Review of

"Assessment of Arctic Sea Ice Thickness Retrieval Ability of the Chinese HY-2B Radar Altimeter"

By Dong et al.

The manuscript presents a feasibility study of using the China Ocean Dynamic Environment Satellite Series 2b (HY-2B) radar altimeter for estimating sea ice thickness. The study compares the retrieved radar freeboards, derived sea ice freeboards (using various snow depth products), and derived sea ice thickness with auxiliary data in the form of either airborne observations (Operation IceBridge) or other satellite observations (using CryoSat-2 (CS2) or ICESat-2 (IS2) observations).

At the current state of the manuscript, I believe that more work is necessary before publication. The study could benefit from additional analysis, since the methodology is not convincing, which the results also illustrate. Furthermore, the readability of the manuscript is low, and I believe the figures require more work, to better present the results. Overall, I believe a manuscript that shows the retrieval ability of HY-2B would be beneficial, as it could provide yet another means of estimating thickness from altimeters. Unfortunately, I do not believe that the paper at its current state has presented the retrieval ability of HY-2B, thus requiring major revisions and additions prior to any publication.

General comments

I am not convinced of the soundness of the freeboard retrieval methodology presented in this study. While they refer to the study of Kwok et al. (2007) for the retrieval methodology, that study used laser observations rather than radar, which are not as impacted by speckle and noise as radar observations. They also refer to the study of Zhang et al. (2021) that applied a similar (although not entirely same processing chain) to Envisat freeboards, and produced reasonable results, however I am in earnest also not fully convinced of their methodology either. There are some assumptions that I question, and I therefore hope that the authors can present some more results regarding the validity of their methodology – and why they decide to use this methodology instead of the more commonly used methodology of classifying waveforms as either leads/floes based on shape (e.g., Tilling et al. 2018, Ricker et al. 2014, Hendricks and Ricker 2020). This includes for example:

- Can you be sure that the lowest 9 points are in truth from leads? What if all points within the segment are from floes (or negative freeboards, e.g., Figure 4)? If you select them as leads, you will bias the SSHA.
- How big of an impact do you estimate this SSHA methodology to make (e.g., the behavior you point out in Figure 5b+d+f+h)? I appreciate the authors work on investigating different 'SSHA determination schemes', but this does not really question the overall method of using the lowest points, which are to most likely noise/speckle.
- And how can you be sure this methodology is applicable for radar freeboards – especially as you are already observing this biasing of the HY-2B SSHA?

I appreciate that the authors have mentioned in the conclusion that using lead/floe discrimination is expected to be future work, but if you are already aware of several limitations of the applied methodology (and your results also show this in the comparison with Operation IceBridge data), then I would expect a more exhaustive discussion on the applicability of this method.

I do believe that this paper would benefit from investigating the actual waveforms of the radar altimeter to assess the sea ice thickness retrieval ability. Other papers that have assessed such for new sensors (e.g., AltiKa or Sentinel-3) have both employed a lead/floe discrimination algorithm and derived freeboard from here (Armitage and Ridout, 2015; Shen et al., 2020). I encourage the authors to conduct such an analysis in case the radar waveforms are available for the authors. It would not require the authors to compute elevation afterwards using a new or different re-tracker but would simply provide a way of filtering and selecting 'true' leads (after validating the analysis with e.g., SAR or optical data). From there, they can use the leads to interpolate between and compute the SSHA in a more robust manner. This will also likely provide more comparable freeboard observations. I think such an analysis is crucial to assess HY-2B's retrieval ability.

The study also does not discuss much regarding why CS2 and HY-2B could be different, but merely states that they are. Several aspects could contribute to this e.g., change in footprint or retrieval methodology (re-tracking and SSHA/freeboard retrieval). I do not believe the authors have discussed this aspect thoroughly enough to truly show that HY-2B is 'reliable and consistent with CS2', even if they make a short sentence regarding re-trackers in the conclusion. However, studies have shown that change of re-tracker alone for CS2 introduce significant differences (e.g., Ricker et al. 2014, Landy et al. 2021 etc.), so simply comparing freeboard products from two different satellites using different methods all together warrants at least a discussion.

Overall, I believe the discussion must be expanded. The results of the study are primarily presented, but not discussed in further detail including the aspects that are causing some of these large discrepancies. Especially the less encouraging results with OIB and discrepancies between both HY-2B/CS2 and HY-2B/IS2 should be discussed rather than only presented. What are causing these differences? What do we expect? What will impact this – spatial/temporal coverage? Measurement modes? There are many differences between CS2 and HY-2B that could be causing the discrepancies, and while there is some consistency between H2-YB and CS2, this should still be discussed. It seems the authors are focusing more on positive/agreeable results that the more disheartening ones.

