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

Review for The Cryosphere Discuss., *https://doi.org/10.5194/tc-2019-192* ***Multidecadal Arctic sea ice thickness and volume derived from ice age***

Liu et al., 2019

This study generates a new product for estimated pan-Arctic sea ice thickness, spanning the full satellite era. Observations of sea ice age are used as a proxy for thickness, with the age-thickness relationship derived incrementally over different periods of the satellite data. A consistent sea ice thickness product covering four decades has considerable novelty, because state-of-the-art ice thickness products from satellite altimeters cover only small chunks of this record, with inter- satellite biases not yet properly reconciled.

Despite the attraction of such a new long-term sea ice thickness record, I have some concerns with the method used to derive ice thickness from age, particularly regarding the verification approach. The value of this record as a tool for further studies (e.g. for model assimilation, forecasting, understanding decadal Arctic climate/ocean trends) depends entirely on its success reproducing the well-validated altimetry observations; however, there is no evidence presented for this. The new product is also lacking robust estimates for random and systematic ice thickness uncertainties.

I have provided a set of general comments on the methodology and recommendations for improving the analysis. I've also made some minor suggestions to improve the readability of the paper and clarify a few confusing statements. I'd recommend this manuscript is reconsidered for publication in The Cryosphere following these major revisions. Please do get in contact if you have questions regarding these comments. Kind regards, Jack Landy

We appreciate the reviewer's critical evaluation and constructive suggestions. All of the reviewer's comments have been addressed, and reasons are given why some of the suggested work by the reviewer cannot be fully carried out at this time. All the responses are included in the revised manuscript. We believe revisions responding to reviewer's comments make the manuscript better. For that, we thank the reviewer.

General comments:

1. *In my view, the derived ice thickness and volume estimates should be described as 'proxies for ice thickness and volume' throughout the paper, as the method uses ice age observations which are a proxy – but not direct replacement for – sea ice thickness observations.*

   We agree with the reviewer that the derived ice thickness and volume are not direct observations or direct replacement for the observations. Thus, we emphasize this in the introduction, and discussion and conclusion section.

   In the "introduction", we added "These ice thickness and ice volume estimates are a proxy based on ice age, thus are not intended as a direct replacement for sea ice thickness observations", and "These ice thickness and volume estimates are a proxy from ice age products, thus they are not a direct replacement for sea ice thickness observations" in the "discussion and conclusion".

2. *This study desperately requires a detailed evaluation against available sea ice thickness observations from state-of-the-art altimeters, e.g. ICESat or CryoSat-2. The authors use ICESat data in their calibration of the ice age-thickness relationship, which essentially discounts their assessing the final product against ICESat data (but still do), and only compare to annual mean estimates of ice thickness and volume from CryoSat-2. Several gridded ice thickness datasets are available (from CPOM, AWI, NASA GSFC, LEGOS) which the authors could compare their derived product to. As they haven't used CS2 to calibrate their relationship, this would represent a valid independent assessment. As a suggestion, can the authors calculate the spread of CS2 ice thicknesses within each ice age category of the NSIDC product? I would recommend showing this as a plot. This would provide an estimate for the random uncertainty and potential bias in using age as a proxy for thickness. If one sigma of the PDF of CS2 thicknesses in an age category crosses another, it suggests ice age does not provide a valid proxy for thickness. Can the authors also provide maps of average November and march thickness for the coincident period of CS2 and IAD results?*

In the revised manuscript, we added the following analysis.

Monthly mean Cryosat2 ice thickness from CPOM, AWI, and NASA GSFC from January to April, and from October to December of 2011 to 2018 are used to calculate the spread of CS2 ice thickness within each ice age categories, and to evaluate the IceAgeDerived ice thickness and volume as the reviewer suggested,.

We collocated the NSIDC weekly ice age from 2011 to 2018 with correspondent Cryosat2 monthly ice thickness, and the spreads in March and November are as in the following figure.

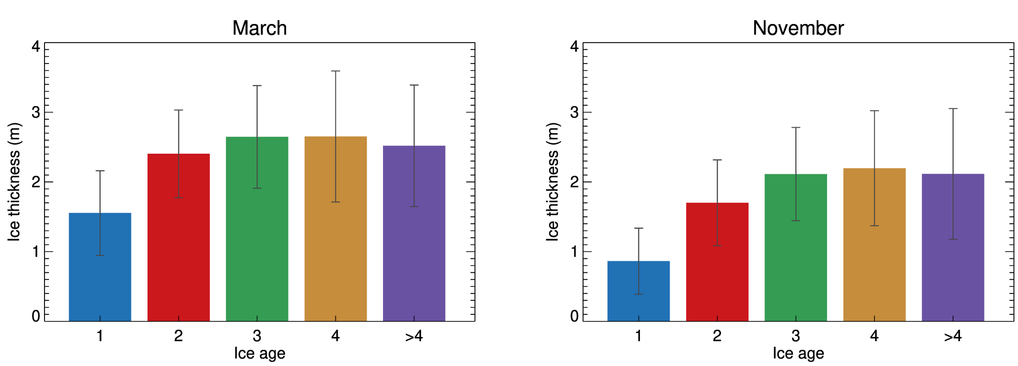


Figure 1: Ice age versus ice thickness from collocated ice age and AWI Cryosat2 ice

thickness. The error bar shows the one standard deviation of ice thickness in each ice age category.

