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.

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.

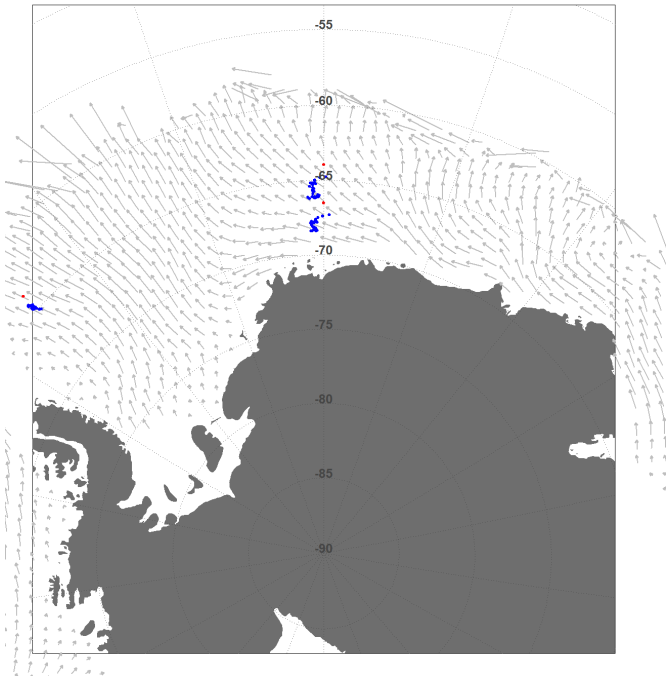

Fig. 1 The 30-day origins of the sea ice passing the three ULS sites in July 2011. The red dots stand for the ULS locations and the blue dots stand for the original locations of the sea ice using backward tracking method.

- 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 uncertainties 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.

- 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.

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 causes 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.**

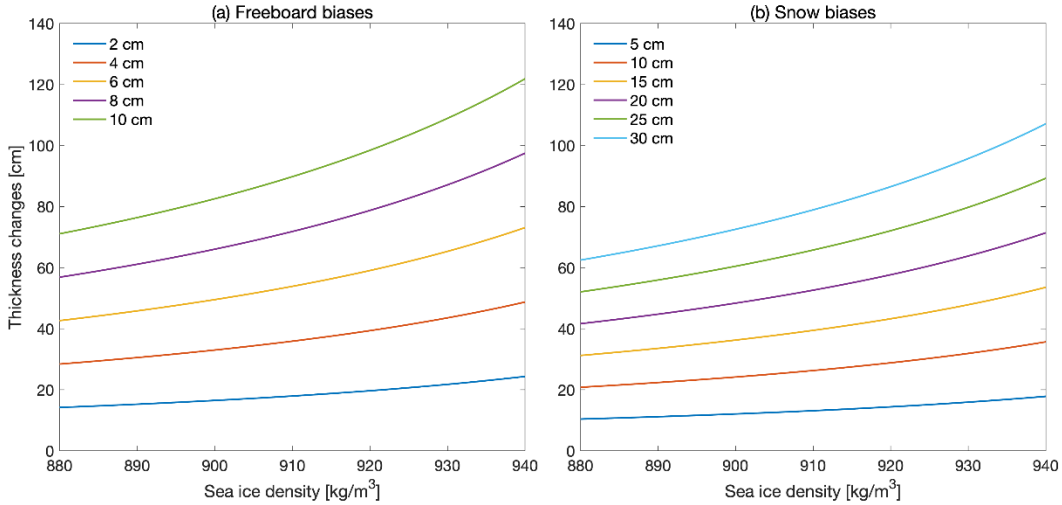

Fig. 2 Sensitivity of sea ice thickness changes to sea ice freeboard biases and snow depth biases as function of sea ice density. (a) SIT changes computed with Eq. (1) for different sea ice freeboard biases (2 cm to 10 cm). (b) Similar to (a) but computed for different snow depth biases (5 cm to 30 cm).

- 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 several investigations.”

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 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 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 to summer, the positive differences between Envisat and ICESat SIT in the western Weddell Sea turn 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.