Reviewer's comment is in italic, response is in normal font.

Review for The Cryosphere Discussions, *https://www.the-cryosphere-discuss.net/tc-2019-192* ***Multidecadal Arctic sea ice thickness and volume derived from ice age***

Liu et al., 2019

The study of Liu et al. (2019) introduces an Arctic-wide sea ice thickness and volume data product and retrieval method derived from sea ice age. Their product extends all the way back to the early 1980s and presents a data set created with a consistent method, thus providing an interesting novel addition to the existing sea ice thickness products. In addition to complementary information to the more recent satellite altimetry based products, the product could bring additional information about the conditions before the more systematic period of satellite altimeter sea ice measurements.

However, the manuscript currently lacks clarity and detail in explanation of some of the implemented methods. In particular the validation of the product should be improved and justified in order to prove the usefulness of the product. In addition there are some minor cases, included in the review comments below, which should be corrected to accomplish a more finished manuscript.

Considering the novelty and added value in extending the satellite based sea ice thickness records, I recommend this manuscript to be considered for publication in the Cryosphere, after addressing the major review comments.

We appreciate the reviewer's critical evaluation and constructive suggestions. All of the reviewer's comments have been addressed. All the responses are included in the revised manuscript. We believe revisions responding to reviewer's comments make the manuscript better.

General comments:

1. The data and methods section lacks clarity. There is a great number of data sets used for creating the product and then those used for validation/comparison. And some of the products are used for both purposes. And not really in a chronological order. It would improve the readability if you could structure this section so that it is clear for which purpose the data sets are used, maybe adding separate sections for datasets used in IceAgeDerived creation and for validation data.

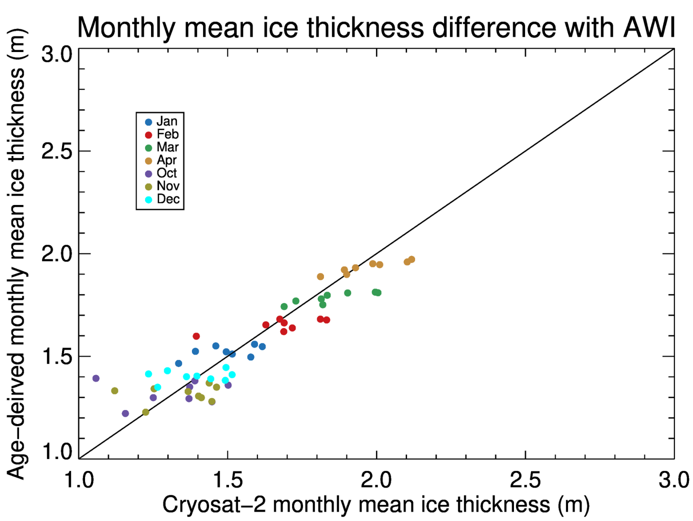
The data and method section is updated and re-structured. Detailed information of all the data sets used in this study is added, with information on which data set is used for algorithm development and which data set is used for evaluation/validation, and which is used for both purposes. Subsections are added as the reviewer suggested.

2. You seem to use ICESat data as one set of validation data, which is always a bit suspicious if you are using it to construct your data set. The same applies to the draft data. You could either remove the comparisons to these completely from the *results or really emphasise and justify more, what comparisonal value these bring.*

Another reviewer raised the same questions. We added text to highlight the limitations of such comparisons the reviewer referred to. In the revised manuscript, besides the comparison to ICESat data and draft data, we added independent validation data sets, Cryosat-2 products from NASA GSFC, AWI, and CPOM. Comparisons of IceAgeDerived ice thickness and ice volume with those from Cryosat-2 are included in the revised manuscript.

A comprehensive assessment of the IceAgeDerived ice thickness and ice volume against Cryosat-2 has been carried out. The IceAgeDerived ice thickness and volume are compared to monthly mean Cryosat-2 ice thickness from AWI, NASA GSFC, and CPOM 2011-2018. The following figures, Figure 1 and 2, show the scattering plots of the comparisons, with statistics shown in Table 1. The monthly mean ice thickness shown in the figures is the mean of ice thickness of all pixels in the Arctic.