I also wondered why the authors did not compare with Sentinel-3, since Sentinel-3 has the same coverage (till 81.5°N/S) as HY-2B? I know that Sentinel-3's dedicated Sea Ice preprocessors have become available in August 2022 (https://scihub.copernicus.eu/dhus/#/home, they can be found by selecting the product types "SR_2_LAN_HY", "SR_2_LAN_SI", "SR_2_LAN_LI" (respectively for Hydrology, Sea Ice and Land Ice), and that they will reprocess whenever possible aiming for end 2022. It might be that it is not available for the period of the study and in case they are not, this could be a suggestion for a future comparison study as well.

Specific comments

Overall

There appears to be some issues with the referencing/citation software used throughout the article. In several instances (e.g., line 150, "Rasmus Tonboe et al. (2016)"; line 158, "Signe et al., (2021)") the citation is written with either both first and last name of the first-author, or by only first name of the first author. This should be corrected – while it may be an issue with LaTeX, I believe it is the responsibility of the author to correct this.

I also encourage you to do a thorough proof-reading and edit, especially in the results, discussion, and conclusion sections. There are many unnecessary repetitions, and by removing these and constraining the text, it would improve the readability of the manuscript.

All acronyms should be checked whether they are properly named and that they are explained when appearing first in the text. I have written some examples in 'technical corrections', but the authors should check the entire manuscript.

Figures and tables

The manuscript includes a lot of figures and tables. I encourage the authors to consider ways to limit the figures/tables and more concisely present the results. Also, some figures could be provided in an appendix or as supporting information.

Not all figures include sub-figure labels (e.g., Figure 4). Also, several figures have sub-figure labels given in the x-axis labels – I encourage the authors to include the sub-figure labels in the top left corner instead. However, in case they keep the current format, they should include a space between the sub-figure label and the x-axis label. Also, the manner of which units are displayed in the labels and in legends do not follow The Cryosphere guidelines (e.g., currently have no spaces between the units and the numbers).

Figure 6, 8, 14, 15: Consider a different colormap to ensure readers with color vision deficiencies can correctly interpret your findings (as per The Cryosphere submission guidelines). 'Jet' or 'rainbow' are not encouraged as colormaps. Also, I think the maps are too small – it's hard to see what is going on. I suggest enlarging the maps and use only 1 color-bar for the entire map. This will reduce the space between the maps, allowing to increase the size too.

Figure 1 + 2, 7 + 9, and 10 + 11: I suggest combining the figures – just to reduce the sheer number of figures. It would just generate several subplots. This would of course require some slight alteration of the text to ensure that the sequence of which figures are referenced are still in the correct order.

Open data and research

I was not able to access the FTP folder with the SDGR HY-2B data; neither using explorer nor filezilla. Therefore, I was not able to investigate the data available nor whether potential lower-level products were available for the authors to conduct a waveform analysis. I encourage the authors to ensure the data is available through the link, and/or provide the actual data used in the publication through a database (raw and processed data if possible) with associated DOI. I also encourage the authors to make the code available for reproducibility and transparency.

However, I thank the authors for providing the link to all other data including date of access! Refreshing to see in a publication and right in line with the expectations of the science community regarding open research.

Abstract

I encourage the authors to shorten the abstract. 7 lines about the general background of satellite altimetry and using more than 10 lines on a summary of the results seems a bit too much. At the same time, I believe only the main results (with one or two numerical results) would be enough to underline the points made. Future work/limitations of study should be provided in the abstract as well.

Introduction

Line 52: "Sea ice thickness, as the third dimension of sea ice (…)": what is meant by sea ice density here? Surely, sea ice thickness can in combination with sea ice area be used to calculate the sea ice volume, but sea ice density cannot be combined with sea ice thickness to compute sea ice volume.

Line 68. "Shen et al. (2020) (…)": Lawrence et al. (2019) showed that this was mostly a result of the processing chain of Sentinel-3 not having included zero-padding or Hamming-weighting. I believe this should be mentioned, since the study of Lawrence et al. (2019) in which these corrections were applied showed greater consistency.

Data

Line 105. What type of re-tracking algorithm is used and what are potential impacts of using this re-tracker (limitations that other studies have shown? etc.)?

Line 113-117: Baseline-D ESA and AWI use different re-tracking algorithms which will have different results. As such, your comparison is not 'consistent' in the way that the products are not the same and any differences you observe between the HY-2B and AWI/ESA could be different depending on re-tracker. A suggestion could be to investigate the Climate Change Initiative (CCI) products, where both trajectory and gridded products are available – and they use the same processing chain (ftp://ftp.awi.de/sea_ice/projects/cci/crdp/v3p0-rc1/cryosat2/nh/).