Ice thickness increases with ice age for ice age from 1 to 4, and then decreases from ice age 4 to 5. This is consistent with what we found based on upward looking sonar data. This trend shows similarity and difference with what were found in Tschudi et al. (2016), where ice thickness increases from ice age from 1 to 5. Similar as those in Tschudi et al. (2016) (Figure 2), one sigma of the PDF of Cryosat2 thickness in an age category crosses another. We respect the Reviewer's view that "*. If one sigma of the PDF of CS2 thicknesses in an age category crosses another, it suggests ice age does not provide a valid proxy for thickness.",* and we think the estimation of ice thickness from ice age can still be valid, however, the crosses lead to uncertainty in the ice thickness estimation based on ice age.

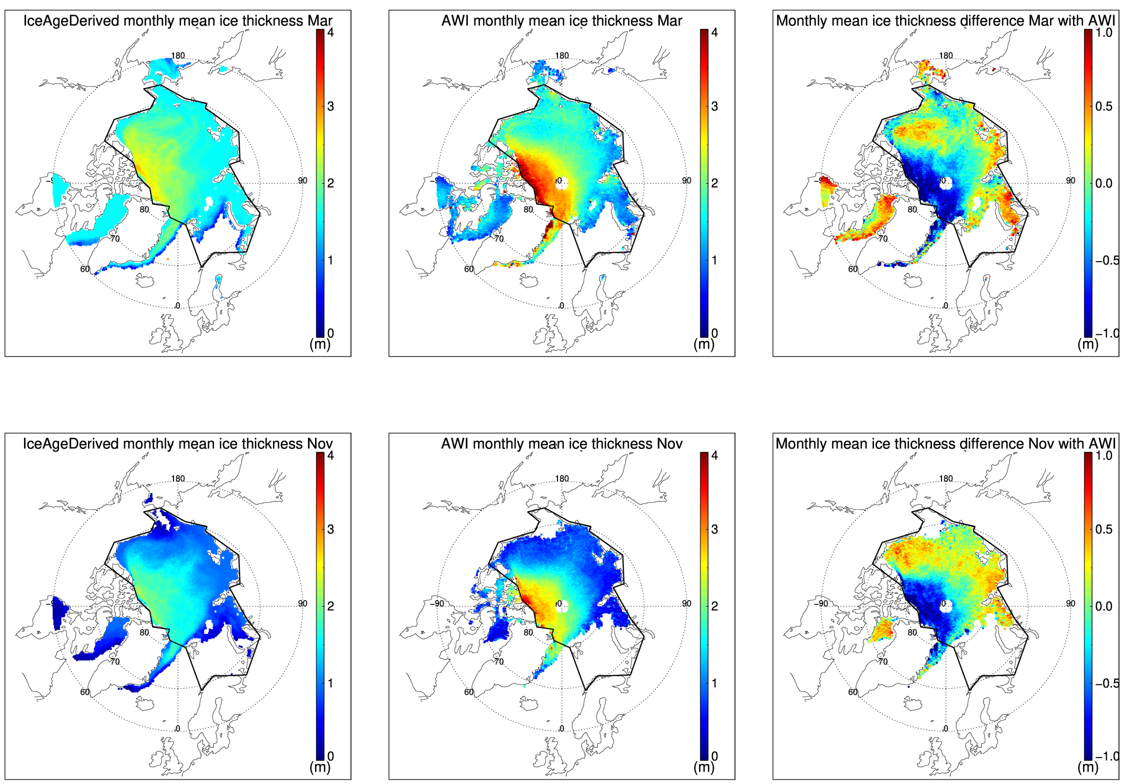


Figure 2: Monthly mean ice thickness from IceAgeDerived (top left), from AWI Cryosat-2 (top middle), and their difference in March 2011-2018; and monthly mean ice thickness from IceAgeDerived (bottom left), from AWI Cryosat-2 (bottom middle), and their difference in November 2011-2018.