It shows the IceAgeDerived has slightly smaller monthly ice thickness and volume compared to AWI Cryosat-2 products from January to April, and from October to December, with overall means (standard deviations) of -0.02 m (0.11) and -0.76 10^3 km^3 (0.86). Comparison to NASA GSFC Cryosat-2 products shows the largest negative bias in those months, with overall means (standard deviations) of -0.27 m (0.15) and -1.79 10^3 km^3 (0.95) for ice thickness and ice volume respectively. The negative biases to CPOM Cryosat-2 products are in between. Please note, both AWI and CPOM have holes surrounding North Pole not filled, while NASA GSFC fills those holes. We only compared where both products have valid values. Also, you can see the spread between the different Cryosat-2 products.



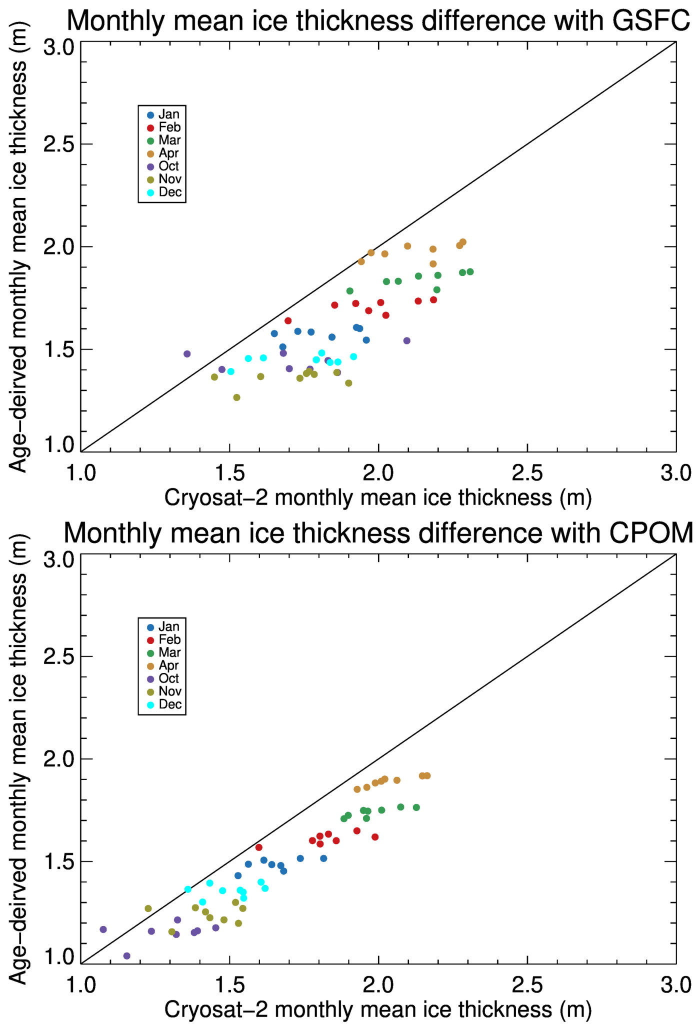
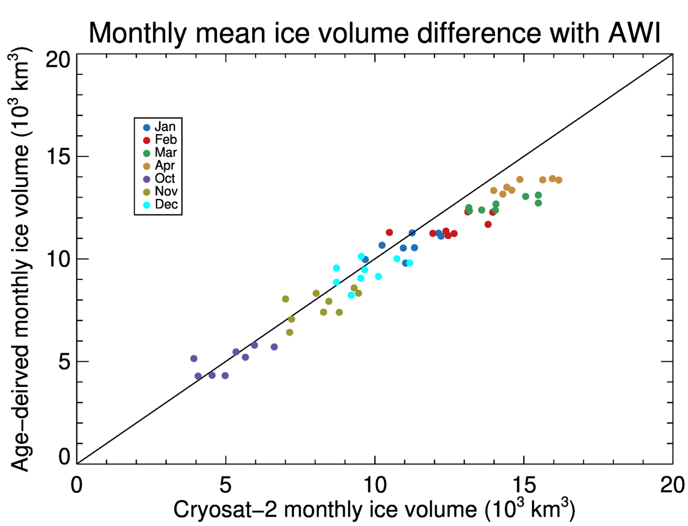