Line 162. How come you use DTU18MSS and not the new (and likely improved) DTU21MSS? Available here (https://doi.org/10.11583/DTU.19383221.v1): https://data.dtu.dk/articles/dataset/DTU21_Mean_Sea_Surface/19383221

Line 164. What is meant by "this model can precisely determine the instantaneous elevation of lead"? I do not believe that the referenced study (Skourup et al., 2017) has stated this. A mean sea surface represents the sea level due to constant phenomena and represents the position of the ocean surface averaged over an appropriate time period to remove annual, semi-annual, seasonal, and spurious sea surface height signals – I am not sure what you mean by the statement. Surely, if the model could determine the instantaneous elevation of a lead, we could just interpolate the model along the floe observations and not care about the individual lead observations along the track? The sentence should be corrected.

Line 190. You already have a lot of tables. Instead of providing a table with the exact densities for each month, instead you could consider just writing the equation in the section.

Methodology

Section 3.1: Figure 3 shows a removal of 'outliers' using ±1 m, but this is not mentioned in the text.

Line 198. What is meant by "invalid values". How are they considered invalid?

Line 198-201. Which exact corrections (providers) are you using? This should be stated.

Line 207-213. The purpose of removing the residuals should be stated. This is not commonly applied for the lead/floe retrieval methodologies, and as such can add to the reasons why the results differ.

Line 215. Why are you assuming the SSHA to be 0? And what is the impact of this? Figure 4+5 show that SSHA is not necessarily zero (zero-elevation is only relative to MSS). Instead, could you consider nan (not-a-number) the data and just interpolate where data is not available?

Line 226. Isn't the 0.22 value depending on snow density and thus varying, even in the AWI product? It is based on slower wave propagation, given by $\left(\frac{c}{c_s} - 1\right)$, where c is the speed of light and c is the speed of light through snow. Here, $c_s = c \cdot \eta_s$, where η_s depends on the snow density (which you vary when using Mallet et al. (2020)). If so, this should be corrected. If not, please explain exactly what this value is and where it comes from.

Line 229. How is the gridding performed?

Results

Line 237-249: "Fig 5(a) (….)." This entire paragraph could benefit from some rephrasing, as it was confusing and long to read. Also, check the manuscript guidelines for The Cryosphere again to ensure that the way you write the different expressions ("0$\pm$0.31m") follows the guidelines.

Line 264-265. I think you need to consider what you mean by 'smallest' and 'largest' deviation. Sure, in actual numbers, -2.4 cm is smallest, and 1.9 cm is largest. But the deviation from zero must mean that the -2.4 cm is largest! So, this needs to be rephrased for clarity.

Line 336 -358. You mention the Kwok snow depth product is used, but in Figure 3 you mention that the AWI snow depth is used for the processing of the radar freeboards. Then, later you mention that the Kwok snow depth is used for the ICESat-2 freeboards to obtain the sea ice freeboard and for the CS2 to obtain sea ice freeboard. How come you decide to use this snow depth product suddenly, and not just use the AWI snow depth throughout?

Line 274-275. The fact that HY-2B modal freeboard is thicker than the mean is not expected – we usually expect a skewed distribution with an abundance of thinner freeboards compared to the thicker freeboards. This should be discussed further.

Line 337-338: Be careful when discussing spatial resolution of IS2. The footprint of IS2 is approximately 11 m, yes, but the ATL10/ATL20 products are processed data that is not of this spatial resolution.

Line 358. You mention sea ice thickness error? What is meant by this – this has not been discussed previously.

Figure 13. For almost all subfigures, we observe HY-2B to generate significantly thicker sea ice thickness than IS2 (CS2 also does not observe this behavior). This seem to showcase that some overestimation of thickness from HY-2B, which I think could be discussed more in the text for the IS2 comparison.

Discussion

Line 371-374. Long and confusing sentence, however it sounds like it is saying that in general HY-2B is observing smaller freeboards than the AWI freeboards – but, when comparing otherwise you said that HY-2B was higher than CS2 freeboard (I assume this is the ESA observations then). I suggest you rephrase this sentence for clarity, and then discuss further the fact that you are comparing observations with different re-trackers – what can you really get from this? And are you comparing along-track or gridded? This should also be mentioned.

Line 386. If you are following Mallet et al. (2020), the 0.22 value should be changing with density.

Line 393, Equation 8. I am intrigued by why you are including the impact of $+0.22\rho_w$, when e.g., the study of Ricker et al. 2014 did not. This may explain my comment further down regarding line 415-425. I suggest you check up on this to make sure that this is the correct partial derivative.

Line 404. The uncertainty of sea ice densities should be cited.