Monthly mean ice thickness from IceAgeDrived and Cryosat-2 (AWI, NASA GSFC, and CPOM) shows similar spatial patterns in March and November. The sea ice thickness has the maximum values north of the Canadian Archipelago, and decreases radially toward the coastal regions of Alaska and Russia. The major differences are over the area north of the Canadian Archipelago, with the IceAgeDrived underestimating the thickness up to 1 m compared to Cryosat-2.

These analysis are added in the revised manuscript.

3. Validation approach. A majority of the comparisons made between ice age derived thickness and independent data (Line 222-227) are not truly independent, as the datasets were originally used to calibrate ice age-thickness relationships. If they are statistically dependent, i.e. data X is used to calibrate Y then Y is compared against X, it doesn't tell us much. Some evaluation of annual mean ice thickness/volume are made against truly independent CS2 observations, but this gives no evaluation of the spatial/regional accuracy. I would recommend including either a comprehensive assessment against CS2 (as described above) or to reserve a selection of the submarine/ICESat data only for assessing the final product, rather than calibrating with AND assessing it against the same thing. In its present form, I don't believe the validation has 'proven the soundness of the IAD thickness' as suggested on lines 337-339.

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 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 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 $\times 10^3$ 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.

Though the comparison to the Cryosat-2 ice products show overall agreement in both thickness and volume, further investigation and analysis shows that there are rather apparent differences in the ice thickness retrieval spatial distributions as shown in Figure 2. It appears the IceAgeDerived ice thickness underestimates the ice thickness for the older ice while overestimates the ice thickness for the first year ice with comparison to Cryosat-2. It should be also noted that Cryosat-2 also has relatively high uncertainties for very thin and very thick sea ice. In total, these underestimates and overestimates may balance off in the overall mean ice thickness and ice volume comparisons. These noted differences are surely a research topic for future studies.