Figure 1: Scattering plot of IceAgeDerived monthly mean ice thickness and Cryosat-2 monthly mean ice thickness from AWI, NASA GSFC, and CPOM.



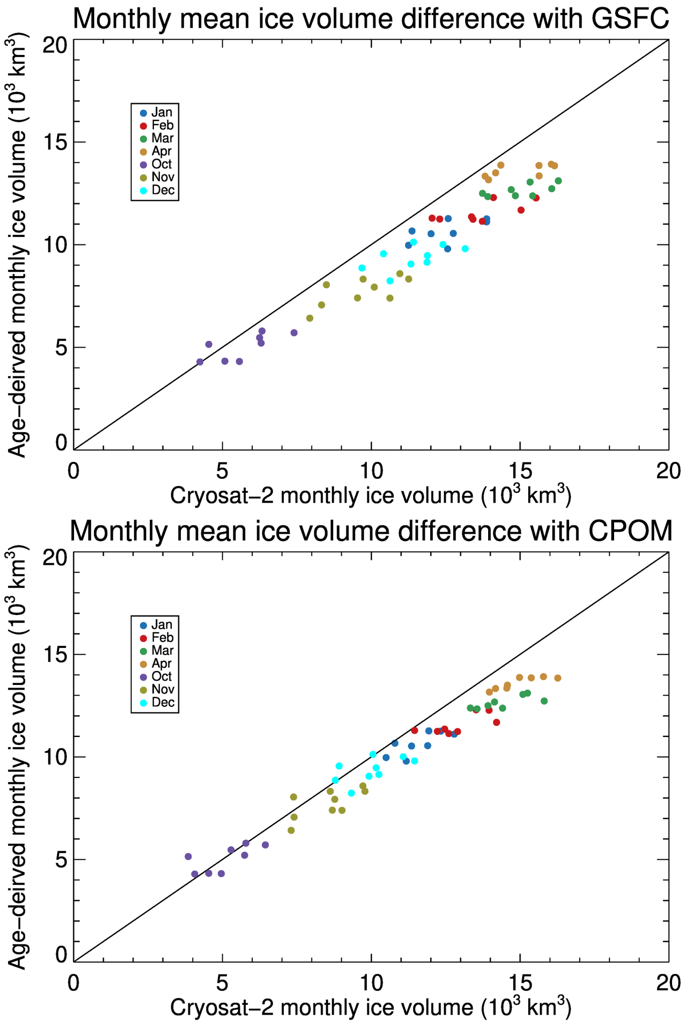


Figure 2: Scattering plot of IceAgeDerived monthly ice volume and Cryosat-2 monthly ice volume from AWI, NASA GSFC, and CPOM.

Table 1: Comparison of monthly ice thickness and ice volume between IceAgeDerived and Cryosat-2.

		AWI	NASA GSFC	CPOM
Comparison of monthly ice thickness of IceAgeDerived and Cryosat-2, 2011-2018 mean (standard deviation) in m	Mean	-0.02 (0.11)	-0.27 (0.15)	-0.18 (0.09)
	January	0.02 (0.09)	-0.24 (0.12)	-0.17 (0.08)
	February	-0.03 (0.11)	-0.27 (0.13)	-0.21 (0.10)
	March	-0.06 (0.09)	-0.30 (0.11)	-0.24 (0.07)
	April	-0.03 (0.08)	-0.14 (0.11)	-0.14 (0.06)
	October	0.00 (0.16)	-0.27 (0.22)	-0.14 (0.12)
	November	-0.03 (0.12)	-0.35 (0.14)	-0.19 (0.11)
	December	0.01 (0.10)	-0.29 (0.14)	-0.18 (0.09)