Line 415-425. Something seems odd in the uncertainty determination. Ricker et al. (2014) determined freeboard uncertainties of 6 cm (12 cm) for FYI (MYI), but for sea ice thickness, they estimate 60 cm (120 cm) (increase by factor 10). This is not observed in your estimation, where the CS2 sea ice thickness uncertainty is <u>only</u> 12.1-15.4 cm. Also, your freeboard uncertainties to thickness uncertainties do not increase by factor 10 for HY-2B observations? Something seems off here. I suggest going through this again and ensuring that the units are correct.

Conclusion

Line 442-44: I am not convinced by this methodology, when the comparisons and correlations with OIB are so low compared with CS2. A correlation of 0.58 for HY-2B but 0.84 for CS2 for freeboards, and 0.41 for HY-2B but 0.80 for CS2 on thickness is quite a big difference – and for me just shows that the method used here is not that applicable for sea ice, and not reliable.

Line 460-468: A lot of new information provided here which is not discussed or even commented upon in the text. I strongly encourage the authors to include a discussion (in the discussion section) on the impact of re-tracking methodology applied to the HY-2B freeboards (and how it differs from the CS2 re-tracking). There are also some considerations for what should be done in future work, which are not discussed in the text but only here. Finally, the authors conclude that they have used a sub-optimal re-tracking algorithm which is <u>only</u> applicable for ocean (not sea ice?). Makes me question why they did this analysis with this algorithm, when it is not suitable, and the results clearly show this too.

Technical corrections

Line 113: "main" – not sure they are the only main products? The CCI product could be mentioned here. Furthermore, the upcoming releases of CryoTEMPO (http://cryosat.mssl.ucl.ac.uk/tempo/downloads.html) are expected to be a favorable product to be used in the future by the science community.

Line 116. What is meant by L2I? The intermediate product is not provided to the public, but rather the L2 product is.

Line 123. “were by “ -> “were provided by”

Line 148. “Denmark Technical University (DTU)” -> “the Technical University of Denmark (DTU)”

Line 171. “W99” – “Warren et al. 1999 (W99)”: reference the publication and explain the acronym.

Line 171. “AMSR-2” -> “Advanced Microwave Scanning Radiometer-2 (AMSR-2)”

Line 202. “were interpolated” -> “were linearly interpolated” (or however it was interpolated)

Line 204. “was not used” -> “was not considered”

Line 206-207. “from the obtained surface elevations of ground objects (…)”: consider rephrasing this sentence, as it is not clear.

Line 215-216. “(…) were subtracted from the SSHA to obtain”: The SSHA are subtracted from the segment height, not the other way around. This sentence should be rephrased.

Line 231: “$\rho_{seawater}$” -> “ρ_w” (following the equation)

Line 298-299. “Because the (…)”: I wouldn’t say that the lower SSHA necessitates higher freeboards, because they are only higher if the elevations of the floes are of the same height as CS2 or higher. If the floe elevations were also lower, HY-2B would not necessarily have higher freeboards. Consider rephrasing it.

Line 451-452. You mention sea ice error again, but you have merely referenced this by the paper of Ricker et al. (2014) and not done more sensitivity studies on this. I suggest removing this sentence or discuss this more in the text if you aim to keep it.

References

Armitage, T. W. K., and Ridout, A. L. (2015), Arctic sea ice freeboard from AltiKa and comparison with CryoSat-2 and Operation IceBridge, Geophys. Res. Lett., 42, 6724– 6731, doi:10.1002/2015GL064823.

Hendricks, S. and Ricker, R. (2020): Product User Guide & Algorithm Specification: AWI CryoSat-2 Sea Ice Thickness (version 2.3)

Lawrence, I.S., Thomas W.K. Armitage, Michel C. Tsamados, Julienne C. Stroeve, Salvatore Dinardo, Andy L. Ridout, Alan Muir, Rachel L. Tilling, Andrew Shepherd, (2019) Extending the Arctic sea ice freeboard and sea level record with the Sentinel-3 radar altimeters, Advances in Space Research, Volume 68, Issue 2, doi: 10.1016/j.asr.2019.10.011.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., and Davidson, M.: Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation, The Cryosphere, 8, 1607–1622, https://doi.org/10.5194/tc-8-1607-2014, 2014.

Shen , X., Chang-Qing Ke, Hongjie Xie, Mengmeng Li & Wentao Xia (2020) A comparison of Arctic sea ice freeboard products from Sentinel-3A and CryoSat-2 data, International Journal of Remote Sensing, 41:7, 2789-2806, DOI: 10.1080/01431161.2019.1698078

Tilling, R.L., Andy Ridout, Andrew Shepherd, Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data, (2018) Advances in Space Research, Volume 62, Issue 6, doi: 10.1016/j.asr.2017.10.051.