satellite datasets, I think this should be looked at more carefully and in more detail. Based on their uncertainty analysis, the authors conclude that most of the bias is probably explained by radar penetration issues. I do not believe that the authors successfully demonstrated this, especially given that the assumptions on snow depth and snow density are different for the two instruments.

Thanks for your comments. We need to clarify that the snow densities used by both Envisat and ICESat SIT retrieval are 300 kg m^{-3}. In addition to the radar penetration biases, we also discuss the impacts of the snow depth product and ICESat uncertainties in section 4. For modification, we conclude the sensitivity of the SIT changes to freeboard biases, snow depth biases and sea ice density in Fig. 2 by analyzing Eq. (1):

$$I = \frac{F\rho_{water} + S\rho_{snow}}{\rho_{water} - \rho_{ice}} \quad (1)$$

The sensitivities to freeboard biases and to snow depth biases are calculated by:

$$\frac{dI}{dF} = \frac{\rho_{water}}{\rho_{water} - \rho_{ice}} \quad (2)$$

$$\frac{dI}{dS} = \frac{\rho_{snow}}{\rho_{water} - \rho_{ice}} \quad (3)$$

From Fig. 2, we can see that though the magnitudes of the resulting thickness changes are quite similar, the SIT changes are more sensitive to sea ice freeboard biases than to snow depth biases. Besides, with the increase of sea ice density, the SIT changes rise. For typical sea ice freeboard biases (7 cm for the Arctic nominal adjustment suggested by Nandan et al. (2017, 2020)), the sea ice density variations induce the thickness changes ranging from ~0.5 m to ~0.8 m. For typical snow depth biases (20 cm for the monthly mean retrieval uncertainty in Kern and Ozsoy-Cicek (2016)), the thickness changes from ~0.4 m to ~0.7 m. **Although this sensitivity analysis is not solid enough for the explanation for the SIT differences in three seasons, it can provide a reasonable conjecture that freeboard biases are the main cause of the positive differences in summer and autumn.** *(please see line 319-330 and 356-368 in the revised manuscript)*

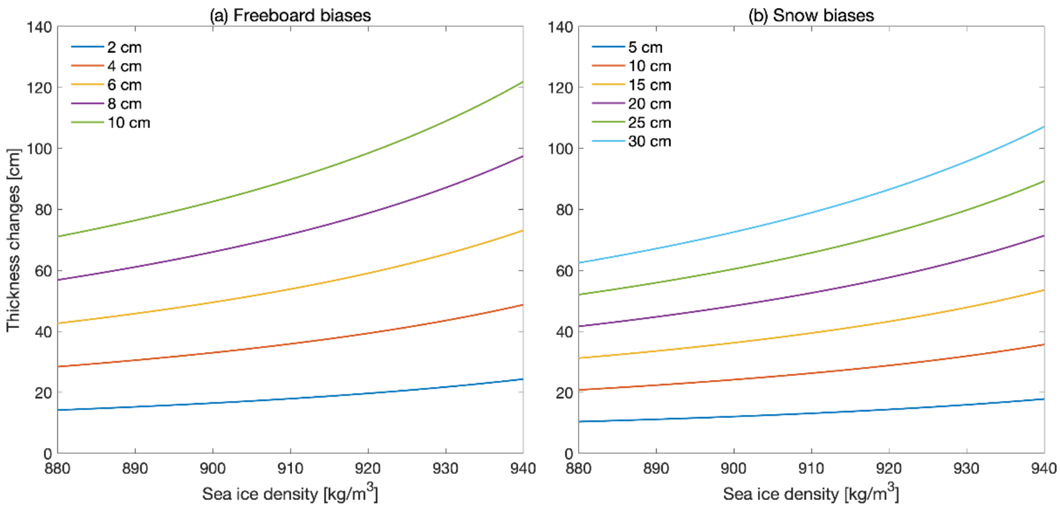

Fig. 2 Sensitivity of sea ice thickness changes to sea ice freeboard biases and snow depth biases as function of sea ice density. (a) SIT changes computed with Eq. (1) for different sea ice freeboard biases (2 cm to 10 cm). (b) Similar to (a) but computed for different snow depth biases (5 cm to 30 cm).

- Some of the phrasing needs to be reviewed carefully. Especially in the introductory part of the paper, some sentences are poorly constructed and lack clarity. It challenges the understanding of the paper.

We apologize for the language problems in the original manuscript. We have carefully amended the phrasing and modified the expressions throughout the paper.

Minor comments:

P1L9: the sentence" The crucial role that Antarctic sea ice plays in the global climate system is strongly linked to its thickness" does not really mean anything. Maybe you mean that thickness is important to evaluate the role of Antarctic sea ice in the global climate system?

Yes, we meant to point out that sea ice thickness is a critical component in assessing the role of Antarctic sea ice in the global climate system. Therefore, we focus on sea ice thickness in this study.

P1L10-11: What do you mean by “on a hemispheric scale, satellite radar altimetry data can be applied with a promising prospect”? Do you mean that large scale estimates of SIT are achievable with radar altimetry? Again revise the wording to make clearer statements.

This sentence means satellite radar altimetry can be used to achieve large-scale and long-term SIT variations, while field observations cannot.

P1L28: Replace “declines” by “decline”.

We corrected this word.

P2L59: Replace “CyroSat-2” by “CryoSat-2”.

We corrected this word.

P2L60-61: I suggest rephrasing this sentence:” The SICCI product covers the entire Antarctic sea ice for the complete annual cycle from 2002 to 2017, and it is finally a combined data set of Envisat and CyroSat-2” to “The SICCI product is derived using measurements from Envisat and CryoSat-2 and covers the entire Antarctic sea ice for the complete annual cycle from 2002 to 2017”.

We rewrote the sentence as you suggested.

P3L76:” This data set has been investigated for many years”. I believe this dataset has been used in several investigations, not investigated.

We modified the sentence: “This dataset has been used in previous studies for many years.”

P4L94:”between the two datasets” please specify that you are referring to the satellite data.

We changed it to “the inter-comparisons between the two satellite data sets”.

P5L127: Replace “are conducted with” by “are characterized by”

We amended this phrase.

P6L163: Replace “derived” by “from”

We removed “derived”.

P6L171: Please revise:” For each period, we choose the corresponding time period during which Envisat monthly data are used”.

We revised the sentence: “For each ICESat operating period, we choose the corresponding Envisat monthly data.”

P6, L175-177: Please revise :” The weighting has taken into account periods where only Envisat SIT of one month are present, i.e., we use this equation for grid cells where we have valid SIT data from both months, while we only use the Envisat SIT of the respective month without weighing for those grid cells where we only have valid data from either month.”

We simplified the sentence: “We use this weighing equation only for grid cells where valid Envisat SIT data exist in both months, while the weighing is not conducted for grid cells where valid data only exist in either one month.”

P8L236: I suggest to replace “Envisat does not show the young ice in the Ross Sea” by “Thin ice in the Ross Sea is not captured by Envisat”.

We rewrote the sentence as you suggested.

P9L244-255: Revise “Compared to summer, the differences in the western Weddell Sea spread to the whole Weddell Sea sector and decrease from west to east.”. The statement is not clear.

We clarified the sentence: “Compared with summer, the positive differences in the western Weddell Sea expand to positive differences over the whole Weddell Sea sector, and the differences decrease from west to east.”

P12L345: Replace “Previous study reveals” by “Previous studies show”.

We corrected this word.

P14L389: Remove “Firstly”. The comparison to ULS data is carried out first.

We removed this word.