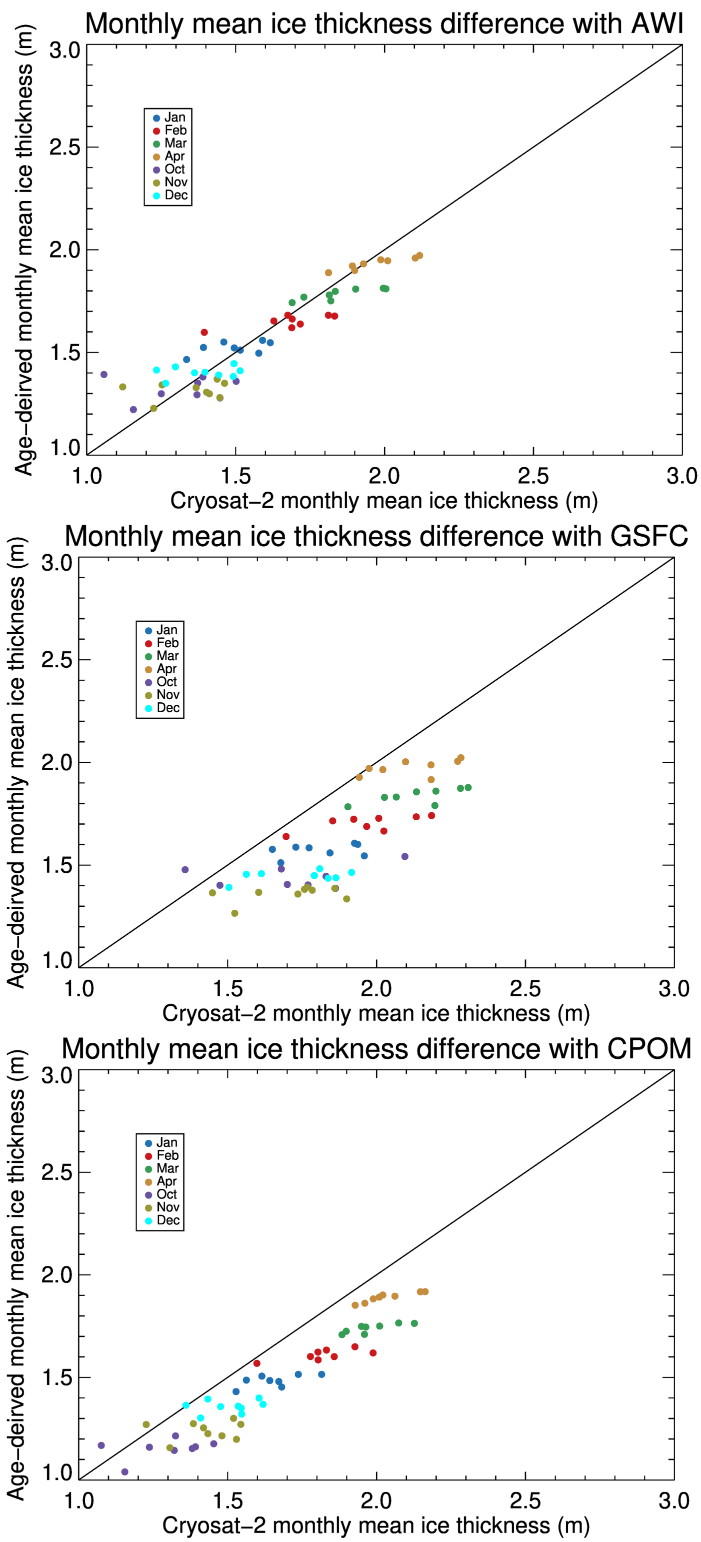


Figure 3: Scattering plot of IceAgeDerived monthly mean ice thickness and Cryosat-2 monthly mean ice thickness from AWI, NASA GSFC, and CPOM.

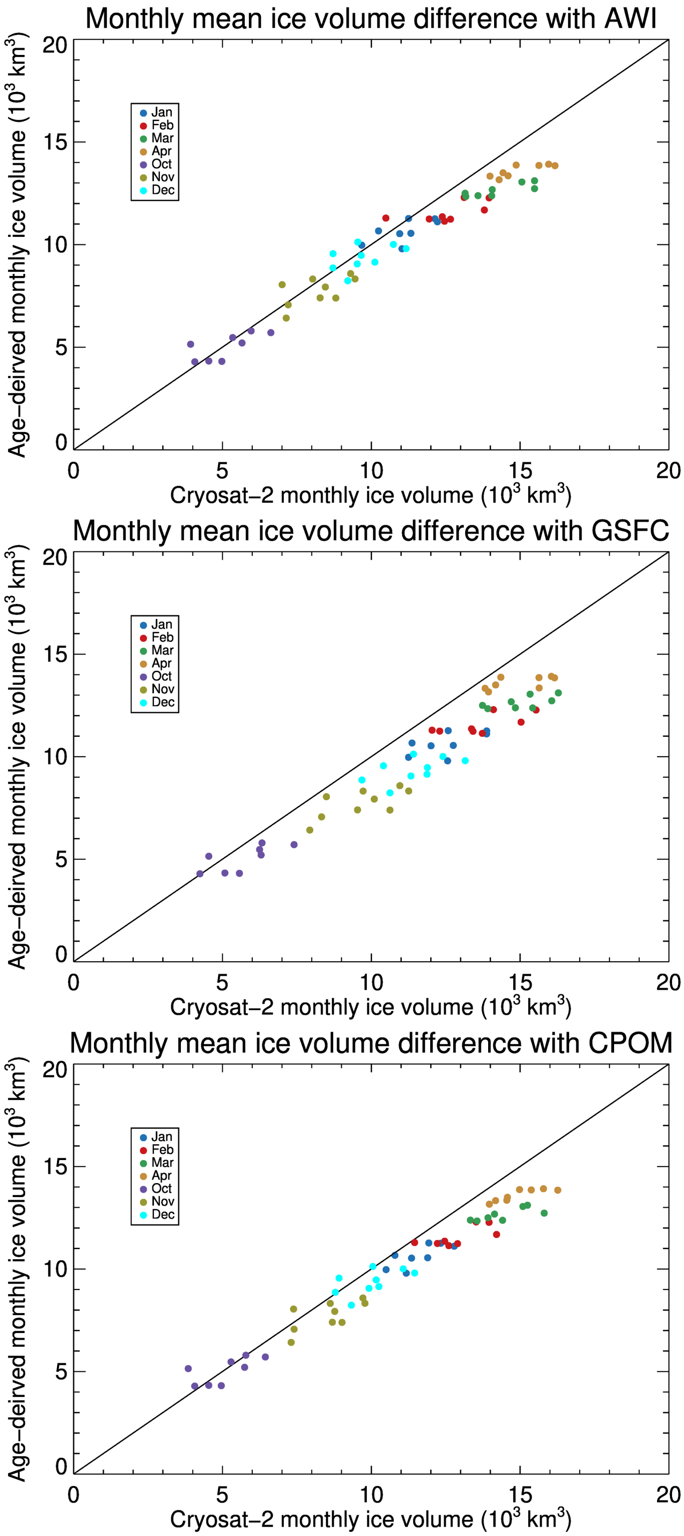


Figure 4: Scattering plot of IceAgeDerived monthly ice volume and Cryosat-2 monthly ice volume from AWI, NASA GSFC, and CPOM.

Table 1: Differences 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)

We also carried out similar evaluation/validation with Envisat from 2003-2010, and got similar results. Please refer to response to another reviewer's comments.

All these analysis and discussions are added in the revised manuscript.