Comparison of monthly ice thickness of IceAgeDerived and Cryosat-2, 2011-2018 mean (standard deviation) in 10^3 km^3	Mean	-0.76 (0.86)	-1.79 (0.95)	-0.98 (0.81)
	January	-0.46 (0.64)	-1.89 (0.80)	-0.95 (0.51)
	February	-1.03 (0.87)	-2.12 (0.94)	-1.35 (0.68)
	March	-1.61 (0.74)	-2.39 (0.76)	-1.79 (0.68)
	April	-1.38 (0.59)	-1.37 (0.83)	-1.35 (0.55)
	October	-0.11 (0.66)	-0.68 (0.73)	-0.05 (0.66)
	November	-0.46 (0.76)	-1.94 (0.87)	-0.80 (0.71)
	December	-0.35 (0.75)	-1.79 (0.95)	-0.98 (0.81)

3. Uncertainties are sometimes painful, but they could be handled more systematically. You mention some, but there is very little analysis. In the data section there are some uncertainty estimates for OTIM, but not really for the other data sets. In results there are brief mentions of ICESat and CryoSat-2 uncertainties. And you mention significance levels for ice thickness and volume trends. Adding more discussion and quantifying the uncertainties in a comparable manner, as well as stating seasonal differences in uncertainties, perhaps adding some discussion on the possible biases from using submarine vs. laser altimeter in the ice age derived thickness, would add a nice touch to the manuscript.

The uncertainties of the ICESAt and Cryosat-2 are added in the revised manuscript through literature review. More quantitative analysis of the uncertainties are also included as detailed in the response to comment #2.

"Each of these ICESat and CryoSat-2 ice thickness products has its uncdertainty. The major contributors of these uncdrtainteis are uncertainties in snow depth and snow density, and overall uncertainty in ice thickness is estimated around 0.7 m for ICESat (Kwok and Cunningham 2008). Kwok and Rothrock (2009) estimated the ICESat ice thickness uncertainty around 0.37 m. Comparions with in situ ice thickness observations show unbiased icd thickness estimation in CPOM CryoSat-2 ice thickness, with uncertainties from 34 cm to 66 cm, and error analysis shows the uncertianteis in Arctic-wide sea ice volume are typically about 13.5% (Tilling et al. 2017). Comparsion of NASA GSFC CryoSat-2 ice freeboard to IceBridge data shows a rms difference range from 7.4 to 11.1cm in ice freeboard retrievals(Kurtz et al. 2014). The percentages of ice thickness uncertainty to the ice thickness from AWI CryoSat-2 monthly mean ice thickness from 2011-2018 range from around 35% at mean thickness at 1.4 m to around 20% at mean thickness at 5 m (Figure A1 in appendix). "

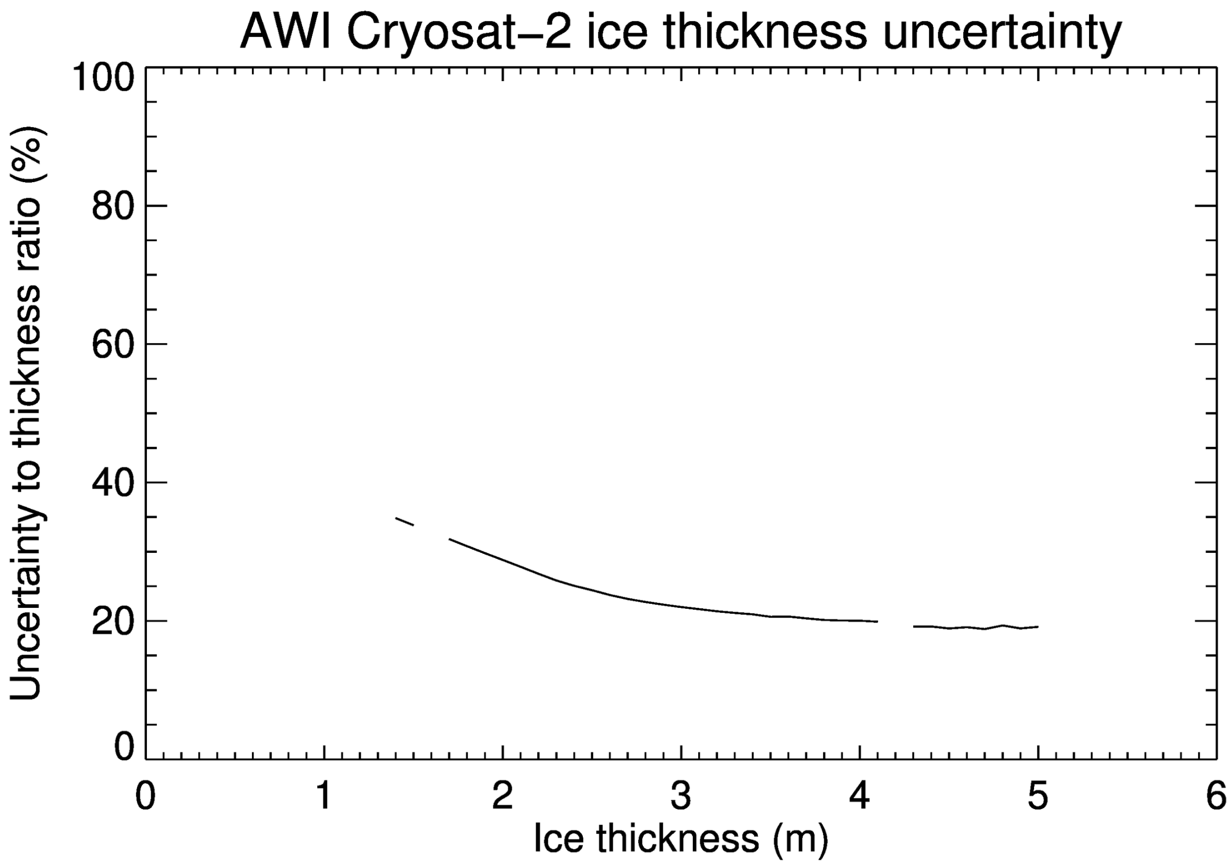


Figure A1: The percentage of uncertainty to sea ice thickness in AWI CryoSat-2 monthly mean ice thickness 2011-2018.

4. *The results, particularly the comparisons with different data sets, should be discussed in detail. Currently the statement about the usefulness of IceAgeDerived is not made that clear. PIOMAS and OTIM seem to be used here as the main comparison sets, and they are good in a sense that both extend to the early 1980s, but to reason the usefulness of IceAgeDerived, you could consider using a satellite altimetry observation based thickness data set with a good temporal extent. I would see some of the main users for the IceAgeDerived being those who are already using altimetry data for sea ice thickness and volume, and thus it would be good to see how these two compare over a longer time period. There are for example datasets combining EnviSat and CryoSat-2, where efforts have been made to bring these to a level of consistency. Such data sets are provided at least by CTOH/LEGOS and ESA CCI. This is only a suggestion for the comparison data, but in case you decide to stay with PIOMAS and OTIM, it would be good to add explanation why you chose these, what are they good for and what do the comparisons really tell about the usefulness of IceAgeDerived.*

We thank the reviewer's good point. Besides the comparison to Cryosat-2 as shown above, we also added similar comparison with EnviSat from 2003 to 2010. The results are shown below and also included in the revised manuscript. Because the spatial coverage of EnviSat and Cryosat-2 are different, different size of hole size without data near the North Pole, we keep these two comparisons separate.