4. Uncertainty. I was surprised to see no estimate of uncertainty for the derived sea ice thickness, particularly as this product is a proxy based on the imperfect relationship between ice age and thickness. The underlying sea ice age data have an uncertainty estimate. There are several empirical equations used in the methodology with derived coefficients that will have uncertainties. Several biases are corrected for and these will also have uncertainties, potentially varying over the annual cycle. A proper comparison with independent observations will additionally produce estimates for random and potential systematic uncertainties. I appreciate the added work required to produce robust uncertainty estimates, and for this proxy product they may be high, but for users to trust the new product they need some idea of its accuracy/precision. I expect the authors to make estimates for both the random uncertainty (errors in coefficients, errors in ice age product, noise in comparison to independent data) and systematic uncertainty (uncertainties in bias corrections, errors in extrapolating beyond your data collection period, potential biases compared to independent data) in a revised *version of the manuscript. These sources of uncertainty also need to be estimated for each month of the year separately, as one would expect the error to vary considerable across the seasonal cycle.*

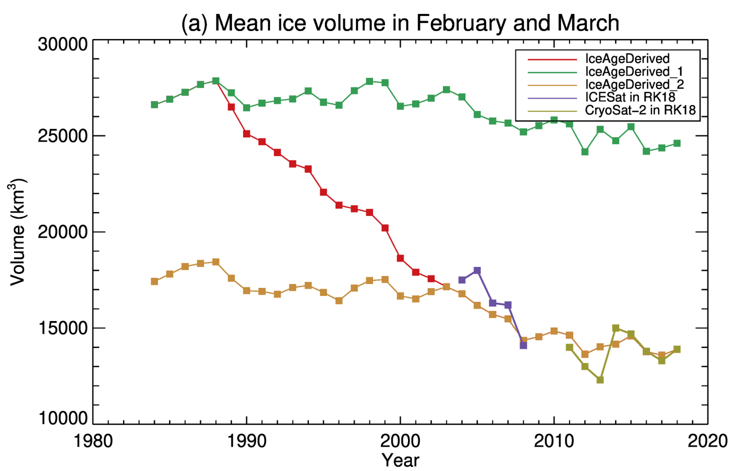
According to Tschudi et al. (2019) and discussion with Dr. Tschudi, there is no explicit uncertainty estimation in the sea ice age data. As shown in Figure 1, there is uncertainty, as the one standard deviation, corresponding to each ice age category, and these estimations are comparable to those shown in Figure 2 in Tschudi et al. (2016). To estimate the random uncertainty of the IceAgeDerived ice volume over the Arctic Ocean we applied the ice thickness uncertainty errors in each ice age category when converting the weekly ice age to ice thickness from 1984 to 2018. The uncertainty in weekly or monthly ice volume over the Arctic Ocean is the sum of the ice volume uncertainty of all grid cells, where the ice volume uncertainty in a cell is the product of the sea ice concentration, the grid cell area, and the ice thickness uncertainty. This provides the upper limit on the random uncertainty in ice volume. The overall uncertainties in ice thickness and ice volume in every month from 1984 to 2018 are derived. The average ratios of ice volume uncertainties to the mean range from 21% to 29% over the period 1984 - 2018.

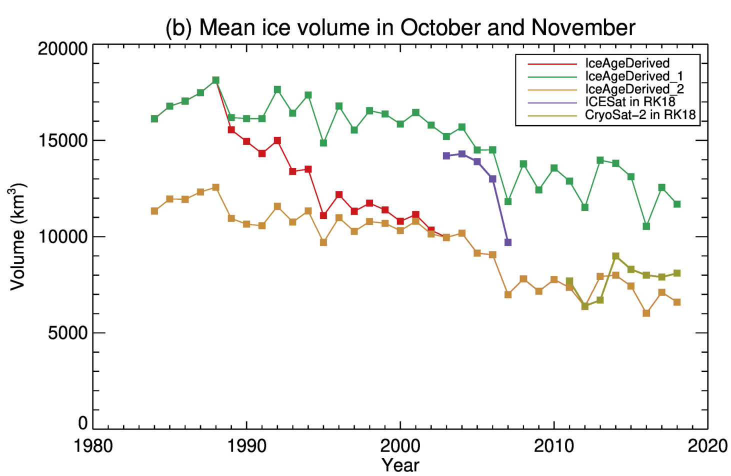
The systematic uncertainties of the IceAgeDerived ice thickness and ice volume are estimated by comparison to independent ice thickness and ice volume data from Cryosat-2, which is shown in the response to reviewer's major comment #3.

All these analysis and discussion are added in the revised manuscript.