Similar as the comparison to Cryosat-2 product, assessment of the IceAgeDerived ice thickness and ice volume again EnviSat ice product has also been carried out. The EnviSat ice product is from the European Space Agency's (ESA) Climate Change Initiative (CCI) version 2 product, http://cci.esa.int/content/cci-sea-ice-dataset-release-sea-ice-thickness-v20. The IceAgeDerived ice thickness and volume are compared to monthly mean EnviSat ice thickness from 2003-2010. The following figures, Figure 3 and 4, show the scattering plots of the comparisons, with statistics shown in Table 1. The monthly mean ice thickness shown in the figures is the mean of ice thickness of all pixels in a month. It shows the IceAgeDerived has comparable monthly ice thickness and volume to ESA CCI EnviSat products in all months, with overall means (standard deviations) of 0.07 m (0.10) and -0.08 10^3 km^3 (0.57).

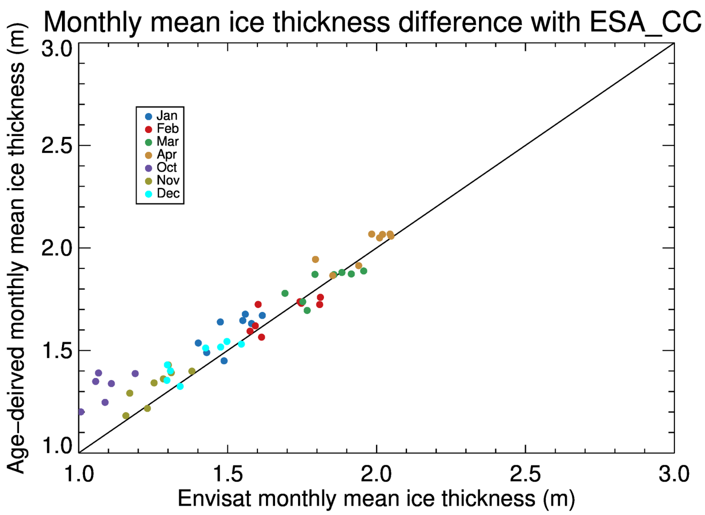


Figure 3: Scattering plot of IceAgeDerived monthly mean ice thickness and Envisat monthly mean ice thickness from ESA CCI.

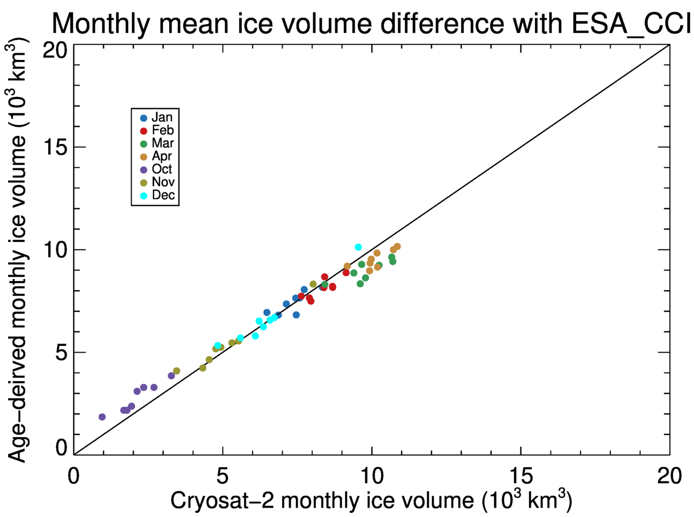


Figure 4: Scattering plot of IceAgeDerived monthly ice volume and EnviSat monthly ice volume from ESA CCI.

Table 2: Comparison of monthly ice thickness and ice volume between IceAgeDerived and Cryosat-2.

		AWI
Comparison of monthly ice thickness of IceAgeDerived and EnviSat, 2003-2010 mean (standard deviation) in m	Mean	0.07 (0.10)
	January	0.08 (0.06)
	February	-0.00 (0.06)
	March	-0.00 (0.06)
	April	0.04 (0.05)
	October	0.24 (0.11)
	November	0.06 (0.05)
	December	0.05 (0.05)
Comparison of monthly ice thickness of IceAgeDerived and EnviSat, 2003-2010 mean (standard deviation) in 10^3 km^3	Mean	-0.08 (0.57)
	January	0.05 (0.34)
	February	-0.23 (0.28)
	March	-0.84 (0.44)
	April	0.67 (0.24)
	October	0.23 (0.23)
	November	0.13 (0.31)
	December	-0.09 (0.57)

Minor comments/edits:

L23: Have declines -> have declined

Corrected.

L27-29: There could be more sources, perhaps making a stronger statement with results based on satellite observations, if possible. And there could be something newer for the model results, as 2002 was almost two decades ago.

Added a new reference.

L34: The "relatively high quality of sea ice concentration retrievals from passive microwave data", relative to what?

Deleted "relatively".

L39-40: Not mentioning EnviSat? It covers almost a decade of historical data. Of course an exhaustive list might be unnecessary here, but you could consider adding "e.g." if only mentioning CryoSat-2 from the radar altimeters as now it sounds like CryoSat-2 is the only source.

Added Envisat, also added a new reference. Analysis using the Envisat products is also added in the revised manuscript.

L52: Maybe a newer source than Wang et al. 2010? In Section 4, Discussion and Conclusions, you mention the new snow products and their remaining uncertainties, so perhaps something from there.

A new references are added.

L55: Does Laxon belong here? And rather many references for PIOMAS?

Deleted this reference.

L58-68: Nice paragraph!

L59: Individual sea ice parcels ?

Added.

L79: Masnalik et al. 2007 -> Maslanik et al. 2007

Done.

L81-83: Each grid cell is tracked as independent parcel, but age of a grid cell of parcels with different ages is assigned to this parcel? The latter sentence in these lines could be more clear.

Changed to "Sea ice thickness can also be derived from sea ice age. An Arctic sea ice age product covering the period from 1984 to the present has been generated based on

Lagrangian tracking of individual sea ice parcels (Tschudi et al. 2019a). Each parcel is tracked independently, and the oldest age of all possible ice parcels within each grid cell is assigned to the cell."

L117-123: Confusing section, it is a bit unclear what you mean with the "interannual change with the annual cycle superimposed in averaged ice thickness". Also, I and A are not explained too well. These equations should be explained better as you mention they will be used in the results section.

Rewrote this part.

L125-126: You reduce 0.29 m from ice thickness of IceAgeDerived when comparing to submarine derived ice thickness e.g. for the statistics in Tables 1 and 2, or which way?

Reduce 0.29 m in the submarine observations. This is added in the revised manuscript.

L127-130: Figures 2 and 3 in RK18 are thickness (Figure 2) and volume (Figure 3), so it would be appropriate to refer to "sea ice thickness and volume" in that order.

Done.

L132: Key, et al., -> Key et al.,   L160-161: Is the 10 km necessary to mention here?

To include the 25 km spatial resolution shows that APPx data spatial resolution is comparable to ice age at 12.5 km polar steoreographic projection, and 25 km resolution of Cryosat-2 data. So, we chose to keep this.

L167-168: You use age classes only up 4+ years, but Tschudi et al. 2016 (Fig. 5) have up to 5+. How did you choose this? Using the same classes would increase the consistency and comparability.

Tschudi et al. used 5+, which including 5 year old sea ice. We used >4 in this manuscript. They are the same.

L168-170: This method needs more reasoning.

We added "However, such information is not available for other months." We agree with the reviewer that such approach leads to uncertainty in the results. We discuss this in the "discussion" section, and propose future fix.

L200-203: Did I understand correctly that you go from weekly to daily to monthly. What is the benefit of doing the daily step?

This makes the monthly mean calculation easier, since the uneven distribution of weeks in a month. We also added such text in the discussion "even though the weekly ice age product is converted to weekly ice thickness and interpolated to daily ice thickness for monthly mean calculation. Such daily product lacks detailed temporal information content of ice thickness, and is not intended for direct comparison to point in situ ice thickness or other daily ice thickness products."

L244-246: Good that you mention this! How about ICESat? That too was used in the development, right?