5. The authors argue the decreasing trends in ice thickness and volume from their new product are consistent with observations of MYI replacement since the mid-2000s (Lines 268-270). However, the trend in their product is imposed by systematically changing the relationship between ice age and thickness throughout the time series (i.e. Fig 2). Comparing negative trends to the ice age product is basically fitting to and comparing against the same dataset. What physical explanation is there for the ice age-thickness relationship to change by such a considerable amount over these 5-yr segments of time? Can you provide citations to support this? Surely if the relationship changes by so much over time, it indicates ice age cannot be used alone as a proxy for thickness. Temporal/spatial sampling biases in the calibration data (especially the submarines) are very likely to have introduced systematic biases in these 5-yr relationships. What do the time series in Figs 9-12 look like if you use a fixed ice age- thickness relationship for the duration of the record? Unreasonable low?

Figure 5 shows the time series of mean ice volume using varying ice age-thickness relationships (as in the manuscript), using relationships in 1984, and in 2004-2008 (ICESat period) respectively. The overall trends are -411, -136, -156 km^3/year from 1984 to 2018 respectively. This indicates in our approach that the replacement of multi-year ice may only accounts for a smaller part of the overall trend (~33% or ~38%, -136/-411 or -156/-411), while the changes in ice age and ice thickness relationship contribute more the overall change. Since the ice age-thickness relationships change is small between the ICESat period and Cryosat-2 period (see both Figure 1 here and Figure 2 in Tschudi et al. (2016)), this ice age-thickness relationship changes may mainly happen between middle 1980s and middle 2000s, that ice thickness decreases in each corresponding ice age category. Sea ice extent in September has been decreasing, with trend from 1997 to 2014 four times as large as that from 1979 to 1996 (Serreze and Stroeve, 2015). More solar heating that the ocean absorbs through the open water area is expected to thin the remaining ice for all ice categories, leading to even less sea ice in the next summer and more solar heating. This may explain the decreasing ice thickness for corresponding ice age. However, it appears the accelerated decrease of ice thickness to corresponding ice age happens before the accelerated decreasing ice extent in September, which needs further investigation.





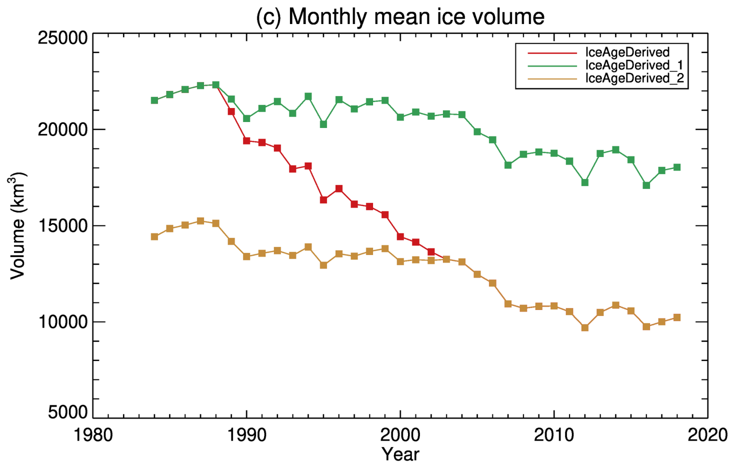


Figure 5: Monthly ice volume over the Arctic Ocean from 1984 to 2018 derived from ice age using varying age-thickness relationship (IceAgeDerived), using age-thickness relationship in 1984 (IceAgeDerived_1) and using age-thickness relationship in 2004-2008 (ICESat) (IceAgeDerived_2) in (a) February and March, (b) October and November, and (c) monthly mean of all months.

Minor comments/edits:

Line 18. Affecting what about the volume?

Added "the sea ice volume trend"

L23. 'declines'. Check spelling errors throughout.

Changed to "declined". Done.

L26. What anomalous ice export? Volume or area export? Needs citations to back up.

Revised and added a reference (Smedstrud et al. 2011)

L28-29. Unclear argument – why does this mean it is more sensitive?

The signal is more apparent with a higher change in percentage, I think.

L37-38. Sentence seems a bit out of place. Is this here just for the citation..?

No. It is a good place to introduce related data set and potential applications.

L48-49. Include the point that the sensor signal must first be sensitive to ice thickness, before modelling or statistical parameters can be used to estimate the thickness.

Added.

L55. Laxon citation is not relevant to this point.

Removed.

L59. Sea ice floes? Grid cells?

Added "parcels" after "sea ice".

L60. Can you comment on the uncertainty of this relationship? Was this reported in the Maslanik paper?

Added "The uncertainty of this relationship appears to increase from new ice to older ice, with values ranging from approximately 0.2 to 1.0 m (Figure 2 in Maslanik et al. 2007, and Figure 2 in Tschudi et al. 2016)."

L64. Not necessarily more robust, but more comprehensive definitely.

Changed to "comprehensive".