Added "ICESat" in the text. We added the comparison to Cryosat-2 as independent validation/evaluation.

L269: How is the partial recovery after summer 2008 visible in these DRA mean ice thicknesses? Particularly in IceAgeDerived?

Older sea ice is generally thicker, and the ice age information from ice age product is utilized in the derivation of ice thickness. So, the partial recovery of multi year sea ice after summer is reflected in the mean ice thickness from IceAgeDerive product. Such discussion is added in the text, "This agreement can be attributed to that the sea ice age information in the ice age product, including intrinsic features of general decreasing and partial recovery of multiyear sea ice after 2008, are utilized to derive the ice thickness."

L295: Arctic sea ice volume for what? Is this still for IceAgeDerived? Maybe add more explanation in the figure caption.

It is still for IceAgeDerived. Revision made in the text and in the figure caption.

L315-328: Interesting analysis! See comment about Fig. 14.

L348-350: This bullet point does not seem as important as the others, as these findings have been shown in other studies. This could be more of a point to state the consistency between methods, IceAgeDerived succeeds in showing this phenomena that the other sea ice thickness products have captured, which would encourage the users to take on IceAgeDerived.

Agree. This point is to show the consistency between method, and the validility of this IceAgeDerived product.

L360-363: Extremely interesting! I missed the information for which area this was done.

It is over the Arctic Ocean. Added this information.

L366-367: Would love to see more analysis on this. Tschudi et al. (2016) seemed to have thicknesses increasing for each age category up to 5 years. It would be a nice addition to see some speculation about the causes.

We do not a good answer to this. We do have some speculations, and these are not included in the revised manuscript. We will do further investigation on this subject.

Ice ages differently, progressing through growth during freeze-up and decay during the melt seasons. Ice growth varies depending on initial thickness, as well as the air and ocean temperatures it is exposed to, and ice dynamics. The older ice gets, the more cycles of variable growth it has passed through. Older ice has been observed to be quite thick, up to 2-3m, in accumulation locations such as the Canadian Archipelago, but has also been observed to be rotten and fairly thin.

Submarines measure the ice freeboard from below with sonar, while space-based sensors such as ICESat are used to estimate thickness based on elevation differences between open water and the ice using snow depth estimates, which introduce the greatest level of uncertainty. It's possible that estimates of snow on the ice and/or localized ice deformation is responsible for the difference in thickness measurements between submarines and spaceborne altimetry in particular years.

L410-411: These references are not used (and maybe never will be)

Deleted. Those are from the journal template.

L475: Malanik -> Maslanik

Done.

L484, L487: Please add a and b for Tschudi 2019.

Done, and revised in the manuscript.

Table 1, and others: SCIEX -> SCICEX

Done.

Table 1, Table 2, Fig. 5, Fig. 6, and where relevant: Cryosat-2 -> CryoSat-2

Done.

Fig. 1 Consider a different latitudinal cut off, now there is quite a lot of uninformative area in the figures and especially it is hard to see the draft observations. Or if wishing to keep similar cut off to your other figures, consider emphasizing the draft points.

Changed the latitude cutoff and also emphasized the draft points.

Fig. 2 to1995 -> to 1995 (space). For consistency, consider having the same colorbar as in the other figures, e.g. Fig. 7 [0,4] instead of [0,~4.5]. Consider as well the choice of colormap if 3 m ice, which now stands out with yellow, does not need extra attention. Also, the unit is missing for sea ice thickness.

All suggested changes are made. Figure is replaced.

Fig. 4a v4.0? This maybe refers to the sea ice age product available at NSIDC, but I did not see the version mentioned elsewhere in the manuscript.

Version of the ice age product is added in the text.

Fig. 9 GORE box? And in general, there are a bit too many names for different areas (Arctic Ocean (as in RK18) , SCICEX box, GORE box, DRA). Use only one, unambiguous name for each area.

Corrected.

Fig. 14 I did not see this figure being referred to. If this is correct, please add it somewhere in L315-328.

Added.

Fig. A1 t0 -> to

Corrected.