L91. Is Cavalieri 1996 the most up to date reference?

Definitely not. But that is the one NSIDC asks for to refer if that specific data set is used.

L93. POP model?

Not sure. That is beyond my knowledge.

L107. Are the ice draft data from submarines analysed entirely by yourselves or do you use statistics produced by others (NSIDC)? How do you do the processing? How do you account for unknown snow depth/density at the ice surface? What is the uncertainty on these estimates?

I used the processed data at NSIDC. Added "Assessment shows the ice thickness has a positive bias of 0.29 m, and the standard deviation if 0.25 m (Rothrock and Wensnahan 2007). The ice thickness from submarine data from 1984 to 2000 from NSIDC are used here"

Eq 1. Please provide explanation for this coefficient and estimate the uncertainty.

This is a equation that Rothrock et al. derived and showed in their 2008 paper. I noted this in the revised manuscript.

L114-115. Should f(tau) not depend on the ice type itself, i.e. accounting for different snow accumulation rates between seasonal and old ice?

Again, such information is available in Rothrock et al. (2008)

L118-124. This section is very confusing and requires a re-write. What exactly is I? What does 'interannual change with the annual cycle' mean? What are these equations used *for? This part seems like method, rather than data.*

I rewrote this section, and emphasized all the equations were derived and details are available in Rothrock et al. (2008).

L124. Determined how?

Added "Details about the bias determination is available in Rothrock and Wensnahan (2007)."

L125-126. You reduce the submarine ice thicknesses by this bias? Or do you reduce both submarine and your final ice-type derived thicknesses by this? Is this bias applicable for the entire seasonal cycle?

Changed to "In this study, we therefore reduce individual ice thickness observations by 0.29 m in all the original submarine observations."

L129. Be more specific about the processing chain used to derive the CS2 data. Is this the JPL product from Kwok and Cunningham, 2014?

To be honest, I tried to figure out what exactly product these values are based by reading Kwok's 2018 paper and his other papers. I could not figure it out. The best I can do is to cite his 2018 paper, stating that the values are from that paper. I can not confirm if that product is from Kwok and Cunningham 2014.

L143-145. The recent comparison paper by Sallila et al 2019 has shown very different results between OTIM and CS2 products. Is it worth comparing both to your independent ice-type product, when there show so much systematic uncertainty? Which are you use as your 'true' reference?

The differences are due to the different retrieval approaches for CS2 and OTIM. As mentioned in the paper by *Sallila et al 2019,* CS2 can only estimate ice thicker than ~0.5m, and OTIM can do for ice thickness between 0 ~ 6m. So both need to be calibrated and validated with in-situ direct measurements from such as submarines and stations for further improvements. So I would say not to take either of them as 'true' reference, just a 'product' reference

L149. Can you comment on the positive bias that may be introduced to the derived relationship from your calibrations against submarine data being focused in the central Arctic Ocean?

In the discussion and conclusion section, we have such discussion: "Third, in deriving the relation of ice age to ice thickness in the years before 2000, only ice draft measurements from submarine ULS over the DRA, e.g. over or near the central Arctic Ocean, are available. The derived relationship may be skewed to higher ice thicknesses. Thus, Arctic ice volume derived in this study before 2004 might be overestimated. Correcting this relationship requires more spatially representative ice thickness measurements, or a well-designed parameterization scheme."

L158-160. Confusing, please reword.

It is changed to "All matched ice thickness and age samples in a month within a 10-year moving window are used to derive the relationship of ice age and ice thickness in that month at the fifth of the ten years."

L169-170. This is a very speculative approach – picking bias corrections from a plot. Would you not expect this relationship to be different between fall and soring, as thinner ice grows more rapidly over winter?

We changed to "However, information of such relationship is not available for other months. According to Figure 2 in RK18, the mean ice thickness in October and November is approximately 0.7 m less than the mean in February and March. Therefore, in October we assign the relationship of ice age and ice thickness the same as that in March except that ice thickness in each age category is 0.70 m less." Also added in the discussion and conclusion section that "The ice age-thickness relationship is not available for months other than in March, and we assumed such relationship is the same in October with ice thickness of 0.7 m less. With CryoSat-2 ice thickness available from October to April, we can derive such relationship in other months, and assess the linear ice thickness growth/decline assumption we made."

L174-175. What is your physical explanation for this?

I do not have a clear physical explanation for this. This can be a research topic for future studies. However, we have some speculations, and they 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.

L177-178. 'keeping the relationship for ice older than four years', what do you mean by this? Extrapolating the thickness for very old ice?

Changed to "As in Tschudi et al. (2016), we use linear regression to derive the relationship between ice age and thickness for ice ages from one to four years, while the relationship for ice older than four years remain unchanged."

L185. Flux of what?

Added "energy flux"

L198-99. Is this realistic? There are so many simplifications and assumptions here that the final result will barely reflect the underlying data.

Since such relationship are not available from the years between, that is all can do. Once observations over those years become available, we will be more than happy to derive those relationship using those observations. Meanwhile, we have to use some assumptions and simple approaches.

L202-3. Weekly to daily to monthly thickness. Why?

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."

L215. You need to explain this above with Eq 2. L232-234. Links to comment 2 above.

Please see the response to comments regarding to Eq1 and 2 before this .

L254-255. PIOMAS is almost being treated as the true reference here. I would urge the authors to consider comparing climatological thickness from CS2 (2010-2019) to the same years of their IAD record.

In response to your major comments, we have carried out evaluation/validation with CS2 from 2011 to 2018. Also, we carried out evaluation/validation with Envisat from 2003 to 2010.

L261. Also the imposed seasonal cycle, with highest ice thickness in May. PIOMAS is highest in April.

This is based on the surface energy annual cycle. This also shows we do not tune our product based on PIOMAS. We generate our product independently, and compare our product with PIOMAS.

L296-298. It looks like the largest decadal volume drop occurred between the 80s and 90s. Does this make sense with respect to the literature? Can you provide citations to support this? Would we not expect largest volume losses in the most recent decades, when concentration has declined strongest? Could this finding perhaps come from the trend in ice age-thickness relationship that you impose yourselves?

All the ice-thickness relationship is based on data. We speculate that the ice thickness decrease may start to accelerate before the ice extent decrease starts to accelerate. This may be a reserch topic that needs further investigation.

L328. Confusing. Please explain in more detail.

Rewrote to "It should be noted that the sum of these two contributions is not 100% because the production of area means of thickness and ice area is only approximately equal to the total ice volume as shown in Eq.7."

L360-363. Although this is simplified, it is a reasonable analysis and I would be interested to see these contributions per Arctic region as well as in total.

I would think there are regional differences because differences in thickness and concentration spatial differences.This would be interesting to seen in future studies.

L373-374. You need to consider and suggest an explanation for this.

I hope I have a simple answer, but I do not. Without detailed and further analysis, I would speculate that this might be related to the linear ice growth/melting model we applied. But how exactly they are related, I am not sure. Added "The annual cycle of trends in ice volume over the Arctic Ocean appears to be opposite to the annual cycle of ice growth, which suggests this trend feature may be related to linear sea ice growth/melting model applied. How they are related and whether a more sophisticated model would remove this feature require further investigation."

L375. Good point.

L378. Have you considered there may be a fundamental limit in the accuracy of ice thickness estimation for which ice age acts as a proxy? Checking the PDFs of CS2 ice thickness within each ice age category for the same month would be a perfect way to evaluate this limit, i.e. the intrinsic uncertainty of the ice age-thickness relationship.

I will consider doing this in the future.

L385-6. Could you have tried evaluating against the entire icebridge thickness archive or for example airborne EMI thickness datasets?

The icebridge thickness observations can be used to derive the ice age-thickness relationship as shown in Tschudi et al. (2016). For the evaluation, IceAgeDerived product assign one single thickness value for sea ice of the same age category, thus lacks spatial changes. For that reason, I do not think the IceAgeDerived ice thickness is suitable for point comparisons. Added " 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 and spatial information content of ice thickness on the daily scale, and is not intended for direct comparison to point in situ ice thickness or other daily ice thickness products, such as CryoSat-2."

Fig 3. Could you not base the shape of this approximation on e.g. the mean seasonal cycle of ice thickness from CryoSat-2 data?

I think you meant "Could you base…?"
Yes. That will be the next step. In the discussion, added "With CryoSat-2 ice thickness available from October to April, we can derive ice age-thickness relationship in all these months, and assess the linear ice thickness growth/decline assumption we made."

Fig 5. Panels a-d are comparing the derived product against in situ observations used to calibrate them. There is evidently much higher scatter versus the CS2 data, that were not used in the calibration. Add r^2, rmse and bias to these plots.

We added r^2, rmse and bias in table 1 and 2.

Fig 6 caption. Volume.

Corrected.

Fig 12 caption. Annual mean ice volume?

Yes, it is monthly ice volume. Corrected.

Fig 13. A great deal of this pattern reflects the annual cycle that was imposed from Fig 3. Can you comment on this?

As the response to the reviewer's previous comment, I agree on this assessment and we noted this in the manuscript. We added the following text: "The annual cycle of trends in ice volume over the Arctic Ocean appears to be opposite to the annual cycle of ice growth, which suggests this trend feature may be related to linear sea ice growth/melting model applied. How they are related and whether a more sophisticated model would remove this feature require further investigation."